

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
Instituto de Ciências Biológicas  
Programa de Pós-graduação em Parasitologia

Victor Freire Ferreira de Aguiar

**DIAGNÓSTICO DIFERENCIAL DA ESQUISTOSSOMOSE POR  
ESPECTROSCOPIA ATR-FTIR INTEGRADA A ALGORITMOS DE  
INTELIGÊNCIA ARTIFICIAL**

**Belo Horizonte  
2026**

Victor Freire Ferreira de Aguiar

**DIAGNÓSTICO DIFERENCIAL DA ESQUISTOSSOMOSE POR  
ESPECTROSCOPIA ATR-FTIR INTEGRADA A ALGORITMOS DE  
INTELIGÊNCIA ARTIFICIAL**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Parasitologia da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Parasitologia

Orientador: Prof. Ricardo Toshio Fujiwara

**Belo Horizonte  
2026**

043

Aguilar, Victor Freire Ferreira de.

Diagnóstico diferencial da esquistossomose por espectroscopia ATR-FTIR integrada a algoritmos de inteligência artificial [manuscrito] / Victor Freire Ferreira de Aguilar. – 2026.

51 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Ricardo Toshio Fujiwara.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Parasitologia.

1. Parasitologia. 2. Doenças Negligenciadas. 3. Esquistossomose. 4. Schistosoma mansoni. 5. Espectroscopia de Infravermelho com Transformada de Fourier. I. Fujiwara, Ricardo Toshio. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 576.88/.89



UNIVERSIDADE FEDERAL DE MINAS GERAIS

**FOLHA DE APROVAÇÃO**

**TÍTULO: "DIAGNÓSTICO DIFERENCIAL DA ESQUISTOSSOMOSE POR ESPECTROSCOPIA ATR-FTIR INTEGRADA A ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL"**

**ALUNO: VICTOR FREIRE FERREIRA DE AGUILAR**

**ÁREA DE CONCENTRAÇÃO: HELMINTOLOGIA**

**DISSERTAÇÃO: 479/2026/01**

Dissertação de Mestrado defendida e aprovada, no dia **nove de fevereiro de 2026**, pela Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação Parasitologia da Universidade Federal de Minas Gerais constituída pelos seguintes doutores:

**Fabiana Martins de Paula**

USP

**Silvio Santana Dolabella**

UFS

**Ricardo Toshio Fujiwara - Orientador**

UFMG

Belo Horizonte, 09 de fevereiro de 2026.

**Observação: Este documento somente terá validade mediante assinatura do(a) Coordenador(a) do PPGPAR.**



Documento assinado eletronicamente por **Ricardo Toshio Fujiwara, Professor do Magistério Superior**, em 10/02/2026, às 09:48, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fabiana Martins de Paula, Usuário Externo**, em 10/02/2026, às 11:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Silvio Santana Dolabella, Usuário Externo**, em 10/02/2026, às 13:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Hudson Alves Pinto, Coordenador(a)**, em 10/02/2026, às 14:18, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4914540** e o código CRC **8A365E72**.

## AGRADECIMENTOS

Agradeço primeiramente a Deus e a Nossa Senhora Aparecida por iluminarem meu caminho, concedendo-me sabedoria, discernimento e força para superar cada etapa dessa trajetória.

Aos meus pais, Marcelo e Laura, minha eterna gratidão por acreditarem em mim e me apoiarem incondicionalmente em todos os momentos da minha vida. Eu amo vocês.

Agradeço aos meus avós, Cristina e Onofrinho, base e alicerce da nossa família, por todos os conselhos, orações e pelo cuidado constante. Aos meus irmãos, Arthur e Marcelinho, agradeço pela parceria e irmandade. Estendo também minha gratidão aos meus tios, em especial à tia Lorena, pela amizade, parceria e pela confidencialidade de todos esses anos.

Aos meus amigos do Norte de Minas, Lais, Heria, Duda, Luma, Betina e João, agradeço por ouvirem minhas angústias, pelos momentos de leveza em meio à correria e por sempre torcerem por mim.

Agradeço aos colegas do Laboratório LICP por me acolherem com carinho, por cada ajuda direta ou indireta e por todos os ensinamentos compartilhados. Em especial, agradeço a Ana Laura, Josy, Jorge, Taty e Andressa pelo suporte diário e pela amizade.

Agradeço às minhas colegas de mestrado, Rachel e Natão, por cada risada compartilhada, pelos momentos de leveza em meio à rotina intensa e por se tornarem meu porto seguro em Belo Horizonte. Sou muito grato pela amizade e pela parceria de todos os dias.

Agradeço também a todos os professores do Programa de Pós-graduação em Parasitologia da UFMG. Sinto-me verdadeiramente honrado por ter aprendido com profissionais que antes conhecia apenas pelos livros e que, ao longo dessa jornada, se tornaram referências, inspirações e, muitos deles, amigos.

Agradeço ao professor Dr. Ricardo Fujiwara e à professora Dra. Lilian Bueno pela confiança depositada em mim, pela oportunidade de fazer parte do laboratório e pelo cuidado com nossa formação. Sou profundamente grato por cada conselho científico e pessoal ao longo dessa jornada.

Agradeço ao professor Dr. Robinson Sabino e ao Dr. Douglas pela paciência em me ensinar a manusear o ATR-FTIR, pela ajuda na interpretação dos resultados e pela recepção atenciosa em Uberlândia. Estendo meu agradecimento ao Laboratório de Nanobiotecnologia da UFU.

Agradeço à Universidade Federal de Minas Gerais (UFMG), Universidade Federal de Uberlândia (UFU), ao Programa de Pós-graduação em Parasitologia da UFMG e a Fundação de Amparo à Pesquisa do Estado de Minas Gerais, instituições fundamentais para a realização deste trabalho.

Por fim, agradeço a todos que, de alguma forma, contribuíram para a execução e conclusão desta pesquisa. A cada um de vocês, meu sincero muito obrigado.

*Porque se chamavam homens*

*Também se chamavam sonhos*

*E sonhos não envelhecem*

**Milton Nascimento, Lô Borges e Márcio Borges (1972)**

## RESUMO

A esquistossomose é classificada como uma Doença Tropical Negligenciada e representa um importante problema de saúde pública no Brasil, tendo *Schistosoma mansoni* como agente etiológico na América do Sul. Os métodos diagnósticos atualmente disponíveis apresentam limitações de acurácia, sensibilidade e especificidade, além de demandarem experiência técnica tanto no processamento quanto na interpretação dos resultados. Neste estudo, empregou-se a metodologia de espectroscopia no Infravermelho por Transformada de Fourier com Reflectância Total Atenuada (ATR-FTIR) com o objetivo de identificar assinaturas espectrais capazes de auxiliar no diagnóstico diferencial da esquistossomose. As amostras de soro utilizadas foram obtidas do repositório do Laboratório de Imunobiologia e Controle de Parasitos da Universidade Federal de Minas Gerais. Para a análise por ATR-FTIR foram empregadas 54 amostras positivas para esquistossomose e 58 amostras provenientes de indivíduos não infectados, residentes em áreas endêmicas e não endêmicas para a esquistossomose. Todos os participantes portadores da esquistossomose apresentavam infecção exclusiva por *Schistosoma mansoni*, confirmada por exames parasitológicos complementares, garantindo a especificidade das análises. Técnicas de inteligência artificial com foco em métodos de aprendizado de máquina foram aplicadas para classificação dos espectros. Os resultados indicaram que o pré-processamento com a primeira derivada (1st derivative) proporcionou o melhor desempenho do modelo, resultando em uma classificação diagnóstica mais eficiente. Utilizando o algoritmo Random Forest foram obtidos valores de acurácia de 80%, sensibilidade de 81%, especificidade de 79%, área sob a curva (AUC) de 0,83 e F1-score de 0,80. As características SHAP extraídas dos espectros de infravermelho destacaram os principais modos vibracionais associados a proteínas, ácidos nucleicos e carboidratos no soro, responsáveis pela discriminação entre indivíduos infectados e não infectados por *S. mansoni*. Desse modo, o uso da técnica ATR-FTIR associada ao aprendizado de máquina mostra-se um método de triagem promissor para áreas endêmicas, principalmente por se tratar de uma ferramenta não invasiva, altamente sensível e livre de reagentes.

**Palavras-chave:** Doença Tropical Negligenciada. Assinaturas espectrais. *Schistosoma mansoni*

## ABSTRACT

Schistosomiasis is classified as a Neglected Tropical Disease and represents an important public health problem in Brazil, with *Schistosoma mansoni* as the etiological agent in South America. The currently available diagnostic methods present limitations in accuracy, sensitivity, and specificity, in addition to requiring technical expertise both in sample processing and in result interpretation. In this study, Fourier Transform Infrared Spectroscopy with Attenuated Total Reflectance (ATR-FTIR) was employed to identify spectral signatures capable of assisting in the differential diagnosis of schistosomiasis. Serum samples were obtained from the repository of the Laboratory of Immunobiology and Parasite Control at the Federal University of Minas Gerais. For ATR-FTIR analysis, 54 serum samples from individuals positive for schistosomiasis and 58 from non-infected individuals living in endemic and non-endemic areas were used. All schistosomiasis-positive participants presented exclusive infection by *Schistosoma mansoni*, confirmed through complementary parasitological examinations, ensuring the specificity of the analyses. Artificial intelligence techniques, focusing on machine learning methods, were applied for spectral classification. The results indicated that preprocessing with the first derivative provided the best model performance, resulting in more efficient diagnostic classification. Using the Random Forest algorithm, accuracy, sensitivity, specificity, area under the curve, and F1-score values of 80%, 81%, 79%, 0.83, and 0.80, respectively, were obtained. SHAP features extracted from the infrared spectra highlighted the main vibrational modes associated with proteins, nucleic acids, and carbohydrates in serum, responsible for discriminating between individuals infected and non-infected with *S. mansoni*. Therefore, the use of ATR-FTIR combined with machine learning proves to be a promising screening method for endemic areas, particularly due to its non-invasive, highly sensitive, and reagent-free nature.

**Keywords:** Neglected Tropical Disease. Spectral signatures. *Schistosoma mansoni*

## LISTA DE FIGURAS

- Figura 1.** Distribuição mundial das espécies do gênero *Schistosoma* (Fonte: Weerakoon, 2015).....13
- Figura 2.** Ciclo da esquistossomose (Fonte: Adaptado de: Centers for Disease Control and Prevention (CDC)).....14
- Figura 3. A)** Princípio da Espectroscopia no Infravermelho por Transformada de Fourier (FTIR). **B)** Princípio do modo de reflexão interna na espectroscopia de infravermelho. Fonte: Neves et al., 2024.....17
- Figura 4.** Espectro biológico típico mostrando a atribuição de picos biomoleculares de 3.000 a 800  $\text{cm}^{-1}$ , onde  $\nu$  = vibrações de estiramento,  $\delta$  = vibrações de flexão, s = vibrações simétricas e as = vibrações assimétricas. Fonte: Baker et al., 2014.....19
- Figura 5.** Fluxograma representativo da metodologia empregada no estudo. Fonte: Autoria própria.....27
- Figura 6.** Espectros ATR-FTIR médios representativos (3700–900  $\text{cm}^{-1}$ ) no soro de indivíduos negativos e positivos para esquistossomose. A cor azul representa o grupo controle e a cor vermelha o grupo positivo (SCH+). As sombras no espectro representam o desvio padrão de cada grupo.....30
- Figura 7.** Análise de componentes principais (PCA) dos espectros de ATR-FTIR de amostras dos grupos controle (azul) e positivos SCH+ (vermelho). A PCA foi aplicada a espectros truncados abrangendo as regiões lipídica (3050–2800  $\text{cm}^{-1}$ ) e de impressão digital (1800–900  $\text{cm}^{-1}$ ).....31
- Figura 8.** Matriz de confusão obtida pelo algoritmo Random Forest na classificação dos espectros ATR-FTIR para distinção entre amostras do grupo controle e grupo positivo (SCH+).....34

**Figura 9.** Gráfico de importância do recurso (valor SHAP) do modelo Random Forest.....35

## LISTA DE TABELAS

<b>Tabela 1.</b> Características dos estudos incluídos na revisão de literatura (em andamento) sobre a precisão diagnóstica da espectroscopia ATR-FTIR aplicada a doenças parasitárias.....	22
<b>Tabela 2.</b> Características dos pacientes por sexo, idade e carga parasitária (número de ovos por grama de fezes/OPG).....	29
<b>Tabela 3.</b> Algoritmos de aprendizado de máquina aplicados em espectros de soro para discriminar indivíduos negativos e positivos para esquistossomose.....	32
<b>Tabela 4.</b> Atribuição das bandas do valor SHAP de cada modo vibracional do modelo Random Forest para discriminar indivíduos positivos para esquistossomose e negativos e suas respectivas atribuições moleculares. ....	35

## LISTA DE ABREVIATURAS E SIGLAS

AUC – Área sob a curva (*Area Under the Curve*)

ATR – Reflectância Total Atenuada (*Attenuated Total Reflectance*)

ATR-FTIR – Espectroscopia de Infravermelho com Transformada de Fourier por Reflectância Total Atenuada (*Attenuated Total Reflectance–Fourier Transform Infrared Spectroscopy*)

°C – Grau Celsius (unidade de temperatura)

CDC – *Centers for Disease Control and Prevention* (Centros de Controle e Prevenção de Doenças)

cm<sup>-1</sup> – Número de onda (centímetro inverso, unidade usada em espectroscopia)

COEP – Comitê de Ética em Pesquisa

DALYs – *Disability-Adjusted Life Years* (Anos de Vida Ajustados por Incapacidade)

