



**SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
PRÓ-REITORIA DE PESQUISA, PÓS GRADUAÇÃO E INOVAÇÃO
TECNOLÓGICA
PROGRAMA DE PÓS GRADUAÇÃO EM BIOCÊNCIAS**

ANTÔNIO ANASTASIS DE OLIVEIRA MARTINS

**UTILIZANDO INTELIGÊNCIA ARTIFICIAL NA RELAÇÃO SOCIOECONOMICA
E SOROLOGIA Anti-PGL-1 DE CASOS DE HANSENÍASE EM UM MUNICÍPIO
HIPERENDÊMICO NO ESTADO DO PARÁ**

**Santarém-Pará
2023**

ANTÔNIO ANASTASIS DE OLIVEIRA MARTINS

**UTILIZANDO INTELIGÊNCIA ARTIFICIAL NA RELAÇÃO SOCIOECONOMICA
E SOROLOGIA Anti-PGL-1 DE CASOS DE HANSENÍASE EM UM MUNICÍPIO
HIPERENDÊMICO NO ESTADO DO PARÁ**

Projeto de Dissertação apresentada ao Programa de Pós-Graduação em Biociências da Universidade Federal do Oeste do Pará como requisito para qualificação.

Orientador: Dr. Guilherme Augusto Barros Conde
Coorientadora: Dra. Valney Mara Gomes Conde

**Santarém-Pará
2023**

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema Integrado de Bibliotecas – SIBI/UFOPA

-
- M386u Martins, Antônio Anastasis de Oliveira
 Utilizando inteligência artificial na relação socioeconomica e sorologia Anti-PGL-1 de casos de hanseníase em um município hiperendêmico no Estado do Pará. / Antônio Anastais de Oliveira Martins. – Santarém, 2023.
 75 p. : il.
 Inclui bibliografias.
- Orientador: Guilherme Augusto Barros Conde
 Coorientadora: Valney Mara Gomes Conde
 Dissertação (Mestrado) – Universidade Federal do Oeste do Pará, Pró-reitoria de Pesquisa, Pós-graduação e Inovação Tecnológica, Instituto de Biodiversidade e Florestas, Programa de Pós-Graduação em Biociências.
1. Hanseníase. 2. Anti-PGL-1. 3. Clusterização. 4. Algoritmos de aprendizagem não supervisionada. 5. Inteligência artificial. I. Conde, Guilherme Augusto Barros, *orient.* II. Conde, Valney Mara Gomes, *coorient.* III. Título.

CDD: 23 ed. 614.546098115

Bibliotecária - Documentalista: Mary Caroline Santos Ribeiro – CRB/2 566

Aos meus familiares pelo apoio e incentivo. E, em especial a minha mãe Professora *Maria Ocy de Oliveira Martins* e “*in memoriam*” a meu pai *Aluizio Lopes Martins*.

AGRADECIMENTOS

À Deus por me proporcionado paciência e discernimento, e ter me permitido concluir mais esta fase em minha vida.

À minha família, por todo apoio, paciência e ajuda tão necessários no decorrer desta jornada acadêmica e em especial a minha mãe Professora Maria Ocy de Oliveira Martins, meu Pai (*in memoriam*) Professor Aluizio Lopes Martins e meus irmãos Aluizio Martins Junior, Alcymar Martins, Alciney Martins, Paulo Martins e Manuel Martins Neto.

As minhas cunhadas Jaqueline Lopes, Andreza Bentes e Floriania Baia.

Aos meus sobrinhos Gabriel, Ítalo (*in memoriam*), Heitor, Yasmin, Layse, Larissa, Igor e Ingrid Martins.

Ao Professor doutor Guilherme Augusto Barros Conde, orientador, por sua dedicação ao projeto e constante disponibilidade para discutir sobre os resultados contidos neste trabalho.

A Professora doutora Valney Mara Gomes Conde, coorientadora, pelo apoio e incentivo com a pesquisa.

Às instituições que diretamente contribuíram para a realização deste estudo: Laboratório de Suporte a Decisão (LSD) da Universidade Federal do Oeste do Pará (UFOPA).

Aos colegas do LSD, conhecidos como “*elissidianos*”, Marcio Bastos e Marcos Baia pelas colaborações durante o desenvolvimento da pesquisa.

Aos membros da coordenação do programa de Biociências, Professor doutor Guilherme Augusto Barros Conde e Professor doutor Paulo Sérgio Taube, que tornou possível a efetividade e o sucesso na realização deste curso.

Aos meus amigos e parceiros de outrora Marlison Repolho, Jairo Franco e Edvan Cruz pelo carinho, atenção, conselhos e a compreensão por vezes ausente em virtude da dedicação aos estudos.

A todos que de alguma forma contribuíram para o crescimento de cada momento para realização deste trabalho.

“Talvez eu seja enganado inúmeras vezes..., mas não deixarei de acreditar que em algum lugar, alguém merece minha confiança”.

Aristóteles

RESUMO

A hanseníase é uma doença infecciosa granulomatosa crônica causada pelo organismo intracelular obrigatório o *Mycobacterium leprae* que afeta principalmente a pele e os nervos periféricos, podendo levar a deficiências físicas graves e deformidades se não for diagnosticada e tratada em seus estágios iniciais. Os exames sorológicos nas últimas décadas vêm se padronizando e aperfeiçoando a capacidade de detecção do anticorpo específico do *M. leprae*, como é o caso do anti-PGL-1 que demonstrou ter mais sensibilidade que o exame baciloscópico, sendo um indicativo para a presença de infecção subclínica que mostra a concentração dos anticorpos séricos e correlaciona com a carga bacilar. A aplicação da Inteligência Artificial (IA) vem sendo destacada em diversas áreas, inclusive no campo da saúde, onde o aprendizado de máquina aplicado na descoberta de padrões vem sendo utilizados como ferramentas inovadoras no processo de tomada de decisão. O objetivo deste estudo foi analisar a epidemiologia da hanseníase através do processo de mineração de dados utilizando técnicas de clusterização nas variáveis clínicas, socioeconômicas e dosagem de anti-PGL-1 de pacientes e seus comunicantes no município de Santarém no oeste do Pará. Foi realizado um estudo epidemiológico, descritivo, analítico e retrospectivo a partir dos dados oficiais do Sistema de Informação de Agravos de Notificação (SINAN) e do Instituto Brasileiro de Geografia e Estatística (IBGE), foi realizada em Santarém no ano de 2014 uma semana de busca ativa por novos pacientes de hanseníase. Nesta semana, uma equipe multiprofissional realizou visitas domiciliares (busca ativa) às famílias de 70 pacientes registrados no SINAN e aqui denominados casos índices (57,1% masculino para 42,9% feminino) para realizar exame dermatológico e teste laboratorial entre todas as pessoas avaliadas. Nos domicílios visitados foram encontradas 280 pessoas (62,5% feminino para 37,5% masculino) convivendo com os casos índices, os chamados comunicantes, totalizando 350 pessoas atendidas. Entre os comunicantes avaliados, 35 (65,7% feminino para 34,3% masculino) foram diagnosticados com casos novos de hanseníase. Entre os 70 casos índices avaliados, 65,7% tinham sido registrados no SINAN como Multibacilar (MB) e entre os 35 casos novos, 88,6% foram diagnosticados como MB. A soropositividade do anti-PGL-1 foi de 39,6% (111/280) nos comunicantes e 37,1% (13/35) nos casos novos. Utilizando a técnica de clusterização nas variáveis socioeconômicas e dados clínicos dos pacientes e seus comunicantes, foi possível encontrar perfis de risco de infecção simbolizados por três clusters: C1 “baixo”, C2 “médio” e C3 “alto”. Ao comparar os clusters encontrados com o nível do resultado do teste anti-PGL-1, obteve-se uma relação diretamente proporcional entre os valores médios do anti-PGL-1 e os perfis de risco dos clusters. Na semana da busca

ativa foram diagnosticados 35 casos novos de hanseníase, sendo a maioria diagnosticados na forma clínica MB, caracterizando assim o diagnóstico tardio e por consequência a cadeia de transmissão ativa da doença. Com a aplicação de técnicas de mineração de dados baseadas em clusterização foi possível potencializar o resultado do teste anti-PGL-1 associando-o a variáveis socioeconômicas e clínicas em diferentes níveis de risco de desenvolver a doença. Esse achado representa um avanço no suporte e na tomada de decisão das equipes de vigilância de contatos, fornecendo ferramentas importantes para melhorar o desempenho do combate da hanseníase.

Palavras-Chave: Hanseníase; Anti-PGL-1; Clusterização; Algoritmos de aprendizagem não supervisionada; Inteligência Artificial.

ABSTRACT

Leprosy is a chronic granulomatous infectious disease caused by the obligatory intracellular organism *Mycobacterium leprae*, which mainly affects the skin and peripheral nerves, and can lead to severe physical deficiencies and deformities if not diagnosed and treated in its early stages. Serological tests in recent decades have been standardizing and improving the detection capacity of the specific antibody of *M. leprae*, as is the case of anti-PGL-1 that has been shown to have more sensitivity than bacilloscopic examination, being an indication for the presence of subclinical infection that shows the concentration of serum antibodies and correlates with bacillary load. The application of Artificial Intelligence (AI) has been highlighted in several areas, including in the field of health, where machine learning applied in the discovery of patterns has been used as innovative tools in the decision-making process. The aim of this study was to analyze the epidemiology of leprosy through the data mining process using clustering techniques in the clinical, socioeconomic and anti-PGL-1 measurement of patients and their communicators in the municipality of Santarém in western Pará. An epidemiological, descriptive, analytical and retrospective study was conducted based on official data from the Notifiable Diseases Information System (SINAN) and the Brazilian Institute of Geography and Statistics (IBGE), was conducted in Santarém in 2014 a week of active search for new leprosy patients. This week, a multidisciplinary team conducted home visits (active search) to the families of 70 patients registered in SINAN and here called index cases (57.1% male to 42.9% female) to perform dermatological examination and laboratory testing among all people evaluated. In the households visited, 280 people were found (62.5% female to 37.5% male) living with the index cases, the so-called communicators, totaling 350 people attended. Among the reported members evaluated, 35 (65.7% female to 34.3% male) were diagnosed with new cases of leprosy. Among the 70 index cases evaluated, 65.7% had been registered in SINAN as Multibacillary (MB) and among the 35 new cases, 88.6% were diagnosed as MB. Anti-PGL-1 seropositivity was 39.6% (111/280) in communicators and 37.1% (13/35) in new cases. Using the clustering technique in the socioeconomic variables and clinical data of patients and their communicators, it was possible to find risk profiles of infection symbolized by three clusters: C1 "low", C2 "medium" and C3 "high". When comparing the clusters found with the level of the anti-PGL-1 test result, a directly proportional relationship was obtained between the mean values of anti-PGL-1 and the risk profiles of the clusters. In the week of active search, 35 new cases of leprosy were diagnosed, most of them diagnosed in clinical form MB, thus characterizing the late diagnosis and consequently the chain of active transmission of the

disease. With the application of clustering-based data mining techniques, it was possible to potentiate the result of the anti-PGL-1 test by associating it with socioeconomic and clinical variables at different levels of risk of developing the disease. This finding represents an advance in the support and decision-making of contact surveillance teams, providing important tools to improve leprosy combat performance.

Keywords: Leprosy; Anti-PGL-1; Clustering; Unsupervised learning algorithms; Artificial Intelligence.

LISTA DE ILUSTRAÇÕES

Figura 1 - Coeficiente de detecção anual de casos novos por 10.000 hab.	16
Figura 2 - <i>Mycobacterium leprae</i> visto em globias e isolados.	21
Figura 3 - Lesões cutâneas de hanseníase lepromatosa.	22
Figura 4 – Micobactérias em uma coloração para organismos acidorresistentes.	22
Figura 5 - Espectro clínico da hanseníase por Ridley-Jopling.	25
Figura 6 - Hanseníase indeterminada. Uma mancha eritematosa mal definida sobre o rosto.	26
Figura 7 - Hanseníase TT, com placa anestésica bem delimitada, eritematosa, sem pelos, seca.	26
Figura 8 - Hanseníase BT maculo anestésica, hipopigmentadas, hipoestética, placa seca com perda de apêndices.	27
Figura 9 - Hanseníase Bordeline Bordeline (BB).	28
Figura 10 - Hanseníase Bordeline Lepromatoso (BL).	28
Figura 11 - Hanseníase Lepromatoso Lepromatoso (LL) A. infiltração difusa da face e B. infiltração e nodulação da orelha.	29
Figura 12 - Distribuição geográfica de novos casos de hanseníase, 2019.	31
Figura 13 - Cenário nacional - ano 2019.	32
Figura 14 - Taxa de detecção geral de casos novos de hanseníase (por 100 mil hab.). Pará, região Norte e Brasil, 2009 a 2018.	34
Figura 15 - Taxa de detecção de casos novos de hanseníase em menores de 15 anos (por 100 mil hab.). Pará, região Norte e Brasil, 2009 a 2018.	34
Figura 16 - Taxa de casos novos de hanseníase com grau 2 de incapacidade física no momento do diagnóstico (por 1 milhão de hab.). Pará, região Norte e Brasil, 2009 a 2018. ..	34
Figura 17 - Modelo da parede/envelope celular de <i>Mycobacterium leprae</i>	36
Figura 18 - Uma Visão Geral dos Passos que compõem o Processo KDD.	40
Figura 19 - Ilustração de dados agrupados em quatro grupos.	41
Figura 20 - Exemplo de aplicação do algoritmo K-means utilizando o software Orange Canvas.	44
Figura 21 - Coeficiente geral de detecção de hanseníase por municípios. Pará, 2010.	45
Figura 22 - Municípios de Santarém, área urbana, no Estado do Pará – Brasil.	46
Figura 23 - Casos Novos de hanseníase registrados entre 2003 a 2014, dados SINAN; * Busca-ativa em 2014 realizada no município de Santarém – Pará.	51

Figura 24- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) dos casos índices, comunicantes e casos novos.	60
Figura 25- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) $\geq 1,1$ dos casos índices e comunicantes.....	61
Figura 26- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) $\geq 1,1$ dos comunicantes.	62
Figura 27 - Anti-PGL-1 por <i>Clusters</i> C1, C2 e C3.....	65
Figura 28 - Densidade óptica (O.D.) dos casos novos por <i>cluster</i>	67

LISTA DE TABELAS

Tabela 1 - Classificação operacional da hanseníase.	24
Tabela 2 - Classificação Hanseníase: Ridley Jopling e OMS.	25
Tabela 3 - Detecção de caso novo, por região da OMS, 2019.	31
Tabela 4 . Características da classificação de Ridley Jopling em todo o espectro.	37
Tabela 5 . Caracterização epidemiológica dos grupos casos índices e comunicantes de acordo com Gênero e Faixa etária.....	50
Tabela 6 . Características epidemiológicas dos pacientes índices e os casos novos detectados na semana de busca ativa no município de Santarém-Pa, 2014, de acordo com o gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica.	52
Tabela 7 . Características epidemiológicas dos comunicantes (280) dos pacientes índices de hanseníase diagnosticados no município de Santarém-Pa, 2014 de acordo com o gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices).....	56
Tabela 8 . Dosagens dos anticorpos anti-PGL-I, <i>cut-off</i> de 0,295) dos comunicantes (n = 280).	60
Tabela 9 . Características epidemiológicas dos comunicantes (280) no município de Santarém-Pa de acordo com os 3 <i>clusters</i> C1, C2, e C3 encontrados e as variáveis: gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices). .	63
Tabela 10 . Características epidemiológicas dos casos novos (35) de hanseníase diagnosticados no município de Santarém-Pa de acordo com os 3 <i>clusters</i> C1, C2, 3 C3 encontrados e as variáveis: gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices).....	65

LISTA DE SIGLAS

MH – Mal de Hansen

PB – Paucibacilar

MB – Multibacilar

BAAR - Bacilo álcool-ácido resistente

TT – Tuberculóide-Tuberculóide

BT – Borderline-Tuberculóide

BB – Borderline-Borderline

BL – Borderline-Lepromatoso

LL – Lepromatoso-Lepromatoso

I– Indeterminada

T– Tuberculóide

V– Virchowiana

OMS/WHO – Organização Mundial de Saúde

SINAN – Sistema de Informação de Agravos de Notificação

IBGE – Instituto Brasileiro de Geografia e Estatística

PGL-I – Glicolípido Fenólico I

QGIS – Quantum Geographic Information System

LSD – Laboratório de Suporte à Decisão

SVS – Sistema de Vigilância em Saúde

GIF – Grau de Incapacidade Física

DTNs – Doenças Tropicais Negligenciadas

ELISA – Enzyme Linked Immunosorbent Assay

DO – Densidade Óptica

LSD – Laboratório de Suporte à Decisão

LDI – Laboratório Dermato-Imunológico

G2I – Grau 2 de incapacidade física

SUMÁRIO

1	INTRODUÇÃO.....	15
2	OBJETIVOS	18
2.1	Objetivo Geral.....	18
2.2	Objetivos Específicos	18
3	REFERENCIAL TEÓRICO	19
3.1	Histórico da Hanseníase	19
3.2	Etiologia	20
3.3	Hanseníase (Infecção, Transmissão, Classificação e Diagnóstico)	22
3.4	Aspectos Epidemiológicos da Hanseníase	30
3.4.1	Cenário mundial	30
3.4.2	Cenário nacional	32
3.4.3	Cenário regional	33
3.5	Glicolipídio Fenólico 1 (PGL-1).....	35
3.6	Mineração de Dados	38
3.6.1	Agrupamento de dados.....	41
3.6.2	Algoritmos de agrupamentos.....	42
4	METODOLOGIA.....	45
4.1	Tipo de estudo	45
4.2	Cenário de Estudo.....	45
4.3	Aquisição de Dados	46
4.4	Escolha das Variáveis	47
4.5	Aspectos Éticos.....	49
4.6	Técnicas de Agrupamento	49
4.7	Análise de Dados	49
5	RESULTADOS E DISCUSSÃO.....	50
5.1	Panorama da Hanseníase no município de Santarém - Pará	50
5.2	Perfil epidemiológico dos comunicantes.....	56
5.3	Soropositividade do anti-PGL-1 dos casos hansênicos e comunicantes	59
5.4	Aplicação de técnicas de clusterização utilizando K-means	62
6	CONCLUSÃO.....	68
	REFERÊNCIAS.....	69

1 INTRODUÇÃO

A hanseníase é uma doença infecciosa granulomatosa crônica causada pelo organismo intracelular obrigatório *Mycobacterium leprae* que afeta principalmente a pele e os nervos periféricos que pode levar a deficiências físicas graves e deformidades se não for diagnosticada e tratada em seus estágios iniciais (WHO, 2022).

Nos últimos 20 anos, 16 milhões de casos de hanseníase foram registrados de acordo com *World Health Organization* – WHO (Organização Mundial da Saúde - OMS); no período de 2013 a 2015, segundo a WHO, mais de 200 mil novos casos foram notificados a cada ano, indicando que a doença ainda está progredindo sem obstáculos (WHO, 2016).

