

UNIVERSIDADE ANHEMBI MORUMBI

CAMILA RIBEIRO DANIEL

**IDENTIFICAÇÃO DE ALTERAÇÕES BIOQUÍMICAS EM AMOSTRAS DE
COÁGULOS SANGUÍNEOS DE PACIENTES ONCOLÓGICOS
PEDIÁTRICOS ATRAVÉS DA ESPECTROSCOPIA RAMAN**

TESE DE DOUTORADO

**DOUTORADO EM ENGENHARIA BIOMÉDICA
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU**

São José dos Campos, Outubro/2023

UNIVERSIDADE ANHEMBI MORUMBI

CAMILA RIBEIRO DANIEL

**IDENTIFICAÇÃO DE ALTERAÇÕES BIOQUÍMICAS EM AMOSTRAS DE
COÁGULOS SANGUÍNEOS DE PACIENTES ONCOLÓGICOS
PEDIÁTRICOS ATRAVÉS DA ESPECTROSCOPIA RAMAN**

TESE DE DOUTORADO

Tese apresentada ao Programa de Pós-Graduação Stricto Sensu em Engenharia Biomédica – Doutorado, da Universidade Anhembi Morumbi, como requisito parcial para obtenção do título de Doutor em Engenharia Biomédica.

Orientador: Prof. Dr. Marcos Tadeu Tavares Pacheco

Coorientador: Prof. Dr. Landulfo Silveira Jr.

São José dos Campos, Outubro/2023

UNIVERSIDADE ANHEMBI MORUMBI

CAMILA RIBEIRO DANIEL**IDENTIFICAÇÃO DE ALTERAÇÕES BIOQUÍMICAS EM AMOSTRAS DE
COÁGULOS SANGUÍNEOS DE PACIENTES ONCOLÓGICOS
PEDIÁTRICOS ATRAVÉS DA ESPECTROSCOPIA RAMAN**

Tese apresentada ao Programa de Pós-Graduação Stricto Sensu em Engenharia Biomédica – Doutorado, da Universidade Anhembi Morumbi, como requisito parcial para obtenção do título de Doutor em Engenharia Biomédica. Aprovada pela seguinte Banca Examinadora:

Prof. Dr. Marcos Tadeu Tavares Pacheco

Orientador
Doutorado em Engenharia Biomédica
Universidade Anhembi Morumbi

Prof. Dr. Pedro Luiz de Brito (externo)
Hospital Municipal Infantil Menino Jesus (IRSSL)

Prof. Dr. Adriano Moraes da Silva (externo)
Abbott Laboratórios do Brasil (ABBOTT)

Prof. Dr. Landulfo Silveira Jr.
Universidade Anhembi Morumbi

Prof. Dr. Leandro Procópio Alves
Universidade Anhembi Morumbi

Prof. Dra. Adriana Barrinha Fernandes Moretti(Suplente)
Universidade Anhembi Morumbi

Prof. Dra. Luciana Lopes Guimarães (Suplente)
Universidade Santa Cecília

São José dos Campos, Outubro/2023

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da Universidade, do autor e do orientador.

CAMILA RIBEIRO DANIEL

Fisioterapeuta formada pela Universidade Estadual do Piauí (2013), especialista em fisioterapia Dermatofuncional (ABRAFIDEF). Mestre em Bioengenharia pela Universidade Brasil (2017).

Ficha Bibliográfica elaborada pela biblioteca UAM
Com os dados fornecidos pelo(a) autor(a)

D185i Daniel, Camila Ribeiro
Identificação de alterações bioquímicas em amostras de coágulos sanguíneos de pacientes oncológicos pediátricos através da espectroscopia Raman - Camila Ribeiro Daniel, 2023.
106f.: 30 cm.

Orientador: Marcos Tadeu Tavares Pacheco.
Tese (Doutorado em Engenharia Biomédica) - Universidade Anhembi Morumbi, São José dos Campos, 2023.
Bibliografia: f. 86-99.

1. Engenharia Biomédica. 2. Células Sanguíneas. 3. Câncer.
4. Diagnóstico. 5. Espectroscopia Raman. 6. Análise Discriminante.
I. Título.

CDD 610.28

DEDICATÓRIA

Dedico essa conquista à minha filha
Olívia.

AGRADECIMENTOS

Agradeço primeiramente à Deus, pois Ele nos deu a vida e é o responsável por guiar nossos passos, sem a presença Dele nada disso seria possível.

À minha mãe Fernanda Daniel, por servir de exemplo e todo apoio, incentivo e amor, sempre acreditando em mim e sendo suporte nos meus momentos de ausência em casa. Aos meus irmãos, Daniel Neto e Guilherme que estão sempre presentes na minha vida compartilhando bons momentos.

Ao meu marido Cosme Filho e minha filha Olívia, pela compreensão nos momentos de ausência e por todo o carinho de sempre.

Agradeço a minha amiga Ana Mara, pelo companheirismo de sempre, obrigada também pela troca de aprendizado e experiências, se esse sonho está se concretizando agora, sem dúvidas, é porque caminhamos juntas!

Agradeço ao Prof. Dr. Landulfo Silveira Jr., pelos ensinamentos e orientações, permitindo que eu realizasse mais uma importante meta na minha vida profissional, contribuindo para meu crescimento como pessoa e profissional. Agradeço também ao Prof. Dr. Marcos Tadeu pelas orientações.

O Prof. Dr. Landulfo Silveira Jr. agradece à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pela aprovação do Processo 2009/01788-5, que permitiu a realização desta pesquisa.

Agradeço à CAPES, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 - Bolsa PAIS, com número do processo 88882.367255/2019-01.

“É muito melhor lançar-se em busca de conquistas grandiosas, mesmo expondo-se ao fracasso, do que alinhar-se com os pobres de espírito, que nem gozam muito nem sofrem muito, porque vivem numa penumbra cinzenta, onde não conhecem nem vitória, nem derrota.”

(Theodore Roosevelt)

RESUMO

O câncer infantil corresponde a um conjunto de doenças caracterizadas pela proliferação descontrolada de células anormais, sendo de natureza predominantemente embrionária. O diagnóstico é feito com exame clínico, exames laboratoriais (obtidos a partir de amostras de sangue periférico) e, para confirmação, é solicitada biópsia do tecido acometido. Este estudo objetivou utilizar a técnica de espectroscopia Raman para analisar amostras de células sanguíneas de pacientes pediátricos com diagnóstico de neoplasias *versus* pacientes não neoplásicos utilizando técnicas multivariadas para a interpretação das diferenças espectrais nos grupos normal e neoplásico como a análise de componente principal (PCA) e análise discriminante por mínimos quadrados parciais (PLS-DA). Para o estudo foram coletadas 121 amostras de sangue de crianças e adolescentes saudáveis, com leucemia e com outros tipos de câncer. Um volume de 100 μL de células sanguíneas foi colocado em um porta-amostra de alumínio e submetido à espectroscopia Raman dispersiva (espectrômetro com excitação de 830 nm e potência de 250 mW acoplado a uma sonda Raman), com tempo de exposição de 30 s. Um total de 308 espectros Raman foram analisados e indicaram diferenças significativas nas bandas com picos de ácidos nucleicos (671, 756 cm^{-1}), aminoácidos (671, 1215, 1544 cm^{-1}), proteínas (756, 1430, 1531 cm^{-1}), β -caroteno (1531 cm^{-1}), amida II (1215, 1548 cm^{-1}) e citocromo c (756 cm^{-1}), apresentando maior intensidade nas amostras normais. A análise de variância (ANOVA) foi aplicada para analisar as variáveis da PCA (componentes principais – PCs) demonstrando que existe uma maior concentração de citocromo c nas amostras normais, também pode-se observar maior concentração de proteínas e fenilalanina (1004 cm^{-1}) no grupo leucêmico com diferenças estatisticamente significativas ($p < 0,05$). A análise discriminante realizada pelo modelo PLS-DA entre os grupos normal, leucêmico e não-leucêmico alcançou sensibilidade de 100%, especificidade de 87,8% e acurácia global de 67,5%, demonstrando a eficácia da espectroscopia Raman para diferenciar amostras saudáveis de amostras neoplásicas. Os resultados demonstraram a aplicabilidade da espectroscopia Raman para a identificação de compostos bioquímicos em células sanguíneas, revelando as diferenças espectrais mais relevantes entre os grupos saudáveis, com leucemia e com outros tipos de câncer, e permitindo a discriminação entre o grupo normal e os grupos com câncer; no entanto, apresentou uma baixa acurácia devido aos erros de classificação quando se buscou a discriminação entre os grupos com câncer (leucêmico e não-leucêmico).

Palavras-chave: Células sanguíneas. Câncer. Diagnóstico. Espectroscopia Raman. Análise discriminante.

IDENTIFICATION OF BIOCHEMICAL CHANGES IN BLOOD CLOTS SAMPLES FROM PEDIATRIC ONCOLOGY PATIENTS USING RAMAN SPECTROSCOPY

ABSTRACT

Childhood cancer corresponds to a set of diseases characterized by the uncontrolled proliferation of abnormal cells, being predominantly embryonic in nature. The diagnosis is made by clinical examination, laboratory tests (obtained from peripheral blood samples) and, for confirmation, a biopsy of the affected tissue is requested. This study aimed to use the Raman spectroscopy technique to analyze blood cell samples from pediatric patients diagnosed with neoplasms *versus* non-neoplastic patients using multivariate techniques for the interpretation of spectral differences in normal and neoplastic groups such as principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). For the study, 121 blood samples were collected from healthy children and adolescents with leukemia and other types of cancer. A volume of 100 μL of blood cells was placed in an aluminum sample holder and submitted to dispersive Raman spectroscopy (spectrometer with excitation of 830 nm and power of 250 mW coupled to a Raman probe), with an exposure time of 30 s. A total of 308 Raman spectra were analyzed and indicated significant differences in the bands with peaks of nucleic acids (671, 756 cm^{-1}), amino acids (671, 1215, 1544 cm^{-1}), proteins (756, 1430, 1531 cm^{-1}), β -carotene (1531 cm^{-1}), amide II (1215, 1548 cm^{-1}) and cytochrome c (756 cm^{-1}), showing higher intensity in normal samples. Analysis of variance (ANOVA) was applied to analyze the variables of PCA (principal components – PCs) demonstrating that there was a higher concentration of cytochrome c in normal samples, it could also be observed a higher concentration of proteins and phenylalanine (1004 cm^{-1}) in the leukemic group with statistically significant differences ($p < 0.05$). The discriminant analysis performed by the PLS-DA model between the normal, leukemic and non-leukemic groups reached 100% sensitivity, 87.8% specificity and 67.5% global accuracy, demonstrating the efficiency of Raman spectroscopy to differentiate healthy samples from neoplastic samples. The results demonstrated the applicability of Raman spectroscopy for the identification of biochemical compounds in blood cells, revealing the most relevant spectral differences between healthy groups, with leukemia and with other types of cancer, allowing the discrimination between the normal group and the groups with cancer; however, it presented a low accuracy due to classification errors when seeking to discriminate between groups with cancer (leukemic and non-leukemic).

Keywords: Red blood cells. Cancer. Diagnosis. Raman spectroscopy. Discriminant analysis.

SUMÁRIO

| | |
|--|----|
| 1. INTRODUÇÃO | 15 |
| 1.1. Problema | 19 |
| 1.2. Hipótese | 19 |
| 1.3. Objetivo geral | 20 |
| 1.4. Objetivos específicos | 20 |
| 2. REVISÃO BIBLIOGRÁFICA | 21 |
| 2.1. Câncer | 21 |
| 2.2. Cânceres pediátricos | 24 |
| 2.3. Câncer leucêmico | 25 |
| 2.3.1. Medula óssea e hematopoiese | 25 |
| 2.3.2. Leucemias | 27 |
| 2.3.2.1. Leucemia linfóide aguda (LLA) | 29 |
| 2.3.2.2. Leucemia linfóide crônica (LLC) | 29 |
| 2.3.2.3. Leucemia mieloide aguda (LMA) | 30 |
| 2.3.2.4. Leucemia mieloide crônica (LMC) | 30 |
| 2.4. Diagnóstico padrão das leucemias | 31 |
| 2.5. Cânceres não leucêmicos | 32 |
| 2.5.1. Tumores do sistema nervoso central | 32 |
| 2.5.2. Linfomas | 33 |
| 2.5.2.1. Linfoma não Hodgkin | 33 |
| 2.5.2.2. Linfoma Hodgkin | 34 |
| 2.5.3. Osteossarcoma | 35 |
| 2.6. Espectroscopia Raman | 35 |
| 2.6.1. Efeito Raman | 36 |
| 2.6.2. Diagnóstico de câncer não leucêmico utilizando espectroscopia Raman | 37 |
| 2.6.3. Diagnóstico de leucemias utilizando espectroscopia Raman | 39 |
| 2.7. Análise estatística | 40 |
| 2.7.1. Teste <i>t</i> de Student | 40 |
| 2.7.2. Análise de variância (ANOVA) | 42 |
| 2.7.3. Análise de componente principal (PCA) | 43 |
| 2.7.4. Análise discriminante utilizando regressão por mínimos quadrados parciais (PLS-DA) | 44 |
| 3. MATERIAIS E MÉTODOS | 47 |
| 3.1. Comitê de Ética | 47 |
| 3.2. Seleção dos pacientes e obtenção das amostras | 47 |
| 3.3. Espectrômetro Raman e coleta dos espectros | 49 |
| 3.4. Coleta dos espectros Raman | 50 |
| 3.5. Identificação dos picos e aplicação do teste ANOVA | 52 |
| 3.6. Análise exploratória por análise de componente principal (PCA) | 52 |
| 3.7. Análise discriminante por mínimos quadrados parciais (PLS-DA) | 53 |
| 4. RESULTADOS | 55 |
| 4.1. Espectroscopia Raman | 55 |
| 4.2. Análise exploratória por PCA | 65 |
| 4.3. Análise discriminante por PLS-DA | 70 |
| 5 DISCUSSÃO | 73 |
| 5.1. Espectroscopia Raman | 73 |

| | |
|--|----|
| 5.2. Análise exploratória por PCA | 75 |
| 5.3. Análise discriminante de mínimos quadrados parciais (PLS-DA)..... | 79 |
| 6. CONCLUSÃO..... | 84 |
| REFERÊNCIAS BIBLIOGRÁFICAS | 86 |

LISTA DE ILUSTRAÇÕES

| | |
|---|----|
| Figura 1: Distribuição proporcional dos tipos de câncer mais incidentes estimados para 2020 por sexo, exceto câncer de pele não melanoma. | 22 |
| Figura 2: Estágios do desenvolvimento do câncer. | 23 |
| Figura 3: Hematopoiese em humanos. | 27 |
| Figura 4: Diagrama de energia demonstrando o efeito Raman: excitação UV, visível e infravermelho próximo. S0, estado eletrônico basal, S1 e S2, estados eletrônicos excitados. | 37 |
| Figura 5: Histograma das fases do tratamento em que os pacientes se encontram. | 48 |
| Figura 6: Histograma da faixa etária dos pacientes oncológicos. | 49 |
| Figura 7: Curvas de resposta do espectrômetro Raman: (a) intensidade obtida com lâmpada de calibração de tungstênio; (b) número de onda obtida com o composto orgânico naftaleno. | 50 |
| Figura 8: Espectros Raman médios normalizados de células sanguíneas de grupos normais, leucêmicos e não leucêmicos. | 55 |
| Figura 9: Espectros das diferenças entre os espectros médios normalizados dos grupos Leucêmico – Normal e dos grupos Não leucêmico – Normal. | 65 |
| Figura 10: Plotagem das variáveis PCA (PCs e Scores) para amostras de células sanguíneas para análise exploratória. NS = não significativo. | 69 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1: Número de amostras de cada grupo histopatológico e número de espectros Raman considerados em cada grupo. | 52 |
| Tabela 2: Posições dos picos nos espectros Raman das células sanguíneas e suas respectivas atribuições de bandas vibracionais e elementos bioquímicos de acordo com a literatura citada. Os picos entre parênteses são os que aparecem na análise exploratória por PCA (Figura 10). | 55 |
| Tabela 3: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em três classes, pacientes normais (saudáveis), pacientes com leucemia e pacientes com outros tipos de câncer (não leucêmico). | 70 |
| Tabela 4: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em duas classes, pacientes normais (saudáveis) e pacientes com câncer. | 71 |
| Tabela 5: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em cinco classes, pacientes normais (saudáveis), pacientes com leucemia e pacientes com sarcoma, blastoma e linfoma + carcinoma. | 71 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|----------|--|
| COVID-19 | doença do coronavírus 2019 |
| INCA | Instituto Nacional de Câncer |
| LDA | análise discriminante linear |
| LDBCL | linfoma difuso de grandes células b |
| LLA | leucemia linfoide aguda |
| LLA-B | leucemia linfoide aguda – doença precursora de células B |
| LLA-T | leucemia linfoide aguda – doença precursora de células T |
| LLC | leucemia linfocítica crônica |
| LMA | leucemia mieloide aguda |
| LMC | leucemia mieloide crônica |
| MO | medula óssea |
| OMS | Organização Mundial de Saúde |
| PLS-DA | análise discriminante por mínimos quadrados parciais |
| PCA | análise de componente principal |
| PC | componente principal |
| SNC | sistema nervoso central |
| SD | desvio padrão |
| SVM | máquina de vetor de suporte |

1. INTRODUÇÃO

O câncer infantojuvenil corresponde a um conjunto de doenças que têm em comum a proliferação descontrolada de células anormais. Ao contrário do câncer adulto, o câncer infantojuvenil é de natureza predominantemente embrionária e geralmente afeta células do sistema sanguíneo e tecidos de suporte. A sua classificação é baseada na morfologia do tumor e não no local de origem (INCA, 2020a). A leucemia é um tipo de doença neoplásica que atinge as células do sangue, tendo como principal característica o acúmulo de células doentes na medula óssea que acabam por substituir as células sanguíneas normais (INCA, 2020a; MITCHELL e STEIDL, 2019). É o tipo de câncer mais comum em crianças menores de 15 anos, mas também ocorre mais frequentemente em adultos com mais de 55 anos. Sendo sua prevalência estimada em 28% dos cânceres pediátricos (BHAKTA et al., 2019; INCA, 2020a; MITCHELL e STEIDL, 2019). Para o ano de 2020, a estimativa mundial mostrou que ocorreram aproximadamente 475 mil casos novos de leucemia (SUNG et al., 2021).

Os quatro principais tipos de leucemia dentre os 12 tipos existentes são: leucemia mieloide aguda (LMA), leucemia mieloide crônica (LMC), leucemia linfocítica aguda (LLA) e leucemia linfocítica crônica (LLC) (INCA, 2020a). As leucemias agudas (LA) correspondem a mais de 95% de todas as leucemias infantis, incluindo LLAs, que são o tipo mais frequente (> 75%), e LMAs que correspondem (16%) na faixa etária ≤ 1 a 14 anos. A LLA precursora de células B é muito mais comum que as células T, sendo responsável por aproximadamente 80% dos casos de LLA (KAPLAN, 2019; MELARAGNO e CAMARGO, 2013; VAN MAELE-FABRY et al., 2019).

O diagnóstico de leucemias agudas normalmente se inicia a partir de suspeitas clínicas, onde o paciente é encaminhado para realização de exames laboratoriais através da coleta do sangue periférico e, se for confirmado alterações importantes nos exames é realizada a punção da medula óssea (MO) para confirmação e diagnóstico diferencial. Na maioria das vezes, a punção da MO é necessária para diagnosticar leucemia, mas o diagnóstico

também pode ser feito a partir sangue periférico, sendo útil nos casos em que a contagem de leucócitos é alta, permitindo um diagnóstico mais rápido, o que é muito vantajoso. Tanto a análise morfológica como a citometria de fluxo são usadas para avaliar a MO. A análise morfológica avalia a medula e blastos leucêmicos ao microscópio, enquanto os diferentes marcadores imunofenotípicos podem ser encontrados através da citometria de fluxo. A avaliação de todos esses parâmetros se torna necessária para diferenciação e definição da conduta terapêutica a ser abordada (KAPLAN, 2019; MELARAGNO e CAMARGO, 2013).

A espectroscopia Raman tem sido amplamente pesquisada e aplicada na literatura para identificar, classificar e até quantificar biomarcadores que podem fornecer informações sobre os vários tipos de câncer. Pesquisadores têm investigado amostras de soro sanguíneo utilizando microespectroscopia Raman (excitação em 830 nm) e métodos estatísticos multivariados para comparar amostras doentes e controles saudáveis, no entanto foram observados pequenos deslocamentos de pico e pequenas diferenças de intensidades de pico, de forma que, não mostrou uma discriminação clara entre pacientes doentes e saudáveis (GONZÁLEZ-SOLÍS et al., 2014; GONZÁLEZ-SOLÍS, 2019; MARTÍNEZ-ESPINOSA et al., 2009). Outros estudos também objetivaram diferenciar células cancerosas de células normais através da análise da saliva, da quantificação dos níveis de água celular, da análise da pele *in vivo* e *ex vivo*, obtendo resultados promissores (ATKINS et al., 2017; BARROSO et al., 2015; CHRISTIAN et al., 2014; FENG et al., 2015; LIMA et al., 2019; RASHID et al., 2014; SILVA et al., 2018). Ressalta-se ainda a possibilidade de utilização da espectroscopia como ferramenta transoperatória, podendo ser útil na determinação de quão radical deve ser a ressecção de determinado tumor ainda no ato cirúrgico (determinação das margens), podendo com isso melhorar o prognóstico do paciente (BARROSO et al., 2015).

A espectroscopia Raman é uma técnica óptica baseada no espalhamento inelástico da luz incidente por moléculas, sendo capaz de fornecer informações bioquímicas sobre os tecidos, gerando impressões

digitais químicas de células, tecidos e fluidos biológicos que podem ser utilizadas para diagnóstico médico (HANLON et al. 2000). Devido a sua alta especificidade bioquímica e pouca ou nenhuma necessidade de preparo de amostras, a espectroscopia Raman ganhou maior visibilidade nos estudos diagnósticos da patologia humana (HANLON et al. 2000). A espectroscopia Raman é capaz de gerar informações detalhadas sobre as alterações bioquímicas das moléculas que ocorrem com a progressão das doenças como uma opção promissora para o diagnóstico de várias patologias usando excitação no infravermelho próximo (por exemplo, 830 nm), com vantagem de reduzir a autofluorescência que ocorre em amostras biológicas (HANLON et al., 2000; KONG et al., 2015; SILVEIRA et al., 2012; SILVEIRA et al., 2015). A espectroscopia Raman tem sido usada para identificar, classificar e quantificar biomarcadores que podem fornecer informações sobre diferentes tipos de câncer (PACHECO et al. 2013), buscando a diferenciação do câncer de tecidos normais por meio de análise de saliva (CHRISTIAN et al. 2014), quantificação dos níveis de água celular (BARROSO et al., 2015), análise de pele *in vivo* e *ex vivo* visando diagnóstico (BODANESE et al., 2012; FENG et al., 2015; LIMA et al., 2019; RASHID et al., 2014; SILVEIRA et al., 2012; SILVEIRA et al., 2015) e diagnóstico em amostras de sangue total, soro e células sanguíneas (ATKINS et al., 2017; BAI et al., 2020; CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011; FÉRÉ et al., 2020; HAPILLON et al., 2015; HASSOUN et al., 2018; KHETANI et al., 2015; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b; MANAGÒ et al., 2018a; MANAGÒ et al., 2018b; NEUGEBAUER et al., 2010a; NEUGEBAUER et al., 2010b; PREMASIRI et al., 2012; SILVA et al., 2018) com resultados promissores.