DNA – Ácido desoxirribonucleico (*Deoxyribonucleic Acid*)

DTN - Doença Tropical Negligenciada

DTNs – Doenças Tropicais Negligenciadas

ES- Excreção e secreção

FTIR – Espectroscopia de Infravermelho com Transformada de Fourier (*Fourier Transform Infrared Spectroscopy*)

GBD – *Global Burden of Disease* (Carga Global de Doenças)

kNN – *k-Nearest Neighbors* (k Vizinhos Mais Próximos)

LDA-SVM – *Linear Discriminant Analysis–Support Vector Machine*

LICP – Laboratório de Imunobiologia e Controle de Parasitos

OMS – Organização Mundial da Saúde

OPG – Ovos por grama de fezes

PCA – Análise de Componentes Principais (*Principal Component Analysis*)

PC1 – Primeiro componente principal (*Principal Component 1*)

PC2 – Segundo componente principal (*Principal Component 2*)

POC-CCA – *Point-of-Care Circulating Cathodic Antigen test* – teste rápido para detecção de antígeno catódico circulante

RNA – Ácido ribonucleico (*Ribonucleic Acid*)

SCH+ – Grupo positivo para esquistossomose

SEA- Soluble Egg Antigens (Antígenos solúveis do ovo)

SHAP – *Shapley Additive Explanations*

SVM – *Support Vector Machine* (Máquina de Vetores de Suporte)

TCLE – Termo de Consentimento Livre e Esclarecido

TPP- Target Product Profile (Perfil de Produto Alvo Diagnóstico)

TPP – Perfis de Produto Alvo Diagnóstico (*Target Product Profile*)

UFMG – Universidade Federal de Minas Gerais

VEs- Vesículas extracelulares

μL – Microlitro (unidade de volume equivalente a  $10^{-6}$  litro)

WHO – *World Health Organization* (Organização Mundial da Saúde)

Zika – Doença causada pelo vírus Zika

ZIKV – *Zika Virus*

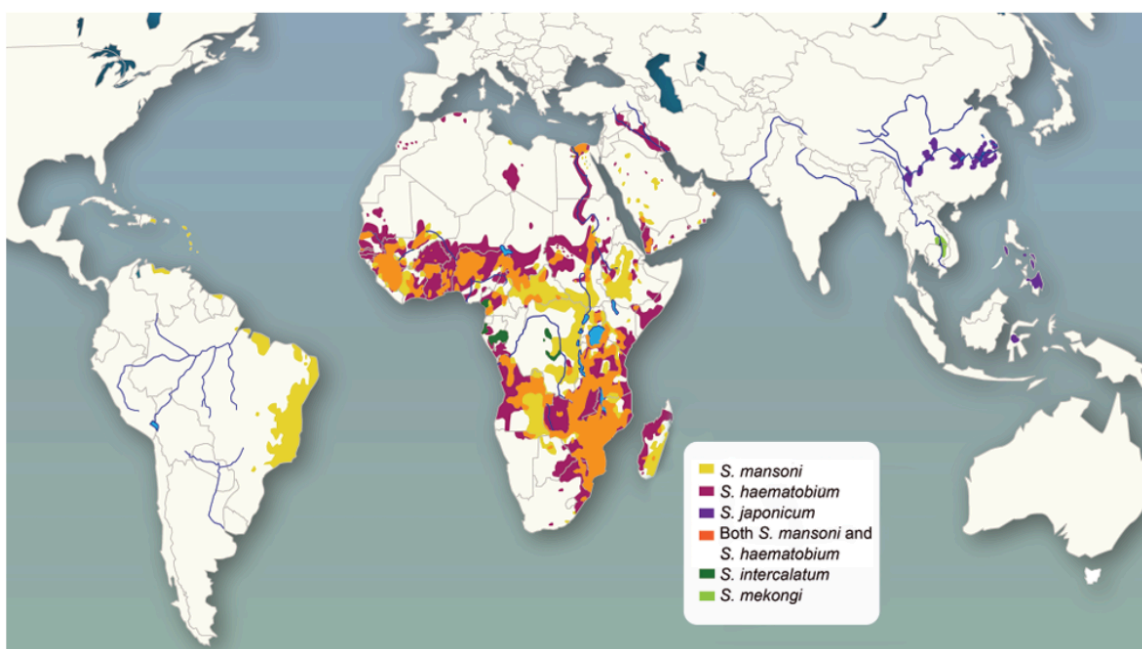
## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>13</b>
1.1 Esquistossomose mansoni.....	13
1.2 Método de Kato-Katz.....	15
1.3. Espectroscopia no Infravermelho por Transformada de Fourier (FTIR).....	16
1.4. ATR-FTIR associada ao aprendizado de máquina no diagnóstico de doenças infecciosas e parasitárias.....	19
<b>2. JUSTIFICATIVA.....</b>	<b>23</b>
<b>3. OBJETIVO GERAL.....</b>	<b>25</b>
3.1. Objetivos específicos.....	25
<b>4. METODOLOGIA.....</b>	<b>26</b>
4.1. Aspectos Éticos.....	26
4.2. Amostras de soro.....	26
4.3. Espectroscopia ATR-FTIR.....	27
4.4. Processamento dos dados espectrais, análise estatística e preditiva.....	27
<b>5. RESULTADOS.....</b>	<b>29</b>
5.1. População de estudo.....	29
5.2. Análise ATR-FTIR de espectros de soro entre grupo controle e grupo positivo.....	29
5.3. Análise de Discriminação.....	31
<b>6. DISCUSSÃO.....</b>	<b>36</b>
<b>7. CONCLUSÃO E PERSPECTIVAS.....</b>	<b>40</b>
<b>8. REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>41</b>

## 1. INTRODUÇÃO

### 1.1 Esquistossomose mansoni

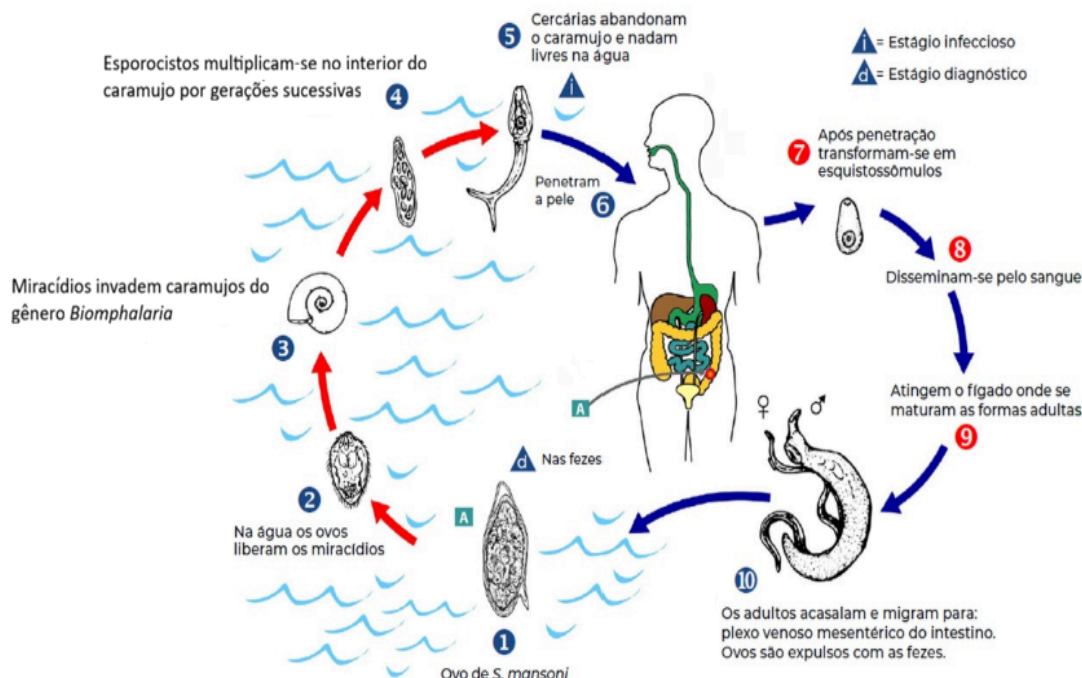
A esquistossomose é uma doença de importância em saúde pública causada principalmente pelos parasitos *Schistosoma mansoni* e *Schistosoma haematobium* (WHO, 2022). Na África Subsaariana e em alguns países da América do Sul, a principal espécie é *S. mansoni*, responsável por manifestações intestinais, enquanto *S. haematobium*, predominante em regiões tropicais e subtropicais da África e do Oriente Médio, está associado a sintomas urogenitais (**Figura 1**) (CDC, 2024).



**Figura 1.** Distribuição mundial das espécies do gênero *Schistosoma* (Fonte: Weerakoon, 2015).

Essa helmintose é responsável pela perda de aproximadamente 1,74 milhão de anos de vida ajustados por incapacidade (DALYs) no mundo (Li et al., 2025). O ciclo biológico do *S. mansoni* é caracterizado pela eliminação de ovos viáveis nas fezes de hospedeiros vertebrados que, em contato com ambientes aquáticos, liberam miracídios. Estes, por sua vez, infectam ativamente hospedeiros intermediários, moluscos do gênero *Biomphalaria*, nos quais ocorre o desenvolvimento e a liberação de cercárias (Nelwan, 2019). As cercárias, estágio infectante

para o hospedeiro vertebrado, penetram ativamente através da pele durante o contato com águas contaminadas (McManaus, 2018) (**Figura 2**).



**Figura 2.** Ciclo biológico do *Schistosoma mansoni* (Fonte: Adaptado de: Centers for Disease Control and Prevention (CDC)).

Por ser uma doença negligenciada, a esquistossomose afeta principalmente populações de baixa renda em regiões com saneamento básico precário. Em um período de 14 anos (2010–2023) foram registrados 528.944 casos de esquistossomose no Brasil, o que a coloca na terceira posição entre as Doenças Tropicais Negligenciadas (DTNs) em número de casos. No mesmo período, as 12 DTNs mais prevalentes somaram 126.503 óbitos na população brasileira, sendo a esquistossomose responsável por 9.597 mortes (7,6%), ocupando a terceira posição também em número de óbitos, atrás apenas da doença de Chagas (69,1%) e da dengue (9,3%) (BRASIL, 2025). Nos últimos anos, a Organização Mundial da Saúde (OMS) lançou um roteiro que visa a eliminação da esquistossomose como um problema de saúde pública em nível global (WHO, 2022). Para isso, testes diagnósticos mais sensíveis e específicos fazem-se necessários.

Testes diagnósticos são fundamentais para a eliminação da esquistossomose, mas muitos métodos ainda apresentam limitações, como baixa sensibilidade e especificidade, além de exigirem procedimentos operacionais complexos, o que dificulta sua aplicação (Fasogbon

et al., 2024). A baixa sensibilidade em áreas endêmicas é uma limitação do método Kato-Katz, método diagnóstico recomendado pela OMS, principalmente em locais com baixa transmissão onde indivíduos infectados tendem a eliminar poucos ovos nas fezes (Assaré et al., 2021; Chieng et al., 2021). Em contrapartida, o método POC-CCA apresenta sensibilidade relativamente maior quando comparado ao Kato-Katz (De Freitas et al., 2021; Fuss; Mazigo; Mueller, 2021; Mewamba et al., 2021), porém a falta de reprodutibilidade entre os lotes e a baixa especificidade são problemas associados a esta técnica (Viana et al., 2019, Graeff-Teixeira et al., 2021). Outros testes parasitológicos, como o Helmintex, apresentam alta sensibilidade, porém são trabalhosos e sua utilização em campo é limitada (Lindholz et al., 2018). Por outro lado, testes moleculares apresentam alta sensibilidade e especificidade, mas o custo é relativamente alto e são limitados como método de triagem (Abdel-Hafeez et al., 2015, Alemu et al., 2025). A utilização do teste ELISA indireto também tem sido discutida na literatura científica como alternativa para identificação inicial de casos, a partir da caracterização de peptídeos imunogênicos ou outros antígenos do parasito. No entanto, essa abordagem ainda apresenta limitações, como maior custo operacional e persistência de anticorpos após a cura (Hinz et al., 2017, Ruas et al., 2025). Diante desse cenário, marcado pela predominância de infecções com baixa carga parasitária e pelas limitações dos métodos diagnósticos atualmente disponíveis, torna-se necessária a investigação de abordagens alternativas, rápidas e sensíveis, que possam ser aplicadas como ferramentas de triagem em áreas endêmicas.

## **1.2 Método de Kato-Katz**

O método diagnóstico preconizado pela OMS para a detecção da esquistossomose é o exame parasitológico de fezes Kato-Katz (WHO, 2023). Esse método quantitativo permite estimar a carga parasitária a partir da contagem de ovos por grama de fezes (OPG), possibilitando classificar a intensidade da infecção em baixa, moderada ou alta (Gomes et al., 2016). O método original foi desenvolvido na década de 1950 por Kato e Miura (Kato & Miura, 1954). Posteriormente, na década de 1970, o pesquisador brasileiro Naftale Katz e seus colaboradores modificaram o método, tornando-o mais prático e adequado para aplicação em campo, dando origem à técnica Kato-Katz (Katz; Chaves; Pellegrino, 1972).