Em 2019, foram reportados à WHO 127.396 casos novos da doença no mundo. Desses, 19.195 (15,1%) ocorreram na região das Américas e 17979 foram notificados no Brasil, correspondendo a 93% do número de casos novos das Américas. Deste modo, o Brasil, a Índia e Indonésia reportaram mais de 100.000 casos novos, correspondendo a 78,5% dos casos novos detectados no ano de 2019, deixando o Brasil, nesse contexto, ocupando o segundo lugar entre os países com maior número de casos no mundo, atrás apenas da Índia (WHO, 2019).

No Brasil, a incidência de casos notificados de hanseníase nos estados é heterogênea, sendo as maiores taxas encontradas nas regiões Norte, Nordeste e Centro-Oeste. A porcentagem de casos em indivíduos <15 anos (um importante indicador de evolução da doença) ainda é alta em alguns estados, destacando a deficiência em vigilância e controle de doenças (CASTRO et al., 2016).

Na rede dos serviços de saúde pública do país que executam as ações do Programa Nacional de Controle da Hanseníase, o controle dos comunicantes tem sido desvalorizado e até mesmo negligenciado, pois enfatiza o serviço para o controle da patologia e do doente, deixando como plano alternativo as ações de controle dos comunicantes (NETO et al., 2014).

A investigação epidemiológica deve incluir o exame das pessoas que convivem ou conviveram no domicílio ou fora dele com o doente de hanseníase, qualquer que seja a sua forma clínica, com o objetivo de descobrir a fonte de infecção e de conhecer outros casos oriundos da mesma fonte (NETO et al., 2014).

Para cada paciente de hanseníase há possibilidade de terem familiares ou algum indivíduo para dar suporte em seu dia a dia, logo, passam a ser conhecidas como “comunicantes” (contatos, contactantes ou conviventes), que podem ser qualquer pessoa que resida ou tenha residido com o doente nos últimos cinco anos. Esta categoria é subdividida em dois tipos: intradomiciliares (pessoas que moram na mesma residência do paciente) e

extradomiciliares (que moram em residências diferentes do paciente, entretanto convivem com o mesmo). Os comunicantes intradomiciliares envolvem um grupo reconhecível com alto risco de infecta-se com a patologia, por viverem no arredor da fonte de infecção e muitas vezes se tornam casos subdiagnosticados (ROCHA, 2016).

Considera-se contato domiciliar toda e qualquer pessoa que resida ou tenha residido com o doente de hanseníase. Contato social é qualquer pessoa que conviva ou tenha convivido em relações familiares ou não, de forma próxima e prolongada. Os contatos sociais, que incluem vizinhos, colegas de trabalhos e de escola, entre outros, devem ser investigados de acordo com o grau e tipo de convivência, ou seja, aqueles que tiveram contato muito próximo e prolongado com o paciente não tratado. Atenção especial deve ser dada aos contatos familiares do paciente (pais, irmãos, avós, tios etc.) (BRASIL, 2016).

Além disso, a falta de informações sobre a hanseníase e suas formas de transmissão coloca o próprio indivíduo como um ser passivo diante do controle da doença, pois muitos contatos domiciliares, a despeito da solicitação de comparecimento ao serviço de saúde para exame dermatoneurológico, acabam não comparecendo. Por outro lado, as equipes de saúde parecem não conseguir cumprir um programa de busca ativa desses contatos (DESSUNTI et al., 2008).

Segundo o Ministério da Saúde, os indicadores de saúde permitem fazer analogia entre diferentes áreas ou momentos e fornecem subsídios ao planejamento das ações de saúde. Os indicadores epidemiológicos mensuram a magnitude/transcendência do problema de saúde pública, indicando a situação verificada na população ou no meio ambiente num dado momento ou num determinado período (BRASIL, 2002, p. 64). A partir desses indicadores epidemiológicos, o coeficiente de detecção anual de casos novos por 10.000 habitantes (Figura 1) traz subsídios e justifica a produção desse trabalho devido a sua condição infectocontagiosa, impacto socioeconômico e repercussão psicológica (BRASIL, 2002; OLIVEIRA et al., 2015).

Figura 1 - Coeficiente de detecção anual de casos novos por 10.000 hab.

Indicadores epidemiológicos	Construção	Utilidade	Parâmetro
Coeficiente de detecção anual de casos novos por 10.000 hab.	$\frac{\text{Casos novos residentes diagnosticados no ano}}{\text{População total residente}} \times 10.000$	Determinar a tendência secular da endemia e medir a intensidade das atividades de detecção de casos	Hiperendêmico: $\geq 4,0/10.000$ hab. Muito Alto: $4,0 - 2,0/10.000$ hab. Alto: $2,0 - 1,0/10.000$ hab. Médio: $1,0 - 0,2/10.000$ hab. Baixo: $< 0,2/10.000$ hab.

Fonte: Oliveira et al., (2015).

Na década de 1980, foram desenvolvidos os primeiros exames sorológicos para hanseníase após a identificação do anti-PGL-1 (antígeno glicolipídio-fenólico 1), biomarcador

específico do *M. leprae*. O anti-PGL-1 produz anticorpos das classes IgG e IgM, cuja pesquisa é o parâmetro sorológico mais padronizado e sendo avaliado na hanseníase, ter sensibilidade maior do que a baciloscopia. A detecção desses anticorpos vem demonstrando um indicativo de infecção subclínica pelo *M. leprae* ou doença ativa. A concentração dos anticorpos séricos correlaciona com a carga bacilar (BRASIL, 2022; MOURA et al., 2008a).

A aplicação da Inteligência Artificial (IA) pode simplificar a detecção de doenças apenas a partir das imagens da pele infectada e pode ajudar na capacidade humana de analisar informações complexas. A IA está assumindo a automação de processos, inclusive diferenciando no campo da saúde; onde a análise preditiva é o aprendizado de máquina (*machine learning*) usado para projetar vários algoritmos e modelos que auxiliam muito o processo de tomada de decisão (PATNAIK et al., 2018).

Dentro da mineração de dados existem técnicas de clusterização, como os algoritmos não supervisionados, que tem como objetivo agrupar um conjunto de objetos de dados em vários grupos ou “clusters” de modo que os objetos dentro de um *cluster* têm alta semelhança/afinidades, mas são muito diferentes dos objetos em outros *clusters* (HAN; KAMBER; PEI, 2014). O uso de técnicas de clusterização podem ser usadas para agrupar pacientes com hanseníase com base nas variáveis utilizadas no estudo, a fim de personalizar as intervenções de tratamento; prever a incidência futura de hanseníase ou pode ajudar a direcionar as intervenções de saúde pública.

Diante deste contexto, este estudo busca analisar a epidemiologia da hanseníase através de dados clínicos, socioeconômicos e da concentração do anti-PGL-1, obtidos através de exame sorológico, de pacientes e seus comunicantes utilizando técnicas de clusterização em um município hiperendêmico na região Oeste do Pará.

2 OBJETIVOS

2.1 Objetivo Geral

Analisar a epidemiologia da Hanseníase através do processo de mineração de dados utilizando técnicas de clusterização nas variáveis clínicas, socioeconômicas e dosagem de anti-PGL-1 de pacientes e seus comunicantes no município de Santarém no oeste do Pará.

2.2 Objetivos Específicos

- Realizar levantamento dos casos de hanseníase registrados na base do SINAN entre os anos de 2003 a 2014 e em uma semana de Busca Ativa na região estudada;
- Analisar a epidemiologia da hanseníase com base nos dados clínicos e socioeconômicos dos pacientes e seus comunicantes na região estudada;
- Analisar a soropositividade anti-PGL-1 em contatos domiciliares dos casos notificados com hanseníase na região estudada;
- Correlacionar a soropositividade do anti-PGL-1, dados clínicos e variáveis socioeconômicas dos pacientes e seus comunicantes através do processo de mineração de dados utilizando técnicas de clusterização.

3 REFERENCIAL TEÓRICO

3.1 Histórico da Hanseníase

A palavra lepra vem do latim *lepros*, que significa ato de sujar ou poluir. A lepra como era conhecida na antiguidade é uma das mais antigas patologias registradas pela humanidade. Mesmo fazendo parte do sofrimento humano desde a antiguidade, a sua identidade etiológica foi descoberta apenas ao final do século XIX, quando o médico norueguês Gerhard Henrik Armauer Hansen, ao analisar material de lesões cutâneas, descobriu o *Mycobacterium leprae*, o bacilo causador da doença e que pertence ao mesmo gênero do bacilo que causa a tuberculose (SANTOS; FARIA; MENEZES, 2008).

Relatos de hanseníase apareceram pela primeira vez por volta de 600 a.C., em papiros antigos que descrevem a existência da hanseníase na Índia, no Egito e na China (MONOT et al., 2005). Dados moleculares indicam que a hanseníase se originou no leste da África ou no Oriente Médio e se espalhou com sucessivas migrações humanas. Europeus e/ou norte-africanos, então, introduziram a hanseníase na África Ocidental e nas Américas nos últimos 500 anos (MONOT et al., 2009).

No continente europeu, o primeiro registro ocorreu na Grécia antiga logo após a chegada do exército de Alexandre, o Grande, vindo da Índia, e o deslocamento das legiões romanas, fizeram com que a Europa mantivesse uma alta prevalência da doença durante muitos anos (MARIOTTI et al., 2005).

Para os hebreus, ela era considerada uma maldição, um castigo divino e não propriamente uma enfermidade corporal. Para esse povo, existiam três enfermidades da pele que podiam se confundir: a sífilis, as micoses e a própria lepra. O estigma, a discriminação com a doença e com quem sofre a ação desta em seu corpo foram construídos pela associação do termo lepra às deformidades causadas ao paciente. Esta junção somada aos conceitos populares, sem fundamento científico algum, e aos religiosos, que eram carregados de piedade cristã, foram os responsáveis por grande parte, senão pela totalidade, dos problemas psicossociais que afetam os doentes até hoje (ALVES; FERREIRA; FERREIRA, 2014, p. 20).

Segundo Opromolla (1981, 2000), nas Américas não existia evidências da existência da moléstia entre as tribos indígenas do novo mundo, então presume-se que a hanseníase deve ter chegado com os colonizadores entre os séculos XVI e XVII.

No Brasil, os primeiros registros de hanseníase ocorreram no ano de 1600, quando se adotou o mesmo método de controle utilizado na Europa medieval, o isolamento dos

enfermos. Após os primeiros casos no Rio de Janeiro, outros focos da doença foram identificados, principalmente na Bahia e no Pará (YAMANOUCHI et al., 1993).

No século XIX, o estado do Pará já existia uma grande prevalência da hanseníase, uma vez que, Santarém encaminhava seus doentes para internamento no asilo Tucunduba (extinto em 1938), de Belém, presumindo que a hanseníase já existisse, nessa época, também em Manaus, devido as relações de comércio entre as cidades (SOUZA-ARAUJO, 1933, p. 172).

Os primeiros relatos sobre medidas profiláticas estabelecidas com o intuito de conter a doença, ocorreram na França, que determinou rigorosamente o isolamento dos enfermos, obrigando-os a morar em locais reservados para este fim (LONGO, 2006).

Os acometidos pela doença tinham de trazer a cabeça coberta e andar calçados para não infectarem os caminhos, passar à beira das estradas e avisar da sua passagem por meio de campainhas ou matracas para avisar os sadios de sua aproximação, logo se afastassem e não corressem risco de contaminar-se (MAURANO, 1950, p. 65).

No entanto o método utilizado não apresentou resultados positivos, em decorrência de contribuir para aumento do medo e o estigma associados à doença, especialmente em relação às deformidades, e às falsas informações sobre a forma de transmissão e possibilidade de cura, o que contribuía e ainda contribui para uma rejeição pela sociedade e, surpreendentemente, por profissionais da área da saúde (EIDT, 2004; LONGO, 2006).

Após a Medida Provisória nº 373/07, e a Lei nº 11.520/2007, os pacientes que foram submetidos ao isolamento e internação compulsória, têm direito a requerer a pensão especial, mensal, vitalícia e intransferível, um direito reconhecido pelo Governo Federal (BRASIL, 2007a, 2007b). Embora não tenha como apagar os momentos de sofrimento e angústia, foi a forma do Estado de reparar uma política baseada em preconceito, que infelizmente ainda deixa suas marcas até os dias atuais.

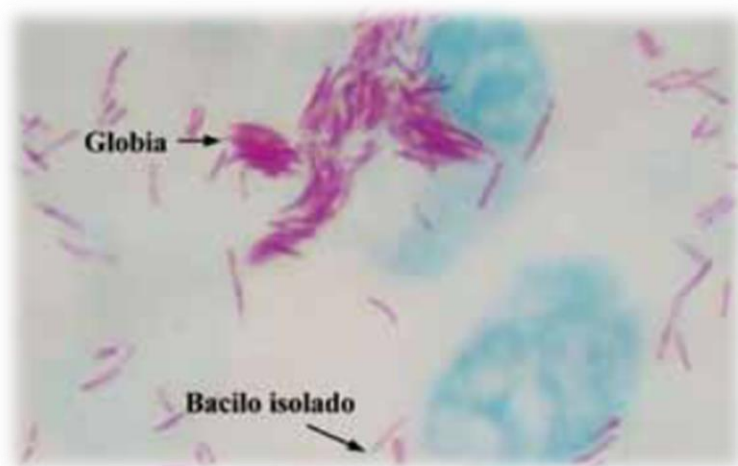
3.2 Etiologia

O *Mycobacterium leprae* causador da hanseníase, taxonomicamente, é uma micobactéria pertencem ao gênero *Mycobacterium*, sendo o único deste gênero dentro da família das Mycobacteriaceae, na ordem Actinomycetales. O nome do gênero *Mycobacterium* foi designado a ser utilizado num grupo de organismos que cresciam sob a forma de película em meio líquido. E, sendo definida como aeróbicas (uma característica geral desta micobactéria, mas que em algumas espécies e linhagens seja microaerofílicas) (RASTOGI; LEGRAND; SOLA, 2001).

As espécies deste grupo são organismos bacilares que, em algum estágio de seu ciclo de crescimento, apresentam uma distintiva propriedade tintorial, denominada acidorresistência. Essa propriedade se deve à presença de lipídeos singulares na superfície da célula, denominados ácidos micólicos, encontrados somente no gênero *Mycobacterium*, na superfície da célula micobacteriana (MANDIGAN et al., 2016, p. 500).

O *M. leprae* é um parasita intracelular obrigatório, com forma de bastonete e de coloração ácido-resistente, com 1–8 μm de comprimento e 0,3 μm de diâmetro, onde pode apresentar devido seu arranjo em centenas de bacilos aglomerados ou “globias” e/ou aglomerados menores onde as células individuais ocorrem em arranjos paralelos, os conhecidos “pacotes de charutos” (Figura 2) (BRASIL, 2010; BRENNAN; SPENCER, 2019).

Figura 2 - *Mycobacterium leprae* visto em globias e isolados.



Fonte: Brasil (2010, p. 24).

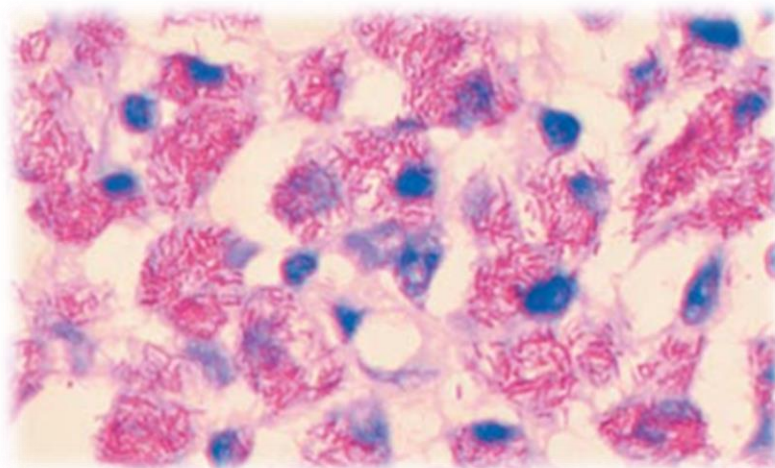
A forma mais agressiva da hanseníase é conhecida como lepromatosa, mostrada na Figura 3, sendo caracterizada por lesões rugosas e nodulares em todo o corpo (decorrentes a multiplicação das células bacterianas), com especificidade nas partes/regiões mais frias do corpo, como face e extremidades [...] “As células de *M. leprae* isoladas das lesões coram-se em vermelho-escuro pela fucsina carbólica, utilizada na técnica de coloração de organismos acidorresistentes, permitindo uma comprovação rápida e definitiva de infecção ativa (Figura 4)” (MANDIGAN et al., 2016, p. 861).

Figura 3 - Lesões cutâneas de hanseníase lepromatosa.



Fonte: Mandigan et al., (2016, p. 861).

Figura 4 – Micobactérias em uma coloração para organismos acidorresistentes.



Fonte: Mandigan et al., (2016, p. 860).

3.3 Hanseníase (Infecção, Transmissão, Classificação e Diagnóstico)

A hanseníase é uma doença crônica, infectocontagiosa, capaz de infectar grande número de indivíduos (alta infectividade), embora poucos adoeçam (baixa patogenicidade). Essas propriedades não ocorrem em função apenas das características intrínsecas do agente etiológico, mas dependem, sobretudo, da relação com o hospedeiro e o grau de endemicidade do meio, entre outros aspectos.

As vias aéreas superiores são consideradas o principal método de entrada do bacilo no organismo do ser humano (os humanos são os hospedeiros naturais, embora o tatu possa corresponder a um reservatório para a infecção humana), desenvolve numa temperatura excelente de crescimento aproximadamente entre 30°C a 35°C (inferior à temperatura corporal), por esse motivo o organismo cresce preferencialmente na pele e em nervos superficiais. A via

de eliminação e infecção do indivíduo pelo bacilo são principalmente as vias aéreas superiores: mucosa nasal e orofaringe (RODRIGUES; LOCKWOOD, 2011). Baseando-se em relatos de transmissão por lesões da pele, a possibilidade de um indivíduo patológico e sem tratamento pode eliminar bacilos através dessas lesões de pele e, provavelmente, infectar indivíduos saudáveis na condição deste indivíduo ter uma porta de entrada, ou seja, que tenha alguma fissura ou que não estejam com a pele íntegra (BRASIL, 2010).

Ao ser infectado pelo *M. leprae*, o hospedeiro pode encubar a patologia inicialmente de 12 a 14 dias, faixa de tempo médio de 2 a 7 anos, pois a hanseníase possui um tempo de incubação longo, apesar de alguns casos particulares foram notificados num período curto, aproximadamente 7 meses ou um mais longo de 10 anos. Geralmente, o surgimento da doença manifesta-se por meio de lesões de pele, que se apresentam como manchas esbranquiçadas ou avermelhadas, com diminuição ou ausência de sensibilidade (BROOKS et al., 2014, p. 323; LEVINSON, 2010, p. 173; TORTORA; FUNKE; CASE, 2016, p. 617).