A espectroscopia Raman tem sido estudada para o diagnóstico diferencial de leucemias (ATKINS et al., 2017; BAI et al., 2020; CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011; FÉRÉ et al., 2020; HAPILLON et al., 2015; HASSOUN et al., 2018; KHETANI et al., 2015; LIMA et al., 2023; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b; MANAGÒ et al., 2018a; MANAGÒ et al., 2018b; NEUGEBAUER et al., 2010a; NEUGEBAUER et al., 2010b; PREMASIRI et al., 2012; SILVA et al., 2018),

pois tem demonstrado grande capacidade de identificar informações espectrais sobre células sanguíneas e compostos bioquímicos presentes no sangue, como proteínas, aminoácidos, ácidos nucleicos, lipídios e carotenóides (ATKINS et al., 2017; NEUGEBAUER et al., 2010a; PREMASIRI et al., 2012; SILVA et al., 2018), permitindo inclusive diferenciar os estágios de maturação celular (BAI et al., 2020; MANAGÒ et al., 2018a; MANAGÒ et al., 2018b). A espectroscopia Raman tem sido utilizada como ferramenta para o diagnóstico e discriminação de células neoplásicas (leucemia) de células normais (CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011; FAZIO et al., 2016; FÉRÉ et al., 2020; GONZÁLEZ-SOLÍS et al., 2014; GONZÁLEZ-SOLÍS, 2019; HAPPELLON et al., 2015; HASSOUN et al., 2018; KHETANI et al., 2015; LIMA et al., 2023; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b; MANAGÒ et al., 2016; MARTÍNEZ-ESPINOSA et al., 2009; NEUGEBAUER et al., 2010a; NEUGEBAUER et al., 2010b; PREMASIRI et al., 2012; SILVA et al., 2018; VANNA et al., 2015; ZHANG et al., 2015) e também para monitorar as alterações espectrais que ocorrem durante o processo de apoptose celular (FAZIO et al., 2016; KHETANI et al., 2015; ZHANG et al., 2015), usando amostras de plasma sanguíneo (SILVA et al., 2018), amostras de soro sanguíneo (GONZÁLEZ-SOLÍS et al., 2014; GONZÁLEZ-SOLÍS, 2019; LIMA et al., 2023; MARTÍNEZ-ESPINOSA et al., 2009), amostras de sangue total (FÉRÉ et al., 2020; HAPPELLON et al., 2015; SILVA et al., 2018), e através de células isoladas, que podem ser cultivadas ou extraídas de amostras de sangue (CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011; FAZIO et al., 2016; HASSOUN et al., 2018; KHETANI et al., 2015; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b; NEUGEBAUER et al., 2010a; NEUGEBAUER et al., 2010b; VANNA et al., 2015; ZHANG et al., 2015). Diferentes variações da espectroscopia Raman têm sido utilizadas no estudo de sangue e células sanguíneas, tais como: Raman dispersivo (LIMA et al., 2023; MANAGÒ et al., 2016; SILVA et al., 2018), microespectroscopia Raman (FAZIO et al., 2016; FÉRÉ et al., 2020; GONZÁLEZ-SOLÍS et al., 2014; GONZÁLEZ-SOLÍS, 2019; HAPPELLON et al., 2015; MANAGÒ et al., 2018a; MARTÍNEZ-ESPINOSA et al., 2009; NEUGEBAUER et al., 2010a;

NEUGEBAUER et al., 2010b; VANNA et al., 2015; ZHANG et al., 2015), espectroscopia Raman intensificada por superfície (SERS) (HASSOUN et al., 2018; KHETANI et al., 2015; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b), pinças ópticas (CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011), espectroscopia Raman ressonante (BARROSO et al., 2015) entre outros. Desta forma, tornou-se possível, por exemplo, a identificação de bandas de biomarcadores como: DNA, proteína e citocromo c que permitem a discriminação entre pacientes saudáveis e com LLC (FÉRÉ et al., 2020), e diferenciação entre pacientes com câncer de mama, leucemia e câncer de colo de útero baseado na composição química (bandas de amida I, amida III, polissacarídeos, fenilalanina e fosfolipídios) (GONZÁLEZ-SOLÍS, 2019).

1.1. Problema

A espectroscopia Raman é uma técnica de diagnóstico óptico capaz de identificar as alterações bioquímicas e as diferenças espectrais em amostras de células sanguíneas de crianças e adolescentes saudáveis e crianças e adolescentes com câncer e, se associada a análises estatísticas, é capaz de diferenciar amostras saudáveis de amostras doentes, permitindo um diagnóstico rápido e preciso?

1.2. Hipótese

A espectroscopia Raman é uma técnica óptica largamente utilizada para análise e diferenciação de amostras normais e doentes, e tem demonstrado ser promissora. Baseado na literatura espera-se que a espectroscopia Raman seja capaz de identificar as alterações bioquímicas nas amostras de células sanguíneas saudáveis e com câncer e, a partir dessas alterações consiga diferenciar os grupos normais e com câncer, de forma a contribuir para a evolução de uma técnica de diagnóstico rápida e precisa.

1.3. Objetivo geral

O objetivo do presente trabalho foi estabelecer padrões de espectroscopia Raman a partir de células sanguíneas de pacientes pediátricos saudáveis e portadores de neoplasias (câncer), detectando as alterações bioquímicas, através da utilização de variáveis obtidas com a análise de componente principal (PCA) e análise discriminante por mínimos quadrados parciais (PLS-DA).

1.4. Objetivos específicos

- Estabelecer um padrão de espectroscopia Raman para amostras de células sanguíneas;
 - Avaliar as características espectrais dos pacientes com leucemia;
 - Identificar os picos Raman que possibilitem a diferenciação entre as amostras saudáveis e os tipos de câncer através da PCA;
- Determinar a significância das diferenças entre os picos através do teste ANOVA;
 - Discriminar as amostras saudáveis das amostras de câncer através da PLS-DA;
 - Avaliar os principais marcadores biológicos alterados em pacientes com leucemias fazendo correlação com as características do tipo de câncer, para diferenciar as leucemias dos outros tipos de neoplasias nos pacientes pediátricos portadores de neoplasia.

2. REVISÃO BIBLIOGRÁFICA

2.1. Câncer

Câncer é o nome dado a um conjunto de mais de 100 doenças caracterizadas pelo crescimento desorganizado de células, que acabam se infiltrando em tecidos e órgãos adjacentes (INCA, 2020b). O câncer é a principal causa de morte e uma importante barreira ao aumento da expectativa de vida em todos os países do mundo. A mais recente estimativa mundial, ano 2020, estima em 19,3 milhões de novos casos (18,1 milhões excluindo câncer de pele não melanoma, exceto carcinoma basocelular) e 10 milhões de mortes por câncer (9,9 milhões excluindo câncer de pele não melanoma, exceto carcinoma basocelular) em todo o mundo (SUNG et al., 2021).

Para o Brasil, a estima-se para o triênio de 2023 a 2025 que ocorrerão 704 mil casos novos de câncer (483 mil se excluídos os casos de câncer de pele não melanoma). O câncer de pele não melanoma será o mais incidente (220 mil), seguido pelos cânceres de mama (74 mil), próstata (72 mil), cólon e reto (46 mil), pulmão (32 mil) e estômago (21 mil). A Figura 1 mostra a distribuição dos tipos de câncer mais incidentes dividida por sexo (INCA, 2022).

Estas estimativas baseiam-se em dados anteriores à pandemia causada pela doença do coronavírus 2019 (COVID-19). Ainda que a extensão total do impacto da pandemia de COVID-19 seja atualmente desconhecida em diferentes regiões do mundo, espera-se uma redução em curto prazo na incidência do câncer, seguida por aumentos nos diagnósticos em estágio avançado e mortalidade por câncer em alguns cenários, pois devido ao cenário pandêmico, houve atrasos nos diagnósticos e tratamentos decorrente do fechamento do sistema de saúde com suspensão de programas de triagem (SUNG et al., 2021).

| Localização Primária | Casos | % | | | Localização Primária | Casos | % |
|-----------------------------|--------|-------|--------|----------|-----------------------------|--------|-------|
| Próstata | 71.730 | 30,0% | Homens | Mulheres | Mama feminina | 73.610 | 30,1% |
| Cólon e reto | 21.970 | 9,2% | | | Célon e reto | 23.660 | 9,7% |
| Traqueia, brônquio e pulmão | 18.020 | 7,5% | | | Colo do útero | 17.010 | 7,0% |
| Estômago | 13.340 | 5,6% | | | Traqueia, brônquio e pulmão | 14.540 | 6,0% |
| Cavidade oral | 10.900 | 4,6% | | | Glândula tireoide | 14.160 | 5,8% |
| Esôfago | 8.200 | 3,4% | | | Estômago | 8.140 | 3,3% |
| Bexiga | 7.870 | 3,3% | | | Corpo do útero | 7.840 | 3,2% |
| Laringe | 6.570 | 2,7% | | | Ovário | 7.310 | 3,0% |
| Linfoma não Hodgkin | 6.420 | 2,7% | | | Pâncreas | 5.690 | 2,3% |
| Fígado | 6.390 | 2,7% | | | Linfoma não Hodgkin | 5.620 | 2,3% |

Figura 1: Distribuição proporcional dos tipos de câncer mais incidentes estimados para 2023 por sexo, exceto câncer de pele não melanoma.

Fonte: INCA (2022).

O crescimento das células cancerígenas ocorre de maneira distinta das células normais, pois as células dividem-se de forma veloz, agressiva e sem controle, espalhando-se para outras regiões do corpo provocando transtornos funcionais (INCA, 2020b).

Carcinogênese ou oncogênese é o nome dado ao processo de formação do câncer e normalmente acontece lentamente, podendo levar anos para que uma célula cancerosa origine um tumor. O período de latência do processo varia de acordo com a intensidade do estímulo carcinogênico, com a existência ou não dos agentes oncoiniciadores, oncopromotores e oncoaceleradores, e com o tipo e localização primária do tumor. Os efeitos cumulativos desses fatores determinam o início, progressão e inibição do tumor, sendo determinada pela frequência, tempo e interação entre esses fatores (INCA, 2020b). Este processo acontece lentamente, existindo vários estágios até a formação do tumor (Figura 2):

- Estágio de iniciação: é primeiro estágio da carcinogênese, onde as células sofrem efeitos dos agentes cancerígenos, provocando algumas modificações genéticas, porém estas alterações dos genes ainda não são suficientes para detecção de um tumor.

- Estágio de promoção: Neste estágio, as células que já sofreram alterações genéticas passam a sofrer efeitos dos oncopromotores, que as transformam gradativamente em células malignas. Para que esse processo ocorra, é necessário um longo e continuado período de contato das células que

sofreram mutações com o agente cancerígeno promotor. Dentre os agentes oncopromotores pode-se citar os hormônios, que podem favorecer a transformação de células normais em células cancerosas.

- Estágio de progressão: fase final do processo de oncogênese, caracterizado pela proliferação descontrolada e irreversível das células modificadas. O câncer já se encontra desenvolvido, evoluindo para os primeiros sinais e sintomas clínicos característicos da doença. Os fatores chamados de carcinógenos ou oncoaceleradores são responsáveis por acelerar esse processo de instalação da patologia (INCA, 2020b).

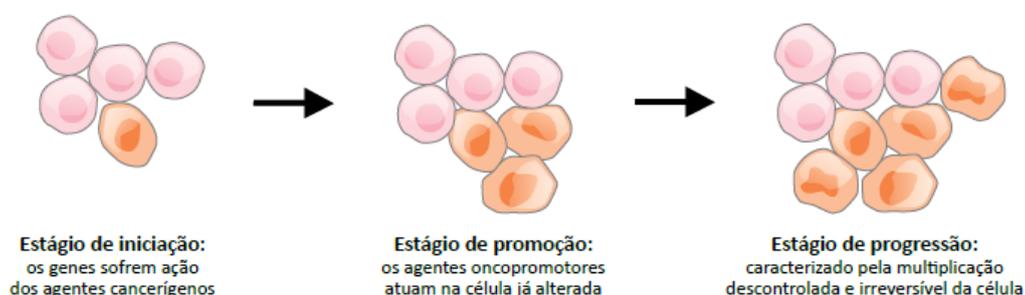


Figura 2: Estágios do desenvolvimento do câncer.

Fonte: INCA (2020b), adaptado.

Os fatores de risco para o câncer podem ser classificados, segundo a possibilidade de modificação, em modificáveis e não modificáveis, a interação entre os fatores modificáveis e os não modificáveis é que vai determinar o risco individual de desenvolver câncer. Os fatores modificáveis são conhecidos também como fatores ambientais, dentre eles o uso de tabaco, o sedentarismo, consumo de bebidas alcoólicas, obesidade, exposição a radiações (radiação ultravioleta, raio-X), causas ocupacionais (produtos químicos), poluição ambiental e até o comportamento sexual. Esses fatores, como a própria classificação diz, são passíveis de alteração, e uma vez modificados, podem, reduzir o risco de desenvolver um câncer. Já os fatores não modificáveis não dependem do comportamento, hábitos e práticas individuais ou coletivas, e são conhecidos como fatores de risco intrínsecos. São eles: idade, gênero, etnia ou raça e herança genética ou hereditariedade (INCA, 2020b). Fatores estes, que

isolados ou concomitantemente, agem promovendo as alterações iniciais, podem regredir nessa fase ou após um período de latência, ou passar à fase de progressão neoplásica propriamente dita (INCA, 2020b).

2.2. Cânceres pediátricos

Crianças e adolescentes diagnosticados com câncer representam uma pequena proporção de, aproximadamente, 1% a 2% de todos os cânceres diagnosticados em todo o mundo a cada ano, sendo consideradas raras. No entanto, não existem estimativas globais de incidência, sobrevivência e mortalidade para crianças com câncer na maioria dos países de baixa e média renda (BHAJTA et al., 2019; KAPLAN, 2019; MELARAGNO e CAMARGO, 2013).

O câncer infantojuvenil (faixa etária entre 0 a 19 anos) corresponde a um conjunto de várias doenças que têm em comum a proliferação descontrolada de células anormais e com a possibilidade e acometerem em qualquer local do organismo. Diferentemente do câncer do adulto, o câncer infantojuvenil é predominantemente de natureza embrionária e, geralmente, afeta as células do sistema sanguíneo e os tecidos de sustentação, sendo sua classificação baseada na morfologia do tumor, ao invés do local de origem. Outra diferença é que os fatores de risco relacionados com o estilo de vida não influenciam o risco de uma criança e adolescente desenvolver câncer, como acontece com o câncer do adulto (BURKHAMER et al., 2017; INCA, 2020a; MELARAGNO e CAMARGO, 2013).

No Brasil, para o triênio 2023-2025, estima-se que o número de novos casos de câncer infantojuvenil seja de 4.230 casos novos para o sexo masculino e de 3.700 para o sexo feminino. Corresponde a um risco estimado de 134,81 casos novos por milhão de crianças e adolescentes (INCA, 2022).

Os tipos prevalentes de cânceres infantojuvenis são leucemia (28%), neoplasias do sistema nervoso central (26%) e linfomas (8%) (INCA, 2020a). Os tumores sólidos extracranianos representam 17% dos cânceres infantis e provavelmente são responsáveis por uma grande proporção dos

cânceres infantojuvenis (BHAKTA et al., 2019). Alguns exemplos destes são: neuroblastoma (tumor de células do sistema nervoso periférico, frequentemente de localização abdominal), tumor de Wilms (tipo de tumor renal), retinoblastoma (afeta a retina, fundo do olho), tumor germinativo (das células que originam os ovários e os testículos), osteossarcoma (tumor ósseo) e sarcomas (tumores de partes moles) (BHAKTA et al., 2019; BURKHAMER et al., 2017).

Atualmente, a taxa de cura em 5 anos de crianças e adolescentes diagnosticados com câncer é de aproximadamente 80% em muitos países. A maior parte deles terá uma boa qualidade de vida após o tratamento adequado. Para isso, é necessária uma avaliação mais precisa da carga global do câncer infantojuvenil, para que seja possível incrementar políticas que melhorem os resultados das estatísticas de sobrevida para crianças e adolescentes diagnosticados com câncer. É de conhecimento geral que quando mais precoce for a detecção do câncer, maior será a sobrevida dos pacientes e com melhor qualidade (BHAKTA et al., 2019; INCA, 2020a).

2.3. Câncer leucêmico

2.3.1. Medula óssea e hematopoiese

A hematopoiese é um processo complexo e altamente organizado responsável pela produção de células sanguíneas. As células do sangue surgem a partir de uma célula hematopoiética pluripotente (células tronco – *stem cells*), ou seja, um precursor comum e indiferenciado, sendo a MO o órgão responsável pela produção dessas células durante a vida após o nascimento. O sangue é um tecido conjuntivo fluido imprescindível à manutenção de todos os tecidos e órgãos do corpo humano. É composto por um conjunto de células: eritrócitos, leucócitos e plaquetas, que circulam em suspensão em um meio líquido, o plasma. O plasma é constituído de componentes bioquímicos como eletrólitos, proteínas, hormônios, aminoácidos, lipídios e as imunoglobulinas, que são proteínas especializadas na função de

imunidade do organismo contra diversos antígenos, além de água, que corresponde a cerca de 90% do seu volume. O volume total do sangue pode ser considerado pela soma dos eritrócitos e plasma, visto que o volume dos outros componentes é desprezível, sua composição é cerca de 55% de plasma, 4% leucócitos e plaquetas e 41% eritrócitos (NAYKA e RAI, 2017; RODRIGUES e OLIVEIRA, 2017).

As principais funções do sangue são de transporte de eletrólitos, metabólitos, gases, nutrientes, hormônios e outras moléculas sinalizadoras, sendo responsável por oxigenar órgãos e tecidos, realizando as trocas gasosas. O processo de hematopoiese produz uma grande quantidade de células por dia, sendo esse processo controlado pelas células e proteínas da matriz extracelular que formam o estroma medular (NAYKA e RAI, 2017; RODRIGUES e OLIVEIRA, 2017).

Cada célula sanguínea é originada a partir de uma célula pluripotente indiferenciada, as células-tronco. As células progenitoras indiferenciadas pluripotentes por sua vez dão origem a duas linhagens de células progenitoras multipotentes, uma linfóide e outra mieloide, denominada como unidades formadoras de colônias. A linhagem de células mielóides dará origem aos eritrócitos, granulócitos, monócitos e megacariócitos. Enquanto as unidades formadoras de colônias linfóides darão origem aos linfócitos T e B e células exterminadoras naturais. A partir da influência dos fatores estimuladores de colônias, cada célula multipotente dará origem a uma população de células unipotentes, que originarão apenas um tipo de célula sanguínea. A Figura 3 mostra o processo de hematopoiese em humanos (NAYKA e RAI, 2017; RODRIGUES e OLIVEIRA, 2017).

Existem fatores estimuladores e fatores moduladores da hematopoiese que são regulados pela expressão de determinados genes. A eritropoietina é um hormônio glicoproteico fundamental para o desenvolvimento de linhagem eritroide, sendo imprescindível para a proliferação, maturação e morte celular programada das células. Age aumentando o número de mitoses, estimulando a síntese de hemoglobina, a maturação das células indiferenciadas e aumentando a taxa de reticulócitos (RODRIGUES e OLIVEIRA, 2017).

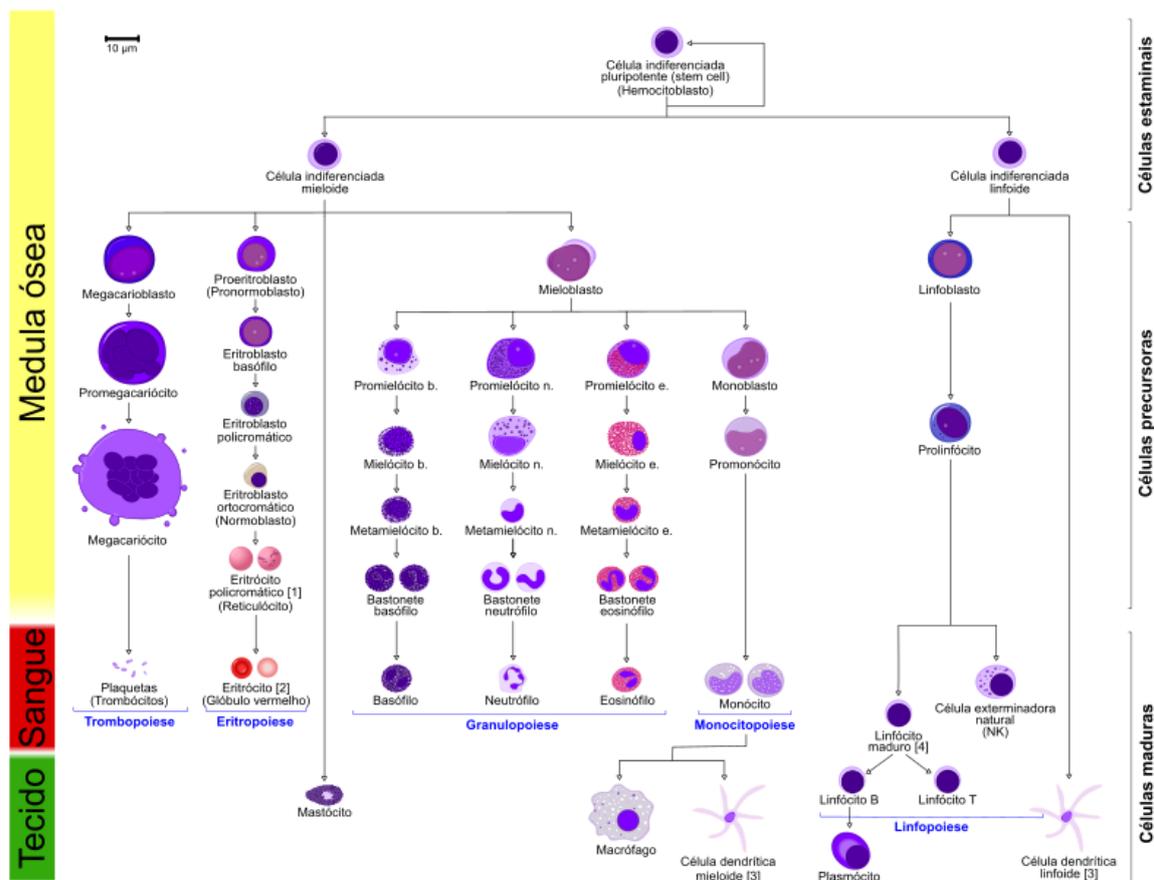


Figura 3: Hematopoiese em humanos.

Fonte: RODRIGUES e OLIVEIRA (2017), adaptado.

2.3.2. Leucemias

A leucemia é um tipo de doença neoplásica que atinge as células do sangue, tendo como principal característica o acúmulo de células doentes na MO que acabam por substituir as células sanguíneas normais. As neoplasias hematológicas relacionadas à leucemia tem origem na MO que levam à expansão anormal dos glóbulos brancos. O tipo da leucemia é classificado de acordo com o tipo de célula sanguínea que se torna cancerosa e com o comportamento dessas células, se crescem rápido ou lentamente, o que caracteriza a doença em aguda ou crônica. É o tipo de câncer mais comum em crianças menores de 15 anos, mas também ocorre mais frequentemente em adultos com mais de 55 anos (INCA, 2020a; MITCHELL e STEIDL, 2020).

Para o Brasil, para cada ano do triênio de 2023 a 2025, é de 11.540 novos casos de leucemia, o que corresponde a um risco estimado de 5,33 por 100 mil habitantes (INCA, 2022). No ano de 2020, a estimativa mundial mostrou que ocorreram aproximadamente 475 mil novos casos de leucemia (SUNG et al., 2021).

Os quatro principais tipos de leucemia dentre os 12 tipos existentes são: LMA, LMC, LLA e LLC (INCA, 2020b). As leucemias agudas correspondem a mais de 95% de todas as leucemias infantis, incluindo leucemias linfoides/linfoblásticas agudas (LLA), que são o tipo mais frequente (> 75%), e leucemias mieloblásticas agudas (LMA) que correspondem (16%) na faixa etária ≤ 1 a 14 anos (KAPLAN, 2019; MELARAGNO e CAMARGO, 2013; VAN MAELE-FABRY et al., 2019).

Nas crianças, as leucemias geralmente se manifestam com sintomas relacionados às citopenias ou infiltração leucêmica da MO, podendo acometer outros órgãos, como o baço, fígado, testículos e sistema nervoso central (KAPLAN, 2019). A MO é substituída por blastos leucêmicos nas crianças com leucemia, gerando uma expansão da cavidade do osso medular causando dor óssea, claudicação ou recusa para deambular, podendo existir sintomas também como dor nas costas. O diagnóstico é feito através de uma contagem completa de células sanguíneas. A leucemia ainda pode se espalhar em outros órgãos fora do osso medular, levando a quadros de linfadenopatia, hepatomegalia, esplenomegalia e lesões renais. As leucemias agudas promovem alterações no tamanho do baço e do fígado, sendo a esplenomegalia com ocorrência mais significativa na LMC (KAPLAN, 2019).