O método de Kato-Katz tem sido amplamente empregado até os dias atuais por sua elevada especificidade e baixo custo, o que o consolidou como a principal ferramenta diagnóstica em inquéritos epidemiológicos e programas de controle da esquistossomose

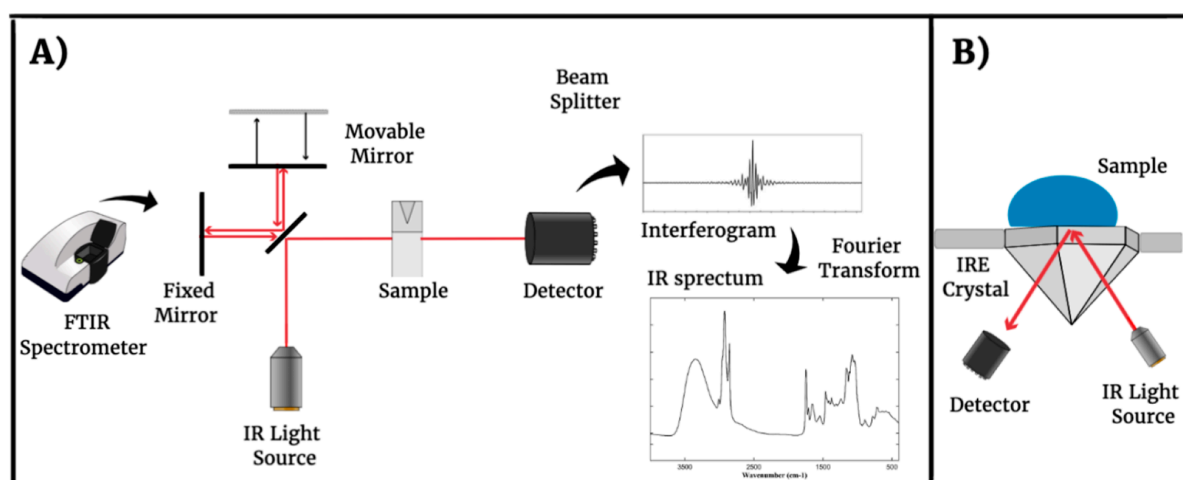
(Worrell et al., 2015, Katz et al., 2018). No entanto, estudos têm demonstrado sua baixa sensibilidade em indivíduos com baixa intensidade de infecção, bem como a necessidade de aumentar o número de lâminas examinadas para otimizar o desempenho diagnóstico (Okoyo et al., 2018, Oliveira et al., 2018). Estudos recentes sugerem que a prevalência da doença esteja subestimada devido à baixa sensibilidade do método de Kato-Katz em infecções leves (Siqueira et al., 2015, Siqueira et al., 2016, Menezes et al., 2023). Além disso, a utilização do Kato-Katz em áreas submetidas a múltiplas rodadas de tratamento pode não ser a abordagem mais adequada, uma vez que a redução da carga parasitária nesses locais compromete a detecção de ovos e, conseqüentemente, a sensibilidade do diagnóstico (Lamberton et al., 2014).

### **1.3. Espectroscopia no Infravermelho por Transformada de Fourier (FTIR)**

A espectroscopia no Infravermelho por Transformada de Fourier (FTIR) baseia-se na interação da radiação infravermelha com a matéria, explorando fenômenos como absorção, reflexão e emissão para investigar a composição das amostras. As frequências mais empregadas situam-se na região do infravermelho médio, entre 4000 e 400  $\text{cm}^{-1}$  (correspondentes a comprimentos de onda de 2,5 a 25  $\mu\text{m}$ ), cujas energias variam de aproximadamente 8 a 40  $\text{kJ mol}^{-1}$  (Chan e Kazarian, 2016). Esse intervalo é suficiente para induzir transições vibracionais nos níveis de energia das moléculas e íons, resultando em modos de vibração característicos (Neves et al., 2024). Dessa forma, as vibrações específicas são influenciadas por esses três fatores: estrutura molecular da amostra, ambiente químico e absorvidade (Davies; Creaser, 1991).

O ensaio de espectroscopia no FTIR é realizado em um espectrômetro de infravermelho equipado com interferômetro de Fourier. Al-Amin et al. (2025) descrevem o princípio dessa técnica da seguinte forma: (a) inicialmente é feita a preparação da amostra que pode ser realizada em diferentes estados físicos, como líquido, sólido ou gasoso; (b) a amostra é então colocada em uma célula de medição adequada, podendo-se utilizar pastilhas de KBr (brometo de potássio) ou placas de cristal que permitem a passagem da radiação infravermelha; (c) a amostra é submetida à incidência de luz infravermelha, abrangendo uma ampla faixa de comprimentos de onda; (d) a interação da radiação com as moléculas da amostra provoca absorções em frequências específicas, correspondentes às vibrações moleculares características de cada ligação química; (e) a radiação transmitida ou refletida pela amostra é coletada pelo espectrômetro e conduzida ao interferômetro de Fourier, que

divide o feixe em dois caminhos: um que atravessa a amostra e outro que não; (f) posteriormente, os dois feixes são recombinados, gerando um interferograma que contém informações sobre a absorção da radiação em diferentes frequências; (g) por fim, o interferograma é submetido à transformada de Fourier, permitindo a obtenção do espectro infravermelho. Este espectro representa a intensidade de absorção em função do número de onda ou do comprimento de onda, fornecendo uma “impressão digital” característica da composição molecular da amostra (**Figura 3 A**)

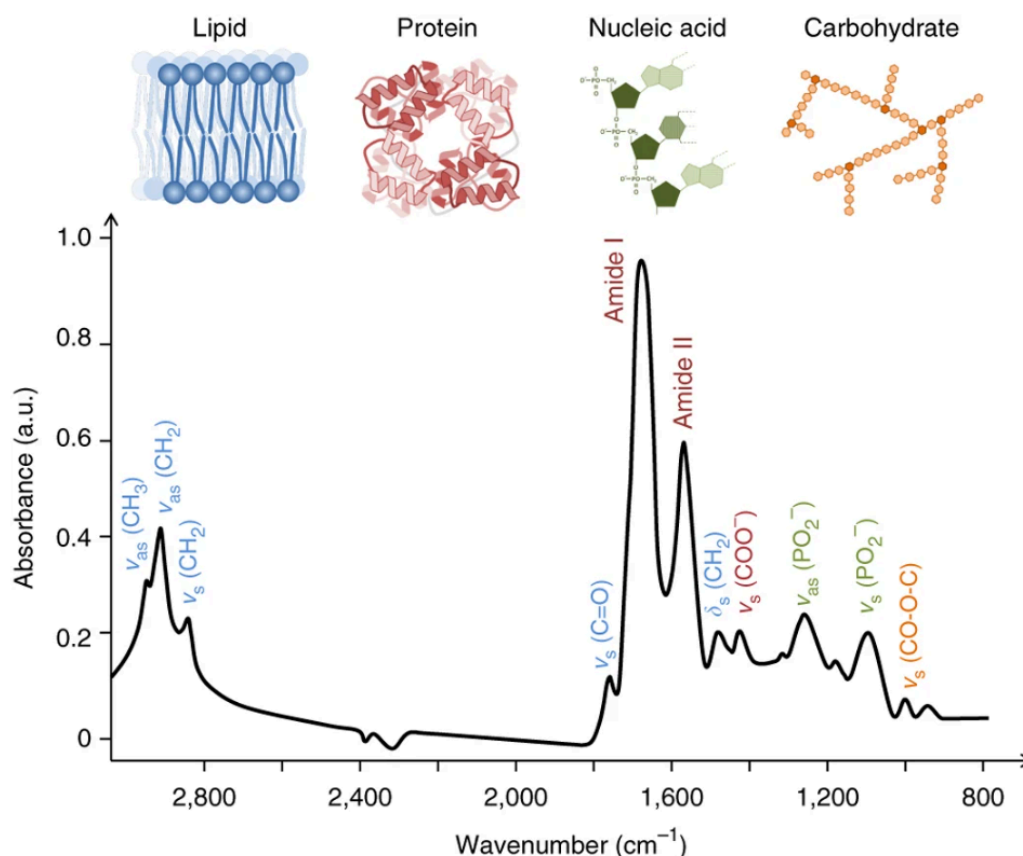


**Figura 3. A)** Princípio da Espectroscopia no Infravermelho por Transformada de Fourier (FTIR). **B)** Princípio do modo de reflexão interna na espectroscopia de infravermelho. Fonte: Neves et al., 2024.

A espectroscopia FTIR analisa as interações entre a matéria e a radiação eletromagnética, resultando em um espectro característico. Esse espectro funciona como uma “impressão digital” molecular exclusiva para cada substância, o que possibilita a sua diferenciação em relação a outras moléculas (Fadlelmoula et al., 2022). Entretanto, a técnica de FTIR apresenta algumas limitações na análise de amostras biológicas, como soro, plasma e urina, destacando-se a sensibilidade superficial e a reprodutibilidade restrita decorrente da variação na espessura do espaçador (Grdadolnik, 2002). Nesse contexto, a Reflectância Total Atenuada (ATR), embora desenvolvida anteriormente como um modo óptico de aquisição, consolidou-se como uma abordagem complementar ao FTIR para a análise de matrizes biológicas, sobretudo após sua integração aos espectrômetros de FTIR (**Figura 3B**). No modo ATR, a radiação infravermelha não é refletida diretamente na superfície da amostra, como ocorre no modo de reflexão convencional (Neves et al., 2024). Em vez disso, a radiação é refletida em um elemento de reflexão interna, o qual deve atender a requisitos específicos, incluindo transparência na faixa de radiação infravermelha, elevado índice de refração e boas

propriedades de transmissão (Larkin, 2011). Esse tipo de técnica busca assinaturas espectrais de moléculas únicas em biofluidos, como plasma, soro e urina, de forma rápida, altamente sensível, portátil e sem a necessidade de reagentes, configurando-se como uma abordagem de baixo custo operacional e adequada para uso em campo, além de ecologicamente sustentável e alinhada às metas da OMS (Honório et al., 2024)

Para compreender os espectros obtidos pela Transformada de Fourier, já foram empiricamente estabelecidas as regiões conhecidas como regiões de impressão digital ou de *biofingerprint*, nas quais é possível analisar de forma mais criteriosa as principais classes de biomoléculas (Baker et al., 2014) (**Figura 4**). Essa técnica é considerada “multiômica” devido à sua capacidade de detectar diversos componentes moleculares (Martins et al., 2025). As regiões espectrais mais relevantes situam-se entre 600–1.450  $\text{cm}^{-1}$ , correspondendo à região de impressão digital, e entre 1.500–1.700  $\text{cm}^{-1}$ , que abrange as bandas das amidas I e II, associadas à proteínas. Já a faixa de 2.550–3.500  $\text{cm}^{-1}$  está relacionada às vibrações de estiramento de grupos funcionais como SH, CH, NH e OH. Em conjunto, essas bandas conferem ao espectro de absorção no infravermelho um padrão característico, frequentemente denominado “impressão digital” molecular da amostra (Ellis; Goodacre, 2006, Movasaghi et al., 2008, Clemens et al., 2014). A capacidade de capturar essas assinaturas bioquímicas complexas tem sustentado o uso da espectroscopia FTIR como uma ferramenta biofísica promissora no diagnóstico de diversas condições patológicas, incluindo câncer, bem como doenças infecciosas e parasitárias (Untereiner et al., 2014, Martinez-Cuazitl et al., 2021, Martins et al., 2025).



**Figura 4.** Espectro biológico típico mostrando a atribuição de picos biomoleculares de 3.000 a 800  $\text{cm}^{-1}$ , onde  $\nu$  = vibrações de estiramento,  $\delta$  = vibrações de flexão, s = vibrações simétricas e as = vibrações assimétricas. Fonte: Baker et al., 2014.

#### 1.4. ATR-FTIR associada ao aprendizado de máquina no diagnóstico de doenças infecciosas e parasitárias

As primeiras aplicações da espectroscopia no infravermelho datam das décadas de 1940 e 1950, quando a técnica, ainda em sua forma dispersiva, era empregada principalmente para a identificação de grupos funcionais e a caracterização estrutural de compostos químicos e materiais biológicos (Randall et al., 1951; Goulden; Sharpe, 1958). O método de ATR foi introduzido posteriormente, no início da década de 1960, com o objetivo de ampliar a aplicabilidade do infravermelho na análise de superfícies e materiais opacos (Fahrenfort, 1961). Entretanto, foi somente a partir das décadas de 1980 e 1990, com o avanço da instrumentação baseada em transformada de Fourier (FTIR) e da capacidade computacional para análise de dados espectrais, que a espectroscopia no infravermelho passou a se consolidar como uma ferramenta robusta para aplicações biomédicas e microbiológicas (Naumann; Helm; Labischinski, 1991; Naumann, 2001). Sendo assim, estudos recentes têm demonstrado a elevada sensibilidade e especificidade da FTIR na identificação e classificação

de doenças, reforçando seu potencial como ferramenta clínica promissora (Al-Kelani; Buthelezi, 2024).

Paralelamente, o avanço da inteligência artificial (IA) tem transformado a forma como dados biológicos são analisados. IA refere-se a sistemas capazes de simular aspectos da cognição humana, como aprendizado, raciocínio e tomada de decisão (Chen; Decary, 2020). Conforme destacado por Taye (2023), “se a inteligência artificial é como um cérebro, então o aprendizado de máquina é o processo pelo qual a IA adquire novas habilidades cognitivas, e o aprendizado profundo é o sistema de autoaprendizagem mais eficaz disponível atualmente”.