Os pacientes no estágio mais avançada da patologia podem excretar grande número de bacilos pelas vias aéreas superiores, em média 109 bacilos viáveis por dia sendo as vias aéreas superiores dos pacientes o principal meio de disseminação de bacilos para o meio ambiente. Sendo também as vias, porta, de entrada mais prováveis do bacilo de Hansen no organismo, o que pode ser facilitado se o indivíduo que entra em contato com o bacilo apresentar pequenas lesões na mucosa nasal, solução de continuidade (MANDIGAN et al., 2016, p. 861).

Segundo Lockwood (2012) deduz que 95% da população mundial não é geneticamente suscetível à hanseníase. E, baseado nesse contexto, o indivíduo estar suscetível devido a sua condição imunológica e a determinados fatores genéticos o risco de adquirir a patologia está vinculado, também, as condições de moradia, número de pessoas que compartilham o mesmo quarto, endemidade do local onde moram, o nível de escolaridade baixo e fatores econômicos desfavoráveis (SALES et al., 2011).

O principal fator considerado neste trabalho, referente a forma clínica, foi a classificação operacional baseada na forma clínica, ou seja, definir as formas clínicas em que o paciente está transmitindo a patologia (multibacilar – MB) e as que não está (paucibacilar – PB)(WHO 1995, p. 21, WHO 2012).

O diagnóstico da hanseníase é exclusivamente clínico, onde o reconhecimento das características clínicas da hanseníase em um paciente detém uma abordagem minuciosa, uma vez que, a patologia manifesta de forma sutil em muitos indivíduos. No ato do exame clínico do paciente com suspeita de hanseníase, além da perícia técnica do profissional de saúde, o

ambiente tem que ter uma boa iluminação para detectar os sinais clínicos da hanseníase. Porém, em consonância com OMS, deve se suspeitar de hanseníase em pessoas com qualquer um dos seguintes sintomas e sinais: manchas hipocrômicas ou avermelhadas na pele, perda ou diminuição da sensação na(s) mancha(s) de pele, dormência ou formigamento de mãos/pés, nervos doloridos ou sensíveis; edema ou nódulos no rosto, ou nos lóbulos das orelhas, feridas ou queimaduras indolores nas mãos, ou pés. (BRASIL, 2021a).

Após o diagnóstico dos pacientes (exame clínico), são agrupados de acordo com a classificação operacional para caso de hanseníase conforme adotado da OMS (Tabela 1), buscando definir o esquema de terapia com poliquimioterapia (PQT), que está associado ao número de lesões cutâneas de acordo com os seguintes critérios: Paucibacilar (PB) – casos com até cinco lesões de pele e Multibacilar (MB) – casos com mais de cinco lesões de pele (Brasil 2016, p. 17).

Tabela 1 - Classificação operacional da hanseníase.

Hanseníase Paucibacilar – PB	PB (I, T e NP) <5 lesões de pele Baciloscopia negativa
Hanseníase multibacilar – MB	MB (BT, BB, BV e V) ≥ 5 lesões de pele Baciloscopia positivo em qualquer local.

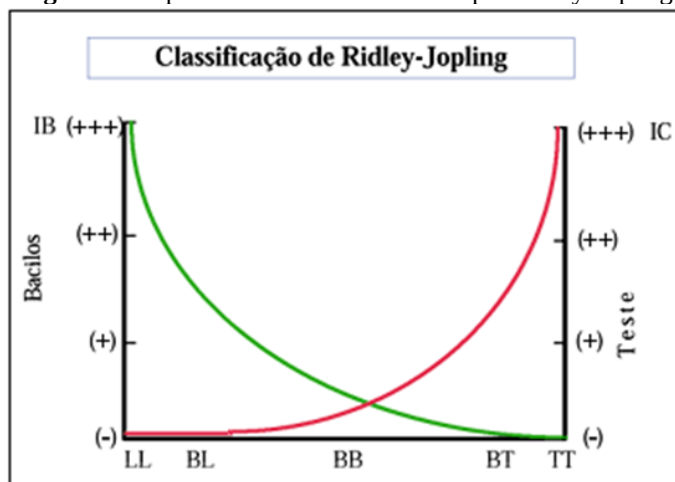
Fonte: Brasil e WHO (2016, p. 17; 2012).

Na década de 60, Ridley e Jopling estabeleceram a classificação espectral da hanseníase, neste sistema classifica a hanseníase como uma patologia espectral mediado pelo sistema imunológico com a hanseníase tuberculóide (TT) numa extremidade e do outro espectro a hanseníase Virchowiana (LL), “essas duas extremidades do espectro são consideradas clinicamente estáveis”. Entre as extremidades ficou subdividido o grupo borderline ou dimorfo em borderline-tuberculóide, borderline-borderline e borderline-Virchowiana, baseada em parâmetros clínico, bacteriológico, histopatológico e do grau de resposta imune mediada por células (GOULART; PENNA; CUNHA, 2002; RIDLEY; JOPLING, 1962).

Através dessa classificação, subdividiu-se a hanseníase em cinco grupos no espectro imunológico: tuberculóide-tuberculóide (TT), que corresponde à forma polar e estável da classificação de Madrid, borderline-tuberculóide (BT), borderline-borderline (BB), borderline-lepromatoso (BL) e lepromatoso-lepromatoso (LL), que é equivalente à forma lepromatosa e mais bacilífera na classificação de Ridley & Jopling do ano de 1966, Figura 5 e Tabela 2. Enquanto, a forma indeterminada (I) foi também proposta como sendo um estágio

inicial e transitório, que posteriormente evolui para uma das cinco formas clinicamente estáveis (KUMAR; UPRETY; DOGRA, 2006; PARDILLO et al., 2007; RIDLEY; JOPLING, 1966).

Figura 5 - Espectro clínico da hanseníase por Ridley-Jopling.



Nota: A imunidade celular (IC) medida pelo teste de Mitsuda é inversamente proporcional a carga bacilar, medida pelo índice baciloscópico (IB).

Legenda: TT – forma polar Tuberculóide (estável); LL – forma polar Virchowiana ou Lepromatosa (estável); BT, BB e BL – Grupo Borderline (Instável).

Fonte: Adaptador de Goulart, Penna e Cunha (2002).

Tabela 2 - Classificação Hanseníase: Ridley Jopling e OMS.

Classificação Ridley Jopling (1966)	Classificação da OMS (1981)
Hanseníase Tuberculóide (TT) Tuberculóide borderline (BT)	Hanseníase Paucibacilar até 5 lesões de pele Baciloscopia negativo em todos os sítios
Borderline borderline (BB) Virchowiana limítrofe (BL) Hanseníase Virchowiana (LL)	Hanseníase Multibacilar 6 ou mais lesões de pele Baciloscopia positivo em qualquer local

Fonte: Kumar, Uprety e Droga (2006).

Hanseníase Indeterminada (I) - forma clínica mais precoce na hanseníase, determinada pela forma (delimitação), cor e pelos. A forma é uma mácula ou mancha hipopigmentadas mal definida, situada na face lateral/externa da coxa, face e face extensora dos membros enquanto a cor vai depender do tipo de pele de cada indivíduo. Os pelos do corpo parecem normais nessas lesões, mas a mácula pode estar ligeiramente seca em comparação com a pele circundante (Figura 6)(KUMAR; UPRETY; DOGRA, 2006).

Figura 6 - Hanseníase indeterminada. Uma mancha eritematosa mal definida sobre o rosto.



Fonte: Kumar, Uprety e Dogra(2006).

Hanseníase Tuberculóide (TT) – forma uma extremidade do espectro na classificação de Ridley-Jopling. O IMC (imunidade mediada por célula) é alto e o teste da lepromina é reagente. Devido a resposta imunológica (IMC), as lesões podem cicatrizar por conta própria na maioria dos pacientes. Os sintomas na TT podem ser cutâneos, neurais ou ambos. Lesões cutâneas em TT não numeram mais de 3, têm até 10 cm de diâmetro e são bem definidos. Apresentam lesões pequenas podendo ocorrer em qualquer parte do corpo, mas muitas vezes há predileção por áreas expostas/descobertas do corpo. As lesões apresentam-se na forma de pápulas ou placas eritematosas com borda externa bem definida (Figura 7) (KUMAR; UPRETY; DOGRA, 2006).

Figura 7 - Hanseníase TT, com placa anestésica bem delimitada, eritematosa, sem pelos, seca.



Fonte: Kumar, Uprety e Dogra (2006).

Bordeline Tuberculóide (BT) – a lesão é muitas das vezes completamente anestésica. Primeiro, a sensibilidade à temperatura é perdida, seguida pela dor e, finalmente, pela perda da sensibilidade tátil. Na face, o comprometimento sensorial é difícil de demonstrar devido às ricas inervações sensoriais que podem compensar os nervos danificados. Em uma lesão individual, a perda sensorial é mais acentuada no centro do que na periferia. Devido ao envolvimento dos nervos autônomos, a superfície fica seca e a sudorese é perdida.

As lesões maculares são hipopigmentadas, mas também podem ser eritematosas, com superfície descamativa leve e seca e bordas bem definidas determinadas com Bordeline Tuberculóide – BT, Figura 8 (KUMAR; UPRETY; DOGRA, 2006).

Figura 8 - Hanseníase BT maculo anestésica, hipopigmentadas, hipostética, placa seca com perda de apêndices.



Fonte: Kumar, Uprety e Dogra (2006).

Borderline Borderline (BB) – caracterizada por ser a “fase/progressão” da infecção da hanseníase por ser imunologicamente mais instável do espectro limítrofe. Os pacientes movem-se rapidamente para os polos lepromatoso ou tuberculóide. Apresentam lesões dimorfas, ou seja, lesões características dos tipos tuberculóide e lepromatoso, podendo apresentar às vezes na forma de pápulas infiltradas, placas e até nódulos, tendem a ser simétricas. A lesão característica BB é uma placa anular com uma margem interna “perfurada” bem demarcada e uma margem externa inclinada conforme a figura abaixo, onde:

Figura 9 (A) – Bordeline Bordeline – demonstra paciente com múltiplas placas eritematosas e secas, algumas delas com a característica morfologia anular com uma borda interna “perfurada”. Notar a presença de placa bem definida remanescente do espectro tuberculóide, sugerindo rebaixamento.

Figura 9 (B) – Bordeline Bordeline – demonstra paciente com múltiplas placas anulares eritematosas, simetricamente distribuídas, mal definidas, com algumas das placas mostrando uma margem interna “perfurada”.

Figura 9 - Hanseníase Bordeline Bordeline (BB).



Fonte: Kumar, Uprety e Dogra (2006)

Bordeline Lepromatoso (BL) – lesões múltiplas e há uma tendência aumentada para a simetria. Com o tempo, as máculas sofrem infiltração a partir do centro e formando placas e nódulos. Uma característica BL é que nos troncos periféricos nervosos são espessados e tendem a ser simétricos. (Figuras 10A, 10B) (KUMAR; UPRETY; DOGRA, 2006).

Figura 10 A – Hanseníase BL rebaixada de BT; múltiplas placas secas bem definidas com apêndices diminuídos e as lesões apresentam simetria em sua distribuição.

Figura 10 B – Hanseníase BL com múltiplas lesões maculo-placas eritematosas mal definidas, mas distintas, com tendência à simetria em sua distribuição.

Figura 10 - Hanseníase Bordeline Lepromatoso (BL).



Fonte: Kumar, Uprety e Dogra (2006).

Lepromatoso Lepromatoso ou Virchowiana (LL/V) – a imunidade é severamente prejudicada nesse espectro devido a multiplicação e disseminação descontrolada dos bacilos da lepra. A LL com aparência de novo é chamada de hanseníase Virchowiana polar (LL p), e quando resultante de rebaixamento, é chamada de hanseníase Virchowiana subpolar (LL s). As lesões cutâneas de LL são múltiplas com disseminação simétrica bilateral em todo corpo exceto em áreas quentes do corpo, como: axilas, virilha, períneo e couro. Clinicamente, a doença LL pode apresentar-se em qualquer um dos estágios descritos abaixo ou com uma mistura dessas lesões (Figura 11A, 11B)(KUMAR; UPRETY; DOGRA, 2006).

Figura 11 - Hanseníase Lepromatoso Lepromatoso (LL) A. infiltração difusa da face e B. infiltração e nodulação da orelha.



Fonte: Kumar, Uprety e Dogra (2006).

Hanseníase Neural Pura (NP) – é uma variante da hanseníase caracterizada pelo envolvimento isolado de troncos nervosos periféricos na ausência de lesões cutâneas. As características clínicas desta variante da hanseníase incluem espessamento do nervo, dor, sensibilidade e comprometimento sensório-motor. A neuropatia sensorial é mais predominante no NP, com perda de calor e sensibilidade à dor. Os troncos nervosos do membro superior são mais comumente envolvidos e pode ter uma apresentação espectral variando de TT a LL. Clinicamente, é provável que o NP seja TT se apenas um ou dois troncos nervosos estiverem espessados e limítrofe se houver vários nervos espessados (KUMAR; UPRETY; DOGRA, 2006).

O diagnóstico da hanseníase tem com padrão ouro o exame clínico, mas pode ser complementado por exames laboratoriais (reação intradérmica de lepromina, teste rápido do bacilo na mucosa nasal e pele, biópsia para exame histopatológico) (BROOKS et al., 2014, p. 323; LEVINSON, 2010, p. 174). Os exames supracitados deverão ser feitos quando o paciente

apresentar um dos seguintes sinais cardinais da patologia: lesão/lesões de pele com alteração de sensibilidade, acometimento de nervo(s) periférico(s), com ou sem espessamento, associado a alterações sensitivas e/ou motoras e/ou autonômicas; e baciloscopia positiva de esfregaço intradérmico (LYON; GROSSI, 2014).

Entre outros exames além da baciloscopia (BAAR), existe teste sorológico do Anticorpo Glicolípido Fenólico-1 (Anti-PGL1) específico do bacilo *M. leprae* (BRENNAN; SPENCER, 2019).

O exame clínico quanto aspectos epidemiológicos são as principais ferramentas para o diagnóstico da hanseníase, mas pode contar com exames de apoio (exames laboratoriais), como raspados de pele ou de mucosa nasal com bisturi ou uma amostra de biópsias corados pela técnica de BAAR – bacilo álcool-ácido resistente, conhecida como Ziehl-Neelsen. Uma curiosidade sobre os testes sorológicos não treponêmicos par sífilis (triagem), induz resultados falsos positivos na hanseníase com frequência (BROOKS et al., 2014).

No entanto, nenhuma dessas abordagens diagnósticas consegue detectar infecções assintomáticas por *M. leprae*, ou seja, além da abordagem clínica, existem técnicas diagnósticas que se baseiam em respostas imunológicas direcionadas contra o *M. leprae*. Em particular, alguns testes de aglutinação lateral de fluxo, fita reagente e partículas que incorporam o epítipo sintético de dissacarídeos ou trissacarídeos de glicolípido fenólico-I (PGL-I) têm sido usados em estudos de campo. O teste ELISA de anticorpo anti-PGL-I baseia-se na detecção de anticorpos IgM contra PGL-I, um componente glicolípido dominante específico do *M. leprae* da parede celular), sendo restrito a detecção da maioria dos pacientes MB (correlaciona com a carga bacilar, ou seja, quanto maior a carga bacilar maior a presença do PGL-I), tornando-se assim limitado na identificação de pacientes com hanseníase PB (VAN HOOIJ; GELUK, 2016).

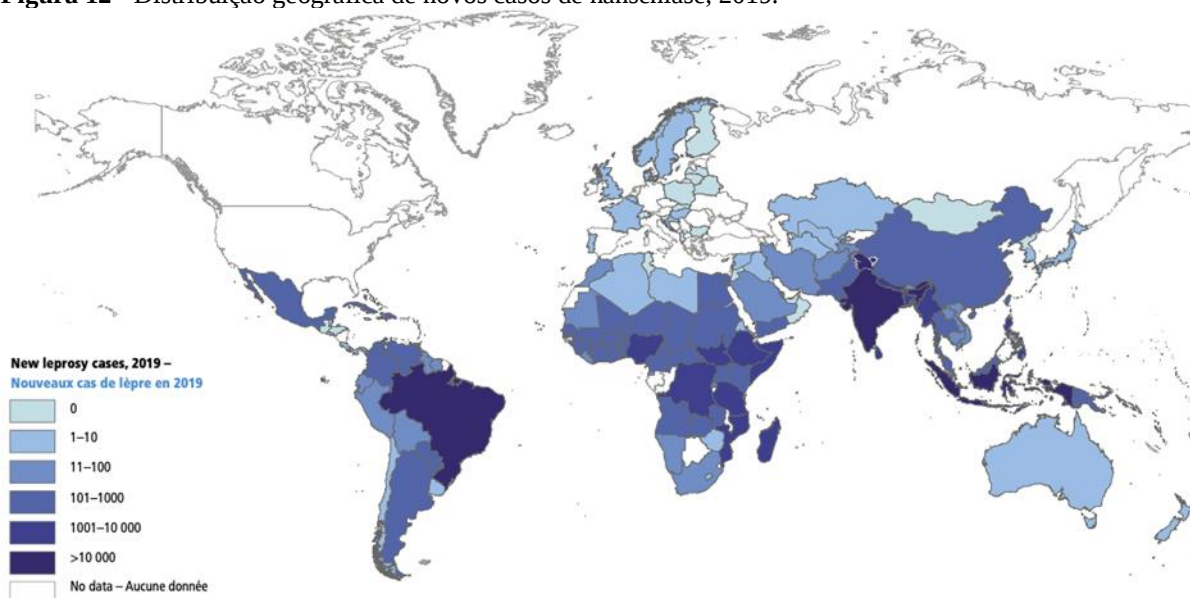
3.4 Aspectos Epidemiológicos da Hanseníase

3.4.1 Cenário mundial

No ano 2020, foram detectados 127558 novos casos de hanseníase, de acordo com dados oficiais de 139 países integrantes a WHO. Novos casos que inclui 8629 crianças com menos de 15 anos, trazendo uma taxa de detecção para novos casos entre a população infantil de 4,4 por milhão de crianças (WHO, 2022).

Em 2019, 161 países relataram um total 202.256 novos casos de hanseníase detectados em 118 países (26,0 por milhão de habitantes). Assim sendo, 96% foram relatados pelos 23 países prioritários globais, incluindo 79% na Índia, no Brasil e na Indonésia e de pôr outro lado 66 países relataram menos de 100 casos, ver Figura 12 e Tabela 3 (OMS, 2021).

Figura 12 - Distribuição geográfica de novos casos de hanseníase, 2019.



Fonte: WHO (2020).

Tabela 3 - Detecção de caso novo, por região da OMS, 2019.

Região	Países que reportaram ≥ 1 caso	Total de casos novos reportados		Casos novos em crianças		Casos novos com G2I	
		Número	Taxa (por milhão de habitantes)	Número	Taxa (por milhão de habitantes)	Número	Taxa (por milhão de habitantes)
África	38	20209	18	2150	5,2	2933	2,6
Américas	24	29936	29,5	1612	7,1	2544	2,5
Mediterrâneo Oriental	15	4271	5,8	149	0,68	254	0,5
Europa	6	42	<0,1	0	0	4	0
Sudeste Asiático	10	143787	70,4	10661	20,4	4817	2,4
Pacífico Ocidental	25	4011	2,1	411	1,1	264	0,14
Total mundial	118	202256	26	14983	7,9	10816	1,4

Legenda: G2I – Grau 2 de Incapacidade Física.