Os fatores de risco para as leucemias ainda não estão bem descritos, mas existe uma suposição de associação entre determinados fatores de risco que gerem uma maior probabilidade de desenvolver a doença: tabagismo, benzeno (encontrado na gasolina), radiação ionizante, quimioterapia, exposição ocupacional ao formaldeído em indústrias, produção de borracha, síndrome de Down e outras doenças hereditárias, desordens sanguíneas, história familiar, exposição a agrotóxicos e infecção por vírus das hepatites B e C (INCA, 2020a).

2.3.2.1. Leucemia linfoide aguda (LLA)

A LLA caracteriza-se por ser uma doença de progressão rápida, causada pelo aumento excessivo de leucócitos imaturos, conhecidos como blastos, que se proliferam na MO, impossibilitando a medula de produzir hemácias, plaquetas e neutrófilos. As células acometidas pelo processo de malignidade são da linhagem de desenvolvimento dos linfócitos, e podem infiltrar o sistema nervoso central, sangue periférico, linfonodo, baço e rins. Na década de 90, o acometimento por LLA na infância era considerado fatal. Hoje, é considerada uma das doenças com maior taxa de cura na oncologia pediátrica (KAPLAN, 2019).

A LLA precursora de células B é muito mais comum que as células T, sendo responsável por aproximadamente 80% dos casos de LLA. Historicamente, os casos de LLA precursor de células T têm um melhor prognóstico em comparação ao LLA precursor de células B, no entanto, atualmente, o tratamento dos dois tipos é muito semelhante, reduzindo esta diferença (KAPLAN, 2019).

2.3.2.2 Leucemia linfoide crônica (LLC)

A LLC é uma doença neoplásica que acomete o sistema linfopoiético, que se diferencia das demais leucemias pelas suas características epidemiológicas. É o tipo de leucemia mais frequente nos países ocidentais, ocupando 22% a 30% de todas as leucemias do adulto, sendo extremamente rara nos países asiáticos. A LLC raramente acomete indivíduos menores de 30 anos, sendo caracterizada pelo predomínio em população idosa, com idade média de diagnóstico entre 64 e 70 anos. É o tipo de leucemia que apresenta maiores índices de sobrevida; no entanto, devido ao fato de muitos pacientes serem assintomáticos ou apresentam evolução clínica indolor e não solicitarem tratamento ou internação acredita-se que exista uma subestimação dos dados epidemiológicos relatados na literatura. Sua

etiologia é desconhecida, e não existem fatores ambientais (contato com produtos químicos, radiação ionizante ou agentes virais) ou genéticos que possam ser atribuídos à sua etiopatogenia (YAMAMOTO e FIGUEIREDO, 2005).

2.3.2.3. Leucemia mieloide aguda (LMA)

A LMA é descrita como uma doença de progressão rápida, pelo excesso de proliferação celular na MO, podendo afetar tanto o sangue periférico como os diversos órgãos e sistemas. As células comprometidas são de origem da linhagem mieloide, podendo ser originados de granulócitos, monócitos, eritrócitos e megacariócitos, o que cria uma variada morfologia para esta doença (KAPLAN, 2019).

A leucostase é um problema comum dos pacientes com LMA, podendo apresentar alguns sinais e sintomas: neurológicos (dor de cabeça, confusão, letargia, tontura, embaçamento da visão, ataxia, hemorragia retiniana), respiratórios (taquipneia, hipóxia, infiltrados radiográficos no tórax, insuficiência respiratória), e vasculares (oclusão vascular periférica, trombose, alterações nos testes de coagulação com protombina elevada e tempos parciais de tromboplastina com diminuição do nível de fibrinogênio). A taxa de sobrevivência livre de recaídas desses pacientes é algo em torno de 60%, dependendo da condução terapêutica (KAPLAN, 2019).

2.3.2.4. Leucemia mieloide crônica (LMC)

A LMC é uma doença hematopoiética maligna caracterizada pela proliferação descontrolada de células da linhagem mieloide, eritroide e plaquetária no sangue periférico e hiperplasia mieloide na MO. É responsável por cerca de 20% das leucemias em adultos e corresponde a menos de 2% das leucemias em crianças. A LMC apresenta-se mais comum na adolescência, suas características morfológicas são únicas, com uma contagem elevada de leucócitos, neutrofilia, aumento de metamielócitos e

mielócitos, eosinofilia, basofilia e trombocitose (KAPLAN, 2019; MELARAGNO e CAMARGO, 2013).

A LMC tem seu curso descrito em três fases: a fase crônica, que pode durar de três a seis anos, progredindo para a fase acelerada e para uma fase terminal, conhecida como crise blástica. Na grande maioria dos casos, o diagnóstico ocorre na fase crônica, onde se observa um número elevado de leucócitos e/ou plaquetas. Os sintomas mais frequentes estão relacionados a cansaço, baixa tolerância a exercícios, anorexia, desconforto abdominal, saciedade precoce (devido ao acometimento do baço), perda de peso excessiva e sudorese excessiva, ao exame físico pode ser observada palidez e esplenomegalia (MELARAGNO e CAMARGO, 2013).

No mielograma, cerca de 75% a 90% do volume da medula óssea está ocupado por tecido hematopoiético, com redução do percentual de células adiposas (MELARAGNO e CAMARGO, 2013).

2.4. Diagnóstico padrão das leucemias

O diagnóstico de leucemias agudas normalmente se inicia a partir de suspeitas clínicas, onde o paciente é encaminhado para realização de exames laboratoriais através da coleta do sangue periférico e, se for confirmado a alterações importantes nos exames é realizada a punção da MO para confirmação e diagnóstico diferencial. Na maioria das vezes, a aspiração da MO é necessária para diagnosticar as leucemias, mas o diagnóstico também pode ser feito a partir sangue periférico caso haja um número suficiente de blastos. Sendo útil nos casos em que a contagem de leucócitos é alta permitindo um diagnóstico mais rápido, o que é muito vantajoso. Tanto a análise morfológica, como a citometria de fluxo são usadas para avaliar a MO. A análise morfológica avalia a MO e blastos leucêmicos ao microscópio, enquanto os diferentes marcadores imunofenotípicos podem ser encontrados através da citometria de fluxo. A avaliação de todos esses parâmetros se torna necessária para diferenciação e definição da conduta terapêutica a ser abordada (KAPLAN, 2019; MELARAGNO e CAMARGO, 2013).

A imunocitoquímica e a citometria de fluxo permitem detectar a presença de proteína p53 em amostras de pessoas com leucemia. A imunofenotipagem por citometria de fluxo permite avaliar e identificar os antígenos celulares por meio de anticorpos monoclonais. Os citômetros de fluxo identificam quatro ou mais tipos de antígenos diferentes ao mesmo tempo, gerando informações detalhadas sobre as células analisadas (RODRIGUES e OLIVEIRA, 2017).

2.5. Cânceres não-leucêmicos

2.5.1. Tumores do sistema nervoso central

Os tumores do sistema nervoso central (SNC) são caracterizados pelo crescimento de células anormais nos tecidos localizados no cérebro e na medula espinal, sendo os novos casos diagnosticados mais frequentemente no cérebro. Embora seja um tumor predominantemente adulto, podem ser encontrados em crianças e correspondem de 15 a 20% dos cânceres infantis (INCA, 2020a; MELARAGNO e CAMARGO, 2013). Os tumores infantis do SNC têm sua etiologia desconhecida, porém, já existem associações com o aumento da incidência e a presença de outras síndromes como neurofibromatose tipo 1 e síndrome de Li-Fraumeni. Os cinco tipos mais comuns de tumores cerebrais infantis são: meduloblastoma, ependioma, astrocitoma, gliomas de tronco cerebral e tumor teratoide rabdoide atípico, podendo levar a déficits neurológicos focais (MELARAGNO e CAMARGO, 2013).

O diagnóstico dos tumores de SNC evoluiu bastante com o avanço da medicina diagnóstica por imagem, a tomografia computadorizada permite detectar facilmente 95% dos casos. A ressonância nuclear magnética se faz necessária para determinar com maior precisão a localização do tumor. O tratamento dos tumores cerebrais, de maneira geral, consiste em cirurgia, radioterapia e quimioterapia, mas a escolha dessas linhas de tratamento dependem do tipo do tumor, localização e da idade do paciente (MELARAGNO e CAMARGO, 2013).

2.5.2. Linfomas

Os linfomas são neoplasias originadas no sistema linfático e são classificados em dois grandes grupos: linfoma ou doença de Hodgkin caracterizado pela presença de células de Reed-Sternberg, e linfomas não Hodgkin originados de linfoblastos de células T e linfoblastos de células B. Os linfomas são o terceiro grupo de doenças mais frequentes em crianças e adolescentes nos Estados Unidos; no Brasil, na maioria dos registros é o segundo mais frequente (MELARAGNO e CAMARGO, 2013). A fisiopatologia, a evolução, o tratamento e o prognóstico dos linfomas são similares em adultos e crianças, sendo os efeitos tardios do tratamento piores em crianças (MELARAGNO e CAMARGO, 2013).

2.5.2.1. Linfoma não Hodgkin

Os linfomas são neoplasias originadas nas células do sistema imunológico, ou seja, neoplasias de linfócitos ou seus precursores que circulam por todo o corpo (MELARAGNO e CAMARGO, 2013). O risco de desenvolver a doença aumenta com o processo de envelhecimento. Estima-se que o número de novos casos de linfoma não Hodgkin esperados para o Brasil, para cada ano do triênio de 2023 a 2025, é de 12.040 casos (INCA, 2022).

O principal meio para classificação do linfoma não Hodgkin é a análise histopatológica associada à imunofenotipagem, e para um diagnóstico mais preciso pode ser feita uma análise molecular (MELARAGNO e CAMARGO, 2013). Os linfomas podem ser classificados a partir da linhagem celular de origem, sendo de células B ou células T. O diagnóstico diferencial é feito através da avaliação das características dos linfonodos neoplásicos: consistências endurecidas, coalescentes, aderidos a planos profundos, sem sinais flogísticos (calor, rubor e dor). Quando essas características permanecem após 10-14 dias de antibioticoterapia deve-se realizar biópsia diagnóstica (MELARAGNO e CAMARGO, 2013). Os principais fatores de risco

para desenvolvimento do linfoma não Hodgkin são pessoas com sistema imunocomprometido em consequência de doenças genéticas hereditárias ou portadores de infecção pelo o vírus da imunodeficiência humana (do inglês *human immunodeficiency virus* – HIV), pacientes que fazem uso de drogas imunossupressoras, portadores do vírus Epstein-Barr e vírus linfotrópico de células T humana do tipo 1 (HTLV1), da bactéria *H. pylori*, bem como algumas substâncias químicas (agrotóxicos, radiação ionizante e ultravioleta) (INCA, 2020a; MELARAGNO e CAMARGO, 2013).

2.5.2.2. Linfoma Hodgkin

O linfoma ou doença de Hodgkin é uma neoplasia que surge no sistema linfático (linfonodos ou gânglios) e tecidos que produzem as células responsáveis pela imunidade do organismo, caracterizado pela presença de células de Reed-Sternberg. A estimativa é de que o número de novos casos de linfoma de Hodgkin para o Brasil, para cada ano do triênio de 2023 a 2025, é de 3.080 casos. Essa doença pode ocorrer em qualquer idade, entretanto, é mais comum entre adolescentes e adultos de 15 a 39 anos, e idosos a partir dos 75 anos, sendo os homens são mais acometidos do que as mulheres (INCA, 2022; MELARAGNO e CAMARGO, 2013).

Os principais fatores de risco para linfoma de Hodgkin, assim como para todos os linfomas, ainda não estão bem estabelecidos, mas sugerem associações para risco aumentado de desenvolvimento da doença: pessoas com sistema imunocomprometido, presença do HIV e pacientes que fazem uso de drogas imunossupressoras. Apesar de ser muito raro, deve-se considerar o histórico familiar, caso uma ou mais pessoas da mesma família apresentar o diagnóstico da doença (INCA, 2020a; MELARAGNO e CAMARGO, 2013). A manifestação clínica mais comum, correspondendo a 80% dos casos, é a linfadenomegalia firme, indolor, cervical ou supraclavicular. A disseminação patogênica ocorre gerando comprometimento hepático, esplênico, ósseo, da medula óssea ou cerebral, estando associada a sintomas sistêmicos (MELARAGNO e CAMARGO, 2013).

2.5.3. Osteossarcoma

O osteossarcoma é o tumor maligno ósseo mais frequente na infância e adolescência, comumente associado à dor local, como também alterações ósseas. É caracterizado por células fusiformes malignas que produzem osteoide ou osso. Para o diagnóstico é avaliada a presença de osteoide; quando presente, caracteriza osteossarcoma, e quando ausente, outro diagnóstico deve ser considerado. Deve-se correlacionar estes achados com os achados radiológicos, pois muitas vezes as biópsias não representam a lesão completa. No diagnóstico, a matriz predominante é identificada e pode incluir osteoide/osso, cartilagem ou tecido fibroso, sendo chamada de osteossarcoma osteoblástico, condroblástico e fibroblástico, respectivamente (MELARAGNO e CAMARGO, 2013).

A queixa mais comum dos pacientes com osteossarcoma é a dor no local da lesão, se iniciando de forma intermitente e profunda, mas evoluindo para uma dor constante. Ainda podem apresentar claudicação, edema e limitações para as atividades diárias. O tumor pode acometer qualquer osso, mas principalmente a região metafisária do esqueleto apendicular, como fêmur, tíbia e úmero (MELARAGNO e CAMARGO, 2013). A estratégia de tratamento consiste em quimioterapia pré-cirúrgica, cirurgia e tratamento adjuvante apresentando melhora no prognóstico dos pacientes. No entanto, quanto maior a idade pior é o prognóstico (MELARAGNO e CAMARGO, 2013).

2.6. Espectroscopia Raman

A espectroscopia Raman é uma técnica de diagnóstico óptico baseada no espalhamento inelástico da luz pelas moléculas, sendo capaz de fornecer informações detalhadas sobre os componentes bioquímicos dos tecidos, gerando impressões digitais químicas das células, tecidos ou fluidos biológicos. Por causa da sua especificidade química elevada, pouca ou nenhuma necessidade de preparação de amostra e por usar tecnologias ópticas

avançadas na faixa espectral visível ou infravermelho próximo, a espectroscopia Raman ganhou uma maior visibilidade nos estudos de diagnóstico de patologias humanas.

Toda doença tem sua progressão acompanhada por alterações químicas das moléculas, e a espectroscopia Raman é capaz de gerar informações detalhadas a respeito destas alterações, por isso surge como uma opção promissora para diagnóstico de diversas patologias. A espectroscopia Raman leva vantagem por reduzir a fluorescência que ocorre em espécies biológicas pelo fato de utilizar a faixa de excitação no infravermelho próximo (ex: 830 nm de lasers de diodo) (HANLON et al., 2000; KONG et al., 2015; SILVEIRA et al., 2012; SILVEIRA et al., 2015).

2.6.1. Efeito Raman

A espectroscopia Raman é baseada em um fenômeno conhecido como efeito Raman, caracterizado pela troca de energia entre a radiação eletromagnética e a matéria. A radiação que incide sobre uma substância pode ser espalhada ou absorvida, sendo que a maior parte da radiação espalhada terá a mesma frequência da radiação incidente, caracterizando o espalhamento elástico (espalhamento Rayleigh). No entanto, parte da radiação pode interagir com moléculas do material em vibração molecular, gerando alteração na frequência emitida (de fato espalhada), caracterizando o espalhamento Raman ou inelástico. Sendo a energia proporcional à frequência da radiação, a alteração da frequência da radiação espalhada deverá ser igual à frequência vibracional do espalhamento das moléculas. Todo esse processo de troca de energia entre a radiação incidente e as moléculas e espalhamento de radiação é conhecido como efeito Raman (HANLON et al., 2000).

Analisando os níveis de energia moleculares, o processo de espalhamento Raman pode ser entendido como a transição de uma molécula do seu estado fundamental (S_0) para o estado vibracional excitado (S_1 , S_2 , ...), devido à absorção do fóton incidente, e emissão de fóton espalhado elástica e inelasticamente (Raman) em seu decaimento. A radiação espalhada é

capturada por um espectrômetro e exibida como um espectro, onde sua intensidade é apresentada como função da mudança de frequência. Cada espécie molecular tem seu conjunto único de vibrações moleculares, portanto o espectro Raman de uma espécie particular consiste de uma série de picos ou bandas, cada pico é deslocado um do outro pelas frequências vibracionais características dessa molécula (HANLON et al., 2000).

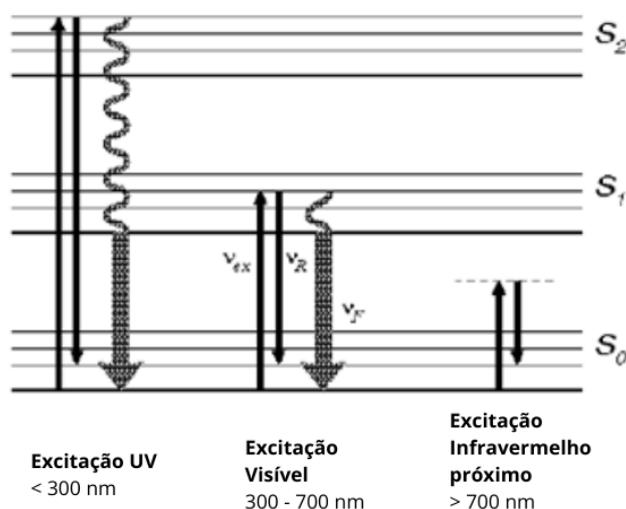


Figura 4: Diagrama de energia demonstrando o efeito Raman: excitação UV, visível e infravermelho próximo. S_0 , estado eletrônico basal; S_1 e S_2 , estados eletrônicos excitados. As linhas horizontais indicam os níveis de energia vibracional. A diferença de comprimento entre as setas pretas para cima e para baixo indica a frequência de vibração molecular.

Fonte: HANLON et al. (2000) (adaptado).

2.6.2. Diagnóstico de câncer não-leucêmico utilizando espectroscopia Raman

A espectroscopia Raman tem sido amplamente pesquisada e aplicada na literatura para identificar, classificar e até quantificar biomarcadores que podem fornecer informações sobre os vários tipos de câncer. Amostras de soro sanguíneo têm sido investigadas utilizando microespectroscopia Raman (830 nm de excitação) e métodos estatísticos multivariados para comparar amostras

doentes e indivíduos saudáveis, no entanto tiveram pequenos deslocamentos de pico e diferenças intensidades de pico, de forma que, não mostrou uma discriminação clara entre pacientes doentes e indivíduos saudáveis (ATKINS et al., 2017).

Christian et al. (2014) utilizaram a espectroscopia Raman com objetivo de diagnosticar câncer oral através da comparação do espectro de amostras do tecido com lesão com o tecido normal *ex vivo* e correlacionando esses achados com os resultados encontrados na análise histopatológica. Rashid et al. (2014) demonstraram em seu estudo que a técnica de microespectroscopia Raman pode ser utilizada para o diagnóstico precoce de câncer de colo de útero, sendo uma importante ferramenta para acompanhamento da evolução da patologia. No estudo de Barroso et al. (2015), foi mostrada a discriminação entre um tecido oral cancerígeno e um tecido sadio, comparando os níveis de água nos tecidos obtidos através da técnica de espectroscopia Raman. Já Feng et al. (2015) utilizaram marcadores encontrados na saliva com o objetivo de diferenciar o tumor de mama benigno do câncer maligno. Ressalta-se ainda a possibilidade de utilização da espectroscopia como ferramenta transoperatória, podendo ser útil na determinação das margens cirúrgicas na ressecção de determinados tumores ainda no ato cirúrgico, podendo melhorar o prognóstico do paciente (BARROSO et al., 2015).

Pesquisadores na China realizaram trabalho com a técnica Raman relacionado à detecção de câncer no plasma. Eles usaram a técnica de deposição revestida por gota em plasma de pacientes com câncer colorretal (15 pacientes) e de sujeitos saudáveis normais (21 voluntários). As informações provenientes dos espectros Raman (excitação de 785 nm) sugeriram um nível aumentado de células de DNA livres e um nível diminuído de β -caroteno no plasma do grupo com câncer. Para explicar essas alterações relativas ao aumento de ácidos nucleicos no sangue de pacientes com câncer, os autores descreveram dois mecanismos: apoptose/necrose, ou liberação de células intactas na corrente sanguínea e sua subsequente lise, enquanto a diminuição do β -caroteno sugere uma deficiência de espécie antioxidante,

sendo esta um importante fator de risco para a progressão de uma fase pré-câncer para câncer efetivamente (ATKINS et al., 2017).

2.6.3. Diagnóstico de leucemias utilizando espectroscopia Raman

A espectroscopia Raman é uma técnica que tem sido amplamente estudada para o diagnóstico diferencial das leucemias, utilizando tanto amostras de plasma como amostras de sangue total. Os espectros gerados através do espalhamento Raman permitem identificar picos que correspondem aos componentes bioquímicos da célula, permitindo a diferenciação das fases de maturação das células (BAI et al., 2020; MANAGÒ et al., 2018b).

Uma característica importante da espectroscopia Raman é sua capacidade de analisar de forma não destrutiva as amostras, classificando-as em subconjuntos distintos de leucócitos sendo significativa tanto para estudos do sistema imunológico *in vivo* quanto *in vitro* em que uma técnica que fornece informações bioquímicas sem rótulo (do inglês *label free*) sobre o tipo de leucócito. O estado bioquímico pode ser muito útil para dar suporte a hemogramas de padrão não específicos padrão, dessa forma Managò et al. (2018b) demonstrou que a espectroscopia Raman pode identificar e classificar com sucesso uma série de subgrupos de leucócitos importantes (granulócitos, monócitos e linfócitos).

A espectroscopia Raman tem demonstrado grande capacidade de identificar informações espectrais referentes a células sanguíneas e elementos presentes no sangue, como aminoácidos, proteínas, lipídios, ácidos nucleicos e carotenóides. No campo de diagnóstico, a técnica tem tido bastante visibilidade na detecção de diferentes tipos de leucemias usando linhagens celulares, esfregaços de sangue e amostras de soro (ATKINS et al. 2017; SILVA et al., 2018).

Diferentes variações da espectroscopia Raman têm sido empregadas no estudo do sangue, cada uma apresentando suas vantagens e desvantagens, como por exemplo: microespectroscopia Raman associada ou não ao uso de

pinças ópticas, SERS, espectroscopia Raman ressonante, dentre outras (ATKINS et al. 2017).

Féré et al. (2020) demonstraram que as bandas de DNA, proteína e citocromo c obtidas através da espectroscopia Raman podem ser marcadores bioquímicos que discriminam entre sujeitos saudáveis e pacientes portadores de LLC (FÉRE et al., 2020).

Outro estudo também demonstrou resultados promissores ao analisar amostras de sangue total e plasma sanguíneo, com o objetivo de diferenciar sujeitos saudáveis de sujeitos com leucemia aguda, através da espectroscopia Raman convencional, observando uma discriminação superior quando se aplicou regressão por mínimos quadrados parciais (do inglês *partial least squares* – PLS) (SILVA et al., 2018). Anteriormente, Neugebauer et al. (2010a) já descreveram resultados promissores com o uso da espectroscopia Raman para discriminação entre células tumorais e células saudáveis, através do cultivo e da análise de células individuais (NEUGEBAUER et al., 2010a).

2.7. Análise estatística

O teste de hipótese serve para auxiliar o pesquisador a tomar decisões referentes a uma população ou conjunto de dados. Os testes estatísticos são aplicados com a finalidade de verificar se os dados obtidos fornecem evidências suficientes para que se considere como verdadeira a hipótese de pesquisa, sendo uma hipótese definida como uma afirmação a cerca do conjunto de dados ou população. Para aplicação do teste estatístico adequado deve-se conhecer a amostra e então definir o teste que melhor se aplica ao conjunto de dados em questão (CAPP e NIENOV, 2020).

2.7.1. Teste t de Student

O teste t de Student, ou simplesmente teste t , é um teste estatístico popularmente utilizado para testar se a diferença entre as médias de dois grupos é estatisticamente significativa. Trata-se de um teste baseado em duas

hipóteses: uma hipótese nula (H_0), que afirma que a média entre os grupos provém do mesmo universo, ou seja se são iguais, e uma hipótese alternativa (H_A), que sugere que as amostras provém de universos distintos e assume que existe um valor de diferença das médias que é estatisticamente significativo (MISHRA et al., 2019).