Dentro desse contexto, o aprendizado de máquina trata do desenvolvimento de programas computacionais capazes de aprimorar seu desempenho em determinada tarefa a partir da experiência adquirida (Armand et al., 2024). O aprendizado de máquina supervisionado ocorre quando existe o objetivo de prever uma variável dependente a partir de variáveis independentes. Para realizá-lo, a base de dados precisa conter exemplos de casos com a resposta desejada. Assim, os algoritmos aprenderão conforme as respostas já conhecidas e farão uma previsão (Sá; Pessanha; Santos Alves, 2024). Entre os algoritmos mais utilizados destacam-se Logistic Regression, k-Nearest Neighbors (kNN), Neural Networks, Random Forest, Naive Bayes e Support Vector Machine, amplamente empregados em análises de espectroscopia ATR-FTIR para fins diagnósticos (Singh; Thakur; Sharma, 2016, Butler et al., 2019, Mannerkorpi et al., 2025)

Nas infecções parasitárias causadas por protozoários e helmintos, ainda são escassos os estudos que avaliam a precisão diagnóstica da espectroscopia ATR-FTIR associada a algoritmos de aprendizado de máquina (**Tabela 1**). Na literatura disponível, considerando investigações conduzidas em modelos animais e em indivíduos adultos ( $\geq 18$  anos) infectados por parasitos de relevância médica e veterinária, e excluindo estudos exclusivamente *in vitro* ou *in silico*, concentra-se principalmente em agentes como *Plasmodium falciparum*, *Plasmodium ovale*, *Plasmodium vivax*, *Leishmania amazonensis*, *Leishmania infantum*, *Trypanosoma cruzi* e *Trypanosoma evansi* (Heraud et al., 2019; Mwanga et al., 2019; Larios et al., 2021; Pacher et al., 2024; Martins et al., 2025).

No contexto da malária, os estudos de Heraud et al. (2019) e Mwanga et al. (2019) enfocaram validações em campo e infecções naturais, obtendo sensibilidade e especificidade superiores a 90% utilizando sangue como biofluido de interesse. De forma distinta, Larios et al. (2021) avaliaram a precisão diagnóstica da técnica em cães, analisando amostras de soro de 20 animais não infectados, 20 infectados por *L. infantum* e oito infectados por *Trypanosoma evansi*, alcançando uma precisão geral superior a 85%. Pacher et al. (2024), por

sua vez, investigaram a acurácia diagnóstica da espectroscopia ATR-FTIR em diferentes tecidos e biofluidos de camundongos infectados por *L. amazonensis*, obtendo acurácia superior a 72% para as diferentes amostras analisadas. Já Martins et al. (2025), utilizando soro humano seco e úmido de indivíduos infectados por *T. cruzi*, relataram sensibilidade, especificidade e acurácia de 93% para soro seco e de 87% para soro úmido.

Em conjunto, esses resultados evidenciam que o uso da espectroscopia ATR-FTIR associada à aprendizagem de máquina representa um avanço significativo no diagnóstico de doenças parasitárias. No entanto, até o momento, não há estudos que explorem o uso dessa abordagem para o diagnóstico diferencial da esquistossomose, caracterizando o presente trabalho como pioneiro nessa aplicação. Nesse sentido, este estudo visa identificar potenciais padrões de modos vibracionais no soro de indivíduos positivos com diferentes intensidades de infecção para esquistossomose analisados por espectroscopia ATR-FTIR que possam ser utilizados para detectar assinaturas específicas para o diagnóstico diferencial.

**Tabela 1.** Características dos estudos incluídos na revisão de literatura (em andamento) sobre a precisão diagnóstica da espectroscopia ATR-FTIR aplicada a doenças parasitárias.

Autor, ano	País	Animal	Biofluido	Parasito	Sensibilidade	Especificidade	Acurácia	Número de ondas	Algoritmo	Grupo positivo (N)	Grupo negativo (N)			
Heraud et al., 2019	Tailândia	Humano	Sangue total	<i>P. falciparum</i> e <i>P. vivax</i>	92%	97%	-	3140-1006 cm <sup>-1</sup>	SVM	112	110			
Mwanga et al., 2019	Tanzânia	Humano	Sangue total	<i>P. falciparum</i> e <i>P. ovale</i>	92,80%	91,70%	92,30%	1730-883 cm <sup>-1</sup>	Logistic Regression	123	173			
Larios et al., 2021	Brasil	Cão	Soro	<i>L. infantum</i> e <i>T. evansi</i>	-	-	85,42% (NI x Li x Te)	1800-800 cm <sup>-1</sup>	KNN	28	20			
					-	-	87,5% (NI x Li)	1700-1450 cm <sup>-1</sup>	Quadratic DA	20	20			
					-	-	100% (Li x Te)	1800-800 cm <sup>-1</sup>	Linear/Cubic SVM	28	0			
					-	-	100% (NI x Te)	1800-800 cm <sup>-1</sup>	(SVM ou KNN)	8	20			
Pacher et al., 2024	Brasil	Camundongo BALB	Soro, lesão de pele e linfonodo	<i>L. amazonensis</i>	-	-	-	1800-700 cm <sup>-1</sup>	-	30	30			
			Soro				72,20%		SVM					
			Lesão de pele				83,30%		SVM					
			Linfonodo				72,00%		SVM					
Martins et al., 2025	Brasil	Humano	Soro seco e úmido	<i>T. cruzi</i>	-	-	-	1700-1500 cm <sup>-1</sup>	-	49	51			
			Soro seco						93%			93%	93%	Logistic regression
			Soro úmido						87%			87%	87%	XGBoos

Nota: O símbolo “-” indica ausência de informação no artigo. As siglas NI: não infectado, Li: *Leishmania infantum* e Te: *Trypanosoma evansi*.

## 2. JUSTIFICATIVA

A esquistossomose é uma doença de importância em saúde pública, estando listada pela Organização Mundial da Saúde (OMS) como uma das doenças tropicais negligenciadas. No Brasil, segundo o inquérito nacional publicado em 2018, em que foram examinados 197.564 escolares de 521 municípios, os resultados mostraram uma proporção de positivos de 0,99% (IC95% - 0,20 a 1,78), evidenciando que a doença ainda persiste em diversas localidades do país (Katz et al., 2018). O teste diagnóstico preconizado pela OMS como referência para o diagnóstico da esquistossomose é o método Kato-Katz, porém sua baixa sensibilidade em indivíduos que apresentam baixa carga parasitária representa uma limitação importante. A confiabilidade diagnóstica desse método aumenta quando múltiplas lâminas são analisadas por amostra de fezes ou quando é empregado em associação com outros testes (Oliveira et al., 2018).

Uma das metas da OMS é a eliminação da esquistossomose como problema em saúde pública até 2030, considerando sua ampla distribuição em regiões tropicais e subtropicais, sendo endêmica em mais de 70 países (WHO, 2022; WHO, 2022). Para isso, a busca por diagnósticos mais sensíveis e específicos e de fácil aplicabilidade se faz necessária, vistas as limitações já descritas dos atuais métodos diagnósticos. A necessidade é ainda mais urgente diante da transição epidemiológica observada em várias regiões brasileiras, caracterizada por focos residuais e infecções de baixa intensidade, que exigem métodos mais robustos para detecção precoce.

Diferentes estudos têm evidenciado a alta sensibilidade e especificidade da espectroscopia no Infravermelho por Transformada de Fourier com Reflectância Total Atenuada (ATR-FTIR) no diagnóstico de diversas doenças. Além disso, a aplicabilidade da técnica em condições de campo já foi comprovada, evidenciando seu potencial como uma ferramenta promissora, portátil e de fácil operação. Outro aspecto relevante é a ausência de necessidade de reagentes químicos, sua flexibilidade para análise de diferentes tipos de amostras, bem como a elevada reprodutibilidade e a rapidez nas análises obtidas (Heraud et al., 2019, Morais et al., 2020).

No entanto, até o momento, não há registros de estudos que avaliem o potencial da espectroscopia ATR-FTIR para o diagnóstico da esquistossomose em amostras de soro de indivíduos com diferentes intensidades de infecção. Nesse contexto, a integração da espectroscopia ATR-FTIR a algoritmos de inteligência artificial surge como uma estratégia inovadora, capaz de extrair padrões espectrais complexos e aprimorar a acurácia diagnóstica.

Essa abordagem pode superar limitações observadas nos métodos parasitológicos tradicionais, oferecendo uma alternativa rápida, reprodutível, de baixo custo operacional e aplicável em campo.

Diante do compromisso global de eliminar a esquistossomose como problema de saúde pública até 2030 e das limitações do teste de referência preconizado pela OMS, torna-se essencial o desenvolvimento e a adoção de novas abordagens diagnósticas que apresentem maior sensibilidade e especificidade. Assim, o presente projeto pretende preencher essa lacuna, avaliando o potencial da espectroscopia ATR-FTIR associada à inteligência artificial como ferramenta complementar para o diagnóstico da esquistossomose, com potencial de aplicação em cenários endêmicos de diferentes intensidades.

### 3. OBJETIVO GERAL

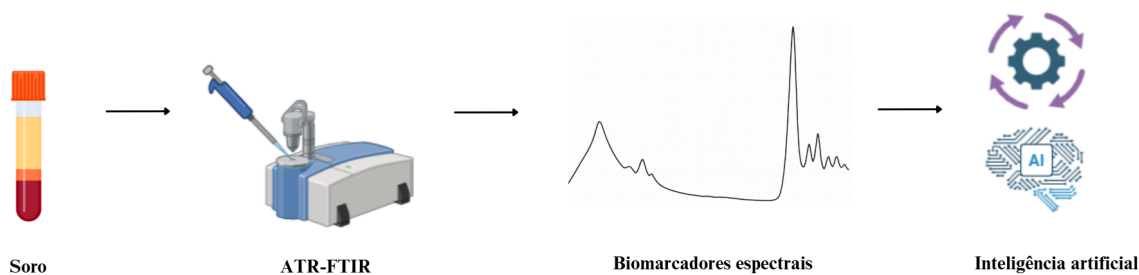
Identificar padrões de modos vibracionais específicos no soro de indivíduos positivos para esquistossomose com diferentes intensidades de infecção, analisados por espectroscopia no infravermelho por transformada de Fourier com reflectância total atenuada (ATR-FTIR), com potencial aplicação como ferramenta de triagem em cenários de baixa intensidade de infecção.

#### 3.1. Objetivos específicos

- Caracterizar o perfil espectral dos soros de indivíduos infectados e não infectados, identificando padrões vibracionais associados à alterações metabólico-imunológicas sistêmicas relacionadas à infecção;
- Caracterizar os padrões de modos vibracionais do soro de pacientes positivos para esquistossomose com diferentes intensidades de infecção;
- Avaliar a capacidade dos diferentes algoritmos em discriminar entre o grupo controle e o grupo positivo infectado por *S. mansoni* (SCH+).
- Identificar potenciais biomarcadores espectrais no soro associados à infecção por *S. mansoni*

## 4. METODOLOGIA

O fluxograma da metodologia utilizada neste estudo está representada na **Figura 5**.



**Figura 5.** Fluxograma representativo da metodologia empregada no estudo.

### 4.1. Aspectos Éticos

O presente estudo foi conduzido em conformidade com as diretrizes do Comitê de Ética em Pesquisa (COEP) da Universidade Federal de Minas Gerais (UFMG), sob o protocolo CAAE nº 55239522.3.0000.5149. Todos os participantes foram informados sobre os objetivos e procedimentos da pesquisa e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) antes de sua inclusão no estudo.

### 4.2. Amostras de soro

As amostras de soro utilizadas neste estudo foram obtidas do repositório do Laboratório de Imunobiologia e Controle de Parasitos (LICP) da Universidade Federal de Minas Gerais (UFMG). Ao todo, foram disponibilizadas 54 amostras provenientes de pacientes positivos para esquistossomose, apresentando diferentes intensidades de infecção com base nos resultados do teste Kato-Katz, além de 34 amostras negativas de indivíduos residentes em área endêmica e 24 negativas provenientes de área não endêmica. Todas as amostras apresentavam soro bem preservado, sem sinais de hemólise, e foram aliquotadas e mantidas a  $-80\text{ }^{\circ}\text{C}$  até o momento das análises. A confirmação do status de infecção foi realizada por meio das técnicas parasitológicas Kato-Katz e Helminex. Para a etapa de análise espectroscópica por ATR-FTIR e com o objetivo de minimizar risco de ajustes excessivos (*overfitting*) nos modelos de aprendizado de máquina supervisionados, todas as amostras negativas foram agrupadas em um único grupo controle, totalizando 58 amostras.

### 4.3. Espectroscopia ATR-FTIR

Os espectros das amostras de soro foram registrados no Laboratório de Nanotecnologia da Universidade Federal de Uberlândia, usando o espectrômetro portátil Agilent Cary 630 FTIR (Agilent Technologies, Santa Clara, EUA) conectado ao modo de refletância total microatenuada (ATR) na faixa de 4000 a 650  $\text{cm}^{-1}$ . Os espectros foram obtidos com resolução espectral de 2  $\text{cm}^{-1}$  e 32 varreduras por amostra. As amostras (1  $\mu\text{L}$ ) foram colocadas diretamente sobre o cristal de diamante ATR com o auxílio de micropipeta e, em seguida, após 3 minutos de secagem, foi realizado o registro dos espectros infravermelhos. Cada amostra teve sua leitura espectral realizada em duplicata, a fim de assegurar a precisão dos resultados. Os espectros foram obtidos em um ambiente controlado (temperatura 22–23  $^{\circ}\text{C}$ ) e em cada leitura, a medição do background foi realizada.

### 4.4. Processamento dos dados espectrais, análise estatística e preditiva

O processamento dos espectros foi realizado com o uso do software Orange 3.3.5. As duas leituras espectrais de cada indivíduo foram combinadas por média aritmética para reduzir o efeito de variações instrumentais. Para gerar um perfil espectral médio, os espectros foram submetidos a ajustes de normalização e linha de base utilizando o método Rubberband. Para a etapa exploratória e de visualização multivariada, foi utilizada a Análise de Componentes Principais (PCA), aplicou-se correção de linha de base pelo método Rubberband seguida de normalização Min-Max, permitindo a inspeção do agrupamento natural das amostras com base nos componentes principais (Guevara-Vega et al., 2024).