Fonte: OMS (2021).

A Hanseníase é uma das vinte Doenças Tropicais Negligenciadas (DTNs) e por esse entendimento, sua ocorrência habitua-se relacionar com as más condições socioeconômicas. Segundo OMS (2021) aponta como falha no diagnostico prévio e a sensibilização da comunidade devido à falta dum diagnostico mais efetivo, pois quando totaliza 10816 de novos casos em 94 países e estão incluídos 370 crianças que apresentaram G2I no momento do diagnostico, ou seja, pode-se dizer que os dados podem estar subnotificados considerando que tais números de criança provavelmente terá uma quantidade significativa maior sobre crianças

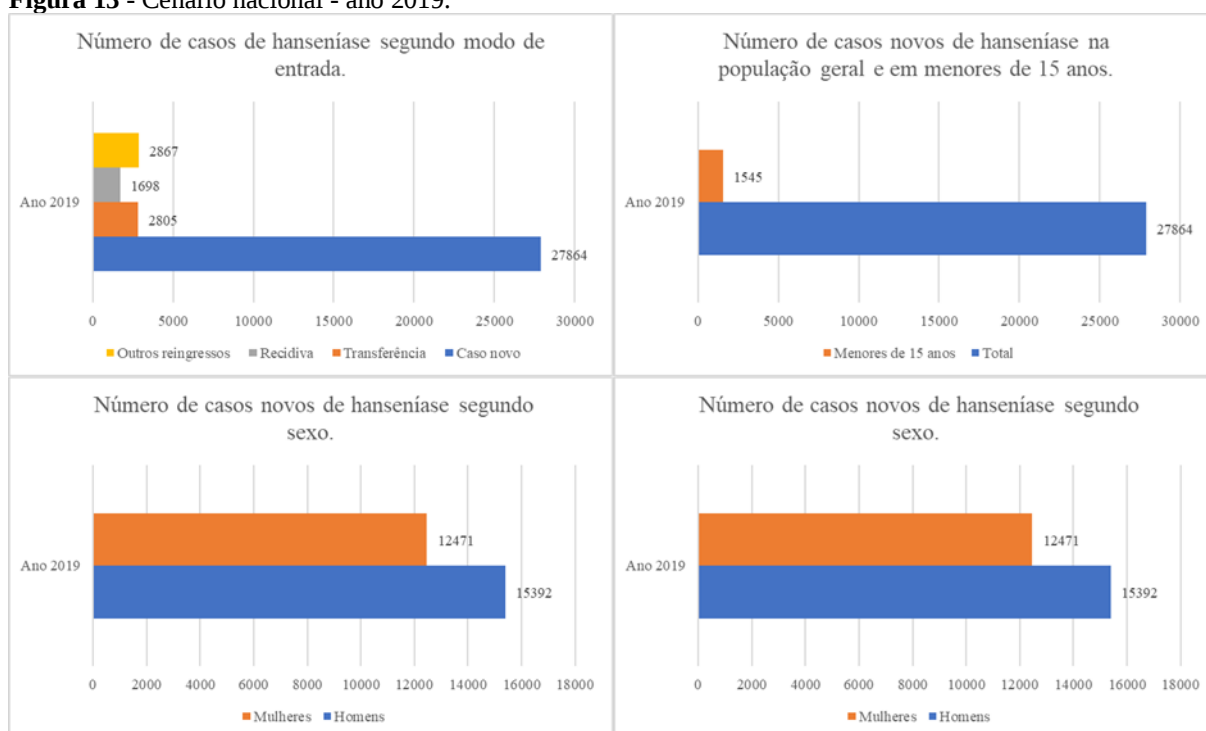
com G2I. A taxa geral de G2I foi de 1,4 por milhões de habitantes sendo inferior de 2,0 assinalados em 2014.

Nesse cenário, o Brasil devido aos casos relatos de hanseníase, ocupa o segundo lugar na relação de países com maior número de casos no mundo, atrás apenas da Índia (BRASIL, 2021a).

3.4.2 Cenário nacional

No Brasil, em 2019, foram notificados 35234 casos de hanseníase (Caso novo, Transferência, Recidiva e Outros reingressos) onde 27863 casos novos, separado por gênero masculino com 15392 (55,2%) e sexo feminino com 12471 (44,8%). Desses, 1545 (5,5%) casos novos de crianças menores de 15 anos, casos multibacilares notificados 21851 (78,4%). E, relacionado ao Grau de Incapacidade Física (GIF) foram notificados 2.351 (9,9%) apresentaram grau de incapacidade física 2 (GIF2/G2I)) ver Figura 13 (BRASIL, 2021b).

Figura 13 - Cenário nacional - ano 2019.



Fonte: Brasil (2021b).

Em um recorte entre os anos de 2014 a 2019, foram diagnosticados 168433 casos novos de hanseníase. Dentre esses casos, observa-se a predominância da doença em indivíduos do gênero masculino (55,2%), na faixa etária de 50 a 59 anos, raça/cor parda (58,4%), seguida de brancos (24,5%) e preta (12,2%) com Ensino Fundamental incompleto de 1ª a 4ª série (20,3%) seguidos com Ensino Fundamental incompleto 5ª a 8ª série (15,2%) (BRASIL, 2021b).

3.4.3 Cenário regional

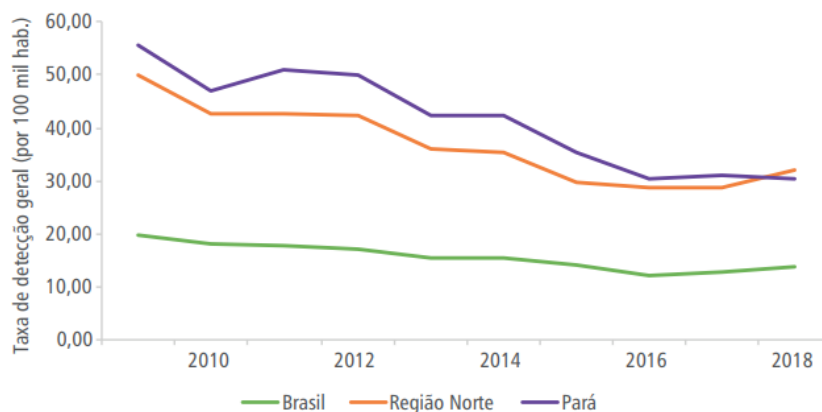
No Estado do Pará entre os anos de 2009 a 2018 foram diagnosticados 32926 casos novos de hanseníase com a maioria de GIF0 com 20546, GIF1 com 6801 e GIF2 com 1885. Pacientes, casos novos, com GIF2 menores de 15 anos foram de 101, de 15 a 29 anos 341, de 30 a 59 anos foram 936 e de 60 anos ou mais com 507 casos novos. Quanto à escolaridade, destaca-se o GIF2 em indivíduos (58%/1885) tendo o ensino fundamental incompleto. E, quando relacionado com a classificação operacional, os casos multibacilares também foram superiores em todas as categorias do GIF, ou seja, para GIF0 54,9%, GIF1 88,6% e GIF2 correspondendo a 95,6% MB dos casos (BRASIL, 2020, p. 51).

Em relação aos casos que receberam alta por cura com avaliação da incapacidade, observa-se que 2014 foi o ano com o maior quantitativo de GIF1 na alta, com 419 casos, e 2013 o ano com o maior quantitativo de GIF2 na alta, com 127 casos. No período da análise, 4.671 indivíduos receberam alta do tratamento com alguma incapacidade física; destes, 1.155 (24,7%) foram avaliados com GIF2 (BRASIL, 2020, p. 52).

No período de 2009 a 2018, a taxa de detecção geral passou de 55,70 para 30,44 casos por 100 mil habitantes, correspondendo a um decréscimo de 45,4% (Figura 14). Essa redução também foi observada na taxa de detecção anual de casos novos de hanseníase na população de zero a 14 anos, que passou de 19,11 para 11,10 casos por 100 mil habitantes, representando declínio de 41,9% (Figura 15). A taxa de casos novos com grau 2 de incapacidade física apresentou oscilação ao longo do tempo; em 2009, foi de 30,28, e em 2018, de 23,53 casos por 1 milhão de habitantes. Durante o período, o decréscimo no indicador foi de 22,3%(Figura 16) (BRASIL, 2020).

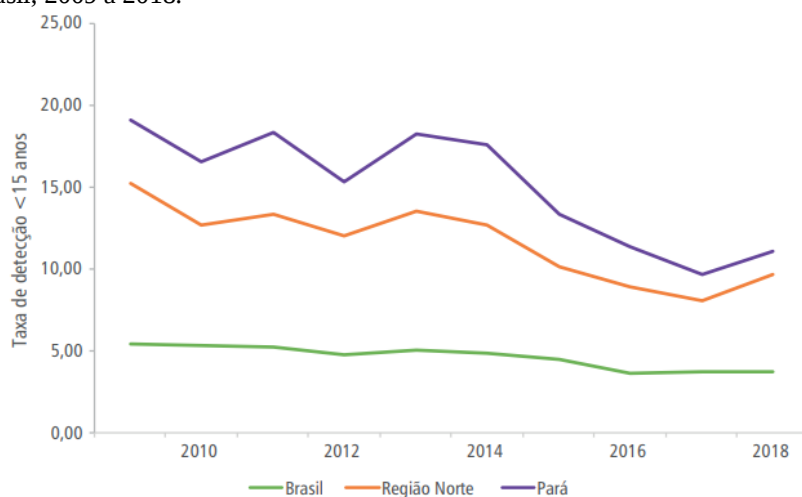
No ano de 2018, o Estado do Pará foi classificado como “hiperendêmico” na população de menores de 15 anos e na população geral como “muito alto” (BRASIL, 2020, p. 52).

Figura 14 - Taxa de detecção geral de casos novos de hanseníase (por 100 mil hab.). Pará, região Norte e Brasil, 2009 a 2018.



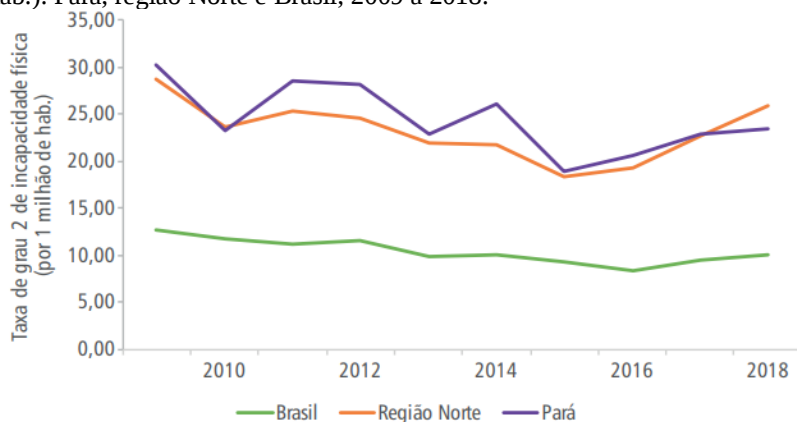
Fonte: Brasil (2020, p. 52).

Figura 15 - Taxa de detecção de casos novos de hanseníase em menores de 15 anos (por 100 mil hab.). Pará, região Norte e Brasil, 2009 a 2018.



Fonte: Brasil (2020, p. 52)

Figura 16 - Taxa de casos novos de hanseníase com grau 2 de incapacidade física no momento do diagnóstico (por 1 milhão de hab.). Pará, região Norte e Brasil, 2009 a 2018.



Fonte: Brasil (2020, p. 52)

3.5 Glicolípido Fenólico 1 (PGL-1)

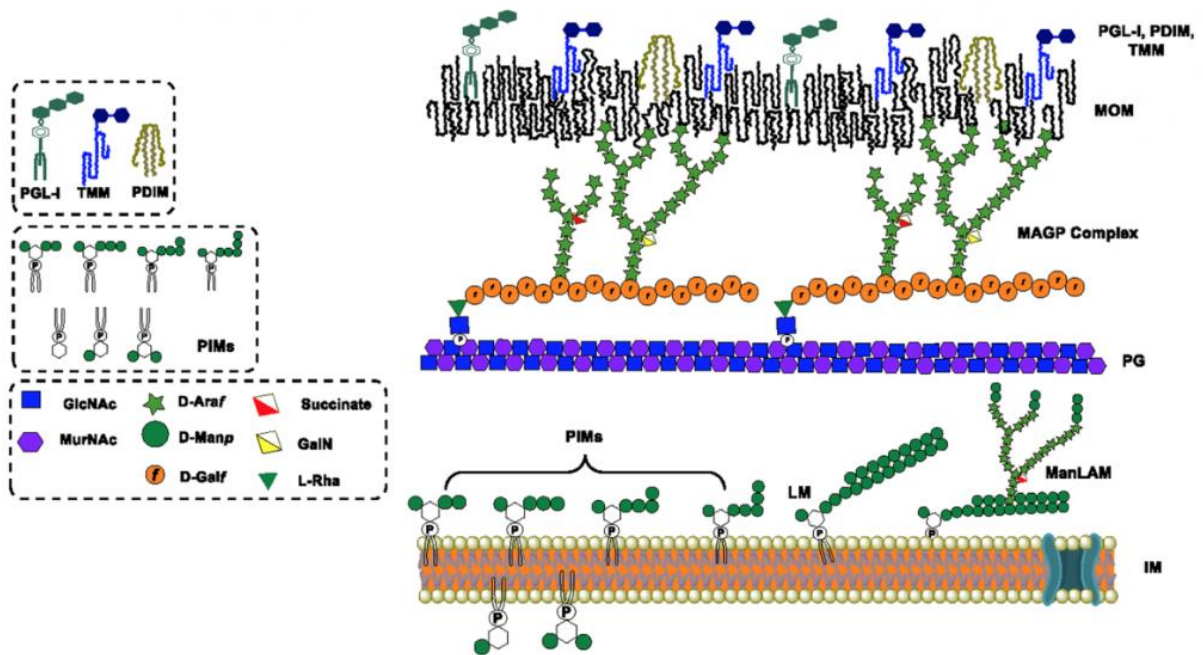
O Glicolípido Fenólico 1 (PGL-1) é um antígeno específico de *M. leprae* que induz a formação de anticorpos de classe IgM (FRADE et al., 2017). A presença desses anticorpos reflete a carga bacilar, com títulos elevados nos multibacilar e baixos ou ausentes nos paucibacilar. A sorologia pode identificar indivíduos com infecção subclínica, demonstrando a soropositividade entre contatos correlação com desenvolvimento das formas multibacilar (LASTÓRIA; ABREU, 2012).

A presença desses anticorpos reflete a carga bacilar, com títulos elevados nos multibacilares e baixos ou ausentes nos paucibacilares. No monitoramento da terapêutica, diminuição dos anticorpos acompanha o clearance do antígeno, enquanto persistência pode representar resistência terapêutica; aumento em indivíduos tratados pode indicar recidiva. Além disso, altas concentrações no início do tratamento indicam risco de desenvolver reação tipo A sorologia pode identificar indivíduos com infecção subclínica, demonstrando a soropositividade entre contatos correlação com desenvolvimento das formas multibacilares (LASTÓRIA; ABREU, 2012).

O lipídeo predominante existente do *M. leprae* além de ser único e específico da espécie é o glicolípido fenólico I (PGL-I; existem alguns subprodutos metabólicos PGL-II e -III). O PGL-I provou ser um antígeno útil para o sorodiagnóstico específico da hanseníase. Também foi demonstrado que danifica indiretamente as mitocôndrias dos axônios, levando à desmielinização dos nervos, a principal causa de danos nos nervos na hanseníase. A Figura 17 fornece uma representação molecular da arquitetura da parede celular do *M. leprae* (BRENNAN; SPENCER, 2019).

O PGL-1 tem sido relatado em vários estudos mostrando que os pacientes com Hanseníase LL produzem altas concentrações de imunoglobulina do tipo IgM em resposta a esse antígeno (soropositividade de 80-100%), enquanto os pacientes BT mostraram Imunoglobulina específica em níveis muito mais baixos (soropositividade de 30-60%). Dentre os estudos mencionados, um deles monitorou os contatos ao longo de um período de 6 anos mostrou que existe um risco 7,2 vezes maior de desenvolver Hanseníase (MB ou PB) em contato soropositivo com anticorpos para PGL-I quando comparado aos contatos soronegativos, aumentando para 24 vezes maior Risco de desenvolvimento de Hanseníase MB. A porcentagem de contatos que evolui para a doença entre os contatos soropositivos sugere que a sorologia com anti-PGL-I poderia ser útil como teste de prognóstico (MOURA et al., 2008b).

Figura 17 - Modelo da parede/envelope celular de *Mycobacterium leprae*.



Legenda: Glicolípido fenólico I (PGL-I); monomicolato de trealose (TMM); dimicocerosato de fticocero (PDIM); manósidos de fosfo-inositol (PIMs); *N*-acetil glucosamina (GlcNAc); D-arabinofuranose (D-Araf); succinato; *N*-ácido acetilmurâmico (MurNAc); D-fosfomanose (D-Man p); galactosamina (GalN); D-galactofuranose (D-Gal f); L-ramnose (L-Rha).

Fonte: Brennan e Spencer (2019)

Tabela 4. Características da classificação de Ridley Jopling em todo o espectro.