Então, o teste t aplicado em dois conjuntos de dados é um teste estatístico usado para determinar se o valor médio de uma amostra é estatisticamente igual (H_0) ou diferente (H_A) do valor médio de sua população-mãe de onde foi retirada a amostra. O valor t calculado é a razão da diferença da média e erro padrão, ou desvio padrão (DP) (MISHRA et al., 2019).

Dessa forma, é um teste utilizado para comparar as médias entre dois grupos onde não existe a necessidade de múltiplas comparações, resultando em valor de p único. Para aplicação do teste é necessário que a amostra seja de variável contínua e distribuição normal, para uma amostra maior que 30 indivíduos (MISHRA et al., 2019).

Dependendo das características da amostra o teste pode ser pareado ou não pareado. O teste t não pareado, também chamado de teste t independente, é um teste estatístico inferencial que determina se há alguma diferença estatística significativa entre as médias de dois grupos não relacionados. Já o teste t pareado, ou teste t dependente, serve para determinar se a diferença nas médias entre duas observações pareadas (anterior e posterior, por exemplo) é estatisticamente significativa. Para esse teste, os mesmos sujeitos são medidos em dois instantes diferentes ou observados por dois métodos diferentes, geralmente são descritos como variáveis emparelhadas (o antes e o depois do mesmo assunto ou intervenção) (MISHRA et al., 2019).

Silva et al. (2018) utilizou teste t na sua pesquisa para avaliar se existiam diferenças estatisticamente significativas nos valores das variáveis obtidas através da análise de componente principal (PCA) com o objetivo de diferenciar grupos de pacientes leucêmicos e saudáveis, observando se há diferenças entre os grupos (DA SILVA et al., 2018).

2.7.2. Análise de variância (ANOVA)

A análise de variância (ANOVA) é um método estatístico utilizado como teste de hipóteses para comparação das médias entre grupos, sendo aplicado para comparar amostras com três ou mais grupos. Um valor p significativo do teste ANOVA indica para pelo menos um par, entre os quais a diferença média foi estatisticamente significativa. Para identificar que par foi significativo realiza-se um teste de comparações múltiplas (*post hoc*) (MISHRA et al., 2019). Para definição do teste a ser aplicado deve-se observar a distribuição da variável de estudo. Quando é possível assumir a normalidade de distribuição da variável, pode ser utilizado um teste paramétrico. Caso contrário utiliza-se o teste não paramétrico. Assim, deve-se primeiramente verificar a hipótese de normalidade da distribuição, para então decidir qual análise estatística utilizar na comparação dos grupos (CAPP e NIENOV, 2020).

Para utilização da ANOVA, além de se considerar a normalidade e o número de grupos a serem comparados (três ou mais), outros dois pré-requisitos devem ser observados: a homogeneidade das variâncias obtida pelo teste de Levene e a independência das observações definida pelo delineamento do estudo. A ANOVA tem como objetivo verificar se existe diferença significativa entre pelo menos duas médias dos grupos. Para tal, a ANOVA considera como hipótese nula (H_0) que não há diferença entre as médias dos grupos (médias = $\mu_A = \mu_B = \dots = \mu_N$), e como hipótese alternativa (H_A) que há diferença entre pelo menos duas das médias dos grupos ($\mu_A \neq \mu_B$). Se H_0 não for rejeitada ($p > 0,05$), conclui-se que não há evidência de diferença significativa entre as médias dos grupos. Em contrapartida, se H_0 for rejeitada ($p \leq 0,05$), aceita-se H_A e passa-se então a se comparar os grupos dois a dois através de testes *post hoc* para identificar quais médias são significativamente diferentes, já que pelo menos duas das médias dos grupos diferem entre si (CAPP e NIENOV, 2020).

A ANOVA gera um valor chamado F, que compara a quantidade de variância sistemática dos dados com a quantidade de variância não sistemática. No entanto, o valor de F indica se houve diferença global entre as

médias, mas não determina quais pares de grupos possuem médias significativamente diferentes. Para isso, precisa-se dos testes *post hoc* (CAPP e NIENOV, 2020; MISHRA et al., 2019).

Quando se utiliza o teste ANOVA com um fator para amostras que seguem distribuição normal, os testes *post hoc* recomendados são o teste de Tukey e de Bonferroni com a comparação par a par das categorias (níveis) do fator para a variável dependente, sendo o teste de Tukey o mais rigoroso. O teste de Kruskal-Wallis substitui o teste ANOVA com um fator quando se tratam de dados não paramétricos e, para verificar qual dos pares se diferencia, o teste de Kruskal-Wallis apresenta uma análise de *post hoc* de Dunn (CAPP e NIENOV, 2020; MISHRA et al., 2019).

2.7.3. Análise de componente principal (PCA)

Pesquisas que trabalham com grande número de dados são cada vez mais comuns e muitas vezes difíceis de interpretar. A análise do componente principal (PCA) é uma técnica matemática desenvolvida para reduzir a dimensionalidade desses conjuntos de dados, aumentando a interpretabilidade dos dados, ao mesmo tempo minimizando a perda de informações (JOLLIFFE e CADIMA, 2016; RINGNÉR, 2008).

A PCA identifica novas variáveis, os componentes principais, que são combinações lineares das variáveis originais do conjunto de dados. Os componentes principais são autovetores normalizados da matriz de covariância e ordenados de acordo com a variação (variância) presente nos dados. O primeiro componente principal (PC1) é descrito como a direção ao longo da qual as amostras mostram a maior variação, já o segundo componente principal (PC2) a direção não correlacionada com o primeiro componente ao longo do qual as amostras mostram a segunda maior variação, e a extração das variações segue sucessivamente (PC3, PC4, etc.) até o componente de menor variância, de forma que nenhuma variação apresentada seja redundante (SILVA et al., 2018). Sendo assim, cada componente principal pode ser interpretado como a direção não correlacionada com os componentes

anteriores, e que possuem ainda a maior variação remanescente nas amostras. Desta forma, conclui-se que a PCA é uma ferramenta capaz de identificar direções de maior variação para separar classes de amostras, mostrando diferentes aspectos dos dados (JOLLIFFE e CADIMA, 2016; RINGNÉR, 2008).

Por tratar-se de uma técnica adaptativa, a PCA é útil para agrupar ou classificar amostras; no entanto, definir quantos e quais componentes usar na análise subsequente é uma grande dificuldade que pode ser abordada de várias maneiras. E os resultados da PCA dependem criticamente do pré-processamento dos dados e da seleção de variáveis (RINGNÉR, 2008).

Silva et al. (2018) estudaram amostras de sangue total e plasma sanguíneo por espectroscopia Raman e descreveram a utilização de PCA para obter as informações mais significativas, baseadas na variância, do conjunto de dados originais. A partir do qual se geraram duas novas variáveis, chamadas de componentes principais (PCs) e vetores de carregamento (Escore – SCs), onde cada vetor de carregamento (SC), que se assemelha ao espectro Raman, apresenta um peso, o componente principal (PC), indicando a intensidade de cada Escore presente nos dados originais. Com essas duas variáveis, as semelhanças e as diferenças entre os grupos podem ser identificadas, sendo as maiores diferenças identificadas no PC1 (SILVA et al., 2018).

2.7.4. Análise discriminante utilizando regressão por mínimos quadrados parciais (PLS-DA)

Historicamente, o algoritmo de regressão por mínimos quadrados parciais (PLS) foi proposto para lidar com dados contendo muitas variáveis, reduzindo os preditores a um conjunto menor de componentes não correlacionados e efetuar a regressão por mínimos quadrados para esses componentes no lugar dos dados originais (RUIZ-PERES et al., 2020). O método de análise discriminante utilizando a regressão por mínimos quadrados parciais (PLS-DA) é uma ferramenta de classificação de amostras em grupos que se utiliza da técnica multivariada PLS visando a redução de dimensionalidade e, fundamentalmente, engloba dois procedimentos principais:

(a) obtenção dos componentes da PLS, que é a redução de dimensão; e (b) construção do modelo de previsão, que é a análise discriminante. Na tarefa de classificação, as variáveis de saída sempre serão categóricas, e para isso, o primeiro passo na modelagem PLS-DA é recodificar as variáveis categóricas (ordinal ou nominal) em variáveis contínuas (ou seja, numérica), geralmente o valor 1 é atribuído à classe de interesse e o valor 0 é atribuído à outra classe (LEE et al., 2018; SANTANA et al., 2020).

Uma etapa importante para a elaboração do modelo PLS-DA é a escolha correta do número de variáveis latentes. A escolha da quantidade de variáveis deve ser adequada, pois quando é escolhido um número insuficiente de variáveis latentes, ocorre uma falta de ajuste no modelo (subajuste), dessa forma, não é utilizada toda a informação útil para construção do modelo. Em contrapartida, quando é escolhido um número maior de variáveis latentes, tem-se o sobreajuste, onde são adicionadas informações não relacionadas ao objeto de interesse na construção do modelo, como por exemplo, ruídos espectrais. Normalmente, a escolha do número de variáveis latentes é realizada através da validação cruzada, na qual uma parte das amostras do conjunto de calibração é separada e utilizada para validação interna. Em sequência, é construído o modelo de treinamento utilizando diferentes números de variáveis latentes e são previstas as amostras de validação interna para cada um dos modelos construídos, os erros obtidos são armazenados e o processo é refeito até que todas as amostras de treinamento sejam previstas. Na PLS-DA analisam-se a porcentagem de amostras classificadas corretamente em cada classe na validação cruzada (SANTANA et al., 2020). Para avaliar a precisão da classificação das amostras foi utilizado o método de validação cruzada deixa-um-de-fora (do inglês *leave-one-out*), que é um caso especial de validação cruzada *k-fold* onde $k = n$, o número de observações (MEIJER e GOEMAN, 2013).

Durante a construção dos modelos multivariados, o pré-processamento dos dados é importante, e a literatura descreve vários tipos de pré-processamentos como normalização dos espectros, correção de dispersão multiplicativa (do inglês *multiplicative scatter correction – MSC*), variação

normal padrão (do inglês standard normal variate – SNV), entre outros. A escolha do pré-processamento deve ser realizada em função da técnica analítica, tipo de amostra e método quimiométrico empregado (SANTANA et al., 2020).

A PLS-DA tem sido utilizado principalmente na modelagem de dados espectrais, como Raman e espectros no infravermelho, pois, os dados espectrais são de alta dimensionalidade e muitas variáveis tendem a se correlacionar entre si (RUIZ-PERES et al., 2020).

3. MATERIAIS E MÉTODOS

3.1. Comitê de Ética

O protocolo utilizado nesse estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Anhembi Morumbi – UAM (Parecer No. 5.170.2013 CAEE 23457019.9.0000.5492 – ANEXO) atendendo às diretrizes brasileiras para pesquisa em seres humanos. Todos os pacientes submetidos à pesquisa assinaram um Termo de Consentimento Livre e Esclarecido (TCLE), que informava todo o procedimento a ser realizado, bem como riscos e benefícios da pesquisa.

3.2. Seleção dos pacientes e obtenção das amostras

Foram incluídos no estudo pacientes voluntários crianças e adolescentes que foram submetidos a tratamento oncológico no Hospital do Centro de Tratamento Infantojuvenil Fabiana Macedo de Moraes (CTFM), mantido pelo Grupo de Apoio à Criança com Câncer (GACC) de São José dos Campos, SP. Também foram incluídos na pesquisa crianças e adolescentes que seriam submetidos a exames de rotina no Hospital Municipal Infantil Menino Jesus de São Paulo, SP, para coleta de amostras de sujeitos saudáveis. Foram coletadas as seguintes informações dos pacientes: anamnese, exame físico e exames subsidiários que permitam a identificação dos pacientes por: sexo, idade, cor e diagnóstico. Nos casos oncológicos foram estabelecidos os diagnósticos anatomopatológicos, estadiamento e a situação clínica no momento da inclusão no estudo: pré-tratamento (indução), em tratamento quimioterápico (QT), pós-tratamento QT, manutenção e se realizou transplante. Os dados da fase do tratamento foram organizados na Figura 5, baseada nas informações obtidas nos prontuários, sendo que alguns prontuários não informavam a fase de tratamento que o paciente se encontrava. Os diagnósticos histopatológicos foram obtidos dos prontuários. A

idade dos pacientes variou de 0 a 23 anos com média de $9,5 \pm 5,39$ anos (Figura 6).

Para o estudo foram coletadas 121 amostras de sangue. Do total de amostras, 107 foram de pacientes em tratamento no CTFM/GACC, onde 50 amostras de pacientes com diagnóstico de leucemia, 43 amostras com diagnóstico de outros tipos de câncer, como: blastomas, carcinomas, linfomas e sarcomas, e um paciente foi diagnosticado com leucemia e sarcoma. Seis pacientes não tinham diagnóstico no prontuário, três amostras eram provenientes de pacientes com outras patologias além do câncer e quatro amostras não foram coletadas por terem material insuficiente para captura do espectro. Um total de 14 amostras de sujeitos saudáveis foi obtido do Hospital Menino Jesus.

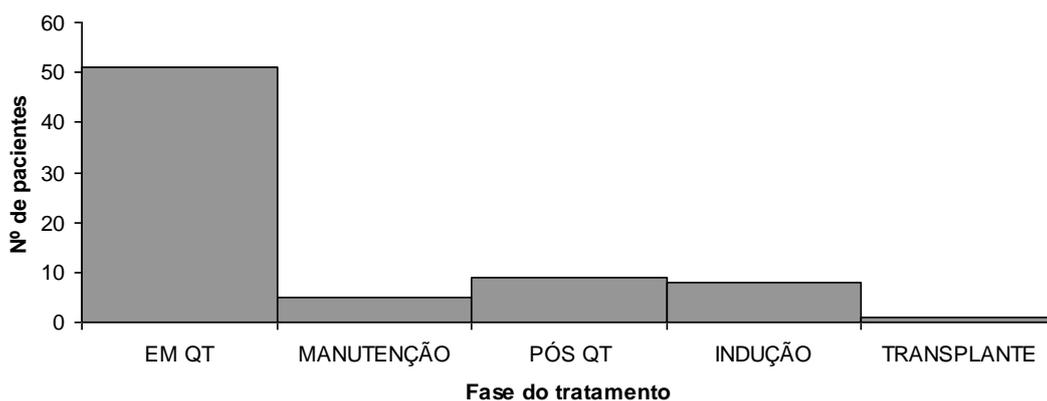


Figura 5: Histograma das fases do tratamento em que os pacientes se encontravam.

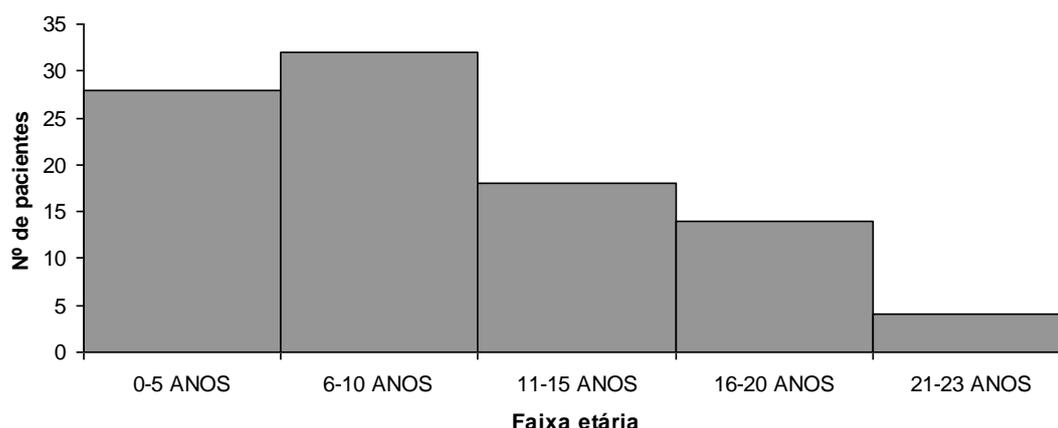


Figura 6: Histograma da faixa etária dos pacientes oncológicos.

A coleta de sangue foi realizada no momento em que os indivíduos foram submetidos à coleta de exames de sangue de rotina. Foram coletados 5 mL de sangue periférico de cada paciente em um tubo com ativador de coágulo mais gel. Após a coagulação do sangue, os tubos foram centrifugados a 3.000 rpm por 15 minutos obtendo-se a separação do soro e das hemácias, ficando o soro na porção superior e as células na porção inferior do gel e então os tubos foram mantidos em freezer (-14°C a -25°C) até o dia da coleta do espectro Raman.

3.3. Espectrômetro Raman e coleta dos espectros

Foi utilizado um espectrômetro Raman dispersivo no infravermelho próximo (modelo Dimension P1, Lambda Solutions Inc., MA, EUA) que utiliza laser de diodo em 830 nm como fonte de excitação; potência: ajustável até 480 mW; resolução: 2,5 cm^{-1} ; faixa de comprimento de onda: de 400 a 1800 cm^{-1} ; câmera CCD de alta sensibilidade refrigerada (-75°C). O software de controle permite ajuste da temperatura, potência do laser, tempo de coleta do sinal Raman e correção da resposta em intensidade do espectrômetro. O equipamento utiliza uma sonda Raman de fibra óptica (modelo Vector probe, Lambda Solutions, Inc.). A potência utilizada na saída da sonda foi ajustada de

200 a 250 mW nas amostras de sangue para preservar a integridade da amostra.

A calibração do número de onda do espectrômetro foi verificada obtendo um espectro de naftaleno e comparando a posição do pico com os da literatura; a resposta de intensidade do espectrômetro foi determinada usando lâmpada de calibração (lâmpada de tungstênio para calibração da curva de resposta de intensidade) e aplicada a cada espectro antes do pré-processamento; esses procedimentos foram descritos com maiores detalhes em Silveira et al. (2002).

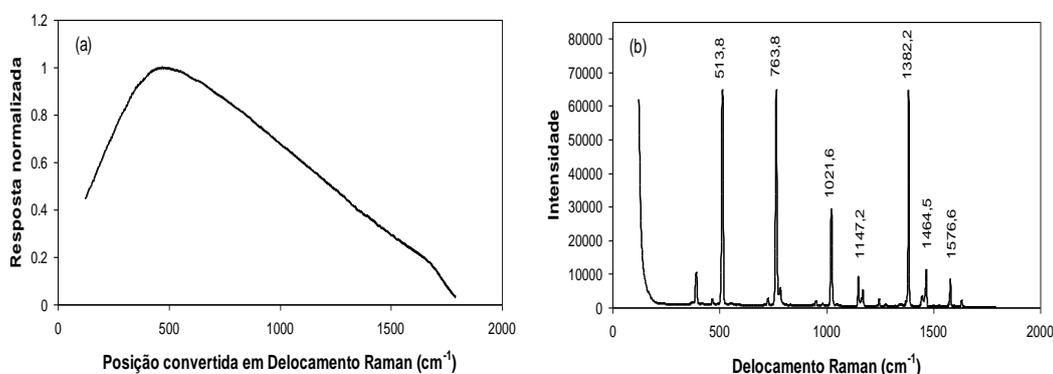


Figura 7: Curvas de resposta do espectrômetro Raman: (a) intensidade obtida com lâmpada de calibração de tungstênio; (b) número de onda obtida com o composto orgânico naftaleno.

3.4. Coleta dos espectros Raman

No dia da coleta dos espectros Raman, as amostras foram descongeladas em temperatura ambiente e a análise realizada em 100 μ L de células sanguíneas. Para a coleta das células sanguíneas, a porção inferior do tubo de ensaio foi cortada com bisturi para que não houvesse mistura de células sanguíneas com fragmentos de soro ou gel; em seguida, 100 μ L de células sanguíneas foram pipetados e colocados no porta-amostras previamente higienizado com água e sabão e álcool 70%, para evitar contaminação entre as amostras. O tempo de exposição do laser para a obtenção dos espectros foi de 3 s com 10 acumulações por espectro (tempo

total de exposição de 30 s), e para cada amostra foram coletados entre 3 e 5 espectros variando a posição do laser na amostra movendo lateralmente o porta-amostras, visando aumentar o número de espectros no modelo de diagnóstico.

Todos os espectros foram pré-processados para análise prévia. A resposta espectral do espectrômetro foi inicialmente corrigida; a fluorescência de fundo foi removida ajustando e subtraindo uma função polinomial de 7^a ordem em uma rotina implementada no software Matlab 7.01 (The Mathworks Inc., Natick, MA, EUA); os picos de raios cósmicos foram removidos manualmente; e a normalização foi realizada pela área sob a curva do espectro (norma 1).

Inicialmente, foram coletadas 107 amostras de sangue de pacientes oncológicos do Hospital CTFM / GACC, quatro amostras não tiveram os espectros coletados por não terem material suficiente para coleta. Do total de amostras disponíveis para coleta, obteve-se 309 espectros provenientes de amostras de sangue de pacientes oncológicos, destes 18 espectros foram excluídos por não terem diagnóstico, 9 espectros foram excluídos por apresentarem outras patologias, 3 espectros foram excluídos pelo fato do paciente apresentar tanto câncer leucêmico com câncer não leucêmico e 12 espectros foram excluídos devido à baixa qualidade do sinal Raman e presença de muito ruído, o que poderia comprometer o cálculo das médias e identificação dos principais picos Raman. Resultando em 267 espectros Raman de amostras de sangue de pacientes oncológicos. Do total de 14 amostras de sangue saudáveis, obteve-se 42 espectros de amostras saudáveis, sendo um espectro excluído devido à baixa qualidade do sinal Raman e presença de muito ruído, resultando em 41 espectros Raman de amostras de sangue de pacientes saudáveis. Totalizaram 308 espectros para análise e interpretação. A Tabela 1 apresenta uma visão geral dos dados apresentados.

Tabela 1: Número de amostras de cada grupo histopatológico e número de espectros Raman considerados em cada grupo.

| Diagnóstico histopatológico | Número de amostras | Número de espectros |
|-----------------------------|--------------------|---------------------|
|-----------------------------|--------------------|---------------------|

| | | | |
|---------------|-----------|--------------|---------------|
| Normal | | 14 amostras | 41 espectros |
| Leucêmico | | 50 amostras | 141 espectros |
| | Sarcoma | 18 amostras | 51 espectros |
| | Blastoma | 16 amostras | 48 espectros |
| Não leucêmico | Linfoma | 6 amostras | 18 espectros |
| | Carcinoma | 3 amostras | 9 espectros |
| Total | | 107 amostras | 308 espectros |

3.5. Identificação dos picos e aplicação do teste ANOVA

Os espectros médios dos grupos normal, leucêmicos e não leucêmicos foram plotados no software Microsoft Excel 2003 com abscissas entre 400 a 1800 cm^{-1} , indicando o deslocamento Raman, para visualizar as diferenças de intensidade dos picos Raman entre os grupos normal, leucêmicos e não leucêmicos. Os picos principais foram identificados e então organizados em uma tabela baseada na composição bioquímica presente nas amostras de células sanguíneas e nas atribuições das bandas Raman. O teste ANOVA foi aplicado na intensidade dos picos do espectro da diferença para determinar se a diferença entre os grupos é estatisticamente significativa, como forma de identificar se existe diferença na composição bioquímica entre os indivíduos saudáveis e doentes. Para a análise estatística, a hipótese nula afirma que as médias dos grupos provêm da mesma população. O nível de significância para rejeitar a hipótese nula foi definido como 5% ($p < 0,05$) (MISHRA et al., 2019).

3.6. Análise exploratória por análise de componente principal (PCA)

A PCA é uma ferramenta matemática capaz de identificar novas variáveis, os componentes principais, que são combinações lineares das variáveis originais do conjunto de dados. Para diminuir a dimensão dos dados e ter melhor condição de identificar as reais diferenças entre os grupos, descartaram-se os componentes principais com as menores variâncias, enfatizando as informações mais relevantes (primeiros PCs) para serem

utilizadas como parâmetros de diagnóstico ou classificação. Com essas duas variáveis PCs e SCs, as semelhanças e diferenças entre os grupos podem ser identificadas (SILVA et al., 2018). Dessa forma, a PCA é uma ferramenta capaz de identificar direções de maiores variações e direções úteis para separar amostras com diferentes características espectrais. Para realizar a análise exploratória, os PCs e SCs foram calculados usando o software Chemoface (NUNES et al., 2012), (<https://www.ufla.br/chemoface/>) e o conjunto de dados de espectros Raman.