Em seguida, todos os espectros foram truncados para as regiões de maior relevância bioquímica: 3050–2800  $\text{cm}^{-1}$  (região lipídica) e 1800–900  $\text{cm}^{-1}$  (região de impressão digital), que foram utilizadas nas análises subsequentes. A análise dos espectros infravermelhos foi conduzida em duas etapas principais: pré-processamento e classificação. Para a construção dos modelos classificatórios, os espectros truncados foram avaliados sob diferentes estratégias de pré-processamento aplicadas de forma independente, com o objetivo de comparar seu impacto no desempenho preditivo: Raw data, Min-Max + Rubberband, 1st derivative, 2nd derivative, ALSS smoothing e PCA denoising.

Foram testados sete algoritmos de aprendizado de máquina supervisionados: Neural Network, Random Forest, Logistic Regression, k-Nearest Neighbors (kNN), AdaBoost, Naive

Bayes e Support Vector Machine (SVM). Para analisar o desempenho preditivo desses algoritmos, foi utilizada a validação cruzada estratificada com 10-folds. As métricas de desempenho empregadas para medir os resultados foram sensibilidade, especificidade e acurácia, calculadas a partir da matriz de confusão (Caixeta et al., 2023)

## 5. RESULTADOS

### 5.1. População de estudo

As características demográficas e parasitológicas dos grupos controle e positivo para esquistossomose (SCH+) estão apresentadas na **Tabela 2**. No grupo SCH+, houve distribuição equilibrada entre os sexos, sendo 50% (n = 27) do sexo feminino e 50% (n = 27) do sexo masculino. A média de idade foi de  $35,8 \pm 18,9$  anos. A carga parasitária apresentou mediana de 20 ovos por grama de fezes (OPG). Entre os indivíduos positivos, 39 apresentaram baixa intensidade de infecção (1-100 OPG), 7 apresentaram intensidade moderada (101-400 OPG) e 8 apresentaram alta intensidade de infecção (>400 OPG), segundo a classificação obtida pelo método de Kato-Katz. No grupo controle, 44% (n = 26) dos participantes eram do sexo masculino e 56% (n = 32) do sexo feminino. A média de idade foi de  $32,1 \pm 16,1$  anos, próxima à observada no grupo SCH+.

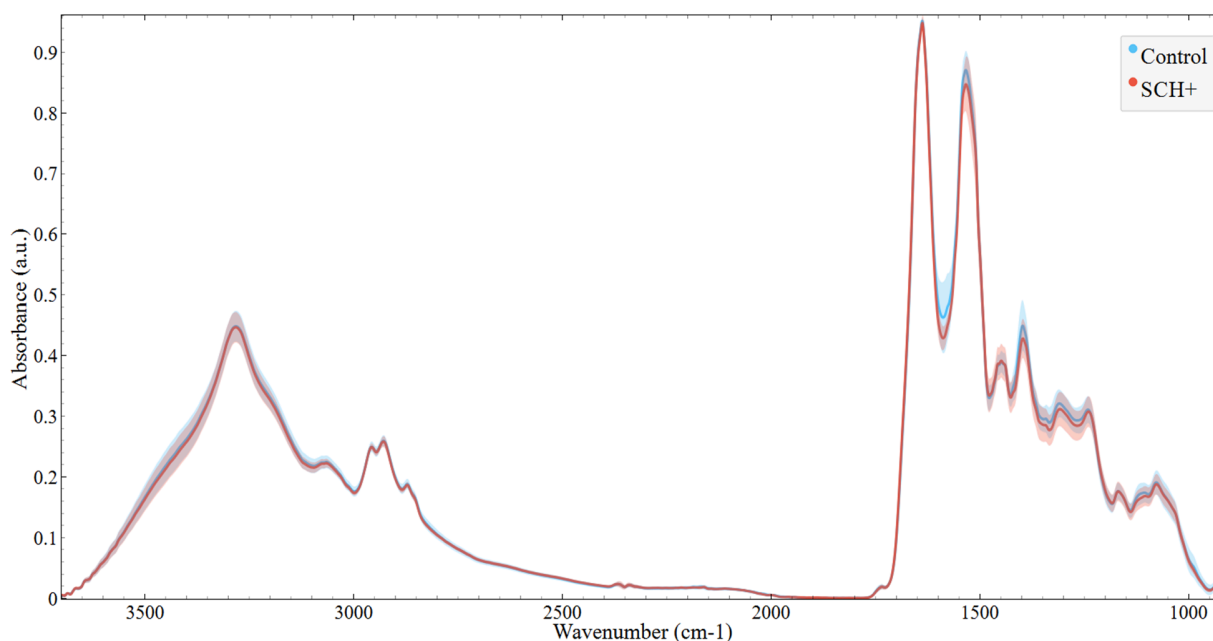
**Tabela 2.** Características dos pacientes por sexo, idade e carga parasitária (número de ovos por grama de fezes/OPG)

Parâmetros	Controle	SCH+
Gênero masculino	44% (n=26)	50% (n=27)
Gênero feminino	56% (n=32)	50% (n=27)
Idade (anos)	$32,1 \pm 16,1$	$35,8 \pm 18,9$
OPG	-	20 OPG

Os dados são apresentados como média  $\pm$  desvio padrão (DP), exceto para a variável ovos por grama de fezes (OPG), que é expressa como mediana.

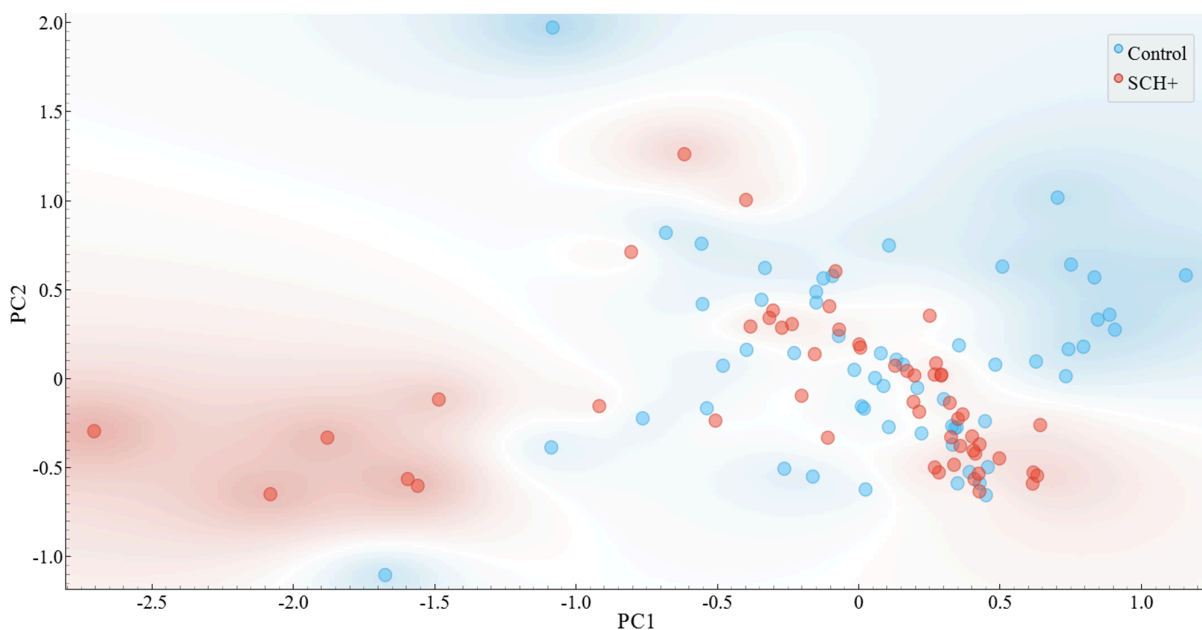
### 5.2. Análise ATR-FTIR de espectros de soro entre grupo controle e grupo positivo

O perfil espectral do infravermelho médio do soro na região de  $3700-900 \text{ cm}^{-1}$  está representado na **Figura 6**. Foi observado uma sobreposição em algumas regiões de absorção de proteínas, lipídios, DNA/RNA e carboidratos, com uma variação moderada que sugere uma resposta semelhante entre as condições, mas com algumas alterações no perfil espectral entre o grupo controle e o grupo positivo (SCH+) (**Figura 6**).



**Figura 6.** Espectros ATR-FTIR médios representativos (3700–900 cm<sup>-1</sup>) no soro de indivíduos negativos e positivos para esquistossomose. A cor azul representa o grupo controle e a cor vermelha o grupo positivo (SCH+). As sombras no espectro representam o desvio padrão de cada grupo.

O PCA foi usado para reduzir a dimensionalidade, comparando as mudanças espectrais no grupo controle e SCH+ com espectros truncados, incluindo a região lipídica (3050–2800 cm<sup>-1</sup>) associada à região da impressão digital (1800-900 cm<sup>-1</sup>). Na etapa de classificação não supervisionada, os resultados de PC1 e PC2 explicaram, respectivamente, 51% e 27% da variabilidade total, totalizando 78%, indicando uma moderada diferença de distribuição entre o grupo controle e o grupo positivo (SCH+) (**Figura 7**)



**Figura 7.** Análise de componentes principais (PCA) dos espectros de ATR-FTIR de amostras dos grupos controle (azul) e positivos SCH+ (vermelho). A PCA foi aplicada a espectros truncados abrangendo as regiões lipídica ( $3050\text{--}2800\text{ cm}^{-1}$ ) e de impressão digital ( $1800\text{--}900\text{ cm}^{-1}$ ).

### 5.3. Análise de Discriminação

Para promover a classificação e discriminação dos espectros de soro entre indivíduos positivos e negativos para *S. mansoni*, foram aplicadas técnicas de inteligência artificial, utilizando algoritmos de aprendizado de máquina (*machine learning*), destacando-se o algoritmo Random Forest com o melhor desempenho geral (**Tabela 3**). Inicialmente, os espectros foram submetidos a diferentes estratégias de pré-processamento, incluindo Raw data, Min-Max + Rubberband, 1st derivative, 2nd derivative, ALSS smoothing e PCA denoising. Entre essas estratégias, o pré-processamento com 1st derivative apresentou o melhor desempenho. Além disso, verificou-se que a combinação das faixas espectrais de  $3050\text{--}2800\text{ cm}^{-1}$  e  $1800\text{--}900\text{ cm}^{-1}$  referentes às regiões lipídica e de *biofingerprint* proporcionou maior discriminação entre os grupos. Dentre os algoritmos testados, o Random Forest obteve os melhores parâmetros de discriminação, com acurácia de 80%, sensibilidade de 81%, especificidade de 79%, área sob a curva (AUC) de 0.83 e F1-score de 0.80, indicando a classificação mais equilibrada entre precisão e robustez.

**Tabela 3.** Algoritmos de aprendizado de máquina aplicados em espectros de soro para discriminar indivíduos negativos e positivos para esquistossomose.

Pré-Processamento (Banda)	Algoritmo	Acurácia	F1	Sensibilidade	Especificidade	AUC
Raw data (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.73	0.74	0.81	0.65	0.81
	kNN	0.72	0.73	0.81	0.63	0.77
	Neural Network	0.67	0.66	0.66	0.69	0.74
	Random Forest	0.64	0.63	0.63	0.65	0.71
	Naive Bayes	0.62	0.61	0.61	0.62	0.65
	AdaBoost	0.59	0.56	0.54	0.64	0.59
	SVM	0.52	0.23	0.15	0.86	0.60
Min-Max + Rubberband (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.74	0.76	0.85	0.64	0.80
	kNN	0.66	0.68	0.76	0.57	0.69
	Neural Network	0.67	0.68	0.72	0.62	0.76
	Random Forest	0.69	0.70	0.74	0.64	0.80
	Naive Bayes	0.56	0.56	0.57	0.55	0.63
	AdaBoost	0.62	0.61	0.61	0.64	0.62
	SVM	0.61	0.70	0.94	0.29	0.53
1st derivative (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.71	0.70	0.70	0.72	0.80
	kNN	0.68	0.70	0.80	0.57	0.75
	Neural Network	0.75	0.74	0.76	0.74	0.81
	<b>Random Forest</b>	<b>0.80</b>	<b>0.80</b>	<b>0.81</b>	<b>0.79</b>	<b>0.83</b>
	Naive Bayes	0.70	0.70	0.74	0.85	0.80
	AdaBoost	0.65	0.65	0.67	0.64	0.66
	SVM	0.60	0.69	0.94	0.28	0.76