	Número e tipo de lesão(ões)	Tamanho da(s) lesão(ões)	Superfície da lesão	Sensações	Nervos periféricos	Cabelo	Baciloscopia	Reação típica	Reatividade e à lepromina	Histopatologia
TT	Uma a três mácula(s) ou placa(s) bem definida(s)	Tamanho variável, geralmente grande	Seco, escamoso e infiltrado	Ausente	Aumento assimétrico de um único nervo. O nervo para o patch pode ser palpável.	Ausente	Negativo	Tipo 1	Fortemente positivo (+++)	Granuloma constituído por linfócitos epitelióides e células gigantes erode a epiderme. Pequenos nervos são completamente apagados por granuloma caseoso. AFB não são vistos.
BT	Máculas ou placas, em número variável, até dez. Ocasionalmente mostrando extensões digitiformes (pseudópodes) ou lesões satélites.	Tamanho variável, algumas lesões são grandes	Seco, escamoso e infiltrado (mas menos do que na lesão TT)	Marcadamente diminuído	Aumento assimétrico de poucos nervos	Marcadamente diminuído	bacilos negativos ou escassos	Tipo 1	Fracamente positivo (+ ou ++)	O granuloma consiste em células epitelióides e uma mistura de macrófagos e linfócitos. Células gigantes também são observadas. Limpar SEZ presente. Os nervos são substituídos por granuloma. Bacilos fragmentados são vistos com coloração AFB.
BB	Máculas ou placas com aspecto anular de 'queijo suíço'. Número da variável (10–30).	Variável	Seco, escamoso levemente brilhante	Moderadamente diminuído	Aumento assimétrico de vários nervos	Moderadamente diminuído	Número moderado de BAAR presente	Tipo 1/Tipo 2	Negativo / fracamente positivo	Granuloma consistindo predominantemente de células epitelióides e macrófagos e linfócitos dispersos também são vistos. A SEZ é sempre clara. A seção transversal do nervo mostra aparência de cebola cortada. AFB estão sempre presentes.
BL	Máculas, pápulas, placas e alguns nódulos, lesões de morfologia variável, de distribuição assimétrica.	Variável, geralmente pequena	Brilhante (menos que LL)	Ligeiramente diminuído	Aumento cada vez mais simétrico dos nervos	Ligeiramente diminuído	Grande número de AFB visualizados	Tipo 1/Tipo 2	Negativo	O granuloma ao redor do feixe neuro vascular do nervo consiste principalmente de macrófagos e aglomerados isolados de células epitelióides. Os linfócitos estão espalhados esparsamente. ZEE é claro. Os nervos mostram infiltração concêntrica de células perineural e aparência de cebola cortada mais conspícua. AFB são abundantes, incluindo pequenos globias.
LL	Máculas, placas e lesões nodulares. Inúmeras lesões mostrando tendência à simetria na distribuição.	Pequena	Brilhante	Pode ser normal	Envolvimento simétrico de troncos nervosos com 'anestesia com luva e meia'. Os nervos podem parecer normais à palpação ou podem estar espessados.	Quase normal (inicialmente)	Abundância de AFB incluindo globias são vistos	Tipo 2	Negativo	O granuloma expansivo e solto é distinto, com muitos macrófagos carregados com BAAR. Poucos linfócitos também são vistos. AFB são vistos em células endoteliais e também em células de Schwann. A SEZ é sempre clara. Os nervos estão minimamente envolvidos.

Fonte: Kumar, Uprety e Dogra (2006 tradução nossa).

A soropositividade em uma população em geral tem uma distribuição uniforme. Isso pode significar que não existe diferença entre indivíduos saudáveis e indivíduos Hansênicos para distinção entre infecção subclínica e doença. Neste caso, testes sorológicos baseados na detecção de anticorpos IgM contra PGL-I não devem ser utilizados como ferramenta única de diagnóstico para rastreio de população a fim de detectar casos de hanseníase (MOURA et al., 2008b)

Além da importância de conseguir um diagnóstico precoce, uma abordagem sistemática e qualificada para monitorar os contatos intradomiciliares é considerada essencial para quebrar a transmissão do *M. Leprae*. Os indivíduos em contato direto com os pacientes não tratados têm um risco vitalício de 10% de sucumbir a hanseníase, particularmente quando são parentes consanguíneos. Embora um título positivo para anti-PGL-I seja um biomarcador de infecção, não é um indicador absoluto de doença. A evidência indica que aumentos substanciais nos títulos do anti-PGL-I são indicadores importantes de progressão da doença, mesmo na ausência de lesões cutâneas ou danos nos nervos. Em indivíduos saudáveis, um alto título anti-PGL-I sugere haver uma carga bacilar escondida que conduz a resposta ao antígeno. Os contatos intradomiciliares que demonstrem aumentos substanciais no título contra PGL-I devem ser monitorados de perto (FRADE et al., 2017)

3.6 Mineração de Dados

Para Fayyad, Piatetsky-Shapiro e Smyth (1996), afirmam que a Mineração de Dados (*Data Mining*) passa a ser um processo de refinamento nos conhecimento, descoberta do conhecimento, que consiste na realização da análise dos dados, , e aplicação de algoritmos, ver limitações computacionais, produzem um conjunto de padrões e/ou agrupamentos de certos dados com mesmos parâmetros. Então, definimos a Mineração de Dados como “o processo não trivial de identificação de padrões válidos, desconhecidos, potencialmente úteis e, no final das contas, compreensíveis em dados”.

Numa abordagem estatística, a mineração de dados serve para análise de grandes conjuntos de dados com a intenção de encontrar relacionamentos inesperados e de concatenar os dados de uma forma que eles sejam tanto utilizáveis quanto compreensíveis ao proprietário da base de dados (HAND; MANNILA; SMYTH, 2001).

A mineração de dados tem um aspecto multidisciplinar devido os campos de atuação, logo, unir técnicas de aprendizagem de máquinas, estatística, reconhecimentos de

padrões, banco de dados e visualização que possibilita extrair informações aglomeradas independentes do tamanho da base de dados (CAMILO; SILVA, 2009).

A Mineração de Dados é comumente classificada pela sua capacidade em realizar determinadas tarefas (CAMILO; SILVA, 2009), como por exemplo:

- **Descrição** (*Description*): é a tarefa utilizada para descrever os padrões e tendências revelados pelos dados. A descrição geralmente oferece uma possível interpretação para os resultados obtidos. A tarefa de descrição é muito utilizada em conjunto com as técnicas de análise exploratória de dados, para comprovar a influência de certas variáveis no resultado obtido.
- **Classificação** (*Classification*): Uma das tarefas mais comuns, a Classificação, visa identificar a qual classe um determinado registro pertence. Nesta tarefa, o modelo analisa o conjunto de registros fornecidos, com cada registro já contendo a indicação à qual classe pertence, a fim de 'aprender' como classificar um novo registro (aprendizado supervisionado).
- **Estimação** (*Estimation*) ou **Regressão** (*Regression*): A estimação é similar à classificação, porém é usada quando o registro é identificado por um valor numérico e não um categórico. Assim, pode-se estimar o valor de uma determinada variável analisando-se os valores das demais.
- **Predição** (*Prediction*): A tarefa de predição é similar às tarefas de classificação e estimação, porém ela visa descobrir o valor futuro de um determinado atributo.
- **Agrupamento** (*Clustering*): A tarefa de agrupamento visa identificar e aproximar os registros similares. Um agrupamento (ou *cluster*) é uma coleção de registros similares entre si, porém diferentes dos outros registros nos demais agrupamentos. Esta tarefa difere da classificação pois não necessita que os registros sejam previamente categorizados (aprendizado não-supervisionado).
- **Associação** (*Association*): A tarefa de associação consiste em identificar quais atributos estão relacionados. Apresentam a forma: SE atributo X ENTÃO atributo Y. É uma das tarefas mais conhecidas devido aos bons resultados obtidos, principalmente nas análises da "Cestas de Compras"(Market Basket), onde identificamos quais produtos são levados juntos pelos consumidores.

Mineração de Dados refere-se à extração de conhecimento de abundância de dados armazenados em banco de dados. Tem como função descrever classes e/ou conceitos, realizar

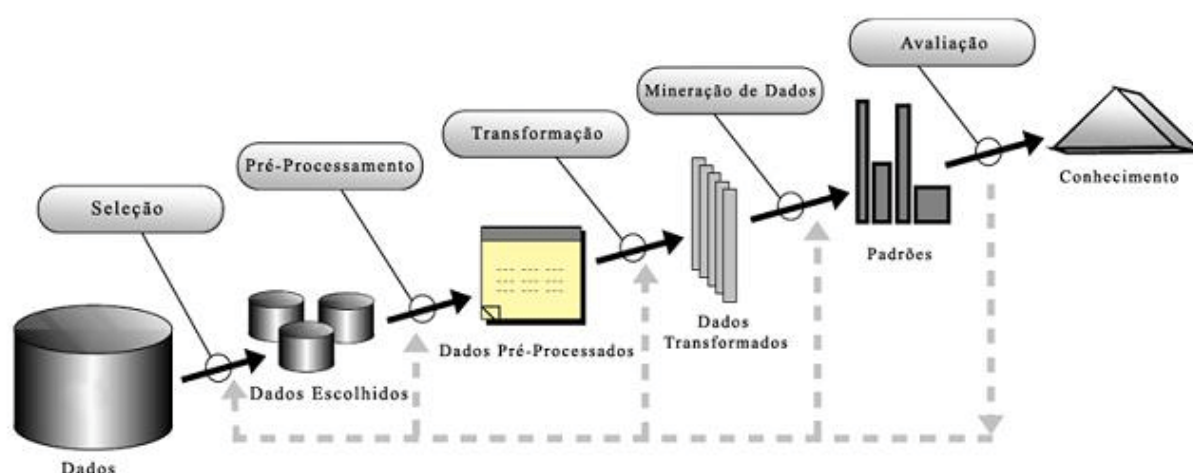
análise associativa, classificação e predição, análise de agrupamento, análise de outliers e análise de evolução de dados (HAN; KAMBER; PEI, 2014).

O processo global de **KDD – Knowledge Discovery in Databases** (Descoberta de Conhecimento em Bancos de Dados) inclui a avaliação e a possível interpretação dos padrões minados para determinar quais padrões podem ser considerados novos conhecimentos (Figura 18) (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996).

Um conjunto de dados pode conter objetos que não estão em conformidade com o comportamento geral ou o modelo dos dados. Esses objetos de dados são excepcionais. Muitos métodos de mineração de dados descartam outliers como ruído ou exceções. No entanto, em alguns aplicativos (por exemplo, detecção de fraude) os raros eventos podem ser mais interessantes do que os mais regulares. A análise de dados outlier são referidas como análise outlier ou anomalias mineração (HAN; KAMBER; PEI, 2014).

Levando em consideração a quantidade de informações associadas ao processo de análise epidemiológica da hanseníase em um período de dez anos além das variáveis associadas ao ambiente, a situação socioeconômica, higiene, quantidades pessoas em contato com pacientes que estão na fase disseminadora da doença, o presente trabalho propõe a utilização técnicas de mineração de dados (*Data Mining*) para identificar padrões e ou tendências que auxiliem o processo de tomada de decisão para o combate na hanseníase (BRITO et al., 2015; HAN; KAMBER; PEI, 2014).

Figura 18 - Uma Visão Geral dos Passos que compõem o Processo KDD.



Fonte: Fayyad, Piatetsky-Shapiro e Smyth (1996).

3.6.1 Agrupamento de dados

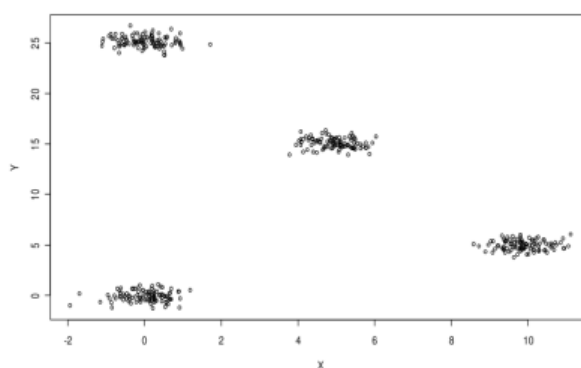
Os algoritmos de *Clustering* (agrupamento) que realizam tarefas de análise automática de dados são excelentes ferramentas para ajudar no processo de mineração de dados, pois consistem em processos de identificação de padrões válidos, novos, com grande potencialidade de serem utilizados e compreendidos, buscando melhorar o entendimento de um problema ou apoiar numa tomada de decisão. Ou seja, dentro do processo se observa o envolvimento de: preparo dos dados, o emprego de algoritmos de mineração de dados, a análise e interpretação de padrões, e saber representa/avaliar o conhecimento extraído. Sendo que, o resultado desse processo possa obter dados que indique casos, cenários, exemplos ou instâncias representativas do domínio a ser tratado (WEBBER; ZAT; LIMA, 2013)

As técnicas de agrupamento consideram as estruturas de dados (tuplas) como objetos. Eles particionam os objetos em grupos, ou *clusters*, de modo que os objetos dentro de um *cluster* são "semelhantes" uns aos outros e "dessemelhantes" a objetos em outros agrupamentos. A similaridade é comumente definida em termos de como "fechar" os objetos estão no espaço, baseado em uma função de distância. A "qualidade" de um *cluster* pode ser representada pelo seu diâmetro, a distância máxima entre dois objetos no *cluster* (HAN; KAMBER; PEI, 2014).

Para Vale (2005) o estudo da análise de agrupamentos, por sua vez, busca encontrar grupos de dados idênticos/semelhantes entre si, ou seja, um agrupamento nada mais é do que um conjunto de dados com características similares. Esses agrupamentos produzem um modelo para a estrutura dos dados e, se analisados adequadamente, podem revelar informações importantes (VALE, 2005).

A análise de agrupamentos pode ser definida como o processo de determinação de k grupos em um conjunto de dados. Para entender o que isso significa, veja a Figura 19.

Figura 19 - Ilustração de dados agrupados em quatro grupos.



Fonte: Vale (2005).

Observado a Figura 19, podemos observar quatro grupos definidos e agrupados separadamente, logo, “o processo automático de descobrimento desses grupos é o principal objetivo da análise de agrupamentos, onde basicamente se buscam grupos de objetos (dados) similares entre si” (VALE, 2005).

3.6.2 Algoritmos de agrupamentos

Após analisar os agrupamentos, o próximo processo passa a ser o agrupamento de dados e nesta etapa o objetivo é dividir determinado conjunto de dados em grupos com as mesmas características entre si. Para tal situação, os algoritmos de agrupamento de dados podem ser classificados em duas grandes classes métodos também conhecidos como: Métodos Hierárquicos (Algoritmos Aglomerativos e Divisivos) e Métodos Particionais (Algoritmos Exclusivos e Não Exclusivos) (VALE, 2005).

Os métodos hierárquicos de agrupamento admitir obter uma família de partições, cada uma associada aos níveis sucessores ou subsequentes de agrupamento entre as observações, calculadas com base nos dados disponíveis. As diferentes famílias de partições podem ser representadas graficamente através de uma estrutura semelhante a uma árvore chamada árvore de agrupamento hierárquico ou dendrograma (FIGINI; GIUDICI, 2009).

No método Algoritmos Aglomerativos, cada elemento dar início a representação de um grupo, e a cada etapa, um grupo/elemento é ligado a outro de acordo com sua similaridade, até o último passo, formando um grupo único com todos os elementos. Observa-se no início do processo, os agrupamentos são menores e os elementos de cada grupo possuem um alto grau de similaridade. No decorrer do processo, parte final, ficam poucos agrupamentos, cada um podendo conter diversos elementos e poucos similares entre si (VALE, 2005).

Enquanto, os métodos Divisivos são poucos entre os métodos hierárquicos por sua baixa ineficácia e precisa de uma capacidade computacional, robusta, maior dos métodos hierárquicos Aglomerativos. Os algoritmos deste método trabalham na direção oposta dos métodos Aglomerativos, consistir em um grupo inicial contendo todos os elementos, depois são divididos em dois subgrupos, de tal maneira que os elementos em um subgrupo estejam afastados dos elementos do outro subgrupo (VALE, 2005).

Os métodos particionais são métodos com embasamento na minimização de uma função de custo, onde os padrões são agrupados num número k de agrupamentos selecionados prioritariamente. Para cada padrão é agrupado na classe em que sua função de custo é minimizada. Dentre as principais vantagens dos métodos particionais em relação aos métodos

hierárquicos é a probabilidade de um padrão capaz de alterar o grupamento com a melhora do algoritmo de se operar com as bases de dados grande (VALE, 2005).

O algoritmo de ***K-means*** é o método mais utilizado de agrupamento não hierárquico, onde ***k*** aponta o número de grupos estabelecidos prioritariamente (***g*** nesta seção). O algoritmo ***k-means*** realiza um agrupamento dos elementos iniciais de ***n***, em grupos ***g*** distintos (com ***g*** previamente fixado), de acordo com o seguinte fluxo operacional: Inicialização, Avaliação de Transferência e Repetição, Figura 20 (FIGINI; GIUDICI, 2009).

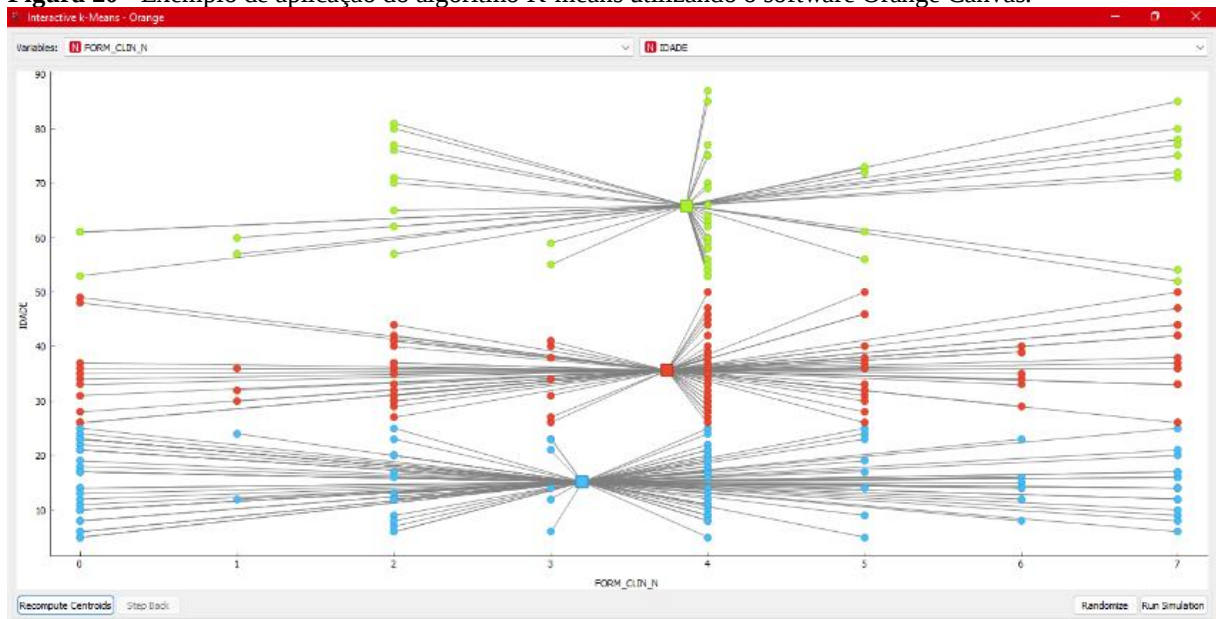
Inicialização – Tendo determinado o número de grupos, os pontos ***g***, chamados sementes (*seeds*), são definidos no espaço p -dimensional. Os *seeds* constituem os centroides (medidas de posição) dos aglomerados na partição inicial. Deve haver distância suficiente entre eles para melhorar as propriedades de convergência do algoritmo. Por exemplo, para o espaço dos centroides adequadamente em R^p , pode-se selecionar observações ***g*** (*seeds*) cuja distância recíproca é maior do que um limiar predefinido, e maior do que a distância entre eles e as observações. Uma vez definidas os *seeds*, uma partição inicial das observações é construída, alocando cada observação ao grupo cujo centroide está mais próximo.

Avaliação de transferência – Calcula-se a distância de cada observação dos centroides dos grupos ***g***. A distância entre uma observação e o centroide do grupo ao qual foi atribuído deve ser mínima; se não for mínimo, as observações serão movidas para o aglomerado cujo centroide é mais próximo. Os centroides do antigo grupo e do novo grupo são então recalculados.