O teste ANOVA foi aplicado nas variáveis PCs para comparar as médias dos valores entre os grupos normais, leucêmicos e não leucêmicos, procurando diferenças estatisticamente significativas entre eles ($p < 0,05$). Antes de realizar o teste ANOVA entre os PCs, foi realizado o teste de normalidade de Kolmogorov-Smirnov para verificar se cada um dos PCs segue uma distribuição normal e o teste de Levene para avaliar a homogeneidade da amostra. Para amostras com distribuição Gaussiana, foi aplicado o teste ANOVA com um fator (paramétrico) com teste *post hoc* Tukey e, para amostras que não seguiram uma distribuição normal, foi aplicado o teste de Kruskal-Wallis (ANOVA não paramétrico), com teste *post hoc* de Dunn. As análises estatísticas foram realizadas no software InStat (GraphPad Software, San Diego, CA, EUA).

3.7. Análise discriminante por mínimos quadrados parciais (PLS-DA)

A análise discriminante utilizando mínimos quadrados parciais (PLS-DA) é uma ferramenta de redução de dimensionalidade multivariada e envolve dois procedimentos principais: a construção do componente PLS, que é a redução de dimensão; e a construção do modelo de previsão e classificação, que é a análise discriminante. A PLS-DA combina redução de dimensionalidade e análise discriminante em um modelo matemático e é preferencialmente aplicável na modelagem de dados de alta dimensão. Tem sido utilizada para modelos espectrais de discriminação e classificação, pois são dados com alta dimensionalidade e as variáveis tendem a ser

correlacionadas (LEE et al., 2018; RUIZ-PERES et al., 2020). Para modelar o problema de discriminação, as variáveis PLS foram calculadas com base nos espectros Raman normalizados por SNV; para análise discriminante, as classes foram denominadas como normais (classe = 1), leucêmicos (classe = 2) e não leucêmicos (classe = 3). A análise discriminante também foi feita utilizando as seguintes classes: normais (classe = 1), leucêmicos (classe = 2), sarcoma (classe = 3), blastoma (classe = 4) e linfoma + carcinoma (classe = 5), com o objetivo de avaliar a capacidade de discriminação dentro do grupo não leucêmico. E também foi feita a análise discriminante considerando apenas duas classes, normais (classe = 1) e cânceres (classe = 2), onde foram agrupados juntos todos os tipos de câncer para obter a discriminação entre saudáveis e cânceres. A análise estatística foi realizada no software Chemoface.

Para avaliar a validade da classificação foi construída uma tabela de confusão para cada agrupamento realizado e calculado os valores de sensibilidade, especificidade e acurácia. A sensibilidade é a probabilidade de o teste ser positivo dado que o indivíduo é doente, representando a habilidade do teste em identificar os verdadeiros doentes. A especificidade é a probabilidade de o teste ser negativo dado que o indivíduo não é doente, ou seja, é a habilidade de identificar corretamente os verdadeiros doentes. E, a acurácia é a habilidade do teste em separar os indivíduos doentes dos saudáveis (PÉREZ et al., 2021).

4. RESULTADOS

4.1. Espectroscopia Raman

A Figura 8 mostra os espectros Raman médios das células sanguíneas dos grupos normais, leucêmicos e não leucêmicos; os picos Raman mais intensos foram marcados em 527, 621, 675, 753, 804, 829, 859, 900, 935, 1003, 1032, 1128, 1159, 1174, 1225, 1312, 1342, 1373, 1450, 1553, 1565, 1585, 1606, 1622 e 1659 cm^{-1} , correspondendo principalmente a picos de vibrações de proteínas e aminoácidos. A Tabela 2 apresenta esses picos organizados por compostos bioquímicos acompanhados de sua respectiva atribuição conforme descrito na literatura; interpretações bioquímicas mais detalhadas são discutidas na análise exploratória por PCA.

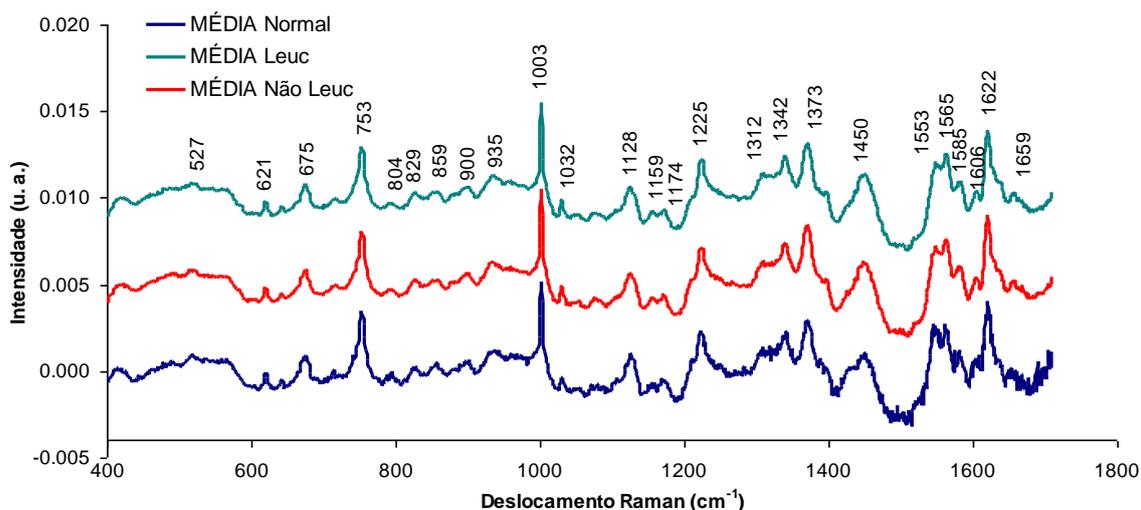


Figura 8: Espectros Raman médios normalizados de células sanguíneas de grupos normais, leucêmicos e não leucêmicos. Os espectros foram deslocados para melhor visualização.

Tabela 2: Posições dos picos nos espectros Raman das células sanguíneas e suas respectivas atribuições de bandas vibracionais e elementos bioquímicos de acordo com a literatura citada. Os picos entre parênteses são os que aparecem na análise exploratória por PCA (Figura 10). Pag. 56 a 64.

| Deslocamento Raman (cm ⁻¹) | Atribuição vibracional | Elemento bioquímico | Referências |
|--|--|-------------------------|--|
| (500) | S–S disulfeto str | Colágeno | (TALARI et al., 2014) |
| (571) | $\nu(\text{Fe–O}_2)$ | Trp, C, G | (BANKAPUR et al., 2010; TALARI et al., 2014) |
| 621 | $\nu_{12}(\text{CC torção})$ | Phe | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014) |
| (631) | $\nu(\text{C–S})$ <i>gauche</i> | Met | (BANKAPUR et al., 2010; TALARI et al., 2014) |
| 671 | ν_7 cis (C–S str) | T, G, Trp | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017) |
| 675 (677-78) | ν_7 cis (C–S str) C–C torção | T,G, proteínas, Met | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; MANAGÒ et al., 2018b) |
| (705) | $\nu(\text{C–S})$ trans | Met | (TALARI et al., 2014) |
| 753 | Sym br | Proteínas, Trp, T, A, C | (NEUGEBAUER |

| | | | |
|--------------|------------------------------|---------------------------------|--|
| (755) | v_{15} (modo br porfirina) | | ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| 829 | O–P–O str DNA bk | Tyr, Trp, glutaciona, lípidios | (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| 856 (860) | $v(C-C)$ | Tyr, proteínas | (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; TALARI et al., 2014) |
| (895) | C–C esquelético | Proteínas, Tyr, Trp, glutaciona | (BANKAPUR et al., 2010; TALARI et al., 2014; DA SILVA et al., 2018) |
| 900 | C–C esquelético | Proteínas, Tyr, Trp, glutaciona | (BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; DA |

| | | | |
|----------------|---|---------------------------------------|---|
| 935 | V_{46} DNA bk C–C str | Ácido nucleico, proteínas | SILVA et al., 2018) (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; MANAGÒ et al., 2018a) (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017) |
| (974) | CH ₃ def | RNA, lípidios | (BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017) |
| (993) | $\nu(C_{\beta}-C_1)_{asym}$ V_{45} | Phe | (ATKINS et al., 2017; DA SILVA et al., 2018; MANAGÒ et al., 2018b; HAUSSON et al., 2018; TALARI et al., 2014) (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI |
| 1003 (1004) | C–C esquelético Sym anel br | Phe; β -caroteno; grupo heme | |
| 1032 | CH def no plano $\delta(=C_bH_2)_{asym}$ | Phe, proteínas | |

| | | | |
|-------------------|---------------------------|---|--|
| 1128 (1125-27) | V_5 C–N str bk | Proteínas, lipídios, fosfolipídios, citocromo c | et al., 2014) (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| 1159 | V_{44} C–C e C–N str | β -caroteno, lipídios, proteínas | (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| 1174 (1173) | V_{30} C–H no plano | T, Tyr | (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017) |
| 1211 | =CH def | Trp, Phe, Tyr, amida III, fosfolipídios | (BANKAPUR et al., 2010; TALARI et al., 2014; |

| | | | |
|----------------------|---|--|--|
| 1225 (1223) | V_{13} Or V_{42} β -folha estrutura | Proteínas, amida III, lipídios | ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) (BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) |
| (1291) | C–H def angular | C | (TALARI et al., 2014) (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; MANAGÒ et al., 2018b) |
| 1312 | V_{21} CH def CH ₂ /CH ₃ torção | G, colágeno, lipídios, amida III | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; MANAGÒ et al., 2018b) |
| 1342 (1341) | V_{41} CH def CH ₂ /CH ₃ torção | Fosfolipídios A, G, proteínas, lipídios, citocromo c | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| 1371-73 (1369-76) | V_4 CH ₃ str sym | Glucosamina (carboidrato), T, A, citocromo c, lipídios | (NEUGEBAUER et al., 2010a; BANKAPUR |

| | | | |
|-------------------|---|---|---|
| | | | et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) |
| (1397) | V_{20} C=O sym str | | (BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014) |
| (1403) | CH def | Glutaciona | (TALARI et al., 2014; DA SILVA et al., 2018) |
| 1450 (1451-54) | CH ₂ /CH ₃ def angular | Proteínas, lípidios, fosfolípidios, aminoácidos | (NEUGEBAU ER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018; HAUSSON et al., 2018) |
| 1521 | V_{38} -C=C- | Carotenóides | (BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017) |
| 1543 (1544-46) | $\nu_{11}(C_{\beta}C_{\beta})$ | Amida II | (BANKAPUR et al., 2010; |

| | | | |
|----------------|--|-------------------------------|--|
| 1553 | V_{11} $\nu(\text{C}=\text{C})$ | Trp, amida II | TALARI et al., 2014; ATKINS et al., 2017) (PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018) (BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014) |
| 1565 | V_{19} COO^- | Proteínas, Trp | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) |
| 1585 (1587) | V_{37} $\text{C}=\text{C}$ olefínico str | Proteínas, Tyr, G, A | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) |
| 1606 (1607) | $\nu(\text{C}=\text{C})_{\text{vinyl}}$ $\text{C}=\text{C}$ | Proteínas, Tyr, Phe, lípidios | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; DA SILVA et al., 2018) |

| | | | |
|----------------|---|--|--|
| (1613-16) | C=C | Proteínas, Tyr, Trp | (MANAGÒ et al., 2018a; MANAGÓ et al., 2018b; BANKAPUR et al., 2010; TALARI et al., 2014) (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018; HAUSSON et al., 2018) |
| 1622 | $\nu(\text{C}=\text{C})_{\text{vinyl}}$ (modo do grupo heme) | Tyr, Trp, lipídios | (BANKAPUR et al., 2010; TALARI et al., 2014; DA SILVA et al., 2018) |
| (1637) | Bandas α -hélice e β -folha | Tyr, Trp, amida I | (BANKAPUR et al., 2010; TALARI et al., 2014; DA SILVA et al., 2018) |
| 1640 (1642) | α -hélice ν_{10} | Amida I | (BANKAPUR et al., 2010; TALARI et al., 2014; ATKINS et al., 2017) |
| 1659 | C=C str (anel aromático) Modos de DNA/RNA | Amida I, proteínas, lipídios, fosfolipídios, T, G, C | (NEUGEBAUER et al., 2010a; BANKAPUR et al., 2010; PREMASIRI et al., 2012; TALARI et al., 2014; |

ATKINS et al., 2017; MANAGÒ et al., 2018b; DA SILVA et al., 2018)

Abreviaturas: Met - Metionina, Phe – fenilalanina, Try – tirosina, Trp – triptofano, T – timina, G – guanina, A – adenina, C – citosina, sym – simétrica, asym – assimétrica, br – respiração, str – alongamento, def – deformação, bk – backbone.

A Figura 9 mostra o espectro da diferença entre as amostras leucêmicas e normais, onde identificou-se os picos 671 cm^{-1} (ácidos nucleicos, aminoácidos – alongamento C-S), 756 cm^{-1} (proteínas e ácidos nucleicos), 1215 cm^{-1} (amida II, triptofano, fenilalanina, tirosina, fosfolipídios – deformação de CH), 1430 cm^{-1} (proteínas), 1531 cm^{-1} (β -caroteno) e 1548 cm^{-1} (triptofano e amida II) atribuídos principalmente a bases nitrogenadas e triptofano, com maior intensidade nas amostras normais, e o pico 1457 cm^{-1} atribuído a aminoácidos, proteínas e lipídios, com maior intensidade no leucêmico. O espectro da diferença entre as amostras não leucêmicas e normais mostrou picos identificados em 753 , 1430 , 1531 cm^{-1} atribuídos principalmente a proteínas, com maior intensidade no grupo normal e o pico 1368 cm^{-1} atribuído a carboidratos, ácido nucleicos e lipídios, com maior intensidade no grupo não leucêmico. Todos os picos identificados no espectro da diferença apresentaram diferença estatisticamente significativa ao aplicar o teste ANOVA, exceto o pico 1430 cm^{-1} .

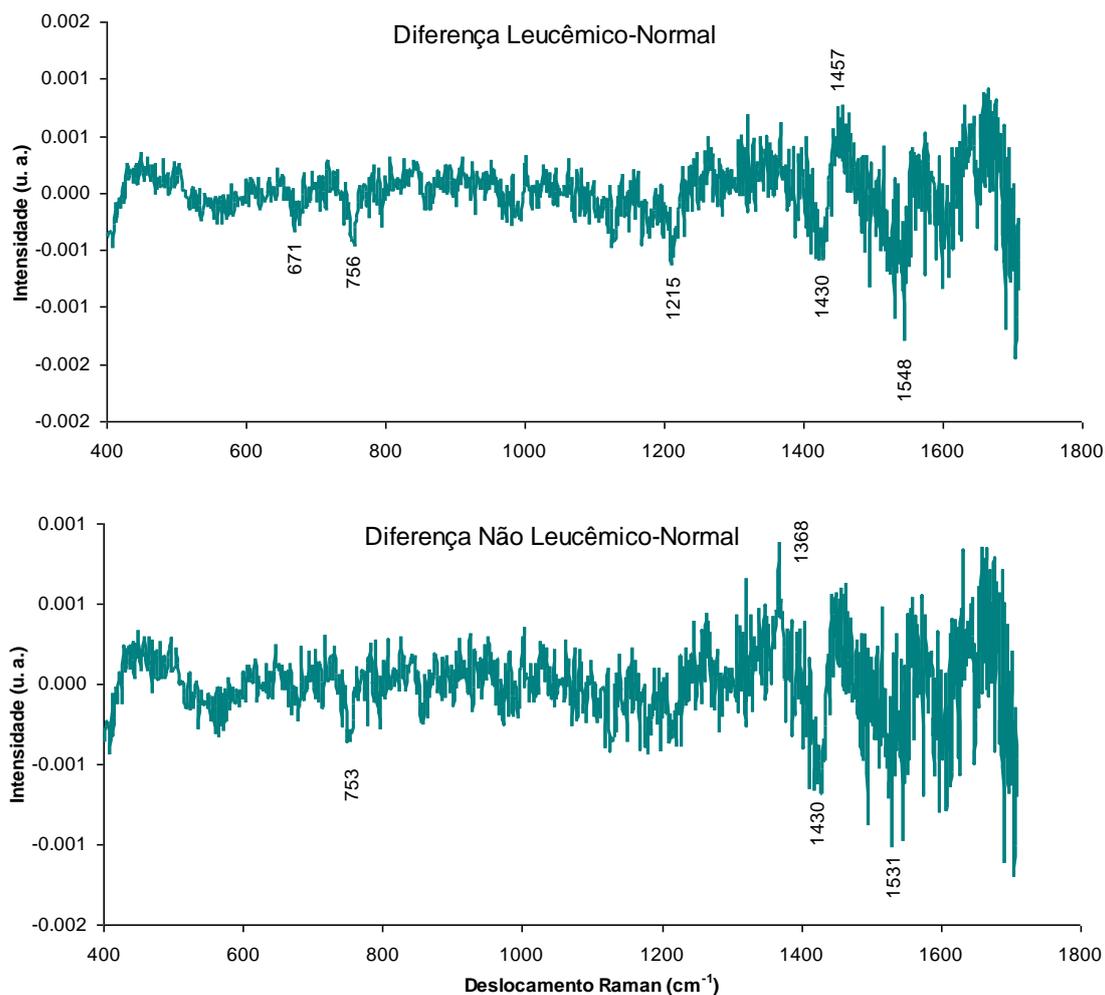


Figura 9: Espectros das diferenças entre os espectros médios normalizados dos grupos Leucêmico – Normal e dos grupos Não Leucêmico – Normal.

4.2. Análise exploratória por PCA

A PCA foi utilizada para identificar diferenças espectrais entre os grupos normal, leucêmico e não leucêmico; essas diferenças estavam relacionadas a alterações na constituição bioquímica das células sanguíneas devido a leucemias e outras neoplasias. A Figura 10 mostra a plotagem das variáveis da PCA (Escores – variâncias espectrais apresentadas nos dados espectrais e PCs – intensidades de cada Escore nos dados originais) das amostras de células sanguíneas, nas quais o teste ANOVA foi aplicado nos PCs para identificar diferenças estatisticamente significativas ($p < 0,05$). As

primeiras 10 variáveis da PCA são vistas na Figura 10. Essas 10 variáveis respondem por 94,2% da variância total observada nos dados espectrais (85,2%, 4,5%, 1,7%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% e 0,2% para PC1 a PC10, respectivamente).

O Escore 1 mostra os picos 621, 675, 753, 829, 943, 1003, 1032, 1128, 1225, 1343, 1373, 1452, 1553, 1564, 1622 cm^{-1} , representando características de células sanguíneas, principalmente eritrócitos, comum a todas as amostras analisadas. O PC1 mostrou maior intensidade para os grupos leucêmicos e não leucêmicos quando comparados ao grupo normal, e apresentou significância com $p < 0,0001$ (ANOVA). A significância foi observada ao comparar o grupo normal com o grupo leucêmico e o grupo normal com o grupo não leucêmico, $p < 0,001$ (teste de Tukey); a comparação entre os grupos leucêmico e não leucêmico não apresentou significância, $p > 0,05$ (teste de Tukey).

No Escore 2 observou-se picos característicos de proteínas que possuem enxofre (S) em sua constituição, presentes em ligações C-S e S-S, representados pelo picos positivos: 570, 678, 755, 900, 942, 1004, 1225, 1401, 1451, 1565, 1587, 1607, 1640 e 1659 cm^{-1} , correspondendo ao grupo leucêmico. E, os picos negativos possuem contribuição tanto do grupo normal como do grupo não leucêmico e revelam características espectrais do citocromo c, através dos picos em 975, 1125, 1217, 1370 e 1544 cm^{-1} . Apesar, das diferenças observadas o PC2 não apresentou significância ($p > 0,05$). O Escore 3 exibiu características espectrais predominantes de aminoácidos, principalmente aminoácidos com ligações C-S e S-S, caracterizados pelos picos em 573, 678, 755, 866, 897, 929, 993, 1110, 1132, 1173, 1223, 1305, 1430 e 1545 cm^{-1} e apresentando maior intensidade para o grupo normal; e, o picos em 1006, 1125, 1247, 1359, 1459 e 1676 cm^{-1} , correspondem principalmente a aminoácidos com anéis aromáticos, e apresentam maior intensidade no grupo leucêmico. O PC3 apresentou significância com $p < 0,0001$ (teste ANOVA), e o teste *post hoc* demonstrou significância entre os grupos normal e leucêmico e entre o grupo normal e não leucêmico (teste de Tukey, $p < 0,001$); a comparação entre os grupos leucêmico e não leucêmico não apresentou significância ($p < 0,05$). No Escore 4 observou-se

características espectrais do citocromo c e aminoácidos, caracterizadas pelos picos em 633, 749, 855, 1004, 1127, 1212, 1433, 1531 e 1614 cm^{-1} com maior intensidade no grupo normal. Os picos em 1226, 1374 e 1563 cm^{-1} caracterizam principalmente proteínas, com maior intensidade nos grupos leucêmico e não leucêmico em relação ao grupo normal. O PC4 apresentou significância com $p < 0,0001$ (ANOVA não paramétrico) e no teste *post hoc*, quando se comparou os grupos entre si houve significância nas comparações entre o grupo normal e leucêmico e entre o grupo normal e não leucêmico (teste de Dunn, $p < 0,001$). O Escore 5 apresentou picos característicos de metionina, tirosina e triptofano, em 631, 705, 860, 1226, 1291, 1381, 1397 e 1616 cm^{-1} , com maior intensidade no grupo leucêmico e, os picos em 750, 1004, 1213, 1430 e 1547 cm^{-1} , representam as características principalmente da fenilalanina, com maior intensidade no grupo normal. O PC5 não apresentou diferença estatisticamente significativa entre os grupos ($p > 0,05$, ANOVA). O Escore 6 mostrou características espectrais de aminoácidos com picos em 755, 1004, 1211, 1427 e 1545 cm^{-1} . O PC6 apresentou significância com $p < 0,001$ (ANOVA não paramétrico), demonstrando significância apenas ao comparar o grupo normal e o grupo leucêmico ($p < 0,001$, teste de Dunn).

O Escore 7 apresentou características espectrais de lipídios, com picos em 1211, 1367 e 1617 cm^{-1} , apresentando maior intensidade no grupo não leucêmico, e o pico 570 cm^{-1} que corresponde ao triptofano, apresentou maior intensidade para o grupo normal. O PC7 demonstrou que houve diferença estatisticamente significativa ($p < 0,0001$, ANOVA não paramétrico), apresentando diferenças significativas no pós teste de Dunn para a comparação entre o grupo normal e leucêmico e grupo normal e não leucêmico ($p < 0,001$); ao comparar grupo leucêmico e grupo não leucêmico obteve-se valor $p > 0,05$. O Escore 8 demonstrou picos referentes triptofano e proteínas caracterizados pelos picos em 576, 752 e 1132 cm^{-1} , com maior intensidade no grupo normal e, os picos em 1003, 1209, 1371 e 1450 cm^{-1} , que representam lipídios e fenilalanina, com maior intensidade no grupo não leucêmico. O PC8 mostrou que houve diferença significativa ($p < 0,0001$, ANOVA paramétrico), apresentando significância na comparação entre o grupo normal e leucêmico e

entre o grupo normal e não leucêmico ($p < 0,001$, teste *post hoc* de Tukey). O Escore 9 apresentou picos característicos com citocromo c, em 1128 e 1370 cm^{-1} , com maior intensidade no grupo normal e o pico em 1453 cm^{-1} (torção CH_2/CH_3) com maior intensidade no grupo leucêmico. O PC9 mostrou significância com $p < 0,0001$ (teste ANOVA paramétrico), demonstrando significância nas comparações entre os grupos ($p < 0,001$, teste *post hoc* de Tukey), exceto entre o grupo leucêmico e não leucêmico ($p > 0,05$). O Escore 10 apresentou picos característicos de proteínas e citocromo c em 675, 755, 1006 e 1371 cm^{-1} , que apresentaram maior intensidade no grupo normal. O PC10 revelou diferença significativa com $p < 0,01$ (ANOVA não paramétrico), e o teste *post hoc* de Dunn, apresentou diferença significativa apenas ao comparar o grupo normal com o grupo leucêmico ($p < 0,01$).