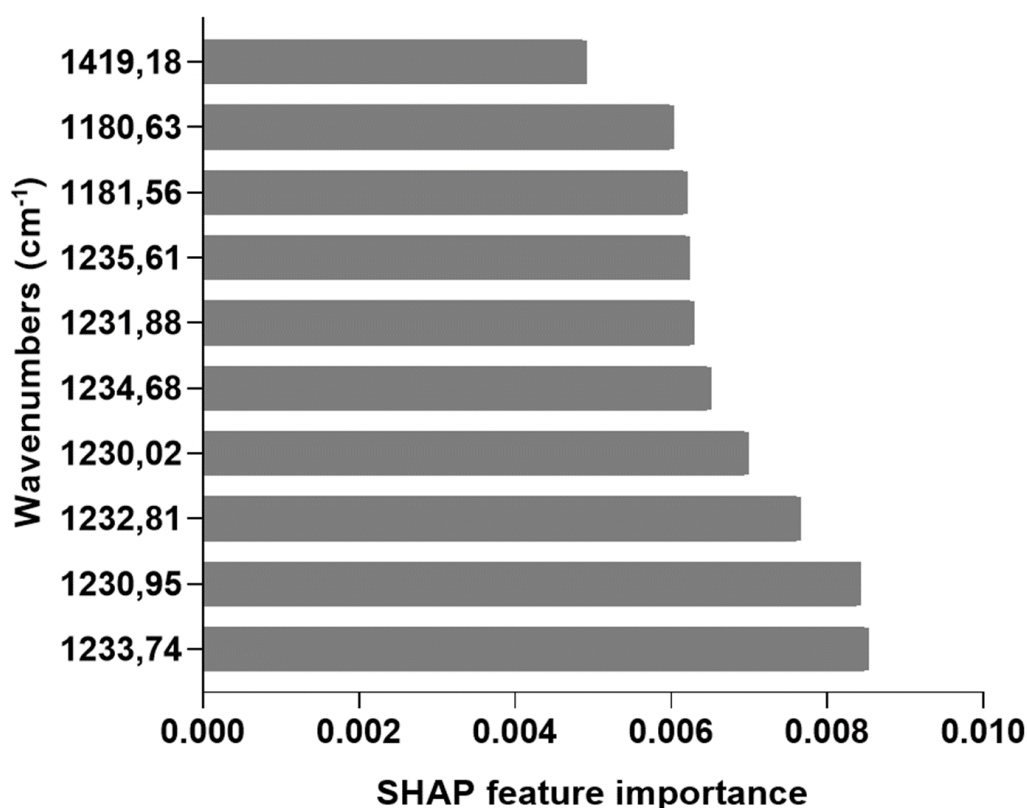
2 nd derivative (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.58	0.43	0.33	0.81	0.74
	kNN	0.70	0.66	0.61	0.78	0.69
	Neural Network	0.58	0.51	0.46	0.69	0.69
	Random Forest	0.67	0.67	0.70	0.64	0.74
	Naive Bayes	0.70	0.71	0.78	0.62	0.75
	AdaBoost	0.55	0.54	0.54	0.57	0.56
	SVM	0.54	0.07	0.04	1.00	0.36
ALSS smoothing (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.74	0.76	0.85	0.64	0.81
	kNN	0.63	0.67	0.76	0.52	0.71
	Neural Network	0.63	0.64	0.68	0.59	0.74
	Random Forest	0.72	0.72	0.72	0.72	0.76
	Naive Bayes	0.70	0.74	0.85	0.57	0.73
	AdaBoost	0.50	0.51	0.54	0.47	0.50
	SVM	0.62	0.71	0.96	0.29	0.74
PCA denoising (1800–900 cm <sup>-1</sup> ; 3050–2800 cm <sup>-1</sup> )	Logistic Regression	0.72	0.74	0.81	0.64	0.80
	kNN	0.70	0.69	0.68	0.72	0.79
	Neural Network	0.73	0.75	0.85	0.62	0.77
	Random Forest	0.72	0.70	0.67	0.78	0.79
	Naive Bayes	0.63	0.61	0.59	0.67	0.66
	AdaBoost	0.65	0.63	0.61	0.69	0.65
	SVM	0.59	0.69	0.96	0.24	0.63

A matriz de confusão (**Figura 8**) gerada pelo algoritmo Random Forest revelou que 44 amostras positivas para *S. mansoni* foram classificadas como verdadeiros positivos e 46 amostras de indivíduos não infectados como verdadeiros negativos. Observou-se ainda que 12 amostras controle foram classificadas como positivas (falsos positivos) e 10 amostras positivas foram classificadas como negativas (falsos negativos). Com base nesses dados, o modelo apresentou valor preditivo positivo (VPP) de 78,6% e valor preditivo negativo (VPN) de 82,1%, indicando que 78,6% das amostras classificadas como positivas correspondem efetivamente a indivíduos infectados, enquanto 82,1% das amostras classificadas como negativas representam corretamente indivíduos não infectados.

	Controle	SCH+	$\Sigma$
Controle	46	12	58
SCH+	10	44	54
$\Sigma$	56	56	112

**Figura 8.** Matriz de confusão obtida pelo algoritmo Random Forest na classificação dos espectros ATR-FTIR para distinção entre amostras do grupo controle e grupo positivo (SCH+).

Para identificar quais recursos espectrais têm maior influência na análise e explicar o modelo de aprendizado de máquina Random Forest, foi gerado o gráfico de pontuações de importância de recursos por meio dos valores SHAP (*Shapley Additive Explanations*), que permitiu identificar os modos vibracionais que mais contribuíram para o desempenho preditivo do algoritmo (**Figura 9**). Como resultado, os principais modos vibracionais para discriminar amostras do grupo controle e positivo (SCH+) foram 1233,74  $\text{cm}^{-1}$ , 1230,95  $\text{cm}^{-1}$ , 1232,81  $\text{cm}^{-1}$ , 1230,02  $\text{cm}^{-1}$ , 1234,68  $\text{cm}^{-1}$ , 1231,88  $\text{cm}^{-1}$ , 1235,61  $\text{cm}^{-1}$ , 1181,56  $\text{cm}^{-1}$ , 1180,63  $\text{cm}^{-1}$ , 1419,18  $\text{cm}^{-1}$ .



**Figura 9.** Gráfico de importância do recurso (valor SHAP) do modelo Random Forest.

As atribuições moleculares de cada modo vibracional indicado pela análise de características SHAP-Random Forest são descritas na **Tabela 4**.

**Tabela 4.** Atribuição das bandas do valor SHAP de cada modo vibracional do modelo Random Forest para discriminar indivíduos positivos para esquistossomose e negativos e suas respectivas atribuições moleculares.

Número de onda	Designação de Banda	Componentes Biomoleculares
1419,18 cm <sup>-1</sup>	$\nu(\text{COO}^-)$	Carboidratos
1180,63 cm <sup>-1</sup>	$\nu(\text{Amida III}) \approx \delta(\text{C-N}), \delta(\text{C=O})$	Proteínas
1181,56 cm <sup>-1</sup>	$\nu(\text{Amida III}) \approx \delta(\text{C-N}), \delta(\text{C=O})$	Proteínas
1235,61 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1231,88 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1234,68 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1230,02 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1232,81 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1230,95 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas
1233,74 cm <sup>-1</sup>	$\nu(\text{P=O})$ ou $\delta(\text{O-P-O})$ ou amida III	Ácidos nucleicos e Proteínas

**Fonte:** Movasaghi; Rehman; Rehman, 2008, Kochan et al., 2020, Bel'skaya et al., 2021

## 6. DISCUSSÃO

A Organização Mundial da Saúde (OMS) recomenda o método Kato-Katz como teste de referência para o diagnóstico da esquistossomose, sendo utilizado em inquéritos epidemiológicos e como ferramenta nas ações de vigilância e controle em áreas endêmicas, como realizado no Brasil (WHO, 1985, WHO, 2023, BRASIL, 2024). Por ser um teste de baixo custo e apresentar alta especificidade, esse método diagnóstico foi fundamental, juntamente com o tratamento quimioterápico e as ações em áreas endêmicas, para a redução da prevalência e da intensidade de transmissão da esquistossomose em diversas regiões do mundo (Burlandy et al., 2003, Fenwick; Webster, 2006, Utzinger et al., 2009).

Com isso, um novo cenário epidemiológico se estabelece na maioria do território brasileiro, onde a esquistossomose é caracterizada por focos residuais e infecções de baixa intensidade que exigem métodos mais robustos para detecção precoce. Em indivíduos com baixa intensidade de infecção e que, conseqüentemente, eliminam poucos ovos nas fezes, o teste de Kato-Katz apresenta sensibilidade reduzida (Bärenbold et al., 2017). Além disso, como requer apenas uma pequena quantidade de fezes para a análise, os ovos podem estar distribuídos de maneira desigual, comprometendo a detecção (Utzinger et al., 2001, Krauth et al., 2012). Embora o método Kato-Katz seja recomendado pela OMS como teste de referência, sua baixa sensibilidade em cenários de baixa carga parasitária impõe desafios diagnósticos importantes. No presente estudo, a predominância de indivíduos com baixa intensidade de infecção reforça esse cenário e justifica a busca por abordagens alternativas capazes de detectar alterações sistêmicas associadas à infecção, mesmo em condições de baixa carga parasitária.

A técnica *Helmintex* surge como uma alternativa para suprir a baixa sensibilidade do Kato-Katz, apresentando elevada capacidade de detecção de ovos de *S. mansoni* (Teixeira et al., 2007). No entanto, trata-se de uma técnica trabalhosa e não aplicável como diagnóstico de rotina e em campo (Lindholz et al., 2018). Sendo assim, tanto o Kato-Katz quanto o *Helmintex* utilizam as fezes como amostra de interesse para o diagnóstico. Como abordagem alternativa aos métodos parasitológicos tradicionalmente empregados na rotina médica, o presente estudo utiliza o soro como biofluido de interesse para investigação diagnóstica usando a espectroscopia ATR-FTIR. O soro constitui uma matriz acessível e informativa para análise diagnóstica, apresentando vantagens como baixo custo, alta disponibilidade e capacidade de refletir estados patológicos precoces através de biomoléculas e metabólitos diferencialmente expressados (Zhang; Sun; Wang, 2012, Zhang et al., 2013). Até o momento,

não há registros de uso da espectroscopia ATR-FTIR utilizando soro ou outros biofluidos aplicada ao diagnóstico da esquistossomose, o que reforça o caráter inovador deste estudo e sua relevância para cenários endêmicos.

É importante destacar que, ao longo do ciclo biológico no hospedeiro definitivo, *S. mansoni* mantém contato direto e contínuo com a corrente sanguínea.. Os esquistossômulos migram pela circulação pulmonar e sistêmica, enquanto os vermes adultos se estabelecem predominantemente nas veias mesentéricas associadas ao sistema porta-hepático, onde se alimentam de sangue e liberam metabólitos, antígenos e moléculas de excreção e secreção (McManus et al., 2020). Esse contato com a circulação sanguínea desencadeia uma resposta imunológica sistêmica, com produção de anticorpos, liberação de citocinas e de antígenos circulantes, além de alterar proteínas plasmáticas, metabólitos e componentes celulares que influenciam diretamente a composição bioquímica do soro (Malaquias et al., 1997, Corrêa-Oliveira et al., 1998, Caldas et al., 2008). Assim, o uso de soro como matriz analítica é biologicamente justificável, pois reflete não apenas a presença de antígenos do parasito, mas também as modificações induzidas no metabolismo e na resposta imune do hospedeiro, induzindo alterações capazes de gerar assinaturas espectrais específicas e detectáveis por ATR-FTIR.

A técnica de espectroscopia no infravermelho por refletância total atenuada com transformada de Fourier (ATR-FTIR) permite a análise de diferentes biofluidos por meio da interação da radiação infravermelha com as ligações químicas presentes na amostra, resultando em espectros capazes de fornecer informações químicas detalhadas para análises subsequentes (Sala et al., 2020, Naseer; Salmann; Qazi, 2021). A técnica apresenta vantagens significativas, incluindo portabilidade do equipamento, aplicabilidade comprovada em campo, baixo custo, ausência de necessidade de reagentes e capacidade de adaptação a distintos tipos de amostras (Bellisola et al., 2012, Heraud *et al.*, 2019, Caixeta et al., 2023). Em cenários endêmicos para a esquistossomose, essas características tornam a técnica especialmente promissora, sobretudo para aplicações em estratégias de triagem. Além disso, há a perspectiva de expansão do seu uso para biofluidos menos invasivos, como a saliva, ampliando ainda mais sua viabilidade em contextos de vigilância e diagnóstico em campo.

A partir dos espectros obtidos por ATR-FTIR aplicam-se análises estatísticas e métodos de aprendizado de máquina, como o Random Forest, um algoritmo baseado em conjuntos de árvores de decisão que pode ser utilizado tanto para classificação quanto para análises de regressão (Breiman, 2001). Esse algoritmo é um método de predição poderoso, conhecido por sua capacidade de capturar padrões complexos de dependência entre o

desfecho e as covariáveis (Acharjee et al., 2020). Dentre todos os algoritmos testados, o Random Forest mostrou uma acurácia de 80%, sensibilidade de 81%, especificidade de 79%, área sob a curva (AUC) de 0.83 e F1-score de 0.80, apresentando o melhor resultado. Entre os indivíduos positivos, a mediana da intensidade de infecção estimada pelo método Kato-Katz foi de 20 ovos por grama de fezes (OPG). Observou-se que a maioria dos indivíduos apresentou baixa carga parasitária (34/54), padrão compatível com o cenário epidemiológico atual da esquistossomose no Brasil, caracterizado por infecções de baixa intensidade em áreas endêmicas.

Os Perfis de Produto Alvo Diagnóstico (TPP) para monitoramento, avaliação e vigilância dos programas de controle da esquistossomose sugerem uma sensibilidade ideal >75% para amostras com pelo menos 100 indivíduos em testes de triagem diagnóstica (WHO, 2021). Em nosso estudo, o pré-processamento com a primeira derivada, aliado ao uso do algoritmo Random Forest, alcançou sensibilidade de 81%, superando o valor recomendado para fins de triagem inicial. Por outro lado, a especificidade apresentou valores inferiores aos 96,5% propostos pelos TPP. Dessa forma, essa técnica pode se mostrar promissora para aplicação em áreas endêmicas como método de triagem inicial, em conjunto com testes diagnósticos de alta especificidade, como o Kato-Katz. Os valores de acurácia, sensibilidade e especificidade obtidos neste estudo são comparáveis aos relatados para outras infecções parasitárias investigadas por ATR-FTIR em diferentes biofluidos e tecidos, incluindo *P. falciparum*, *P. ovale*, *P. vivax*, *L. amazonensis*, *L. infantum* e *T. cruzi* (Heraud et al., 2019; Mwanga et al., 2019; Larios et al., 2021; Pacher et al., 2024; Martins et al., 2025). Esses achados reforçam que a combinação entre espectroscopia ATR-FTIR e algoritmos de aprendizado de máquina apresenta desempenho consistente na detecção de distintos parasitos.