Repetição – Repetimos o passo 2 até chegarmos a uma estabilização adequada dos grupos.

O algoritmo de aprendizado não supervisionado, ***k-means***, é um dos métodos de agrupamento de dados não hierárquico, que tenta minimizar a distância entre os dados de um *dataset* e um centro escolhido ou dado por um valor aleatório. O processo do algoritmo é considerado rápido, onde com poucas interações ele faz agrupamentos contínuos até encontrar o ponto ***k***, conhecido também como centroide, e, para cada *cluster* tem seu ponto central unindo elementos próximos a este ponto dando formato/característica desse *cluster* (FIGINI; GIUDICI, 2009; RICARDO LINDEN, 2009).

Figura 20 - Exemplo de aplicação do algoritmo K-means utilizando o software Orange Canvas.



Fonte: Martins, 2022.

4 METODOLOGIA

4.1 Tipo de estudo

Foi realizado um estudo epidemiológico, descritivo, analítico e retrospectivo a partir dos dados oficiais do Sistema de Informação de Agravos de Notificação (SINAN) e do Instituto Brasileiro de Geografia e Estatística (IBGE).

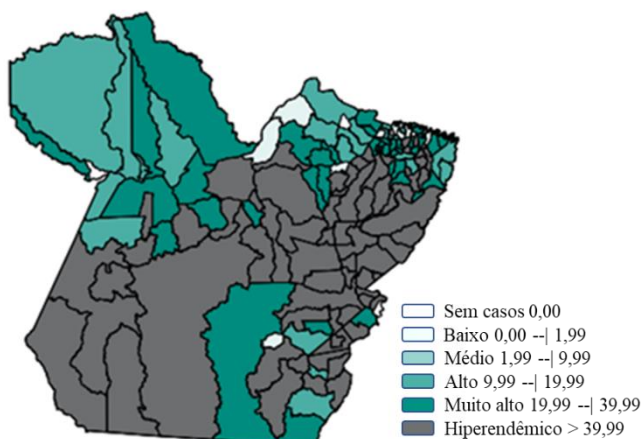
Para Rouquayrol e Gurgel (2018), a epidemiologia tem como objetivo global de analisar, a ocorrência e distribuição de eventos relacionados com a saúde da população (seres humanos), incluindo fatores determinantes (fenômenos e processos associados).

No estudo descritivo detém-se a descrever a ocorrência de uma patologia numa população, que determina o primeiro passo de uma investigação epidemiológica, enquanto o estudo analítico aponta, com mais detalhes, as relações entre o estado de saúde e as outras variáveis (BONITA; BEAGLEHOLE; KJELLSTROM, 2010, p. 39).

4.2 Cenário de Estudo

O município de Santarém no Pará, localizado na mesorregião do Baixo Amazonas, foi escolhido para o estudo deste trabalho como município-sede e devido sua importância estratégica na região, servindo de referência para outros municípios (18) quanto à saúde; ao mesmo tempo, o município em relação ao estado, tem o coeficiente muito alto para hanseníase (Figura 21), como de sua posição estratégica na região.

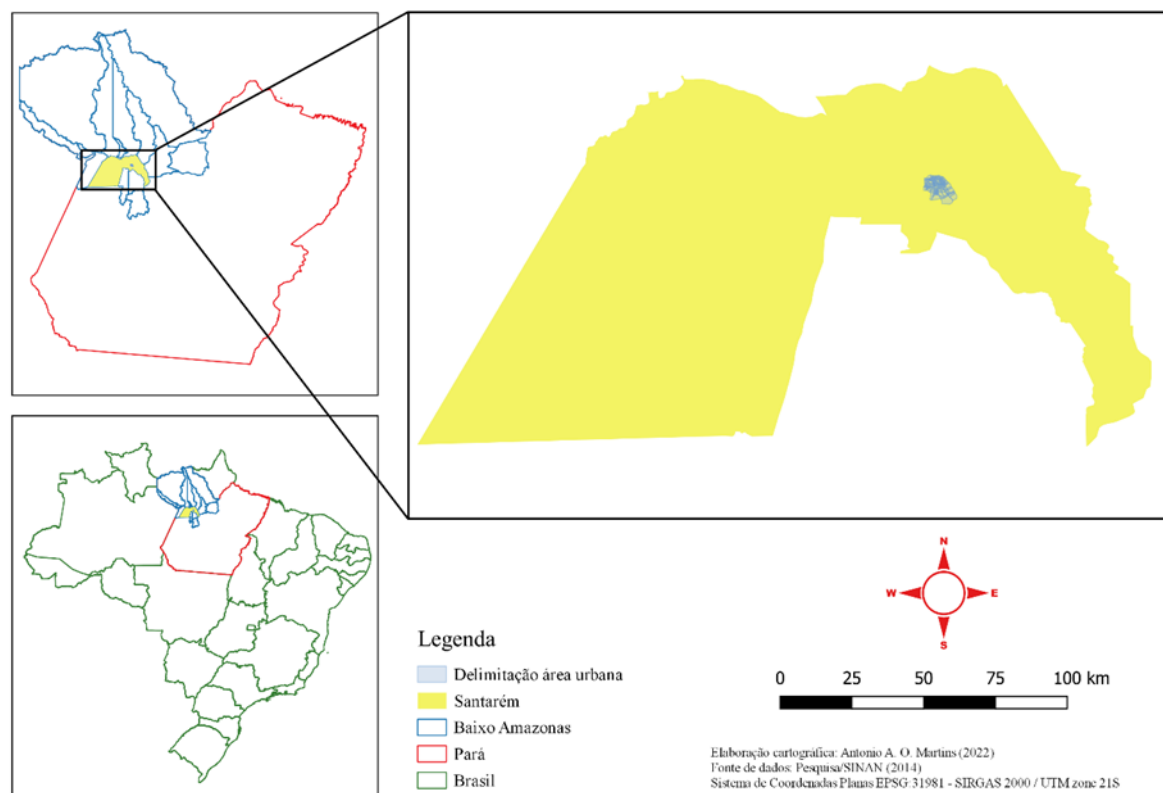
Figura 21 - Coeficiente geral de detecção de hanseníase por municípios. Pará, 2010.



Fonte: MM/SVS/SINAN (2010).

O município de Santarém, tem aproximadamente 294.580 habitantes e uma área de 22.886,624 km² (IBGE, 2010), esta cidade é o centro polarizador da mesorregião do Baixo Amazonas (Figura 22), devido a sua infraestrutura (escolas, universidades, hospitais, aeroporto, comércio, etc.).

Figura 22- Municípios de Santarém, área urbana, no Estado do Pará – Brasil.



Fonte: Martins (2021).

4.3 Aquisição de Dados

Este trabalho está inserido em uma cooperação técnico-científica entre a Universidade Federal do Oeste do Pará – UFOPA através do Programa de Pós-Graduação em Biociências com seu Laboratório de Suporte à Decisão (LSD), da Universidade Federal do Pará – UFPA com seu Laboratório de Dermato-Imunologia (LDI) e da Universidade do Estado do Pará – UEPA com o Laboratório de Semiologia e Estudos Epidemiológicos (LSE), com o objetivo de estudar marcadores biológicos de infecção e progressão da hanseníase: sorologia e resposta linfocitária celular de pacientes de uma região hiperendêmica da Amazônia Brasileira. Portanto, a proposta inicial desta cooperação é avaliar os indicadores epidemiológicos da Hanseníase em municípios endêmicos do Estado do Pará, visitando/coletando dados de alguns

municípios da região oeste do Pará como: Oriximiná (2009), Senador José Porfírio (2014), Santarém (2014), Breves (2013), Redenção (2013), Distrito de Mosqueiro (Belém/PA-2015), Belterra (2015).

Para realizar as semanas de busca ativa nas cidades supracitadas fez necessário criar uma equipe multiprofissional com médicos especialistas, fisioterapeutas, técnicos de laboratório, biomédicos e discentes do programa de iniciação científica do curso de ciências da computação da UFOPA.

Os dados coletados e gerados pela busca ativa no município de Santarém foram utilizados neste trabalho. O município de Santarém foi escolhido considerando sua infraestrutura e coincidir com o fato de o LSD estar localizado no Campus da UFOPA neste município.

A seleção dos casos índices foram retiradas da base do SINAN de maneira aleatória; sendo posteriormente visitados pela equipe multiprofissional para avaliação clínica dos casos índices e seus contatos domiciliares. Todos comunicantes, após avaliação clínica e diagnosticado com hanseníase, foram denominados de casos novos; embora mantenha seu histórico como comunicante.

Para a execução deste estudo, todas as pessoas envolvidas na pesquisa, assinaram o termo de consentimento livre e esclarecido – TCLE, permitindo assim a realização dos exames clínicos, laboratoriais com coleta de linfa dos lóbulos das orelhas para obter resultados da Baciloscopia e coleta de amostras sanguínea que foram encaminhados para o LDI-UFPA (Marituba/Belém) onde foram processadas amostras para detecção do anti-PGL-1 (quantitativa) dos casos índices e os seus comunicantes. Todos esses dados foram inseridos na base de dados do sistema HANSYS (NOGUEIRA et al., 2019)(software desenvolvido pelo grupo de trabalho) que também registrou dados pessoais, domiciliares, laboratoriais, neurológicos e diagnósticos dos pacientes mais comunicantes.

4.4 Escolha das Variáveis

O sistema de coleta de dados utilizado neste trabalho possui dados dos pacientes e comunicantes tais como: dados pessoais, domiciliares, laboratoriais, neurológicos, diagnóstico, etc., contudo, para esse modelo proposto utilizamos como base a forma clínica dos pacientes aliadas a algumas variáveis socioeconômicas (domiciliares) e o nível soropositividade do anti-PGL-1. Foram selecionadas variáveis que diretamente pudessem mapear uma relação

socioeconômica e clínica dos pacientes e seus comunicantes (SILVA et al., 2020, 2019) na facilitação da disseminação da hanseníase. Segue abaixo o detalhamento das variáveis:

- **Gênero:** informa o sexo do paciente ou comunicante;
- **Faixa Etária:** indica a faixa etária do paciente ou comunicante;
- **Grau de Parentesco:** indica o grau de parentesco dos comunicantes com o seu paciente índice;
- **Período de Convivência:** indica o tempo de convivência dos comunicantes com o seu paciente índice;
- **Vacina BCG:** indica o número de cicatrizes da vacina BCG do paciente ou comunicante;
- **Quantidade de Comunicantes:** representa a quantidade de comunicantes que um paciente possui;
- **Densidade de Dormitórios:** indica se há dormitórios em que dormem mais de 2 pessoas;
- **Quantidade de Cômodos:** apresenta informações sobre o tamanho da casa;
- **Renda Familiar:** indicador socioeconômico de renda da família do paciente;
- **Grau de Escolaridade:** indicador do nível de escolaridade do paciente;
- **Água para Consumo:** indica qual o tratamento aplicado na fonte de água consumível na residência do paciente;
- **Forma Clínica:** indica em qual estágio da doença o paciente foi diagnosticado e iniciou o tratamento: I, NP, T, BT, BB, BL e LL;
- **Classificação Operacional:** baseia-se em critérios clínicos para alocar os pacientes para fins terapêuticos, dividindo em PB (I, NP, T) e MB (BT, BB, BL e LL);
- **Grau de incapacidade física:** indica a existência de perda da sensibilidade protetora e força muscular nos membros superiores e inferiores do paciente, com graduação que varia entre 0, 1 e 2;
- **Anti-PGL-1:** resultado do exame laboratorial que indica o nível de soropositividade em relação à carga bacilar da hanseníase que o organismo do paciente ou comunicante possui;

4.5 Aspectos Éticos

Esta pesquisa está obedecendo todos os princípios éticos e preceitos da Resolução nº 466/12 do conselho Nacional de Saúde, que regulamenta as normas para a realização de pesquisa que envolve seres humanos e foi aprovado pelo comitê de ética em Pesquisa em Seres Humanos do Centro de Ciências Biológicas da Saúde da Universidade Estadual do Pará (CEP-CCBS/UEPA) sob o protocolo nº 1.971.621.

4.6 Técnicas de Agrupamento

Para realizar o agrupamento foi adotado o algoritmo *K-means*, executado por meio do software não proprietário e “*open source* – Orange Canvas”. Após a execução do algoritmo, foram analisados os *clusters* encontrados utilizando as informações dos elementos pertencentes a cada *cluster* e suas variáveis, extraídas em formas de tabelas.

4.7 Análise de Dados

Os dados foram tabulados e organizados primariamente através do Microsoft Excel© 2019 (software proprietário); análise estatística descritiva com IBM® SPSS® Statistics 2014 (software proprietário); criação de infográficos com GraphPad 9 Prism©2022 (software proprietário); para a aplicação de algoritmo não supervisionado de aprendizado por clusterização com Orange©2022 (software não proprietário e *open source*) e posteriormente anexados em Microsoft Word© 2019.

5 RESULTADOS E DISCUSSÃO

5.1 Panorama da Hanseníase no município de Santarém - Pará

A partir das informações obtidas do Sistema de Informação de Agravos de Notificação – SINAN, do acompanhamento dos dados de hanseníase no estado do Pará, foram obtidos a soma de 792 casos novos diagnosticados e notificados no período de 2003 a 2013 no município de Santarém. No último ano do período analisado, 2013, foram 47 casos novos, porém o coeficiente de incidência se manteve acima dos parâmetros de referência do Ministério da Saúde de 1/10.000 habitantes com o valor de 1,6, caracterizando assim, um município como de alta endemicidade para hanseníase. Entretanto, no ano da busca-ativa realizada no ano de 2014 no município em estudo foram notificados 67 casos novos, elevando o coeficiente de 1,6 do ano anterior para 2,3, deixando de ser alta para muito alta a endemicidade quanto a hanseníase.

Para realizar a semana de busca ativa por novos pacientes de hanseníase em Santarém, foram selecionados aleatoriamente 70 pacientes já tratados e registrados no SINAN (dos 792 casos) no período de 2003 a 2013. Através da visita domiciliar destes 70 pacientes denominados casos índices, foi possível avaliar 280 comunicantes destes e diagnosticar 35 casos novos de hanseníase, ver Tabela 5.

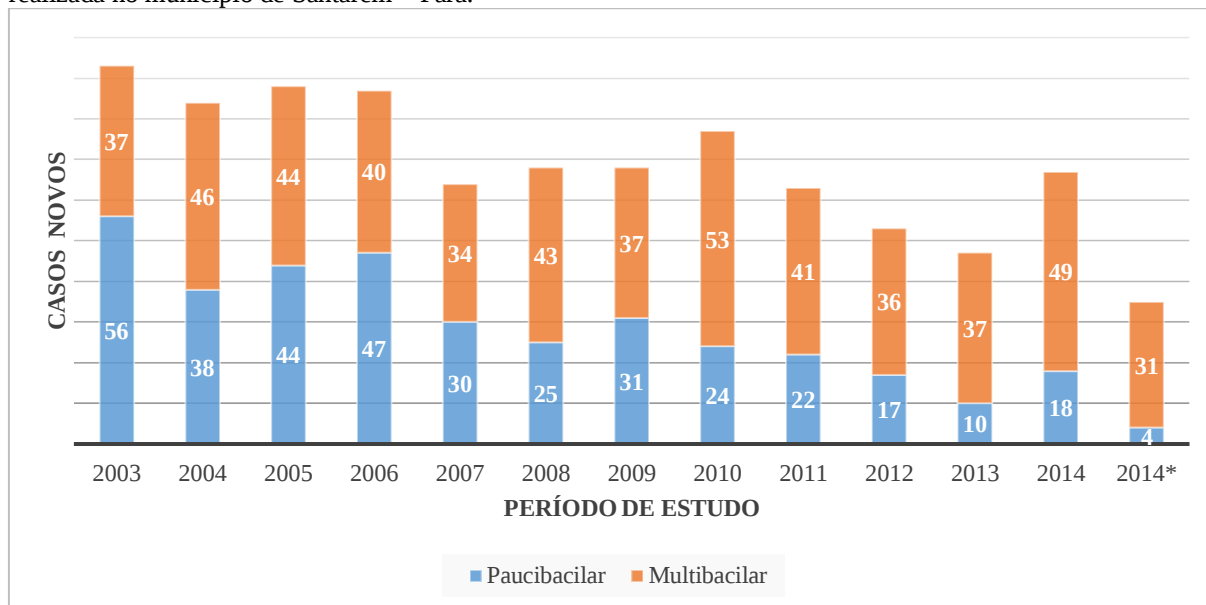
Na Figura 23, pode ser visto a série temporal dos casos de hanseníase registrados no SINAN entre os anos de 2003 a 2014, além dos casos novos diagnosticados na semana de busca ativa, analisada neste trabalho, que ocorreu em 2014.

Tabela 5. Caracterização epidemiológica dos grupos casos índices e comunicantes de acordo com Gênero e Faixa etária.

Variáveis	Casos índices (n = 70)		Comunicantes (n = 280)			
			Negativos (n = 245)		Positivos (n = 35)	
	n	%	n	%	n	%
SEXO						
Feminino	30	42,90%	152	62%	23	65,70%
Masculino	40	57,10%	93	38%	12	34,30%
Total	70	100%	245	100%	35	100%
FAIXA ETÁRIA						
0 - 14 anos	1	1,40%	52	21,20%	9	25,70%
15 - 39 anos	17	24,30%	126	51,40%	15	42,90%
40 -59 anos	25	35,70%	37	15,10%	7	20,00%
Acima de 60 anos	27	38,60%	30	12,20%	4	11,40%
Total	70	100%	245	100%	35	100%

Fonte: Base de dados.

Figura 23- Casos Novos de hanseníase registrados entre 2003 a 2014, dados SINAN; * Busca-ativa em 2014 realizada no município de Santarém – Pará.



Fonte: Brasil (2021) e Base de Dados.

No período de 2003 a 2013 que subsidia a pesquisa, segundo SINAN (2022), observa-se que entre os anos de 2011 a 2013 os casos de hanseníase vinham diminuindo, quando em 2014 a série apresenta uma elevação de casos semelhante aos níveis anteriores a 2010. Os resultados sugerem que não vinha ocorrendo um recuo da doença no município, mas uma diminuição na detecção de casos novos, haja vista que a elevação de casos foi muito corroborada pelos casos novos identificados na semana de busca ativa de 2014. Além disso, observa-se que a maioria dos casos diagnosticados tanto na série histórica como na semana de busca ativa estavam na forma clínica MB, o que caracteriza o estado de disseminação ativa da doença.

Ainda focando nos casos de hanseníase do período estudado, foram registrados 792 casos novos classificados operacionalmente em PB e MB, onde, de 2003 até o ano de 2009, a relação entre a classificação operacional pode ser observado certa equivalência entre os achados (Figura 23); em seguida, no período de 2010 a 2013, se percebe que os casos novos classificados como MB foram superiores ao PB em média de 2 a 3 vezes a mais. Os dados apresentados expressam que há ocorrência de diagnóstico tardio tendo em vista que o período de incubação é longo, ou seja, o indivíduo já vem transmitindo a doença há muitos anos, evidenciando que o decréscimo de casos da doença possa estar relacionado com subnotificação de casos e consequentemente a realidade possa não estar corretamente representada.