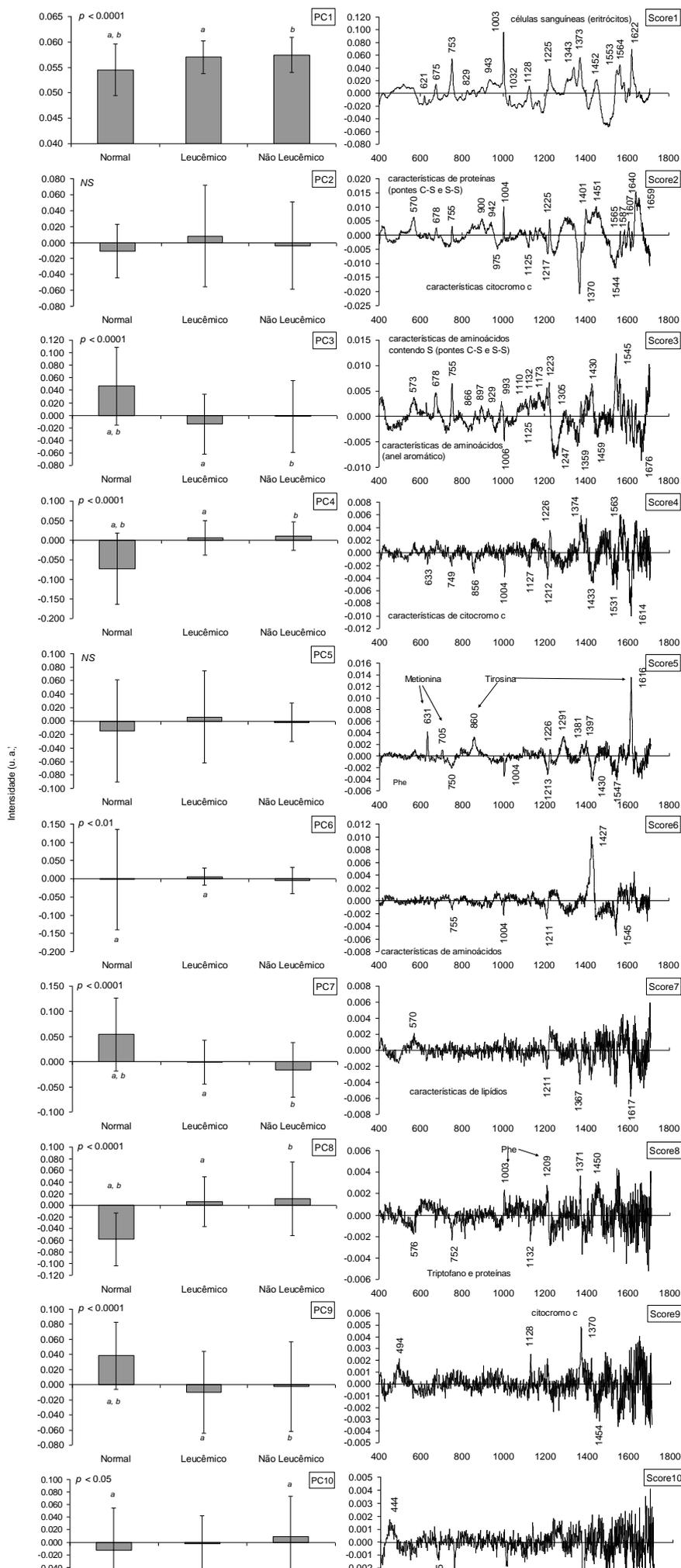


Figura 10: Plotagem das variáveis PCA (PCs e Escores) para amostras de células sanguíneas para análise exploratória. NS = não significativo. Letras sobrescritas a e b nos PCs indicam significância estatística.

4.3. Análise discriminante por PLS-DA

A PLS-DA foi aplicada em todo o conjunto de dados espectrais para diferenciar amostras de células sanguíneas normais (saudáveis), leucêmicas e não leucêmicas (sarcomas, blastomas, linfomas e carcinomas), onde o número de variáveis latentes a serem modeladas foi definido de acordo com o maior acerto de discriminação (precisão) com menor número de variáveis. Para aplicação do modelo de discriminação, utilizou-se o pré-processamento SNV. A Tabela 3 mostra os resultados de classificação para discriminação de três classes: grupo normal, grupo leucêmico e grupo não leucêmico, obtidos a partir de oito variáveis latentes da PLS, alcançando sensibilidade de 100%, especificidade de 87,8% e acurácia global de 67,5%. A Tabela 4 apresenta o resultado de classificação a partir da organização de duas classes: grupo normal e grupo câncer (incluindo todos os pacientes com câncer), sendo utilizadas oito variáveis latentes da PLS para a discriminação. Dessa forma, obteve-se sensibilidade de 100%, especificidade de 90,2% e acurácia global de 98,7%. Para a construção da Tabela 5 considerou-se cinco classes para discriminação, que foram: grupo normal, grupo leucêmico, grupo sarcoma, grupo blastoma e grupo linfoma + carcinoma. Para aplicação da discriminação utilizou-se seis variáveis latentes da PLS, alcançando sensibilidade de 100%, especificidade de 90,2% e acurácia global de 58,1%.

Tabela 3: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em três classes: pacientes saudáveis (normal), pacientes com leucemia (leucêmico) e pacientes com outros tipos de câncer (não leucêmico).

| Classificação histopatológica | Classificação por PLS-DA | | |
|-------------------------------|--------------------------|-----------|---------------|
| | Normal | Leucêmico | Não leucêmico |
| Normal ($n = 41$) | 36 | 2 | 3 |
| Leucêmico ($n = 141$) | 0 | 93 | 48 |
| Não leucêmico ($n = 126$) | 0 | 47 | 79 |
| Sensibilidade* | | 100% | |
| Especificidade* | | 87,8% | |
| Acurácia | | 67,5% | |

*Valores de sensibilidade, especificidade e acurácia foram calculados considerando dois grupos, normal *versus* câncer (leucêmico + não leucêmico).

Tabela 4: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em duas classes, pacientes saudáveis (normal) e pacientes com câncer.

| Classificação histopatológica | Classificação por PLS-DA | |
|-------------------------------|--------------------------|--------|
| | Normal | Câncer |
| Normal ($n = 41$) | 37 | 4 |
| Câncer ($n = 267$) | 0 | 267 |
| Sensibilidade | | 100% |
| Especificidade | | 90,2% |
| Acurácia | | 98,7% |

Tabela 5: Tabela de confusão com os resultados de sensibilidade, especificidade e acurácia (classificação correta) para o modelo de discriminação por PLS utilizando o espectro Raman de células sanguíneas agrupando as amostras em cinco classes, pacientes saudáveis (normal), pacientes com leucemia (leucêmico) e pacientes com sarcoma, blastoma e linfoma + carcinoma.

| Classificação histopatológica | Classificação por PLS-DA | | | | | % de acertos |
|-------------------------------|--------------------------|-----------|---------------|--------|---------|--------------|
| | Normal | Leucêmico | Não leucêmico | | | |
| | | | Sarcom | Blasto | Linfoma | |

| | | | a | ma | + Carcino ma | | |
|------------------------------------|--|-----|----|-------|--------------------|-------|-------|
| Normal ($n = 41$) | 37 | 4 | 0 | 0 | 0 | 90.2% | |
| Leucêmico ($n = 141$) | 0 | 128 | 5 | 8 | 0 | 90.7% | |
| | Sarcoma ($n = 51$) | 0 | 42 | 4 | 5 | 0 | 7.8% |
| Não leucêmi co ($n = 126$) | Blastoma ($n = 48$) | 0 | 36 | 2 | 10 | 0 | 22.8% |
| | Linfoma + Carcino ma ($n = 27$) | 0 | 22 | 2 | 3 | 0 | 0% |
| | Sensibilidade* | | | 100% | | | |
| | Especificidade* | | | 90,2% | | | |
| | Acurácia | | | 58,1% | | | |

*Valores de sensibilidade, especificidade e acurácia foram calculados considerando dois grupos, normal *versus* câncer (leucêmico + não leucêmico).

5. DISCUSSÃO

5.1. Espectroscopia Raman

As bandas representadas pelos picos em 675, 753, 829, 1003, 1128, 1159, 1174, 1225, 1342, 1373, 1450, 1553, 1565, 1585, 1622 e 1659 cm^{-1} (Figura 8) encontradas nos espectros médios de células sanguíneas do grupo normal (saudável), grupo leucêmico e grupo não leucêmico apresentam características predominantes de hemoglobina (ATKINS et al., 2017). As células sanguíneas são caracterizadas pela presença de eritrócitos que contêm hemoglobina, importante proteína responsável pelo transporte de oxigênio no sangue (NAYAK e RAI, 2017). Premasiri et al. (2012) demonstraram que os espectros Raman do sangue total e das hemácias (eritrócitos) são praticamente idênticos, devido à contribuição espectral da hemoglobina. Sato et al. (2001) analisaram hemácias e concluíram que as características obtidas pelo espectro Raman derivam da presença de hemoglobina. Dochow et al. (2011) mostraram que os picos em 1002, 1374, 1564, 1584 e 1621 cm^{-1} nos espectros Raman das hemácias contêm contribuições Raman de ressonância devido ao grupo heme da hemoglobina. Segundo Silva et al. (2018), os picos em 570, 678 e 755 cm^{-1} , picos na faixa entre 820 e 920 cm^{-1} e picos em 1004, 1130, 1160, 1212, 1225, 1344, 1378, 1401, 1452, 1552, 1567, 1586, 1623, 1640 e 1658 cm^{-1} podem ser atribuídos principalmente a proteínas, aminoácidos, lipídios e β -caroteno, ao analisar amostras de sangue total em pacientes leucêmicos em comparação com pacientes saudáveis, demonstrando que o espectro Raman médios das células sanguíneas também sofre influências das bandas espectrais descritas anteriormente.

Analisando a Figura 9, o espectros da diferença entre o grupo leucêmico e o grupo normal apresenta os picos negativos em 671 cm^{-1} (ácidos nucleicos, aminoácidos – alongamento C–S), 756 cm^{-1} (proteínas e ácidos nucleicos), 1215 cm^{-1} (amida II, triptofano, fenilalanina, tirosina, fosfolipídios – deformação de CH), 1430 cm^{-1} (proteínas – deformação de CH_2), 1531 cm^{-1} (β -caroteno) e 1548 cm^{-1} (triptofano e amida II) atribuídos principalmente a bases

nitrogenadas e triptofano, com maior intensidade nas amostras normais e o pico positivo em 1457 cm^{-1} atribuído a aminoácidos, proteínas e lipídios, com maior intensidade no grupo leucêmico (HASSOUN et al., 2018; MANAGÒ et al., 2018b; NEUGEBAUER et al., 2010a; PREMASIRI et al., 2012; SILVA et al., 2018). Happillon et al. (2015) descreveram que o pico em 1456 cm^{-1} representa maior concentração de proteínas no grupo leucêmico ao comparar amostras células de linfócitos com leucemia linfocítica crônica e linfócitos saudáveis. Silva et al. (2018) mostraram que os picos em 678, 755, 1212, 1378, 1552 e 1623 cm^{-1} , atribuídos principalmente a proteína, aminoácidos e amida, foram maiores para o grupo saudável no espectro de diferença entre grupos leucêmicos e saudáveis. Neugebauer et al. (2010b) apresentaram o espectro de diferença entre os espectros médios de células saudáveis e tumorais e observaram picos positivos em 711, 756, 792, 1000, 1376, 1543 e 1602 cm^{-1} que correspondem a células saudáveis (leucócitos). No espectro da diferença entre as amostras não leucêmicas e normais (Figura 9), os picos negativos identificados foram 753, 1430, 1531 cm^{-1} atribuídos principalmente a proteínas, com maior intensidade no grupo normal e o pico positivo 1368 cm^{-1} atribuído a carboidratos, ácidos nucleicos e lipídios, com maior intensidade no grupo não leucêmico (HASSOUN et al., 2018; MANAGÒ et al., 2018b; NEUGEBAUER et al., 2010a; PREMASIRI et al., 2012; SILVA et al., 2018). Observou-se que os picos positivos em 1368 e 1457 cm^{-1} no espectro de diferença caracterizavam o grupo não leucêmico e leucêmico, respectivamente, demonstrando maior concentração de proteína e aminoácidos nestes grupos, com diferença estatisticamente significativa entre os grupos (ANOVA, $p < 0,05$).

Os picos em 671 e 756 cm^{-1} no espectro de diferença (Figura 9) estão em maior intensidade no grupo normal e podem ser atribuídos principalmente a proteínas e ácidos nucleicos (DNA). Chan et al. (2008) atribuíram os picos em 785, 1093, 1373 e 1575 cm^{-1} ao DNA e mostraram que eles são de menor intensidade em células B leucêmicas quando comparados a amostras de células normais avaliadas por pinças ópticas Raman. As pinças ópticas têm sido amplamente estudadas para análise de células sanguíneas (CHAN et al., 2008; NEUGEBAUER et al., 2010b), por serem uma técnica pontual, onde o

feixe de laser incide diretamente sobre a células, o espectro Raman reflete apenas a composição bioquímica específica da região incidida, não podendo se considerar uma representação da composição bioquímica total da célula (NAVAS-MORENO e CHAN, 2018). Para garantir que o espectro reflita com precisão toda a composição bioquímica da célula, deve-se garantir que as células sejam aproximadamente do mesmo tamanho do volumes focais do laser (CHAN, 2012). Contrastando a informação encontrada no espectro da diferença, autores descrevem que o aumento nos níveis de DNA são alterações características de células neoplásicas (CHAN et al., 2008; NEUGEBAUER et al., 2010b; ATKINS et al., 2017). O aumento dos níveis de ácidos nucleicos em amostras de sangue de pacientes com câncer pode ser explicado por mecanismos de apoptose, que aumentam os níveis de DNA circulante (ATKINS et al., 2017). Happillon et al. (2015) mostraram no espectro de diferença entre células linfócitos LLC e linfócitos saudáveis que os picos positivos em 796, 1106, 1380, 1498 e 1585 cm^{-1} sugeriram maior concentração de DNA em células neoplásicas de linfócitos B em comparação com as saudáveis.

5.2. Análise exploratória por PCA

A interpretação das informações obtidas através da PCA consiste na identificação do picos presentes nos primeiros PCs e relacioná-los com os componentes bioquímicos de cada grupo baseado na literatura e em espectros de padrões. Picos positivos com PC positivos, assim como picos negativos com PC negativo, mostram que o elemento bioquímico específico apresenta maior concentração no grupo ao qual corresponde. A Figura 10 mostra os gráficos de PCs e Escores para as amostras de células sanguíneas. O Escore 1 apresentou picos característicos de amostras de células sanguíneas, principalmente hemoglobina (ATKINS et al., 2017; BANKAPUR et al., 2010; NEUGEBAUER et al., 2010b) e o PC1 mostra que o grupo normal apresentou menor concentração de hemoglobina/eritrócitos, com diferença significativa ($p < 0,0001$). O Escore 2 apresentou características principalmente de proteínas

com ligações C–S e S–S, sem significância estatística entre os grupos ao analisar a intensidade dos picos pelo PC2 ($p > 0,05$). O Escore 3/PC3 (Figura 10) mostrou picos positivos correspondendo a aminoácidos com ligações C–S e S–S como metionina e cisteína (573 e 755 cm^{-1}), com maior intensidade no grupo normal, e picos negativos relacionados a aminoácidos com anéis aromáticos, como a fenilalanina (1006 cm^{-1}), com maior intensidade no grupo leucêmico ($p < 0,0001$). Silva et al. (2018) mostraram diferença estatisticamente significativa na presença do pico do aminoácido fenilalanina ($1003 - 1007\text{ cm}^{-1}$), apresentando maior intensidade do grupo leucêmico ao comparar amostras de sangue total de pacientes leucêmicos e saudáveis. Managò et al. (2018b) mostraram uma redução na intensidade do pico de fenilalanina em células B leucêmicas (MN60) de pacientes tratados com quimioterapia em comparação com células não tratadas, sugerindo que essa redução de intensidade se deve ao efeito do tratamento por quimioterapia, pois foi observada uma redução gradativa nas intensidades dos picos Raman relacionados aos ácidos nucleicos e proteínas proporcional ao aumento da concentração da dose dos quimioterápicos.

O Escore 4 apresentou picos negativos característicos do citocromo c, com maior intensidade no grupo normal ao avaliar o PC4. Em Féré et al. (2020), os picos em 745 , 755 , 1120 , 1134 , 1370 e 1577 cm^{-1} foram atribuídos ao citocromo c. Abramczyk et al. (2022) avaliaram alterações no estado redox dos citocromos mitocondriais em fragmentos de tecido tumoral de câncer de mama e células mamárias humanas in vitro de células normais, células malignas leves e células altamente agressivas por meio de espectroscopia Raman e de imagens com excitação de 532 nm ; encontraram correlação entre o aumento do pico do citocromo c em 1584 cm^{-1} e a agressividade do câncer de mama (este pico apresentou alta intensidade devido ao efeito de aumento da ressonância). O aumento na intensidade do pico indica mudanças progressivas no estado redox à medida que o câncer se torna mais agressivo. O citocromo c, que é uma proteína essencial para o transporte de elétrons mitocondrial (HUTTEMANN et al., 2011; MANAGÒ et al., 2018b), pode ser um marcador bioquímico que discrimina entre estados saudáveis e doentes, pois

desempenha um papel central na apoptose e no ciclo respiratório celular, de modo que variações em suas bandas podem mostrar disfunções nos níveis de apoptose (FÉRÉ et al., 2019; HUTTEMANN et al., 2011). O Escore 9 com picos 1128 e 1370 cm^{-1} e o Escore 10 com picos 755 e 1370 cm^{-1} são característicos do citocromo c e possuem maior intensidade no grupo normal ao avaliar seus respectivos PCs, ambos com significância estatística ($p < 0,05$, ANOVA). Cheng et al. (2022) analisaram células B normais e células B de leucemia linfóide aguda (LLA) por microespectroscopia Raman (785 nm) e encontraram diferenças principais nos picos em 714-773 cm^{-1} e 1220-1360 cm^{-1} , onde o picos em 755, 1310, 1337 e 1370 cm^{-1} foram atribuídos ao citocromo c. Eles encontraram características do citocromo c mais altas nos grupos de células LLA-B quando comparadas ao grupo normal (Cheng et al., 2022), um resultado diferente do encontrado no presente estudo, onde as características do citocromo c foram encontradas em células sanguíneas normais. O aumento do citocromo c em amostras de células sanguíneas normais encontradas no presente estudo pode ser objeto de estudos futuros; o citocromo c está envolvido na cadeia de transporte de elétrons e na apoptose celular e sua regulação afeta a atividade metabólica da célula, tornando a célula mais resistente a sinais de estresse que, em condições normais, induziriam a morte celular (HUTTEMANN et al., 2011).

O Escore 5 apresentou picos positivos em 631, 705, 860, 1226, 1291, 1381, 1397 e 1616 cm^{-1} que podem ser atribuídos principalmente aos aminoácidos metionina, tirosina e triptofano, e o PC5 apresentou maior intensidade no grupo leucêmico, e apresentou picos negativos em 750, 1004, 1213, 1430, 1547 cm^{-1} , atribuídos principalmente a fenilalanina, e ao analisar o PC5 evidenciou-se maior intensidade no grupo normal, porém sem significância estatística ($p > 0,05$). O Escore 6 apresentou características de aminoácidos nos picos em 755, 1004, 1211 e 1545 cm^{-1} ; apesar da significância estatística encontrada no PC6 ($p < 0,001$), as intensidades são muito próximas, demonstrando que são características comuns a todas as amostras e podem não ser tão relevantes para diferenciação entre os grupos. O Escore 7 apresentou picos negativos com características predominantes de lipídios, com

picos em 1211, 1357 e 1617 cm^{-1} , com maior intensidade no grupo não leucêmico ao analisar o PC7. Os picos positivos, como o 570 cm^{-1} , são atribuídos principalmente a proteínas e apresentaram maior intensidade no grupo normal a partir do PC7 com ($p < 0,0001$). Abramczyk et al. (2022) encontraram que as células cancerosas epiteliais e o lúmen do ductos mamários com câncer são ricos em lipídios saturados e apresentaram uma proporção anormal entre ácidos graxos saturados e insaturados nas células epiteliais do ducto canceroso, que gera uma instabilidade mecânica na membrana celular. Dessa forma, demonstraram que as membranas das células são o alvo primário de lesões e, a funcionalidade celular depende da composição lipídica e do ambiente celular (ABRAMCZYK et al., 2022). O Escore 8 apresentou picos positivos em 1003, 1209, 1371 e 1450 cm^{-1} que são característicos de fenilalanina e lipídios, apresentando, pela análise do PC8, maior intensidade no grupo não leucêmico com significância estatística ($p < 0,0001$), assim como no PC7. Neugebauer et al. (2010a) encontraram um pico negativo na banda de 1004 cm^{-1} (fenilalanina), que apresentou maior intensidade para células leucêmicas quando comparadas com células leucocitárias. Silva et al. (2018) sugeriram que a fenilalanina pode desempenhar um papel importante nas leucemias agudas devido a sua presença nos primeiro carregamento do PC, além disso, esse aminoácido é descrito como um dos ativadores do gene BCR-ABL, que se apresenta em uma classe de doenças como leucemia mieloides crônicas e outros distúrbios mieloproliferativos. O picos negativos em 576 e 752 cm^{-1} mostrados no Escore 8 correspondem ao triptofano e, ao analisar o PC8 apresentam maior intensidade no grupo normal ($p < 0,0001$).

Os picos negativos apresentados do Escore 9, correspondem a ligações com deformação CH_2 característicos de proteínas, com maior intensidade no grupo leucêmico pelo PC9 ($p < 0,0001$). Chan et al. (2008) mostraram picos em 1126, 1605 e 1615 cm^{-1} , associados a proteínas, com aumento de intensidade nos espectros de células T e pré-B leucêmicas quando comparadas com células normais. Silva et al. (2018) mostraram que o pico em 1658 cm^{-1} apresentou diferença significativa e maior intensidade no grupo

leucêmico quando comparado ao grupo saudável, caracterizando um aumento na concentração de proteínas, e indica que alterações nesses picos correspondem a características alterações das leucemias. O PC10 apresentou diferença estatisticamente significativa apenas ao se comparar o grupo normal e o grupo não leucêmico pelo pós teste de Dunn, com $p < 0,01$.

5.3. Análise discriminante de mínimos quadrados parciais (PLS-DA).

A PLS-DA foi aplicada em todo o espectro para diferenciar amostras de células sanguíneas normais (saudáveis), leucêmicas e não leucêmicas (outros tipos de câncer), onde o número de variáveis latentes foi modelado de acordo com a maior precisão de discriminação com menor número de variáveis. Foram construídas três tabelas de confusão a partir de diferentes agrupamentos entre as amostras a fim de avaliar a discriminação entre os grupos. A Tabela 3 mostra os resultados de classificação para discriminação de três classes, normal, leucêmico e não leucêmico, obtidos a partir de oito variáveis latentes, alcançando sensibilidade de 100%, pois o método foi capaz de classificar todos os doentes como doentes, sem confundir com o grupo normal. A especificidade foi 87,8%, pois houve erros de classificação, onde três amostras normais foram classificadas como leucêmica e duas amostras normais foram classificadas como não leucêmicas. E acurácia global de 67,5%, pelo fato de terem ocorrido erros de classificação entre os grupos leucêmico e não leucêmico. Neugebauer et al. (2010a), que classificaram células de leucemia (OCI-AML3), células de câncer de mama (BT-20, MCF-7) e cultura de células de leucócitos por máquina de vetor de suporte (do inglês support vector machine - SVM) aplicada aos espectros Raman, apresentaram acurácia de 99,7% e obteve apenas nove erros de classificação entre as células de câncer, enquanto o agrupamento realizado na Tabela 3 obteve-se 95 erros de classificação entre o grupo leucêmico e não leucêmico. Dochow et al. (2011) relataram um modelo de classificação entre leucócitos, eritrócitos e células tumorais OCI-AML3, MCF-7 e BT-20 por análise discriminante linear (LDA) nos espectros Raman e mostraram uma acurácia geral de 92,2%. Os elevados valores de acurácia

global de Dochow et al. (2011) e Neugebauer et al. (2010a) podem ser justificadas por terem utilizado células cultivadas, de forma que o espectro Raman não sofre influência de outros hemocomponentes, explicando o menor número de erros de classificação. Além disso, no presente trabalho o número de erros nas classificações entre os grupos leucêmico e não leucêmico foi alto, demonstrando que as amostras de células sanguíneas leucêmicas e não leucêmicas apresentam semelhanças na composição e consequentemente nos espectros Raman, o que dificulta a discriminação com maior precisão. Conforme descrito na análise exploratória por PCA, as principais características espectrais correspondem aos eritrócitos/hemoglobina se sobrepõem às características espectrais dos leucócitos (ATKINS et al., 2017; PREMASIRI et al., 2012 SATO et al., 2001), isso pode explicar a dificuldade de discriminação entre as amostras, uma vez que as principais alterações leucêmicas ocorrem nos leucócitos (KAPLAN, 2019; SETH e SINGH, 2015). Managò et al. (2016) apresentou modelo de classificação foi validado usando o método de validação cruzada validação cruzada deixa-um-de-fora (do inglês *leave-one-out*) onde buscou discriminar células B normais e células tumorais (RS4;11, REH e MN60) através de espectros obtidos por microespectroscopia Raman confocal. As células MN60 foram completamente separadas das demais, o que mais uma vez confirma que este modelo de discriminação é mais diferenciado, o que resultou em uma eficiência de 97,3% (MANAGÒ et al. 2016; 2018b).