A análise de características pelo SHAP-Random Forest identificou bandas vibracionais associadas principalmente à proteínas, ácidos nucleicos (DNA/RNA) e carboidratos, indicando alterações biomoleculares relevantes para a distinção entre indivíduos infectados e não infectados. Esses achados são coerentes com a interação parasito-hospedeiro, que envolve a liberação contínua de moléculas de excreção e secreção (ES) ao longo do ciclo biológico. Diferentes estágios de *S. mansoni* secretam proteínas específicas, como IPSE/ $\alpha$ 1 (ovo), o antígeno de 23 kDa (esquistossômulo) e SmCalp2 (verme adulto), que podem contribuir diretamente para os padrões espectrais observados (Chen et al., 2002; Schramm et al., 2003; Braschi et al., 2006).

Além disso, vesículas extracelulares (VEs) produzidas pelo próprio parasito durante diferentes estágios do ciclo são liberadas apenas em indivíduos infectados, contribuindo para

alterações nos perfis proteicos e lipídicos detectados nos espectros (Gonçalves-Silva et al., 2022). Esse processo, associado às mudanças induzidas pelo parasito nos tecidos, reforça a presença de sinais vibracionais atribuídos a proteínas e lipídios. Outro aspecto importante é o metabolismo predominantemente glicolítico de *S. mansoni*, caracterizado pela conversão acelerada de glicose em lactato, independentemente da disponibilidade de oxigênio (Bueding, 1950). As glicoproteínas abundantes dos antígenos SEA (*Soluble Egg Antigens*) também contribuem para essa assinatura. Essa intensa demanda energética e a liberação de produtos derivados do metabolismo do parasito alteram o microambiente bioquímico do hospedeiro, justificando a predominância de bandas associadas a carboidratos observadas nos espectros.

Com isso, a aplicação da espectroscopia ATR-FTIR associada ao aprendizado de máquina no diagnóstico da esquistossomose configura-se como uma abordagem promissora, considerando os resultados consistentes de sensibilidade, especificidade e acurácia, além de suas vantagens ecológicas, econômicas e operacionais para uso em campo. Ressalta-se, entretanto, que os achados apresentados correspondem a uma análise preliminar de uma metodologia ainda não explorada no contexto da esquistossomose. Dessa forma, os resultados devem ser interpretados com cautela, sendo necessária validação em diferentes cenários epidemiológicos, incluindo a avaliação do padrão vibracional em casos de coinfeção e em comparação com outras helmintíases. Estudos futuros com amostras mais numerosas e maior diversidade de biofluidos, como saliva e urina, bem como validações adicionais em condições reais de campo, serão fundamentais para avaliar a estabilidade dos modelos, sua capacidade de generalização e o real impacto operacional da técnica em programas de vigilância.

## 7. CONCLUSÃO E PERSPECTIVAS

Conclui-se que a técnica ATR-FTIR, quando associada a modelos de aprendizado de máquina, demonstra potencial como ferramenta promissora para triagem diagnóstica em áreas endêmicas, considerando os níveis satisfatórios de sensibilidade, especificidade e acurácia observados. A discriminação entre os grupos foi fortemente influenciada por componentes biomoleculares, especialmente proteínas, ácidos nucleicos e carboidratos, cujas assinaturas espectrais contribuíram de maneira determinante para a construção dos modelos classificatórios. O emprego de abordagens baseadas em inteligência artificial mostra-se relevante, uma vez que pode reduzir a dependência de mão de obra especializada, além de otimizar o processamento e a interpretação das amostras em comparação aos métodos diagnósticos convencionais. Entretanto, para a consolidação da aplicabilidade clínica da técnica, faz-se necessária a ampliação do número de amostras, visando o fortalecimento da robustez estatística e da confiabilidade dos modelos propostos. Além disso, recomenda-se a realização de validação externa multicêntrica, a fim de avaliar a reprodutibilidade dos resultados em diferentes populações e contextos epidemiológicos. A investigação em outros biofluidos, como urina e saliva, também representa um passo estratégico para expandir as possibilidades diagnósticas. Ademais, torna-se essencial a análise da robustez frente à variabilidade pré-analítica, incluindo condições de coleta, armazenamento e preparo das amostras. Por fim, a validação da metodologia em condições reais de campo constitui etapa fundamental para confirmar seu desempenho e viabilidade operacional em cenários endêmicos.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

- ABDEL-HAFEEZ, E. H. et al. Polymerase chain reaction: a better method for diagnosing chronic *Schistosoma mansoni* infections. **Tropical Medicine and Health**, [S.l.], v. 43, n. 4, p. 205–209, 2015.
- ACHARJEE, A. et al. A random forest based biomarker discovery and power analysis framework for diagnostics research. **BMC Medical Genomics**, [S.l.], v. 13, n. 1, p. 178, 2020.
- ACHARYA, S.; DA'DARA, A. A.; SKELLY, P. J. Schistosome immunomodulators. **PLoS Pathogens**, [S.l.], v. 17, n. 12, p. e1010064, 2021.
- AL-AMIN, K.; KAWSAR, M.; MAMUN, M. T. R. B.; HOSSAIN, M. S. Fourier transform infrared spectroscopic technique for analysis of inorganic materials: A review. **Nanoscale Advances**, 2025.
- ALEMU, G et al. Comparative performance of Kato–Katz, POC-CCA and real-time PCR in detecting *Schistosoma mansoni* infection at different endemicity settings in northwest Ethiopia: a cross-sectional study. **Tropical Medicine and Health**, [S.l.], v. 53, n. 1, p. 103, 2025.
- AL-KELANI, M.; BUTHELEZI, N. Advancements in medical research: Exploring Fourier Transform Infrared (FTIR) spectroscopy for tissue, cell, and hair sample analysis. **Skin Research and Technology**, v. 30, n. 6, e13733, 2024.
- ANDREW CHAN, K. L.; KAZARIAN, S. G. Attenuated total reflection Fourier transform infrared (ATRFTIR) imaging of tissues and live cells. *Chemical Society Reviews*, v. 45, n. 7, p. 1850-1864, Apr. 2016. <https://doi.org/10.1039/C5CS00515A>
- ASSARÉ, R. K. et al. Accuracy of two circulating antigen tests for the diagnosis and surveillance of *Schistosoma mansoni* infection in low-endemicity settings of Côte d'Ivoire. **The American Journal of Tropical Medicine and Hygiene**, v. 105, n. 3, p. 677, 2021.
- BAKER, M. J. et al. Using Fourier transform IR spectroscopy to analyze biological materials. **Nature Protocols**, [S.l.], v. 9, n. 8, p. 1771–1791, 2014.
- BÄRENBOLD, O. et al. Estimating sensitivity of the Kato-Katz technique for the diagnosis of *Schistosoma mansoni* and hookworm in relation to infection intensity. **PLoS Neglected Tropical Diseases**, [S.l.], v. 11, n. 10, p. e0005953, 2017.
- BELLISOLA, G.; SORIO, C. Infrared spectroscopy and microscopy in cancer research and diagnosis. **American Journal of Cancer Research**, [S.l.], v. 2, p. 1–21, 2012.
- BEL'SKAYA, L. et al. Application of FTIR spectroscopy for quantitative analysis of blood serum: a preliminary study. **Diagnostics**, [S.l.], v. 11, p. 2391, 2021. DOI: 10.3390/diagnostics11122391.
- BRASCHI, S.; BORGES, W. C.; WILSON, R. A. Proteomic analysis of the schistosome tegument and its surface membranes. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 101, supl. 1, p. 205–212, 2006.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde e Ambiente. **Neglected Tropical Diseases: Impact on child morbidity and mortality in Brazil 2010 to 2023**. Boletim Epidemiológico, Brasília, v. 2025, 2025. ISSN 2358-9450.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde e Ambiente. **Vigilância da esquistossomose mansoni: diretrizes técnicas**. Brasília: Ministério da Saúde, 2024.

BREIMAN, L. Random forests. **Machine Learning**, [S.l.], v. 45, p. 5–32, 2001.

BUEDING, Ernest. Carbohydrate metabolism of *Schistosoma mansoni*. **The Journal of general physiology**, v. 33, n. 5, p. 475-495, 1950.

BUTLER, H. J. et al. Development of high-throughput ATR-FTIR technology for rapid triage of brain cancer. **Nature Communications**, [S.l.], v. 10, n. 1, p. 4501, 2019.

BURLANDY-SOARES, Lanny et al. Schistosomiasis mansoni: follow-up of control program based on parasitologic and serologic methods in a Brazilian community of low endemicity. **Memórias do Instituto Oswaldo Cruz**, v. 98, p. 853-859, 2003.

CAIXETA, D. C. et al. Salivary ATR-FTIR spectroscopy coupled with support vector machine sorting for type 2 diabetes mellitus screening. **Diagnostics**, [S.l.], v. 13, n. 8, p. 1396, 2023.

CALDAS, I. R. et al. Human schistosomiasis mansoni: immune responses during acute and chronic phases of the infection. **Acta tropica**, v. 108, n. 2-3, p. 109-117, 2008.

CENTERS FOR DISEASE CONTROL AND PREVENTION. **Clinical Overview of Schistosomiasis**. Atlanta: CDC, 2024. Disponível em: <https://www.cdc.gov/schistosomiasis/hcp/clinical-overview/index.html>. Acesso em: 25 set. 2025.

CHEN, L.; RAO, K. V.; HE, Y. X.; RAMASWAMY, K. Skin-stage schistosomula of *Schistosoma mansoni* produce an apoptosis-inducing factor that can cause apoptosis of T cells. **Journal of Biological Chemistry**, [S.l.], v. 277, n. 37, p. 34329–34335, 2002.

CHEN, Mei; DECARY, Michel. Artificial intelligence in healthcare: An essential guide for health leaders. In: Healthcare management forum. Sage CA: Los Angeles, CA: Sage Publications, 2020. p. 10-18.

CHIENG, B. et al. Comparison of quantitative polymerase chain reaction, Kato-Katz and circulating cathodic antigen rapid test for the diagnosis of *Schistosoma mansoni* infection: A cross-sectional study in Kirinyaga County, Kenya. **Current Research in Parasitology & Vector-Borne Diseases**, v. 1, p. 100029, 2021.

CORREA-OLIVEIRA, R. et al. Cytokines as determinants of resistance and pathology in human *Schistosoma mansoni* infection. **Brazilian journal of medical and biological research**, v. 31, p. 171-177, 1998.

CLEMENS, Graeme et al. Vibrational spectroscopic methods for cytology and cellular research. **Analyst**, v. 139, n. 18, p. 4411-4444, 2014.

DAVIES, A. M. C.; CREASER, C. S. **Aplicações Analíticas da Espectroscopia II: Anais da Segunda Conferência Internacional realizada de 9 a 12 de julho de 1990 em Hatfield, Inglaterra**. Londres: Royal Society of Chemistry, 1991.

DE FREITAS BEZERRA, D. et al. Diagnostic comparison of stool exam and point-of-care circulating cathodic antigen (POC-CCA) test for schistosomiasis mansoni diagnosis in a high endemicity area in northeastern Brazil. **Parasitology**, v. 148, n. 4, p. 420-426, 2021.

ELLIS, David I.; GOODACRE, Royston. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. **Analyst**, v. 131, n. 8, p. 875-885, 2006.

FADLELMOULA, A. et al. Fourier transform infrared (FTIR) spectroscopy to analyse human blood over the last 20 years: a review towards lab-on-a-chip devices. **Micromachines**, v. 13, n. 2, p. 187, 2022.

FAHRENFORT, J. Attenuated total reflection: A new principle for the production of useful infra-red reflection spectra of organic compounds. **Spectrochimica Acta**, v. 17, n. 7, p. 698-709, 1961.

FASOGBON, I. V. et al. A critical review of the limitations of current diagnostic techniques for schistosomiasis. **All Life**, [S.l.], v. 17, n. 1, p. 2379305, 2024. DOI: 10.1080/26895293.2023.2379305.

FERREIRA, I. C. et al. Attenuated Total Reflection-Fourier Transform Infrared (ATR-FTIR) spectroscopy analysis of saliva for breast cancer diagnosis. **Journal of Oncology**, v. 2020, n. 1, p. 4343590, 2020.

FENWICK, A.; WEBSTER, J. P. Schistosomiasis: challenges for control, treatment and drug resistance. **Current Opinion in Infectious Diseases**, [S.l.], v. 19, n. 6, p. 577–582, 2006.

FUSS, A. et al. Detection of *Schistosoma mansoni* DNA using polymerase chain reaction from serum and dried blood spot card samples of an adult population in north-western Tanzania. **Infectious Diseases of Poverty**, [S.l.], v. 10, n. 1, p. 15, 2021.

GOMES, A. C. L. et al. Prevalência e carga parasitária da esquistossomose mansônica antes e depois do tratamento coletivo em Jaboatão dos Guararapes, Pernambuco. **Epidemiologia e Serviços de Saúde**, Brasília, v. 25, p. 243–250, 2016.