A análise epidemiológica dos casos índices e casos novos identificados na semana de busca ativa pode ser vista na Tabela 6, onde são mostrados os dados epidemiológicos

estratificados com as seguintes variáveis: sexo, faixa etária, classificação operacional, forma clínica, vacina BCG, grau de incapacidade física – GIF, densidade dormitório, quantidade de cômodos, renda familiar, grau de escolaridade e qualidade d'água para consumo humano.

Tabela 6. Características epidemiológicas dos pacientes índices e os casos novos detectados na semana de busca ativa no município de Santarém-Pa, 2014, de acordo com o gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica.

Variáveis	Categoria	Caso Índice (n = 70)		Caso Novo (n = 35)	
		n	%	n	%
SEXO	Feminino	30	42,9%	23	65,7%
	Masculino	40	57,1%	12	34,3%
	Total	70	100,0%	35	100,0%
FAIXA ETÁRIA	0 - 14 ANOS	1	1,4%	9	25,7%
	15 - 39 ANOS	17	24,3%	15	42,9%
	40 - 59 ANOS	25	35,7%	7	20,0%
	ACIMA DE 60 ANOS	27	38,6%	4	11,4%
	Total	70	100,0%	35	100,0%
FORMA CLÍNICA	Indeterminada (I)	12	17,1%	0	0,0%
	Neural Pura (NP)	0	0,0%	3	8,6%
	Tuberculóide (T)	12	17,1%	1	2,9%
	Dimorfa (D)	4	5,7%	0	0,0%
	Borderline Tuberculóide (BT)	22	31,4%	14	40,0%
	Borderline Borderline (BB)	4	5,7%	17	48,6%
	Borderline Virchowiana (BV)	4	5,7%	0	0,0%
	Virchowiana (V)	12	17,1%	0	0,0%
	Total	70	100,0%	35	100,0%
CLASSIFICACAO OPERACIONAL	Paucibacilar (PB)	24	34,3%	4	11,4%
	Multibacilar (MB)	46	65,7%	31	88,6%
	Total	70	100,0%	35	100,0%
VACINA BCG, CICATRIZ	0	21	30,0%	5	14,3%
	1	47	67,1%	28	80,0%
	2	1	1,4%	2	5,7%
	Duvidosa	1	1,4%	0	0,0%
	Total	70	100,0%	35	100,0%
GRAU INCAPACIDADE FÍSICA - GIF	0	52	74,3%	31	88,6%
	1,0	15	21,4%	3	8,6%
	2,0	3	4,3%	1	2,9%
	Total	70	100,0%	35	100,0%
NUMEROS DE COMODOS	1,0	5	7,1%	4	11,4%
	2,0	5	7,1%	1	2,9%
	3,0	9	12,9%	4	11,4%
	4,0	18	25,7%	12	34,3%
	5,0	14	20,0%	7	20,0%
	6,0	10	14,3%	5	14,3%
	7,0	5	7,1%	0	0,0%
	8,0	1	1,4%	1	2,9%
	9,0	1	1,4%	1	2,9%
	12,0	1	1,4%	0	0,0%
	14,0	1	1,4%	0	0,0%
	Total	70	100,0%	35	100,0%
DORMINTORIOS > 2 PESSOAS	Não	45	64,3%	17	48,6%
	Sim	25	35,7%	18	51,4%
	Total	70	100,0%	35	100,0%
RENDIA FAMILIAR	Sem renda	2	2,9%	0	0,0%

Variáveis	Categoria	Caso Índice (n = 70)		Caso Novo (n = 35)	
		n	%	n	%
	Menor que um salário-mínimo	3	4,3%	1	2,9%
	Um salário-mínimo	23	32,9%	17	48,6%
	Até dois salários-mínimos	30	42,9%	10	28,6%
	Até três salários-mínimos	7	10,0%	4	11,4%
	Maior que três salários-mínimos	5	7,1%	3	8,6%
	Total	70	100,0%	35	100,0%
GRAU DE ESCOLARIDADE	Nenhum	44	62,9%	18	51,4%
	Ensino Fundamental	11	15,7%	9	25,7%
	Ensino Médio	11	15,7%	7	20,0%
	Ensino Superior	4	5,7%	1	2,9%
	Total	70	100,0%	35	100,0%
ÁGUA PRA CONSUMO	Outro	1	1,4%	0	0,0%
	Poço artesiano	21	30,0%	7	20,0%
	Poço boca larga	1	1,4%	0	0,0%
	Rede de água encanada	47	67,1%	28	80,0%
	Total	70	100,0%	35	100,0%
TRATAMENTO D'ÁGUA	Coadá ou nenhum outro método	24	34,3%	16	45,7%
	Filtrada, fervida ou clorada	38	54,3%	17	48,6%
	Mineral	8	11,4%	2	5,7%
	Total	70	100,0%	35	100,0%

Fonte: Base de dados.

Em 2014, ano de realização da busca ativa, 67 casos novos foram diagnosticados e notificados pelo SINAN com 26,9% (18/67) PB e 73,1% (49/67) MB. Das 350 pessoas avaliadas neste estudo, casos índices (n=70) e comunicantes (n=280), 58,6% (205/350) são do sexo feminino para 41,4% (145/350) masculino (Tabela 5); onde, dos casos índices, 42,9% (30/70) feminino, para 57,1% (40/70) masculino. Dos casos novos diagnosticados, 8,1% (23/280) foram do sexo feminino e 4,3% (12/280) do sexo masculino (Tabela 6).

Na tabela 6, apresentam-se as características clínicas e epidemiológicas da doença no município a partir dos casos índices visitados e dos casos novos encontrados durante a avaliação desses contatos. Podem ser observadas diferenças nas avaliações epidemiológicas dos casos índices e dos casos novos encontrados. Com relação ao gênero, a maioria dos casos índices era do sexo masculino 57,1% (40/70) enquanto que dos casos novos a maioria era do sexo feminino 65,7% (23/35). Resultados similares foram observados noutros estudos com maior frequência para sexo feminino, 54,3% e 63,3% (BOIGNY et al., 2019; REIS et al., 2019); enquanto no estudo de Peixoto et al. (2011), os casos novos de contatos demonstrou-se inversamente, ou seja, 63,2% (24/38) do sexo masculino.

De acordo com a faixa etária, casos índices, foi observado maior ocorrência na faixa acima de 60 anos com 38,6% (27/70), sendo que os casos novos tiveram maior ocorrência na faixa de 15 - 39 anos, com 42,9% (15/35). A faixa etária de 15 a 59 anos compreende o período economicamente produtivo e/ou ativo do indivíduo sadio. Do mesmo modo, observado em trabalhos correlacionados que a maior frequência entre as faixas etárias concatena entre os 30

e 49 anos. Aliado ao alto potencial incapacitante da patologia, a falta de diagnóstico ou o diagnóstico tardio, pode interferir no trabalho e na vida social do paciente, levando a perdas e traumas psicológicos (BOIGNY et al., 2019; DUARTE; AYRES; SIMONETTI, 2007; REIS et al., 2019).

Quando se analisa um indicador importante para a hanseníase como os casos da doença em menores de 15 anos, observa-se que na faixa de 0 – 14 anos, os casos índices apresentaram 1,4% (1/70) contra 25,7% (9/35) dos casos novos. No estudo de Barreto et al. (2012), considera que menores de 15 anos (crianças) devem ser avaliados ativamente para detecção da hanseníase precocemente em suas formas iniciais, quebrando a cadeia de transmissão e posteriormente impedindo reações traumáticas decorrida pela infecção tardia, bem como os preconceitos sociais, problemas psicológicos e disseminação da patologia. O estudo de Pierneef et al. (2021) demonstra que em menores de 15 anos a transmissão da hanseníase possui uma forte associação com a fonte de transmissão de origem familiar. Ressalta-se ainda, que na infância, devido à maior dificuldade no diagnóstico precoce, existe a possibilidade dos indivíduos evoluírem gradativamente com maiores complicações e deformidades, tornando-se difícil a resolução do problema (PIRES et al., 2012).

Para Ridley e Joplin (1966), utiliza a forma clínica como variável para indicar qual estágio da doença o paciente foi diagnosticado ou teve início ao tratamento. Ao se analisar a forma clínica observa-se maior predominância da forma BT com 31,4% (22/70) para os casos índices e a forma BB com 48,6% (17/35) para os casos novos. Em contrapartida, 17,1% (12/70) para fases polares da hanseníase, forma clínica indeterminada e Virchowiana (polo hansênico disseminador da doença) dos casos índices para nenhum indivíduo, respectivamente, dos casos novos. Estes dados mostram que o diagnóstico da hanseníase na forma clínica indeterminada, 17,1% (12/70), está sendo diagnosticado tardiamente, favorecendo a transmissão ativa da patologia no município, o que demonstra um indicador alarmante (OMS, 2021).

É interessante notar que para a classificação operacional, ambos os tipos de casos avaliados a maioria foi classificada como MB, casos índices 65,7% (46/70) e casos novos 88,6% (31/35), um achado que não se difere entre outros estudos (BASSO; ANDRADE; SILVA, 2021; RODRIGUES et al., 2010). Os dados quanto a classificação operacional mostra uma disparidade entre a quantidade de casos PB e MB, o que reflete que existe uma situação frequente de diagnóstico tardio gerando consequentemente uma subnotificação de casos de hanseníase, ou seja, a quantidade de casos diagnosticados está majoritariamente na fase de disseminação da doença. Esses dados sugerem possíveis falhas na vigilância de contatos, no acompanhamento e controle da doença através de diagnóstico tardio (BRASIL, 2020; 2022).

Quando se avalia os casos pela quantidade de doses da vacina BCG, ambos os tipos de casos avaliados apresentaram maior ocorrência de pelo menos uma dose, 67,1% (47/70) para os casos índices e 80,0% (28/35) para os casos novos. O trabalho de Setia et al. (2006), discutiu o uso da vacina BCG, inicialmente desenvolvida para conferir proteção contra a tuberculose, na proteção contra a hanseníase, sendo que a magnitude desse efeito poderia variar. Uma dose adicional de BCG pode ser necessária para contatos de pacientes com hanseníase em áreas onde a hanseníase continua a ser um problema de saúde pública.

Quanto ao grau de incapacidade física no diagnóstico (GIF), a maioria em ambos os tipos de casos analisados foi GIF 0, com 74,3% (52/70) nos casos índices e 88,6% (31/35) nos casos novos. É interessante notar que entre os casos novos foram registrados 8,6% (3/35) pacientes com GIF 1 (quando há diminuição da força muscular e/ou diminuição de sensibilidade) e 2,9% (1/35) paciente com GIF 2 (pacientes apresentaram alguma deformidade visível). A presença do indicador GIF 2, condiz com um tempo prolongado de adoecimento, de maior possibilidade de transmissão e de potencializar o risco de contaminação entre os contatos intradomiciliares de pacientes MB. A melhor maneira de prevenir incapacidades mais graves (GIF 1 e 2) é o diagnóstico precoce (BRASIL, 2018; CHAVES et al., 2017; LASTÓRIA; ABREU, 2012).

Para a variável dormitório com mais de 2 pessoas, os casos índices apontaram 35,7% (25/70), enquanto os casos novos com 51,4% (18/35). Quando se analisa outro indicador socioeconômico, como o fator renda, observa-se que a maioria dos casos novos tinham renda per capita de até 1 salário mínimo por domicílio, 48,6% (17/35). O estudo de Monteiro et al. (2015) realizado no estado de Tocantins, seus resultados foram proporcionais para domicílios com duas ou mais pessoas por dormitórios mais a renda per capita de meio salário mínimo. É muito provável que em razão da quantidade de pessoas que dividem o mesmo dormitório, independentemente do tipo de relacionamento interpessoal, estes apresentem maior suscetibilidade para a contaminação da hanseníase, devido ao tempo de exposição em ambiente fechado com alguém infectado (SOUZA, 2012).

Outro fator socioeconômico para a hanseníase é o grau de escolaridade, onde a maior ocorrência foi nenhum grau (nenhum tipo de formação escolar) para 62,9% (44/70) dos casos índices e 51,4% (18/35) casos novos. Na mesma perspectiva Chaves et al. (2017), em seus estudos concentram-se majoritariamente em populações esquecidas e baixo nível de escolaridade. Da mesma forma, no estudo de Barreto et al. (2012) na região amazônica traz uma amostra de 37% de seus estudos com pacientes analfabetos ou analfabetos funcionais. No estudo de Leano et al. (2019), de revisão integrativa da literatura sobre fatores socioeconômicos

relacionados a hanseníase, relata que o analfabetismo e poucos anos de estudo foram apontados como fatores de risco por dificultar o reconhecimento das manifestações clínicas da doença, o acesso ao sistema de saúde e a compreensão das orientações de educação em saúde.

Nosso estudo mostra que os casos índices, tem cobertura de fornecimento de água pela rede pública encanada com 67,1% (47/70) e os casos novos de 80% (28/35) com maior frequência em ambos da utilização do tratamento de filtragem, fervura ou cloração de água, 54,3% (38/70) e 48,6% (17/35) respectivamente. A qualidade da água para consumo humano é um indicador importante na prevenção da propagação de DTNs, incluindo hanseníase (EMERSON et al., 2020).

A região norte do Brasil citado nos estudos de Chaves et al. (2017) e Pescarini et al.(2021), descrevem um cenário de alto para muito alto de endemicidade, corroborando nossos achados que demonstram o problema da subnotificação de casos.

5.2 Perfil epidemiológico dos comunicantes

Além da análise epidemiológica dos casos índices e casos novos, também foi realizada uma análise dos 280 comunicantes avaliados na semana de busca ativa. Nesta análise foram avaliadas as seguintes variáveis: sexo, faixa etária, grau de parentesco, período de convivência, tempo de convívio, densidade dormitório, renda familiar, grau de escolaridade, qualidade d'água para consumo humano, vacina BCG, forma clínica – índice e classificação operacional – índice, ver Tabela 7.

Tabela 7. Características epidemiológicas dos comunicantes (280) dos pacientes índices de hanseníase diagnosticados no município de Santarém-Pa, 2014 de acordo com o gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices).

Variáveis	Categoria	Comunicantes (n = 280)	
		n	%
SEXO	Feminino	175	62,5%
	Masculino	105	37,5%
	Total	280	100,0%
FAIXA ETÁRIA	0 - 14 ANOS	61	21,8%
	15 - 39 ANOS	141	50,4%
	40 - 59 ANOS	44	15,7%
	ACIMA 60 ANOS	34	12,1%
	Total	280	100,0%
GRAU DE PARENTESCO	Cônjuge	26	9,3%
	Filho(a)	71	25,4%
	Irmão	25	8,9%
	Não parente	19	6,8%
	Pai/Mãe	29	10,4%
	Parente consanguíneo	45	16,1%

Variáveis	Categoria	Comunicantes (n = 280)	
		n	%
	Parente não consanguíneo	65	23,2%
	Total	280	100,0%
PERIODO DE CONVIVENCIA	> 1 ano	24	8,6%
	> 10 anos	191	68,2%
	1 - 2 anos	10	3,6%
	2 - 5 anos	15	5,4%
	5 - 10 anos	40	14,3%
	Total	280	100,0%
TIPO DE CONVIVIO	Extradomiciliar	45	16,1%
	Intradomiciliar	235	83,9%
	Total	280	100,0%
DORMITORIOS > 2 PESSOAS	Não	131	46,8%
	Sim	149	53,2%
	Total	280	100,0%
RENDIA FAMILIAR	Sem renda	6	2,1%
	Menor que um salário-mínimo	20	7,1%
	Um salário-mínimo	92	32,9%
	Até dois salários-mínimos	107	38,2%
	Até três salários-mínimos	27	9,6%
	Maior que três salários-mínimos	28	10,0%
	Total	280	100,0%
ÁGUA PRA CONSUMO	Poço artesiano	90	32,1%
	Poço boca larga	5	1,8%
	Rede de água encanada	185	66,1%
	Total	280	100,0%
TRATAMENTO D'ÁGUA	Coadá ou nenhum outro método	116	41,4%
	Filtrada, fervida ou clorada	143	51,1%
	Mineral	21	7,5%
	Total	280	100,0%
VACINA BCG, CICATRIZ	0	37	13,2%
	1	216	77,1%
	2	26	9,3%
	Duvidosa	1	,4%
	Total	280	100,0%
FORMA CLINICA - INDICE	Indeterminada (I)	45	16,1%
	Neural Pura (NP)	7	2,5%
	Tuberculóide (T)	43	15,4%
	Dimorfa (D)	15	5,4%
	Borderline Tuberculóide (BT)	90	32,1%
	Borderline Bordeline (BB)	27	9,6%
	Bordeline Virchowiana (BV)	15	5,4%
	Virchowiana (V)	38	13,6%
	Total	280	100,0%
CLASSIFICACAO OPERACIONAL - INDICE	Multibacilar (MB)	185	66,1%
	Paucibacilar (PB)	95	33,9%
	Total	280	100,0%

Fonte: Base de dados.

Durante a semana de busca ativa foram avaliados 280 comunicantes dos 70 pacientes visitados, encontrando-se 35 casos novos, 12,5% do total de comunicantes avaliados (35/280). Quando comparamos o número de casos novos da semana de busca ativa com o total de casos novos registrados no SINAN para todo o ano de 2014 verificamos que ele representa

mais da metade, 52,23% (35/67), de todos os casos registrados em 2014. A avaliação dos comunicantes é o método de detecção de casos com melhor base de evidências e que demonstra grande eficiência no diagnóstico de novos casos de hanseníase (BRASIL. 2021a).

A análise epidemiológica da Tabela 7, evidencia algumas características das pessoas que convivem com os casos índices do SINAN e que deveriam estar sendo acompanhadas pela vigilância de contatos para a identificação de novos casos de hanseníase tão logo eles aparecessem. Na Tabela 7, observa-se que a maior frequência quanto ao gênero foi o feminino com 62,5% (175/280). Observando os indicadores sociais, foram encontradas maiores frequências: para a faixa etária entre 15- 59 anos, concatenação de duas faixas representando os economicamente ativos, num total de 66,1% (185/280); no grau de parentesco 53,92% (151/280) foram contatos que normalmente possuem uma interação mais próxima e direta (cônjuge, filho, filha, irmão, irmã e pai/mãe). Resultados similares foram encontrados no estudo realizado sobre aspectos epidemiológicos dos contatos de hanseníase em São Luís – MA por Peixoto et al. (2011).

Quanto analisado o tipo e tempo de convívio entre comunicantes e os casos índices, observou-se que 83,9% (235/280) são intradomiciliares e 68,2% (191/280) tem convivência acima de 10 anos. Ainda sobre a convivência dentro das residências, 53,2% (149/280) dividem o mesmo dormitório com duas ou mais pessoas, o que representa maior risco de acometimento pela patologia (adoecimento) dado maior tempo de exposição em local fechado do que a população em geral (BRASIL. 2005; PEIXOTO et al., 2011).