A Tabela 4 apresenta o resultado de classificação a partir da organização de duas classes, um grupo normal e o outro grupo que incluiu todos os pacientes com câncer e, obteve-se sensibilidade de 100%, especificidade de 90,2% e acurácia global de 98,7%. Silva et al. (2018), ao analisar amostras de sangue total de grupo saudável e grupo leucêmico através da PLS-DA, obteve classificação semelhante, com 91,9% de sensibilidade, 100% de especificidade e 96,5% de acurácia. A Tabela 4 mostra que houveram apenas 4 erros de classificação, nos quais as amostras normais foram confundidas com amostras doentes, justificando a especificidade não ser 100%, enquanto no trabalho de Silva et al. (2018) todas as amostras normais foram classificadas corretamente, apresentando 5 amostras leucêmicas

classificadas erroneamente como normais. Neugebauer et al. (2010b) construiu um tabela de contingência para classificar em duas classes a partir da aplicação de SVM, separando células normais de células tumorais e obteve valor de sensibilidade de 99,7% e especificidade de 99,4%, mostrando apenas quatro erros de classificação: duas amostras de células tumorais foram classificadas como doentes e duas amostras de células normais foram classificadas como células tumorais, o que resultou em uma acurácia global de 92,4%. Managò et al. (2018a) procurou discriminar células B normais de células leucêmicas MN60 sem tratamento quimioterápico e células MN60 submetidas a três diferentes doses do medicamento quimioterápico. A partir dos resultados encontrados através da PCA, obteve cerca de 91% de acurácia, com uma especificidade de 100%, pois todas as amostras de células B normais foram classificadas corretamente. Recentemente, Lima et al. (2023) usaram PLS-DA para discriminar amostras de soro sanguíneo de crianças e adolescentes normais daqueles com leucemia e cânceres não leucêmicos, mostrando precisão de 66%, sensibilidade de 99% e especificidade de 57%; apresentaram mais erros de classificação do que os encontrados no presente estudo.

A Tabela 5 mostra cinco classes para discriminação, que foram: grupo normal, grupo leucêmico, grupo sarcoma, grupo blastoma e grupo linfoma e carcinoma, alcançando sensibilidade de 100%, especificidade de 90,2% e acurácia global de 58,1%. Analisando a Tabela 5 observou-se que todas as amostras de linfoma foram confundidas com outros tipos de câncer, sendo em sua maior parte confundido com as amostras leucêmicas, isso pode ser explicado porque os linfomas assim como as leucemias acometem linfócitos B e T, por isso podem apresentar alterações bioquímicas semelhantes (MELARAGNO e CAMARGO, 2013).

Uma limitação deste trabalho foi a falta de homogeneidade na quantidade de amostras nos grupos de outros tipos de câncer e no grupo normal, de forma que o grupo leucêmico se torna maior quando comparado com os demais. Esse fator pode justificar a dificuldade de distinção entre os tipos de câncer e os baixos valores de acurácia.

A busca por métodos diagnósticos precisos é constante e a literatura mostra que a espectroscopia Raman é uma ferramenta promissora para análise de tecidos biológicos. sangue para diagnóstico de câncer (ATKINS et al., 2017; GONZÁLEZ-SOLÍS, 2019; LIMA et al., 2019; LIMA et al., 2023; SILVA et al., 2018). Muitos benefícios têm sido observados ao comparar técnicas convencionais com a análise de padrões espectrais obtidos através da Raman espectroscopia para o diagnóstico de câncer. É um método não invasivo ou minimamente invasivo, que não requer preparo de amostras com pequenas quantidades de sangue e pode gerar um espectro Raman capaz de diferenciar pacientes saudáveis de pacientes com câncer com base na composição bioquímica das amostras (GOURLAT et al., 2021; LIMA et al., 2019; NEUGEBAUER et al., 2010a; SILVA et al., 2018; SILVEIRA et al., 2012). A espectroscopia Raman é uma técnica promissora para diagnosticar e monitorar o câncer, pois é uma técnica não destrutiva e rica em informações, oferecendo perspectivas para complementar as técnicas de diagnóstico existentes, detectando estágios iniciais do câncer e monitorando os medicamentos usados em o tratamento (CHAN et al., 2008; CHAN et al., 2009; DOCHOW et al., 2011; FAZIO et al., 2016; FÉRÉ et al., 2020; GONZÁLEZ-SOLÍS et al., 2014; GONZÁLEZ-SOLÍS, 2019; HAPPILLON et al., 2015; HASSOUN et al., 2018; KHETANI et al., 2015; LIMA et al., 2023; MACLAUGHLIN et al., 2013a; MACLAUGHLIN et al., 2013b; MANAGÒ et al., 2016; MANAGÒ et al., 2018a; MARTÍNEZ-ESPINOSA et al., 2009; NEUGEBAUER et al., 2010a; NEUGEBAUER et al., 2010b; SILVA et al., 2018; VANNA et al., 2015; ZHANG et al., 2015). A partir dos resultados apresentados e da literatura existente, pode-se concluir que a espectroscopia Raman e a discriminação PLS podem ser uma técnica importante para ajudar na diferenciação e diagnóstico de amostras cancerígenas usando amostras de células sanguíneas, mostrando resultados promissores ao diferenciar indivíduos normais (saudáveis) daqueles com leucemia e outros tipos de câncer (sarcoma, blastoma e linfoma + carcinoma).

Um ponto negativo que pode ser destacado é o fato das alterações bioquímicas decorrentes do processo de oncogênese ocorrerem de maneira

semelhante em diferentes órgãos e tecidos do corpo, as alterações na constituição do sangue decorrentes do câncer podem se assemelhar bioquimicamente e, em última análise, comprometer a discriminação entre os diferentes tipos de câncer, a literatura tem demonstrado uma discriminação mais eficaz entre os tipos de câncer ao utilizar pinças ópticas(DOCHOW et al., 2011). Leszczenko et al. (2021) sugere melhorias como aumentar o conjunto de dados com o objetivo de eliminar a variabilidade da amostra individual, minimizando assim os erros. Em contrapartida, o trabalho demonstra a real possibilidade de aplicação clínica da técnica de espectroscopia Raman para diferenciar amostras de sangue de pacientes saudáveis e com câncer.

Dessa forma, sugere-se que essa pesquisa seja ampliada para outras faixas etárias e diferentes tipos de câncer, com o objetivo de aumentar ainda mais o banco de dados de espectros Raman de amostras de sangue de pacientes tanto saudáveis como pacientes oncológicos, de forma a tornar o método ainda mais confiável.

6. CONCLUSÃO

O presente estudo teve como objetivo identificar as alterações bioquímicas que ocorrem em crianças e adolescentes com leucemia e outros tipos de câncer (não leucêmicos) e compará-los ao grupo normal (saudáveis) e diferenciar os grupo normal, leucêmico e não leucêmico, utilizando PCA para análise exploratória e PLS para análise discriminante. Através do espectro da diferença entre as amostras leucêmicas e normais, identificou-se os picos em 671 cm^{-1} (ácidos nucleicos, aminoácidos – alongamento C–S), 756 cm^{-1} (proteínas e ácidos nucleicos), 1215 cm^{-1} (amida II, triptofano, fenilalanina, tirosina, fosfolipídios – deformação de CH), 1430 cm^{-1} (proteínas), 1531 cm^{-1} (β -caroteno) e 1548 cm^{-1} (triptofano e amida II) atribuídos principalmente a bases nitrogenadas e triptofano, com maior intensidade nas amostras leucêmicas e o pico em 1457 cm^{-1} atribuído a aminoácidos, proteínas e lipídios, com maior intensidade no grupo normal. O espectro da diferença entre as amostras não leucêmicas e normais, os picos identificados foram em 753 , 1430 e 1531 cm^{-1} atribuídos principalmente a proteínas, com maior intensidade no grupo não leucêmico, e o pico em 1368 cm^{-1} atribuído a carboidratos, ácidos nucleicos, lipídios e citocromo c, com maior intensidade no grupo normal. O Escore 1 mostra os picos em 621 , 675 , 753 , 829 , 943 , 1003 , 1032 , 1128 , 1225 , 1343 , 1373 , 1452 , 1553 , 1564 , 1622 cm^{-1} , que representam os eritrócitos (hemoglobina), comum a todas as amostras analisadas. A análise da plotagem dos PCs e Escores permitiu concluir que existe uma maior concentração de citocromo c nas amostras normais, sendo o citocromo c um importante marcador bioquímico que discrimina entre estados saudáveis e doentes, por ser uma proteína essencial para o transporte de elétrons mitocondrial. Também pode-se observar maior concentração de proteínas e fenilalanina (1004 cm^{-1}) no grupo leucêmico.

A análise discriminante realizada pelo modelo PLS-DA entre os grupos Normal, Leucêmico e Não Leucêmico alcançou acurácia de 67,5%, com sensibilidade de 100% e especificidade de 87,8% quando considerada a

classificação de Normal versus Câncer. Apesar do maior número de erros de classificação na diferenciação dos tipos de câncer (leucêmicos e não leucêmicos), os resultados sugerem a eficácia da espectroscopia Raman para diferenciar amostras saudáveis de cancerosas usando amostras de células sanguíneas com alta precisão.

Considerando os recentes avanços da espectroscopia Raman visando o diagnóstico de câncer (AUNER et al., 2018; GONZALEZ-SOLIS, 2019; LIMA et al., 2023; SILVA et al., 2018), a aplicação direta dessa ferramenta na prática clínica requer estudos adicionais em diferentes amostras e condições para validar sua eficácia. O presente estudo propôs uma abordagem inovadora para a discriminação entre pacientes saudáveis e com câncer usando amostras isoladas de células sanguíneas; portanto, foi necessária uma preparação mínima da amostra antes da análise. Esta abordagem pode contribuir para o desenvolvimento de novas estratégias não invasivas e de alta precisão para o diagnóstico clínico do câncer. No entanto, é importante ressaltar que mais pesquisas são necessárias para validar e aprimorar essa técnica, utilizando adultos e idosos, a fim de torná-la uma ferramenta confiável e amplamente aplicável na prática médica.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMCZYK, H.; BROZEK-PLUSKA, B.; KOPEĆ, M. Double face of cytochrome c in cancers by Raman imaging. **Sci Rep.** v. 12, n. 1, p. 2120, 2022. DOI: <http://dx.doi.org/10.1038/s41598-022-04803-0>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/35136078/>. Acesso em: 22 jun. 2022.

ATKINS, C. G.; BUCKLEY, K.; BLADES, M. W.; TURNER, R. F. B. Raman spectroscopy of blood and blood components. **Appl Spectrosc.** v. 71, n. 5, p. 767-793, 2017. DOI: <http://dx.doi.org/10.1177/0003702816686593>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28398071/>. Acesso em: 06 jun. 2021.

AUNER, G. W. *et al.* Applications of Raman spectroscopy in cancer diagnosis. *Cancer Metastasis Rev.* v. 37, p. 691-717, 2018. DOI: <https://doi.org/10.1007/s10555-018-9770-9>. Disponível em: <https://link.springer.com/article/10.1007/s10555-018-9770-9#citeas>. Acessado em: 20 jan. 2023.

BAI, Y.; YU, Z.; YI, S.; YAN, Y.; HUANG, Z.; QIU, L. Raman spectroscopy-based biomarker screening by studying the fingerprint characteristics of chronic lymphocytic leukemia and diffuse large B-cell lymphoma. **J Pharm Biomed Anal.** v. 190, 113514, 2020. DOI: <http://dx.doi.org/10.1016/j.jpba.2020.113514>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32827998/>. Acesso em: 11 jan. 2020.

BANKAPUR, A.; ZACHARIAH, E.; CHIDANGIL, S.; VALIATHAN, M.; MATHUR, D. Raman tweezers spectroscopy of live, single red and white blood cells. **PLoS One.** v. 5, n. 4, e10427, 2010. DOI: <http://dx.doi.org/10.1371/journal.pone.0010427>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/20454686/>. Acesso em: 11 jan. 2020.

BARROSO, E. M., *et al.* Discrimination between oral cancer and healthy tissue based on water content determined by Raman spectroscopy. **Anal Chem.** v. 87, n. 4, p. 2419-2426, 2015. DOI: <http://dx.doi.org/10.1021/ac504362y>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25621527/>. Acesso em: 20 out. 2021.

BHAKTA, N., *et al.* Childhood cancer burden: a review of global estimates. **Lancet Oncol.** v. 20, n. 1, p. e42-e53, 2019. DOI: [http://dx.doi.org/10.1016/S1470-2045\(18\)30761-7](http://dx.doi.org/10.1016/S1470-2045(18)30761-7). Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30614477/>. Acesso em: 10 fev. 2020.

BODANESE, B., *et al.* Discrimination of basal cell carcinoma and melanoma from normal skin biopsies in vitro through Raman spectroscopy and principal component analysis. **Photomed Laser Surg.** v. 30, n. 7, p. 381-387, 2012. DOI: <http://dx.doi.org/10.1089/pho.2011.3191>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22693951/>. Acesso em: 15 jun. 2019.

BURKHAMER, J.; KRIEBEL, D.; CLAPP, R. The increasing toll of adolescent cancer incidence in the US. **PLoS ONE.** v. 12, n. 2, e0172986. 2017. DOI: <http://dx.doi.org/10.1371/journal.pone.0172986>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28235028/>. Acesso em: 25 mar. 2020.

CAPP, E.; NIENOV, O. H. **Bioestatística quantitativa aplicada**. Porto Alegre: UFRGS, 2020. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/213509/001117624.pdf?sequence=1>. Acesso em: 23 jun. 2022.

CHAN, J. W. Recent advances in laser tweezers Raman spectroscopy (LTRS) for label-free analysis of single cells. **J Biophotonics.** v. 6, n. 1, p. 36-48, 2013. DOI: <http://dx.doi.org/10.1002/jbio.201200143>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/jbio.201200143>. Acesso em: 15 jun. 2022.

CHAN, J. W., *et al.* Nondestructive identification of individual leukemia cells by laser trapping Raman spectroscopy. **Anal Chem.** v. 80, n. 6, p. 2180-2187, 2008. DOI: <http://dx.doi.org/10.1021/ac7022348>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/18260656/>. Acesso em: 25 abril 2020.

CHAN, J. W.; TAYLOR, D. S.; THOMPSON, D. L. The effect of cell fixation on the discrimination of normal and leukemia cells with laser tweezers Raman spectroscopy. **Biopolymers.** v. 91, n. 2, p. 132-139, 2009. DOI: <http://dx.doi.org/10.1002/bip.21094>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bip.21094>. Acesso em: 25 abr. 2020.

CHENG, X., *et al.* Raman spectroscopy differ leukemic cells from their healthy counterparts and screen biomarkers in acute leukemia. **Spectrochim Acta A Mol Biomol Spectrosc.** v. 281, p. 121558, 2022. DOI: <https://doi.org/10.1016/j.saa.2022.121558>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1386142522007077?via%3Dihub#preview-section-cited-by>. Acesso em: 22 jan. 2023.

CHRISTIAN, K., *et al.* Raman difference spectroscopy: a non-invasive method for identification of oral squamous cell carcinoma. **Biomed Opt Express.** v. 5, n. 9, p. 3252-3265, 2014. DOI: <http://dx.doi.org/10.1364/BOE.5.003252>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25401036/>. Acesso em: 26 abr. 2020.

DOCHOW, S., *et al.* Tumour cell identification by means of Raman spectroscopy in combination with optical traps and microfluidic environments. **Lab Chip.** v. 11, n. 8, p. 1484-1490, 2011. DOI: <http://dx.doi.org/10.1039/c0lc00612b>. Disponível em: <https://pubs.rsc.org/en/content/articlelanding/2011/lc/c0lc00612b>. Acesso em: 26 abr.2020.

FAZIO, E., *et al.* A micro-Raman spectroscopic investigation of leukemic U-937 cells in aged cultures. **Spectrochim Acta A Mol Biomol Spectrosc.** v. 159, p. 21-29, 2016. DOI: <http://dx.doi.org/10.1016/j.saa.2016.01.032>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/26827173/>. Acesso em: 26 abr. 2020.

FENG, S., *et al.* Surface-enhanced Raman spectroscopy of saliva proteins for the noninvasive differentiation of benign and malignant breast tumors. **Int J Nanomedicine.** v. 10, p. 537-5-47, 2015. DOI: <http://dx.doi.org/10.2147/IJN.S71811>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25609959/>. Acesso em: 26 abr. 2020.

FÉRE, M., *et al.* Implementation of a classification strategy of Raman data collected in different clinical conditions: application to the diagnosis of chronic lymphocytic leukemia. **Anal Bioanal Chem.** v. 412, n.4, p. 949-962, 2020. DOI: <http://dx.doi.org/10.1007/s00216-019-02321-z>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31853604/>. Acesso em: 26 abr. 2020.

GONZÁLEZ-SOLÍS, J.L. Discrimination of different cancer types clustering Raman spectra by a super paramagnetic stochastic network approach. **PLoS One.** v. 14, n. 3, e0213621, 2019. DOI: <http://dx.doi.org/10.1371/journal.pone.0213621>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30861043/>. Acesso em: 12 maio 2020.

GONZÁLEZ-SOLÍS, J.L., *et al.* Monitoring of chemotherapy leukemia treatment using Raman spectroscopy and principal component analysis. **Lasers Med Sci.** v. 29, n. 3, p. 1241-1249, 2014. DOI: <http://dx.doi.org/10.1007/s10103-013-1515-y>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/24407964/>. Acesso em: 12 maio 2020.

GOULART A. C. C.;ZÂNGARO, R. A.;CARVALHO, H. C.;SILVEIRA JR., L. Diagnosing COVID-19 in human sera with detected immunoglobulins IgM and

IgG by means of Raman spectroscopy. **J Raman Spectrosc.** v. 52, p. 2671, 2021. DOI: <https://doi.org/10.1002/jrs.6235>. Disponível em: <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/jrs.6235>.

Acesso em: 02 set. 2022.

HANLON, E. B., *et al.* Prospects for in vivo Raman spectroscopy. **Phys Med Biol.** v. 45, n. 2, p. R1-59, 2000. DOI: <http://dx.doi.org/10.1088/0031-9155/45/2/201>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10701500/>. Acesso em: 17 jun. 2018.

HAPPILLON, T., *et al.* Diagnosis approach of chronic lymphocytic leukemia on unstained blood smears using Raman microspectroscopy and supervised classification. **Analyst.** v. 140, n. 13, p. 4465-4472, 2015. DOI: <http://dx.doi.org/10.1039/c4an02085e>. Disponível em: <https://pubs.rsc.org/en/content/articlelanding/2015/an/c4an02085e>. Acesso em: 12 maio 2020.

HASSOUN, M., *et al.* A droplet-based microfluidic chip as a platform for leukemia cell lysate identification using surface-enhanced Raman scattering. **Anal Bioanal Chem.** v. 410, n. 3, p. 999-1006, 2018. DOI: <http://dx.doi.org/10.1007/s00216-017-0609-y>. Disponível em: <https://link.springer.com/article/10.1007/s00216-017-0609-y>. Acesso em: 12 maio 2020.

HÜTTEMANN, M., *et al.* The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis. **Mitochondrion.** v. 11, n. 3, p. 369-381, 2011. DOI: <http://dx.doi.org/10.1016/j.mito.2011.01.010>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/21296189/>. Acesso em: 22 jan. 2022.

INCA (Instituto Nacional de Câncer José Alencar Gomes da Silva). Estimativa 2020: incidência de câncer no Brasil / Instituto Nacional de Câncer José

Alencar Gomes da Silva. **Coordenação de Prevenção e Vigilância**. Rio de Janeiro, 2020a. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/estimativa-2020-incidencia-de-cancer-no-brasil.pdf>. Acesso em: 21 abr. 2022.

INCA (Instituto Nacional de Câncer José Alencar Gomes da Silva). ABC do câncer : abordagens básicas para o controle do câncer. / **Instituto Nacional de Câncer José Alencar Gomes da Silva**. 6. ed. rev. atual. Rio de Janeiro: INCA, 2020b. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/livro-abc-6-edicao-2020.pdf>. Acesso em: 29 jun. 2022.

INCA (Instituto Nacional de Câncer José Alencar Gomes da Silva). Estimativa 2023: incidência de câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva. Rio de Janeiro, 2022. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf>. Acesso em: 25 out. 2023.

JOLLIFFE, I. T., CADIMA, J. Principal component analysis: a review and recent developments. **Philos Trans A Math Phys Eng Sci**. v. 374, n. 2065, 20150202, 2016. DOI: <http://dx.doi.org/10.1098/rsta.2015.0202>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/26953178/>. Acesso em: 14 maio 2022.

KAPLAN, J. A. Leukemia in children. **Pediatr Rev**. v. 40, n. 7, p. 319-331, 2019. DOI: <http://dx.doi.org/10.1542/pir.2018-0192>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31263040/>. Acesso em: 10 fev. 2022.

KHETANI, A., *et al.* Hollow core photonic crystal fiber for monitoring leukemia cells using surface enhanced Raman scattering (SERS). **Biomed Opt Express**. v. 6, n. 11, p. 4599-4609, 2015. DOI: <http://dx.doi.org/10.1364/BOE.6.004599>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/26601021/>. Acesso em: 10 fev. 2022.

KONG, K., *et al.* Raman spectroscopy for medical diagnostics - From in-vitro biofluid assays to in-vivo cancer detection. **Adv Drug Deliv Rev.** v. 89, p. 121-134, 2015. DOI: <http://dx.doi.org/10.1016/j.addr.2015.03.009>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25809988/>. Acesso em: 15 set. 2021.

LASCH, P. Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging. **Chemometr Intell Lab.** v. 117, p. 100-114, 2012. DOI: <http://dx.doi.org/10.1016/j.chemolab.2012.03.011>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169743912000561>. Acesso em: 15 jul.2021.

LEE, L. C., LIONG, C. Y., JEMAIN, A. A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. **Analyst.** v. 143, n. 15, p. 3526-3539, 2018. DOI: <http://dx.doi.org/10.1039/c8an00599k>. Disponível em: <https://pubs.rsc.org/en/content/articlelanding/2018/an/c8an00599k>. Acesso em: 22 fev. 2022.

LESZCZENKO, P. *et al.* Towards Raman-Based Screening of Acute Lymphoblastic Leukemia-Type B (B-ALL) Subtypes **Cancers (Basel)**. v. 13, n. 21, p. 5483, 2021. DOI: <https://doi.org/10.3390/cancers13215483>. Disponível em: <https://www.mdpi.com/2072-6694/13/21/5483>. Acesso em: 15 jan. 2023.

LIMA, A. M. F., *et al.* Discrimination of non-melanoma skin cancer and keratosis from normal skin tissue in vivo and ex vivo by Raman spectroscopy. **Vib Spectrosc.** v. 100, p. 131-141, 2019. DOI: <http://dx.doi.org/10.1016/j.vibspec.2018.11.009>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0924203118300808>. Acesso em: 20 out. 2020.