GOMES, E. C. S. et al. Urban schistosomiasis: an ecological study describing a new challenge to the control of this neglected tropical disease. **The Lancet Regional Health – Americas**, [S.l.], v. 8, p. 100144, 2022.

GONÇALVES-SILVA, G. et al. Profiling the serum proteome during *Schistosoma mansoni* infection in BALB/c mice: a focus on the altered lipid metabolism as a key modulator of host–parasite interactions. **Frontiers in Immunology**, [S.l.], v. 13, p. 955049, 2022.

GOULDEN, J. D.; SHARPE, M. E. The infra-red absorption spectra of lactobacilli. **Journal of General Microbiology**, v. 19, p. 76–86, 1958.

GRAEFF-TEIXEIRA, C.; FAVERO, V.; PASCOAL, V. F. et al. Low specificity of point-of-care circulating cathodic antigen (POC CCA) diagnostic test in a non-endemic area for schistosomiasis mansoni in Brazil. **Acta Tropica**, v. 217, p. 105863, 2021.

GRDADOLNIK, J. ATR-FTIR spectroscopy: Its advantages and limitations. **Acta Chimica Slovenica**, v. 49, p. 631–642, 2002.

GUEVARA-VEGA, M. et al. Salivary detection of Chikungunya virus infection using a portable and sustainable biophotonic platform coupled with artificial intelligence algorithms. **Scientific Reports**, [S.l.], v. 14, n. 1, p. 21546, 2024.

- HERAUD, P. et al. Infrared spectroscopy coupled to cloud-based data management as a tool to diagnose malaria: A pilot study in a malaria-endemic country. **Malaria Journal**, v. 18, n. 1, p. 348, 2019.
- HINZ, R. et al. Serological approaches for the diagnosis of schistosomiasis – A review. **Molecular and Cellular Probes**, v. 31, p. 2-21, 2017.
- HONÓRIO-SILVA, G. et al. Development of a novel sustainable, portable, fast, and non-invasive platform based on ATR-FTIR technology coupled with machine learning algorithms for *Helicobacter pylori* detection in human saliva. **Talanta Open**, v. 10, p. 100383, 2024.
- KATO, K. On the comparison of some stool examination methods. **Japanese Journal of Parasitology**, [S.l.], v. 3, p. 35, 1954.
- KATZ, N. et al. **Inquérito Nacional de Prevalência da Esquistossomose mansoni e Geo-helmintoses**. Belo Horizonte: CPqRR, v. 22, 2018.
- KATZ, N.; CHAVES, A.; PELLEGRINO, J. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. **Revista do Instituto de Medicina Tropical de São Paulo**, São Paulo, v. 14, p. 397–400, 1972.
- KOCHAN, K. et al. Infrared spectroscopy of blood. **Applied Spectroscopy**, [S.l.], v. 75, p. 611–646, 2020. DOI: 10.1177/0003702820985856.
- KRAUTH, S. J. et al. An in-depth analysis of a piece of stool: distribution of *Schistosoma mansoni* and hookworm eggs in human stool. **PLoS Neglected Tropical Diseases**, [S.l.], v. 6, n. 12, p. e1969, 2012. DOI: 10.1371/journal.pntd.0001969.
- LAMBERTON, P. H. L. et al. Sensitivity and specificity of multiple Kato-Katz thick smears and a circulating cathodic antigen test for *Schistosoma mansoni* diagnosis pre- and post-repeated praziquantel treatment. **PLoS Neglected Tropical Diseases**, [S.l.], v. 8, n. 9, p. e3139, 2014.
- LARIOS, G. et al. A new strategy for canine visceral leishmaniasis diagnosis based on FTIR spectroscopy and machine learning. **Journal of Biophotonics**, v. 14, e2021100141, 2021. DOI: 10.1002/jbio.202100141.
- LARKIN, P. Instrumentation and sampling methods. **Infrared and Raman Spectroscopy**, p. 27-54, 2011.
- LI, Q. et al. Global trends of schistosomiasis burden from 1990 to 2021 across 204 countries and territories: findings from GBD 2021 study. **Acta tropica**, v. 261, p. 107504, 2025.
- LINDHOLZ, C. G. et al. Study of the diagnostic accuracy of the Helmintex, Kato-Katz and POC-CCA methods for the diagnosis of intestinal schistosomiasis in Candéal, an area of low transmission in northeastern Brazil. **PLoS Neglected Tropical Diseases**, v. 12, n. 3, p. e0006274, 2018.
- MALAQUIAS, L. C. C. et al. Cytokine regulation of human immune response to *Schistosoma mansoni*: analysis of the role of IL-4, IL-5 and IL-10 on peripheral blood mononuclear cell responses. **Scandinavian journal of immunology**, v. 46, n. 4, p. 393-398, 1997.
- MANNERKORPI, M. et al. Diagnostic performance of attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy for detecting osteoarthritis and rheumatoid

arthritis from blood serum. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, [S.l.], v. 338, p. 126206, 2025.

MARTINEZ-CUAZITL, A. et al. ATR-FTIR spectrum analysis of saliva samples from COVID-19 positive patients. **Scientific Reports**, v. 11, p. 19980, 2021. DOI: 10.1038/s41598-021-99529-w.

MARTINS, M. et al. A New Approach for Chagas Disease Screening Using Serum Infrared Spectroscopy and Machine Learning Algorithms. **ACS Infectious Diseases**, 2025.

McMANUS, D.P et al. Schistosomiasis. **Nature Reviews Disease Primers**, v. 4, p. 13, 2018.

McMANUS, D. P. et al. Schistosomiasis—from immunopathology to vaccines. **Seminars in Immunopathology**, Berlin, v. 42, n. 3, p. 355–371, 2020.

MENEZES, D. L. et al. Accuracy study of Kato-Katz and helmintex methods for diagnosis of Schistosomiasis mansoni in a moderate endemicity area in Sergipe, Northeastern Brazil. **Diagnostics**, v. 13, n. 3, p. 527, 2023.

MEWAMBA, E. M. et al. Field assessment in Cameroon of a reader of POC-CCA lateral flow strips for the quantification of *Schistosoma mansoni* circulating cathodic antigen in urine. **PLoS Neglected Tropical Diseases**, v. 15, n. 7, p. e0009569, 2021.

MORAIS, C. L. M.; LIMA, K. M. G.; SINGH, M.; MARTIN, F. L. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. **Nature Protocols**, [S.l.], v. 15, n. 7, p. 2143–2162, 2020.

MOVASAGHI, Z.; REHMAN, S.; REHMAN, I. U. Fourier transform infrared (FTIR) spectroscopy of biological tissues. **Applied Spectroscopy Reviews**, [S.l.], v. 43, n. 2, p. 134–179, 2008.

NASEER, K.; ALI, S.; QAZI, J. ATR-FTIR spectroscopy as the future of diagnosis: a systematic review of the approach using biofluidics. **Applied Spectroscopy Reviews**, v. 56, n. 2, p. 85-97, 2021.

NAUMANN, D.; HELM, D.; LABISCHINSKI, H. Microbiological characterizations by FT-IR spectroscopy. **Nature**, v. 351, p. 81–82, 1991.

NAUMANN, D. FT-infrared and FT-Raman spectroscopy in biomedical research. **Applied spectroscopy reviews**, v. 36, n. 2-3, p. 239-298, 2001.

NELWAN, M. L. Schistosomiasis: life cycle, diagnosis, and control. **Current Therapeutic Research**, v. 91, p. 5-9, 2019.

NEVES, M. M. et al. Perspectives of FTIR as promising tool for pathogen diagnosis, sanitary and welfare monitoring in animal experimentation models: a review based on pertinent literature. **Microorganisms**, v. 12, n. 4, p. 833, 2024.

OLIVEIRA, S. W. et al. Salivary detection of Zika virus infection using ATR-FTIR spectroscopy coupled with Machine learning algorithms and univariate analysis: A proof-of-concept animal study. **Diagnostics**, v. 13, n. 8, p. 1443, 2023.

OLIVEIRA, W. J. et al. Evaluation of diagnostic methods for the detection of intestinal schistosomiasis in endemic areas with low parasite loads: saline gradient, Helmintex, Kato-Katz and rapid urine test. **PLoS Neglected Tropical Diseases**, [S.l.], v. 12, n. 2, p. e0006232, 2018.

- OKOYO, C. et al. Comparing the performance of circulating cathodic antigen and Kato-Katz techniques in evaluating *Schistosoma mansoni* infection in areas with low prevalence in selected counties of Kenya: a cross-sectional study. **BMC Public Health**, [S.l.], v. 18, n. 1, p. 478, 2018.
- RANDALL, H. M.; SMITH, D. W.; COLM, A. C.; NUNGESTER, W. J. Correlation of biologic properties of strains of *Mycobacterium* with infra-red spectrums. I. Reproducibility of extracts of *M. tuberculosis* as determined by infra-red spectroscopy. **American Review of Tuberculosis**, v. 63, p. 372–380, 1951.
- RUAS, A. C. L. et al. Immunoinformatic predictions and characterization of *Schistosoma mansoni* peptides as candidates for immunodiagnostic. **Diagnostic Microbiology and Infectious Disease**, v. 111, n. 3, p. 116632, 2025.
- SALA, A. et al. Biofluid diagnostics by FTIR spectroscopy: A platform technology for cancer detection. **Cancer letters**, v. 477, p. 122-130, 2020.
- SCHRAMM, G. et al. Molecular characterization of an interleukin-4-inducing factor from *Schistosoma mansoni* eggs. **Journal of Biological Chemistry**, [S.l.], v. 278, n. 20, p. 18384–18392, 2003.
- SINGH, Amanpreet; THAKUR, Narina; SHARMA, Aakanksha. A review of supervised machine learning algorithms. In: **2016 3rd international conference on computing for sustainable global development (INDIACom)**. Ieee, 2016. p. 1310-1315.
- SIQUEIRA, L. M. et al. Performance of POC-CCA® in diagnosis of schistosomiasis mansoni in individuals with low parasite burden. **Revista da Sociedade Brasileira de Medicina Tropical**, Uberaba, v. 49, n. 3, p. 341–347, 2016.
- SIQUEIRA, L. M. et al. Evaluation of parasitological and molecular techniques for the diagnosis and assessment of cure of schistosomiasis mansoni in a low transmission area. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 110, n. 2, p. 209–214, 2015.
- TEIXEIRA, C. F. et al. Detection of *Schistosoma mansoni* eggs in feces through their interaction with paramagnetic beads in a magnetic field. **PLoS Neglected Tropical Diseases**, [S.l.], v. 1, n. 2, p. e73, 2007. DOI: 10.1371/journal.pntd.0000073.
- THEODORE ARMAND, Tagne Poupi et al. Applications of artificial intelligence, machine learning, and deep learning in nutrition: a systematic review. **Nutrients**, v. 16, n. 7, p. 1073, 2024.
- UNTEREINER, V. et al. Bile analysis using high-throughput FTIR spectroscopy for the diagnosis of malignant biliary strictures: a pilot study in 57 patients. **Journal of Biophotonics**, v. 7, n. 3-4, p. 241-253, 2014.
- UTZINGER, J. et al. Relative contribution of daily and intra-specimen variation in *Schistosoma mansoni* fecal egg count before and after treatment with praziquantel. **Parasitology**, [S.l.], v. 122, pt. 5, p. 537–544, 2001.
- UTZINGER, J. et al. Schistosomiasis and neglected tropical diseases: towards integrated and sustainable control and a word of caution. **Parasitology**, 136(13), 1859–1874, 2009
- VIANA, A. G. et al. Discrepancy between batches and impact on the sensitivity of point-of-care circulating cathodic antigen tests for *Schistosoma mansoni* infection. **Acta Tropica**, v. 197, p. 105049, 2019.

WORLD HEALTH ORGANIZATION. Diagnostic target product profiles for monitoring, evaluation and surveillance of schistosomiasis control programmes. Geneva: World Health Organization, 2021.

WORLD HEALTH ORGANIZATION. Guideline on control and elimination of human schistosomiasis. Geneva: World Health Organization, 2022. Licence: CC BY-NC-SA 3.0 IGO.

WORLD HEALTH ORGANIZATION. Schistosomiasis. Geneva: World Health Organization, 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/schistosomiasis>. Acesso em: 6 nov. 2025.

WORLD HEALTH ORGANIZATION. Schistosomiasis: WHO guideline development group. Geneva: World Health Organization, 2023. Disponível em: <https://www.who.int/teams/control-of-neglected-tropical-diseases/schistosomiasis/who-guideline-development-group>. Acesso em: 6 nov. 2025.

WORLD HEALTH ORGANIZATION. The control of schistosomiasis: report of a WHO expert committee. Geneva: World Health Organization, 1985. Disponível em: [http://apps.who.int/iris/bitstream/10665/39529/1/WHO\\_TRS\\_728.pdf](http://apps.who.int/iris/bitstream/10665/39529/1/WHO_TRS_728.pdf). Acesso em: 6 nov. 2025.

WORRELL, C. M. et al. Cost analysis of tests for the detection of *Schistosoma mansoni* infection in children in western Kenya. **The American Journal of Tropical Medicine and Hygiene**, [S.l.], v. 92, n. 6, p. 1233–1239, 2015.

ZHANG, A. H. et al. Serum proteomics in biomedical research: a systematic review. **Applied Biochemistry and Biotechnology**, [S.l.], v. 170, n. 4, p. 774–786, 2013.

ZHANG, A.; SUN, H.; WANG, X. Serum metabolomics as a novel diagnostic approach for disease: a systematic review. **Analytical and Bioanalytical Chemistry**, [S.l.], v. 404, n. 4, p. 1239–1245, 2012.