LIMA, A. M. F., *et al.* Discrimination of leukemias and non-leukemic cancers in blood serum samples of children and adolescents using a Raman spectral

model. **Lasers Med Sci.** v. 38, n. 22, 2023. DOI: <https://doi.org/10.1007/s10103-022-03681-2>. Disponível em: <https://link.springer.com/article/10.1007/s10103-022-03681-2>. Acesso em: 02 fev. 2023.

MACLAUGHLIN, C. M., *et al.* Surface-enhanced Raman scattering dye-labeled Au nanoparticles for triplexed detection of leukemia and lymphoma cells and SERS flow cytometry. **Langmuir.** v. 29, n. 6, p. 1908-1919, 2013b. DOI: <http://dx.doi.org/10.1021/la303931c>. Disponível em: <https://pubs.acs.org/doi/10.1021/la303931c>. Acesso em: 11 jul. 2022.

MACLAUGHLIN, C. M., *et al.* Evaluation of SERS labeling of CD20 on CLL cells using optical microscopy and fluorescence flow cytometry. **Nanomedicine.** v. 9, n. 1, p. 55-64, 2013a. DOI: <http://dx.doi.org/10.1016/j.nano.2012.04.003>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1549963412001761>. Acesso em: 11 jul. 2022.

MANAGÒ, S., *et al.* Raman detection and identification of normal and leukemic hematopoietic cells. **J Biophotonics.** v. 11, n. 5, p. e201700265, 2018a. DOI: <http://dx.doi.org/10.1002/jbio.201700265>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29239116/>. Acesso em: 13 ago. 2021.

MANAGÒ, S., *et al.* A reliable Raman-spectroscopy-based approach for diagnosis, classification and follow-up of B-cell acute lymphoblastic leukemia. **Sci Rep.** v. 6, p. 2482.1, 2016. DOI: <http://dx.doi.org/10.1038/srep24821>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27089853/>. Acesso em: 13 ago. 2021.

MANAGÒ, S.; ZITO, G.; DE LUCA, A. C. Raman microscopy based sensing of leukemia cells: A review. **Opt Laser Technol.** v. 108, p. 7-16, 2018b. DOI: <http://dx.doi.org/10.1016/j.optlastec.2018.06.034>. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0030399218307825>.

Acesso em: 13 ago. 2021.

MARTÍNEZ-ESPINOSA, J. C., *et al.* Detection of leukemia with blood samples using Raman spectroscopy and multivariate analysis. **AIP Conf Proc.** v. 1142, n. 99, 2009. DOI: <http://dx.doi.org/10.1063/1.3175637>. Disponível em: <https://www.researchgate.net/publication/252859262>. Acesso em: 24 nov. 2021.

MEIJER, R. J.; GOEMAN, J. J. Efficient approximate k-fold and leave-one-out cross-validation for ridge regression. **Biom J**, v. 55, n. 2, p. 141–155, 24 jan. 2013. DOI: <http://dx.doi.org/10.1002/bimj.201200088>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23348970/>. Acesso em: 13 dez. 2022.

MELARAGNO, R.; CAMARGO, B. **Oncologia Pediátrica: Diagnóstico e Tratamento**. São Paulo: Editora Atheneu, 2013.

MISHRA, P.; SINGH, U.; PANDEY, C.M.; MISHRA, P.; PANDEY, G. Application of student's t-test, analysis of variance, and covariance. **Ann Card Anaesth.** v. 22, n. 4, p. 407-411, 2019. DOI: http://dx.doi.org/10.4103/aca.ACA_94_19. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31621677/>. Acesso em: 24 jun. 2022.

MITCHELL, K.; STEIDL, U. Targeting immunophenotypic markers on leukemic stem cells: How lessons from current approaches and advances in the leukemia stem cell (LSC) model can inform better strategies for treating acute myeloid leukemia (AML). **Cold Spring Harb Perspect Med.** v. 10, n. 1, p. a036251, 2020. DOI: <http://dx.doi.org/10.1101/cshperspect.a036251>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31451539/>. Acesso em: 24 nov. 2021.

NAVAS-MORENO, M., CHAN, J.W. Laser Tweezers Raman Microspectroscopy of Single Cells and Biological Particles. **Methods Mol Biol.** 1745:219-257, 2018. DOI: http://dx.doi.org/10.1007/978-1-4939-7680-5_13. Disponível em:

https://link.springer.com/protocol/10.1007/978-1-4939-7680-5_13. Acesso em: 15 junho 2022.

NAYAK, R.; RAI, S. **Essentials in hematology and clinical pathology**. Jaypee Brothers Medical Publishers: New Delhi; 2017. DOI: <http://dx.doi.org/10.5005/jp/books/12955>.

NEUGEBAUER, U., *et al.* Identification and differentiation of single cells from peripheral blood by Raman spectroscopic imaging. **J Biophotonics**. v. 3, n. 8-9, p. 579-587, 2010a. DOI: <http://dx.doi.org/10.1002/jbio.201000020>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/jbio.201000020>. Acesso em: 24 nov. 2021.

NEUGEBAUER, U., *et al.* Towards detection and identification of circulating tumour cells using Raman spectroscopy. **Analyst**. v. 135, n. 12, p. 3178-3182, 2010b. DOI: <http://dx.doi.org/10.1039/c0an00608d>. Disponível em: <https://www.researchgate.net/publication/47404719>. Acesso em: 24 nov. 2021.

NUNES, C. A.; FREITAS, M. P.; PINHEIRO, A. C. M.; BASTOS, S. C. Chemoface: a novel free user-friendly interface for chemometrics. **J Braz Chem Soc**. v. 23, n. 11, p., 2003–2010, 2012. DOI: <http://dx.doi.org/10.1590/s0103-50532012005000073>. Disponível em: <https://www.scielo.br/j/jbchs/a/XkQMMnpQZrBHMJr3XDtkwYQ/?lang=en>. Acesso em: 30 jan. 2022.

PACHECO, M. T.; MOREIRA, L. M. Raman spectroscopy: new perspectives for its clinical application in diagnosis. **Photomed Laser Surg**. v. 31, n. 10, p. 463-465, 2013. DOI: <http://dx.doi.org/10.1089/pho.2013.9873>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/24032552/>. Acesso em: 24 nov. 2021.

PÉREZ, I., *et al.* How to interpret diagnostic tests. **Medwave**. v. 21, e8432, 2021. DOI: <http://dx.doi.org/10.5867/medwave.2021.07.8432>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34369918/>. Acesso em: 25 jun. 2022.

PREMASIRI, W. R.; LEE, J. C.; ZIEGLER, L. D. Surface-enhanced Raman scattering of whole human blood, blood plasma, and red blood cells: cellular processes and bioanalytical sensing. **J Phys Chem B**. v. 116, n. 31, p. 9376-9386, 2012. DOI: <http://dx.doi.org/10.1021/jp304932g>. Disponível em: <https://pubs.acs.org/doi/10.1021/jp304932g>. Acesso em: 24 nov. 2021.

RASHID, N., *et al.* Raman microspectroscopy for the early detection of pre-malignant changes in cervical tissue. **Exp Mol Pathol**. v. 97, n. 3, p. 554-564, 2014. DOI: <http://dx.doi.org/10.1016/j.yexmp.2014.10.013>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0014480014001737>. Acesso em: 24 nov. 2021.

RINGNÉR, M. What is principal component analysis? **Nat Biotechnol**. v. 26, n. 3, p. 303-304, 2008. DOI: <http://dx.doi.org/10.1038/nbt0308-303>. Disponível em: <https://www.nature.com/articles/nbt0308-303>. Acesso em: 10 fev. 2022.

RODRIGUES, A. B.; OLIVEIRA, P. P. **Hemoterapia e hematologia**: conceitos essenciais para assistência. São Paulo: Rideel, 2017.

RUIZ-PEREZ, D., *et al.* So you think you can PLS-DA? **BMC Bioinformatics**. v. 21, p. 2, 2020. Suppl. 1. DOI: <http://dx.doi.org/10.1186/s12859-019-3310-7>. Disponível em: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-3310-7>. Acesso em: 10 fev. 2022.

SANTANA, F. B., *et al.* 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. **Quim. Nova.** v. 43, n. 3, p. 371-381, 2020. DOI: <http://dx.doi.org/10.21577/0100-4042.20170480>. Disponível em: <https://www.scielo.br/j/qn/a/Bq7xdFNQ8dsKrxhPCbsHYPc/abstract/?lang=en>. Acesso em: 25 mar. 2022.

SATO, H., *et al.* Excitation wavelength-dependent changes in Raman spectra of whole blood and hemoglobin: comparison of the spectra with 514.5-, 720-, and 1064-nm excitation. **J Biomed Opt.** v. 6, n. 3, p. 366-370, 2001. DOI: <http://dx.doi.org/10.1117/1.1380668>. Disponível em: <https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-6/issue-03/0000/Excitation-wavelength-dependent-changes-in-Raman-spectra-of-whole-blood/10.1117/1.1380668.full?SSO=1>. Acesso em: 24 nov. 2021.

SETH, R.; SINGH, A. Leukemias in children. **Indian J Pediatr.** v. 82, n. 9, p. 817-824, 2015. DOI: <http://dx.doi.org/10.1007/s12098-015-1695-5>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25680783/>. Acesso em: 20 fev. 2022.

DA SILVA, A. M., *et al.* Spectral model for diagnosis of acute leukemias in whole blood and plasma through Raman spectroscopy. **J Biomed Opt.** v. 23, n. 10, p. 1-11, 2018. DOI: <http://dx.doi.org/10.1117/1.JBO.23.10.107002>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30350494/>. Acesso em: 26 abr. 2020.

SILVEIRA, F.L., *et al.* Discrimination of non-melanoma skin lesions from non-tumor human skin tissues in vivo using Raman spectroscopy and multivariate statistics. **Lasers Surg Med.** v. 47, n. 1, p. 6-16, 2015. DOI: <http://dx.doi.org/10.1002/lsm.22318>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lsm.22318>. Acesso em: 24 nov. 2021.

SILVEIRA JR., L., *et al.* Discriminating model for diagnosis of basal cell carcinoma and melanoma in vitro based on the Raman spectra of selected

biochemicals. **J Biomed Opt.** v. 17, n. 7, 077003, 2012. DOI: <http://dx.doi.org/10.1117/1.JBO.17.7.077003>. Disponível em: <https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-17/issue-7/077003/Discriminating-model-for-diagnosis-of-basal-cell-carcinoma-and-melanoma/10.1117/1.JBO.17.7.077003.full>. Acesso em: 20 set. 2021.

SILVEIRA, L.; SATHAIAH, S.; ZÂNGARO, R. A.; PACHECO, M. T.; CHAVANTES, M.C.; PASQUALUCCI, C. A. Correlation between near-infrared Raman spectroscopy and the histopathological analysis of atherosclerosis in human coronary arteries. **Lasers Surg Med.** v. 30, n. 4, p. 290-297, 2002. DOI: <http://dx.doi.org/10.1002/lsm.10053>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lsm.10053>. Acesso em: 20 set. 2021.

SUNG, H., *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA Cancer J Clin.** v. 71, n. 3, p. 209-249, 2021. DOI: <http://dx.doi.org/10.3322/caac.21660>. Disponível em: <https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21660>. Acesso em: 20 fev. 2022.

TALARI, A. C. S., *et al.* Raman spectroscopy of biological tissues. **Appl Spectrosc Rev.** v. 50, n. 1, p. 46-111, 2014; DOI: <http://dx.doi.org/10.1080/05704928.2014.923902>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/05704928.2014.923902>. Acesso em: 13 mar. 2022.

VAN MAELE-FABRY, G.; GAMET-PAYRASTRE, L.; LISON, D. Household exposure to pesticides and risk of leukemia in children and adolescents: Updated systematic review and meta-analysis. **Int J Hyg Environ Health.** v. 222, n. 1, p. 49-67, 2019. DOI: <http://dx.doi.org/10.1016/j.ijheh.2018.08.004>. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S143846391830364X>.

Acesso em: 15 fev. 2022.

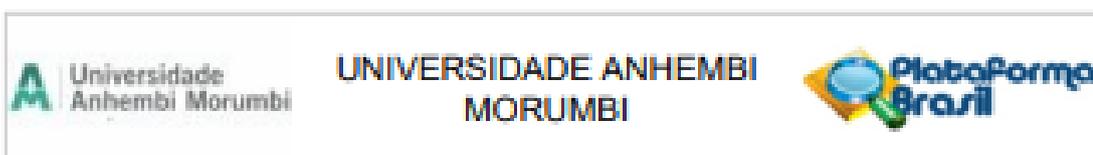
VANNA, R., *et al.* Label-free imaging and identification of typical cells of acute myeloid leukaemia and myelodysplastic syndrome by Raman microspectroscopy. **Analyst**. v. 140, n. 4, p. 1054-1064, 2015. DOI: <http://dx.doi.org/10.1039/c4an02127d>. Disponível em: <https://pubs.rsc.org/en/content/articlelanding/2015/an/c4an02127d>. Acesso em: 15 fev. 2022.

YAMAMOTO, M.; FIGUEIREDO, V. L. P. Epidemiology of chronic lymphocytic leukemia and familial chronic lymphocytic leukemia. **Rev. bras. hematol. hemoter.** v. 27, n. 4, p. 229-232, 2005. Disponível em: <https://www.researchgate.net/publication/262700120>. Acesso em: 25 mar. 2022.

ZHANG, D., *et al.* Raman spectrum reveals the cell cycle arrest of Triptolide-induced leukemic T-lymphocytes apoptosis. **Spectrochim Acta A Mol Biomol Spectrosc.** v. 141, p. 216-222, 2015. DOI: <http://dx.doi.org/10.1016/j.saa.2015.01.037>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1386142515000566>. Acesso em: 15 fev. 2022.

ANEXO

ANEXO 1: Parecer consubstanciado do Comitê de ética e pesquisa (CEP)



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: ESTUDO COM PACIENTES PEDIÁTRICOS E ONCOLÓGICOS - ALTERAÇÕES SANGUÍNEAS VISUALIZADAS ATRAVÉS DA ESPECTROSCOPIA RAMAN

Pesquisador: Landolfo Silveira Junior

Área Temática:

Versão: 3

CAAE: 23457019.9.0000.5492

Instituição Proponente: Universidade Anhembi Morumbi/ Instituto Superior de Comunicação

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 5.170.213

Apresentação do Projeto:

EXERTO EXTRAÍDO DO PROJETO: ESTUDO COM PACIENTES PEDIÁTRICOS E ONCOLÓGICOS - ALTERAÇÕES SANGUÍNEAS VISUALIZADAS ATRAVÉS DA ESPECTROSCOPIA RAMAN

Pesquisador Responsável: Landolfo Silveira Junior

CAAE: 23457019.9.0000.5492

O diagnóstico das neoplasias na faixa etária pediátrica tem aumentado significativamente nos últimos anos e a possibilidade de cura das mesmas

também¹. Vários são os desafios que existem no acompanhamento desses pacientes dentre os quais o diagnóstico precoce, a identificação dos casos de recidiva bem como a individualização do tratamento. Para responder a esses desafios tem se recorrido aos estudos de biologia molecular.

No entanto, tais estudos geralmente demandam a coleta de material dos tumores por vezes através procedimentos cirúrgicos invasivos e arriscados.

A utilização de amostras de sangue para obtenção dessas informações permite uma abordagem menos invasiva, no entanto os marcadores

identificados até o momento ou carecem de especificidade ou não apresentam a sensibilidade que seria desejável, por isso a procura por outros

continua tão atual e relevante. A espectroscopia Raman baseia-se nas características de

Endereço: Rua Dr. Almeida Lima, 1.134 - 2º andar - sala 207

Bairro: Mooca

CEP: 03.164-000

UF: SP

Município: SAO PAULO

Telefone: (11)2792-4658

E-mail: cep@anhembibr

Continuação do Parecer: S.170.213

espalhamento inelástico da luz incidente sobre os materiais. O progresso tecnológico ocorrido tanto em relação às fontes de luz bem como em relação aos equipamentos de detecção e métodos de análise dessa reflexão ampliaram em muito a aplicação desse fenômeno, levando a identificação de uma série de espectros de espalhamento que permitem a caracterização de praticamente todos os materiais orgânicos. Nos últimos anos esse princípio tem sido aplicado também à medicina e particularmente no diagnóstico de uma série de neoplasias²⁻⁴. Embora mais utilizada no estudo de tecidos, inclusive *in vivo*, a espectroscopia Raman tem sido empregada no estudo do sangue periférico⁵⁻⁸, ao que parece possibilitando o diagnóstico de inúmeras condições fisiológicas e patológicas. Em oncologia pediátrica existem relatos da utilização desse método no diagnóstico de algumas patologias⁹⁻¹¹. Tendo em vista a necessidade de métodos diagnósticos mais eficazes, que permitam um melhor acompanhamento e tratamento das crianças com neoplasias o presente estudo tem por finalidade estabelecer padrões de espectroscopia Raman no soro de crianças portadoras de neoplasias que estejam ou venham a ser atendidas no Hospital do Centro de Tratamento Infantojuvenil Fabiana Macedo de Moraes, mantido pelo Grupo de Apoio à Criança com Câncer (CTFM / GACC) de São José dos Campos, procurando estabelecer o quanto esse método pode auxiliar no tratamento dessas crianças.

Objetivo da Pesquisa:

Objetivo Primário:

- Estabelecer padrões de espectroscopia Raman no sangue de pacientes pediátricos portadores de neoplasia

Objetivo Secundário:

- Identificar marcadores que possam diferenciar os pacientes pediátricos portadores de neoplasia dos pacientes controle bem como auxiliar na individualização do tratamento desses pacientes;
- Avaliar as diferenças espectrais entre pacientes com doença neoplásica e pacientes saudáveis;
- Identificar os picos Raman que possibilitem a diferenciação entre pacientes com doença neoplásica e pacientes saudáveis;
- Avaliar os principais

Endereço: Rua Dr. Almeida Lima, 1.134 - 2º andar - sala 207

Bairro: Mooca

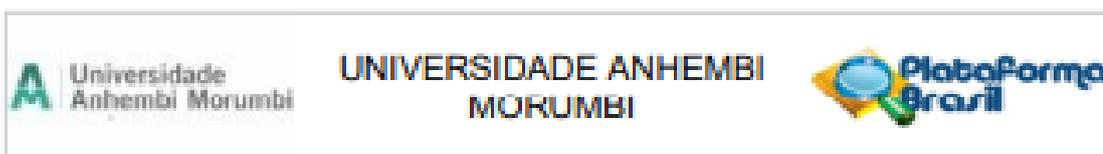
CEP: 03.164-000

UF: SP

Município: SAO PAULO

Telefone: (11)2792-4658

E-mail: cop@anhemb.br



Continuação do Parecer: S.170.213

marcadores biológicos alterados em pacientes com neoplasia em comparação com pacientes saudáveis.

Avaliação dos Riscos e Benefícios:

Riscos:

Os possíveis riscos apresentados na pesquisa são mínimos, como: sangramento e desconforto apresentado devido a coleta sanguínea. Os dados serão coletados de forma a garantir o anonimato e privacidade dos indivíduos. Será realizada em local tranquilo, com total privacidade dos indivíduos e estando o pesquisador responsável e médico presentes no local da coleta e no local do experimento. Se houver descumprimento dos termos do TCLE ou qualquer dano for gerado ao paciente, este será ressarcido. Para evitar complicações serão realizadas, no máximo, três tentativas de punção e compressão do local da coleta.

Benefícios:

A principal contribuição da pesquisa é o desenvolvimento de uma técnica óptica visando a identificação de marcadores do soro sanguíneo que são alterados devido ao desenvolvimento de neoplasia, bem como promove a expansão da aplicabilidade da espectroscopia Raman no diagnóstico das neoplasias na faixa etária pediátrica.

Comentários e Considerações sobre a Pesquisa:

É um projeto com condições de realização, definido em termos metodológicos e logísticos, caracterizando exequibilidade na proposta.

Considerações sobre os Termos de apresentação obrigatória:

Estão adequados e contemplam as exigências da resolução 466/12.

Recomendações:

Sem recomendações.

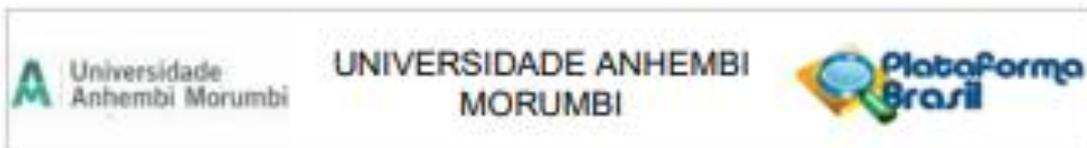
Conclusões ou Pendências e Lista de Inadequações:

Sem pendências.

Considerações Finais a critério do CEP:

1. Apresentar relatório parcial da pesquisa, semestralmente, a contar do início da mesma.
2. Apresentar relatório final da pesquisa até 30 dias após o término da mesma.
3. O CEP UAM deverá ser informado de todos os efeitos adversos ou fatos relevantes que alterem o

Endereço: Rua Dr. Almeida Lima, 1.104 - 2º andar - sala 207
Cidade: Mooca **CEP:** 03.164-000
UF: SP **Município:** SAO PAULO
Telefone: (11) 2790-4838 **E-mail:** cep@anhembibr



Continuação do Parecer: S.170.213

curso normal do estudo.

4. Quaisquer documentações encaminhadas ao CEP UAM deverão conter junto uma Carta de Encaminhamento, em que conste o objetivo e justificativa do que esteja sendo apresentado.
5. Caso a pesquisa seja suspensa ou encerrada antes do previsto, o CEP UAM deverá ser comunicado, estando os motivos expressos no relatório final a ser apresentado.
6. O TCLE deverá ser obtido em duas vias, uma ficará com o pesquisador e a outra com o sujeito de pesquisa.
7. Em conformidade com a Carta Circular nº. 003/2011CONEP/CNS, faz-se obrigatório a rubrica em todas as páginas do TCLE pelo sujeito de pesquisa ou seu responsável e pelo pesquisador.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

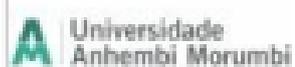
| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|---|---|------------------------|--------------------------|----------|
| Informações Básicas do Projeto | PB_INFORMAÇÕES_BÁSICAS_185479_5_E1.pdf | 06/11/2021 10:32:59 | | Aceito |
| Outros | TCLE_ALTERADO.pdf | 13/11/2019 19:16:13 | Landulfo Silveira Junior | Aceito |
| Outros | CARTA_DE_ENCAMINHAMENTO_TCLE_ALTERADO.pdf | 13/11/2019 19:15:30 | Landulfo Silveira Junior | Aceito |
| Projeto Detalhado / Brochura Investigador | PROJETO_COMPLETO.pdf | 14/10/2019 17:07:13 | Landulfo Silveira Junior | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | TCLE.pdf | 14/10/2019 17:05:54 | Landulfo Silveira Junior | Aceito |
| Declaração de Instituição e Infraestrutura | DECLARACAO_AUTORIZACAO.pdf | 14/10/2019 17:05:40 | Landulfo Silveira Junior | Aceito |
| Declaração de Pesquisadores | DECLARACAO_DE_COMPROMISSO_DOS_PESQUISADORES.pdf | 14/10/2019 17:05:24 | Landulfo Silveira Junior | Aceito |
| Brochura Pesquisa | BROCHURA.pdf | 14/10/2019 17:05:12 | Landulfo Silveira Junior | Aceito |
| Folha de Rosto | FOLHA_DE_ROSTO.pdf | 14/10/2019 17:04:58 | Landulfo Silveira Junior | Aceito |

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Endereço: Rua Dr. Almeida Lima, 1.134 - 2º andar - sala 207
 Bairro: Mooca CEP: 03.164-000
 UF: SP Município: SAO PAULO
 Telefone: (11)2790-6658 E-mail: cep@anhembibr



UNIVERSIDADE ANHEMBI
MORUMBI



Continuação do Processo: S-170.213

Não

SÃO PAULO, 16 de Dezembro de 2021

Assinado por:
CARLOS ROCHA OLIVEIRA
(Coordenador(a))

Endereço: Rua Dr. Almeida Lima, 1.134 - 2º andar - sala 207
Bairro: Mooca **CEP:** 03.164-000
UF: SP **Município:** SÃO PAULO
Telefone: (11)2793-4658 **E-mail:** cep@anhembibr