A classificação operacional dos casos índices, ou seja, dos casos visitados os quais foram o motivo da visita naquela residência, mostrou que os comunicantes estiveram expostos em sua maioria a 66,1% (185/280) com casos MB. No contexto dos comunicantes dos casos índices, a cicatriz da vacina BCG é um indicador importante e estava presente com uma única cicatriz em 77,1% (216/280) dos comunicantes. Embora esta vacina não seja específica para a hanseníase, ela demonstra um efeito protetor contra a doença, reduzindo a morbidade e possibilitando manifestações clínicas mais brandas em caso de doença (BRASIL, 2019). o Ministério da Saúde recomenda a aplicação de uma segunda dose da BCG nos contatos prolongados de portadores da doença, que não apresentem sinais e sintomas no momento da investigação, independente da classificação operacional do caso índice. Sua aplicação depende da história vacinal e/ou da presença de cicatriz vacinal (BRASIL, 2019).

Estudos apontam que a principal estratégia para redução da carga em hanseníase e alcance do controle da doença enquanto problema de saúde pública em nível nacional se dá necessariamente pela vigilância e abordagem qualificada de contatos de casos de hanseníase,

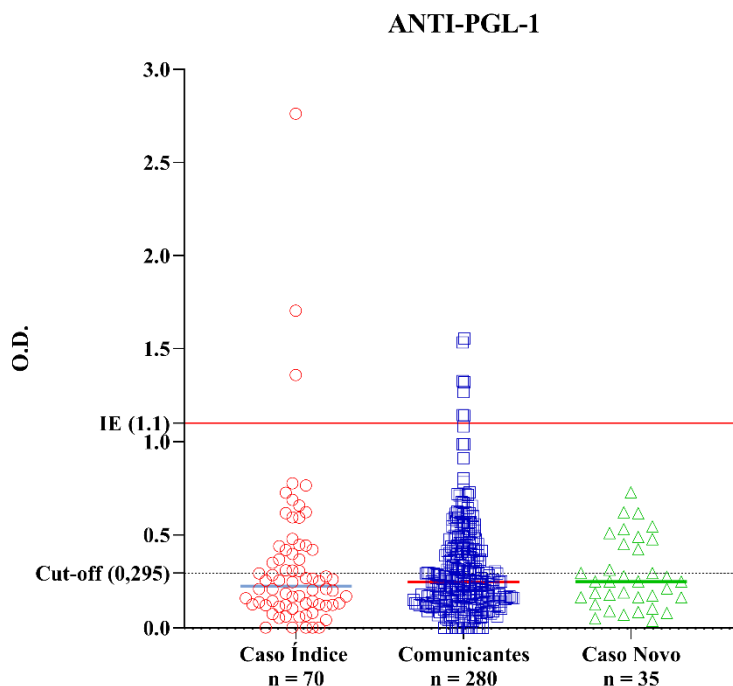
haja vista que esses possuem maior risco de infecção por hanseníase, especialmente em domicílios com casos MB existentes e contatos mais antigos, maior período de convivência (TEIXEIRA et al., 2020). Entretanto, a vigilância de contatos tem sido pouco valorizada e até mesmo negligenciada nos serviços de saúde pública que desenvolvem as ações do Programa Nacional de Controle da Hanseníase, pois privilegia o espaço para controle da doença e doente, reelegendo a um plano secundário as ações de controle dos contatos (SANTOS et al., 2019).

5.3 Soropositividade do anti-PGL-1 dos casos hansênicos e comunicantes

Para a análise do anti-PGL-1, todos os casos índices e seus comunicantes realizaram a coleta sanguínea em suas residências durante a semana de busca ativa. O teste utilizado para a dosagem do PGL-1 foi o teste ELISA, teste sorológico baseado na reações de antígeno e anticorpo através de reações enzimáticas (SILVA GAMA et al., 2020). Utilizamos o *cut-off* do estudo de Barreto et al. (2012), onde o valor do *cut-off* estabelecido do anti-PGL-1 foi de 0,295 (leitura por densidade ótica - *optical density* - O.D.). Considerando todos os indivíduos comunicantes testados para anti-PGL-1 ($n = 280$), 39,6% foram soropositivos (valores acima do *cut-off*) nos comunicantes (111/280) com uma O.D. mediana de 0,546, enquanto 37,1% foram soropositivos nos casos novos (13/35) com um O.D. mediana de 0,484, ver Figura 24 e Tabela 8.

Entre os 111 comunicantes soropositivos, apenas 11,7% (13/111) destes foram diagnosticados como caso novo. Os outros casos novos (22/35) não apresentaram soropositividade para o anti-PGL-1, assim como os outros 98 comunicantes soropositivos não foram diagnosticados como casos novos pela avaliação clínica do dermatologista. Ainda entre os 111 comunicantes soropositivos, 6,30% (7/111) apresentaram leitura O.D. $\geq 1,1$, valor limítrofe para definir a soropositividade independente do *cut-off* utilizado (SPENCER et al., 2012), porém, foram descartados como caso novo após avaliação clínica. O risco de adoecimento de contatos perpassa pela vulnerabilidade social, pela susceptibilidade individual e por dificuldades no acesso aos serviços de saúde. Apesar da influência das respostas imunológicas na suscetibilidade à doença, fatores de risco modificáveis parecem ser os principais determinantes do adoecimento por hanseníase entre contatos (FABRI, 2015; LOBATO et al., 2006; NIITSUMA et al., 2021; VIDAL et al., 2019).

Figura 24- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) dos casos índices, comunicantes e casos novos.



Fonte: Base de dados.

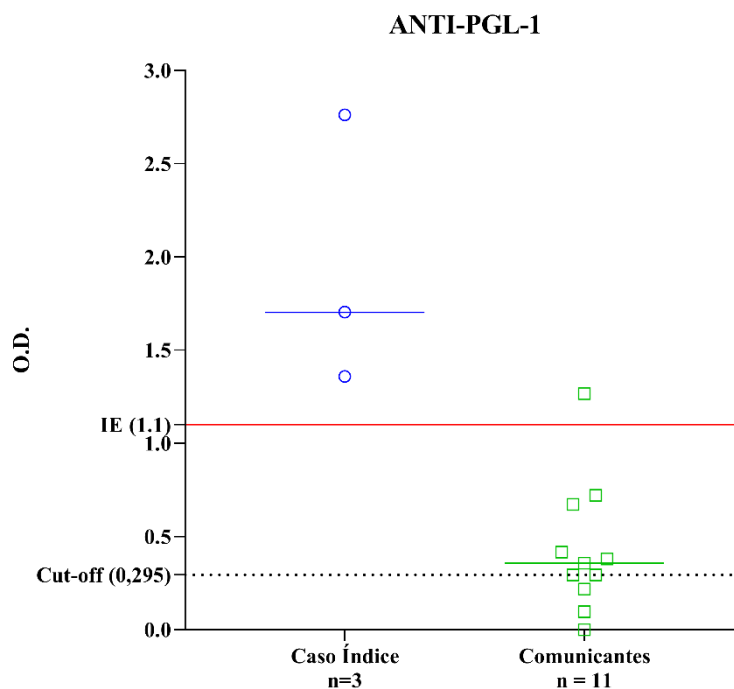
Tabela 8. Dosagens dos anticorpos anti-PGL-I, *cut-off* de 0,295) dos comunicantes (n = 280).

	Caso Índice		Comunicantes					
	Total		Total		Comunicantes		Caso Novo	
	n = 70	%	n = 280	%	n = 245	%	n = 35	%
anti-PGL-I < 0,295 (soronegativo)	45	64,3	169	60,4	147	60,0	22	62,9
anti-PGL-I ≥ 0,295 (soropositivo)	25	35,7	111	39,6	98	40,0	13	37,1
0.3 --- 0.5	13	18,6	63	22,5	56	22,9	7	20,0
0.5 --- 1.0	9	12,9	40	14,3	34	13,9	6	17,1
1.0 --- 1.5	1	1,4	6	2,1	6	2,4	0	0,0
1.5 --- 2.0	1	1,4	2	0,7	2	0,8	0	0,0
> 2.0	1	1,4	0	0	0	0	0	0

Fonte: Base de dados.

Um achado adicional ocorreu entre os casos índices, onde 4,2% (3/70) foram classificados como soropositivos com leitura (O.D.) $\geq 1,1$ (*limítrofe*), ver Tabela 8. Esses casos índices tinham sido classificados como MB no momento do seu diagnóstico e notificação no SINAN, sendo que 66,6% (2/3) desses eram da forma clínica Virchowiana. Ao se avaliar os comunicantes desses 3 casos índices, obtivemos a identificação de 3,9 % (11/280) do total de comunicantes com as leituras O.D. observadas na Figura 25, sendo que nenhum desses comunicantes apresentaram sinais sugestivos de hanseníase pela avaliação clínica.

Figura 25- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) $\geq 1,1$ dos casos índices e comunicantes.

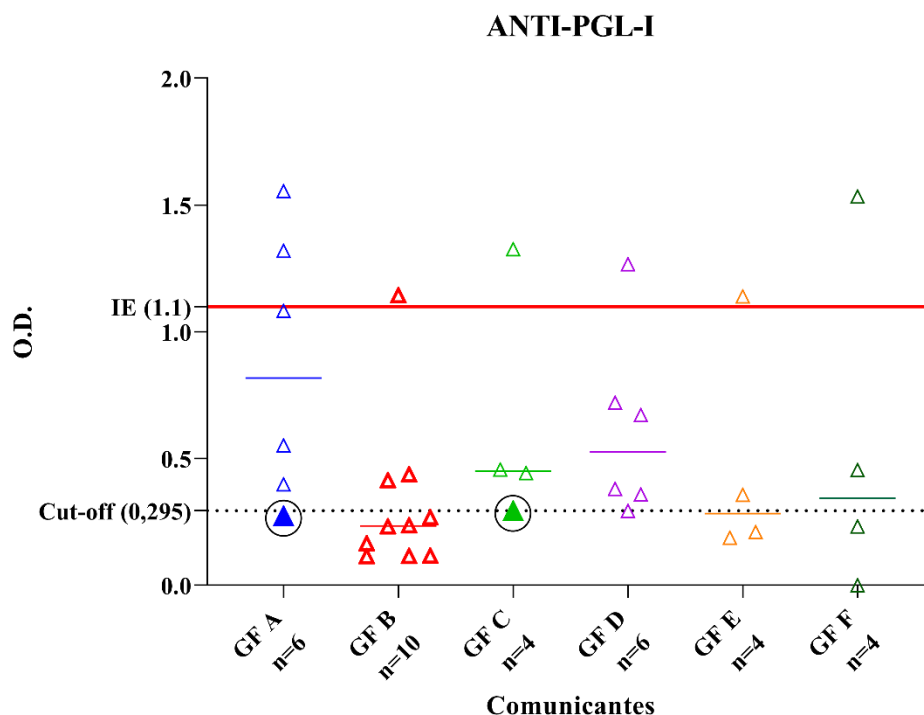


Fonte: Base de dados.

Os casos de hanseníase tratados regularmente com esquemas oficiais padronizados e corretamente indicados e que receberam alta por cura, isto é, saíram do registro ativo da doença no SINAN, que voltam a apresentar novos sinais e sintomas clínicos de doença infecciosa ativa são denominados de recidiva. As chamadas recidivas ocorrem quando pacientes que concluíram o tratamento da hanseníase apresentam novas lesões ou agravamento de lesões antigas, com aumento da quantidade de bactérias no organismo. “A volta das lesões pode ocorrer por bactérias latentes, que ficaram no organismo (SPENCER et al., 2012). Os 3 casos índices que apresentaram um valor muito alto para o anti-PGL-1, possuem um risco de reinfecção eminente, haja vista a alta carga bacilar desses. Deste modo, é recomendado que estes ex-pacientes sejam monitorados frequentemente pela vigilância de contatos.

Uma análise mais criteriosa sobre os 7 comunicantes que apresentaram leitura (O.D.) $\geq 1,1$, identificou que estes pertenciam a seis grupos familiares que juntos representavam 12,1% (34/280) dos comunicantes; sendo que 5,9% (2/34) dos integrantes desses grupos foram diagnosticados como caso novo e classificados como MB (Figura 26). Desses 34 comunicantes, 64,70% (22/34) foram classificados como soropositivos e apenas 1 deles foi diagnosticado como caso novo. Os outros 35,30% (12/34) foram classificados como soronegativos e um deles foi diagnosticado como caso novo. Encontramos nestas famílias um risco muito elevado da maioria destes vir a tornar-se um caso novo devido à alta carga bacilar desses.

Figura 26- Comparativo do ANTI-PGL-1 por densidade óptica (O.D.) $\geq 1,1$ dos comunicantes.



Legenda: IE – índice ELISA; GF – Grupo Familiar; Pontos circulados representam os 5,9% (2/34) que se tornaram casos novos dentro do grupo familiar analisado.

Fonte: Base de dados.

Estes dados evidenciam uma possível infecção-subclínica e a necessidade permanente do acompanhamento dos comunicantes com elevada carga bacilar, principalmente os que tem casos multibacilares diagnosticados na família.

5.4 Aplicação de técnicas de clusterização utilizando K-means

Com os dados epidemiológicos dos comunicantes ($n = 280$) encontrados na semana de busca ativa foram realizadas as simulações de clusterização, algoritmo de aprendizado não supervisionado – **K-means**, utilizando o software não proprietário e de código aberto – **Orange Canvas**. Foram realizadas simulações de formação dos *clusters* com conjuntos de variáveis representando dados socioeconômicos e clínicos dos casos índices e seus comunicantes. A melhor composição de variáveis para a formação dos *clusters*, ou seja, as variáveis que participaram ativamente no processo de definição dos *clusters* foram as seguintes: faixa etária, gênero, grau de parentesco (com o caso índice), período de convivência (com o caso índice), renda familiar, tempo de residência (na moradia da família), número de cômodos (da residência), número total de moradores (da residência), existência de dormitório com mais de 2 pessoas, abastecimento de água, forma clínica do caso índice, classificação operacional do

caso índice. A base de dados foi composta com todas essas variáveis definidas para os 280 comunicantes, assim formamos uma base de dados para a simulação com 280 linhas e 12 colunas (variáveis). Após finalizar o processo de obtenção dos *clusters*, foi gerado um estudo epidemiológico dos 3 *clusters* (C1, C2 e C3) encontrados para se avaliar a composição destes com suas respectivas distribuições dos comunicantes, como pode ser visualizado na Tabela 9.

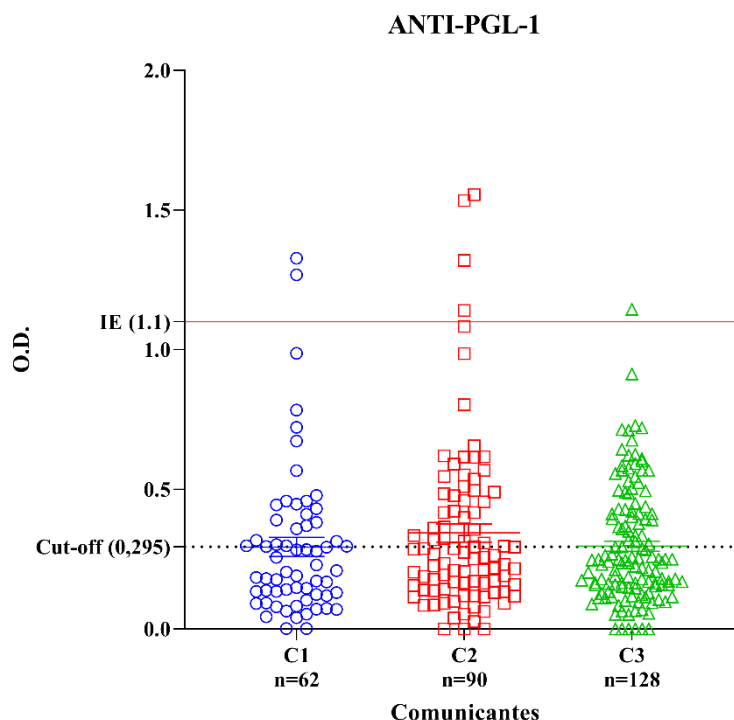
Tabela 9. Características epidemiológicas dos comunicantes (280) no município de Santarém-Pa de acordo com os 3 *clusters* C1, C2, e C3 encontrados e as variáveis: gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices).

		C1		C2		C3	
		n	%	n	%	n	%
SEXO	Feminino	44	71,0%	55	61,1%	76	59,4%
	Masculino	18	29,0%	35	38,9%	52	40,6%
	Total	62	100,0%	90	100,0%	128	100,0%
FAIXA ETARIA	0 - 14 ANOS	13	21,0%	25	27,8%	23	18,0%
	15 - 39 ANOS	31	50,0%	48	53,3%	62	48,4%
	40 - 59 ANOS	12	19,4%	9	10,0%	23	18,0%
	ACIMA 60 ANOS	6	9,7%	8	8,9%	20	15,6%
	Total	62	100,0%	90	100,0%	128	100,0%
GRAU DE PARENTESCO	Cônjuge	7	11,3%	9	10,0%	10	7,8%
	Filho(a)	19	30,6%	28	31,1%	24	18,8%
	Irmão	7	11,3%	9	10,0%	9	7,0%
	Não parente	0	0,0%	2	2,2%	17	13,3%
	Pai/Mãe	10	16,1%	8	8,9%	11	8,6%
	Parente consanguíneo	9	14,5%	12	13,3%	24	18,8%
	Parente não consanguíneo	10	16,1%	22	24,4%	33	25,8%
	Total	62	100,0%	90	100,0%	128	100,0%
PERIODO DE CONVIVENCIA	Entre cinco a dez anos	8	12,9%	13	14,4%	19	14,8%
	Entre dois a cinco anos	3	4,8%	6	6,7%	6	4,7%
	Entre um a dois anos	2	3,2%	3	3,3%	5	3,9%
	Mais que dez anos	46	74,2%	61	67,8%	84	65,6%
	Mais que um ano	3	4,8%	7	7,8%	14	10,9%
	Total	62	100,0%	90	100,0%	128	100,0%
VACINA BCG, CICATRIZ	0	5	8,1%	7	7,8%	25	19,5%
	1	50	80,6%	72	80,0%	94	73,4%
	2	6	9,7%	11	12,2%	9	7,0%
	Duvidosa	1	1,6%	0	0,0%	0	0,0%
	Total	62	100,0%	90	100,0%	128	100,0%
DORMITÓRIOS > 2 PESSOAS	Não	38	61,3%	39	43,3%	54	42,2%
	Sim	24	38,7%	51	56,7%	74	57,8%
	Total	62	100,0%	90	100,0%	128	100,0%
RENDIA FAMILIAR	Sem renda	0	0,0%	4	4,4%	2	1,6%
	Menor que um salário-mínimo	0	0,0%	8	8,9%	12	9,4%
	Um salário-mínimo	4	6,5%	26	28,9%	62	48,4%
	Até dois salários-mínimos	36	58,1%	34	37,8%	37	28,9%
	Até três salários-mínimos	9	14,5%	12	13,3%	6	4,7%
	Maior que três salários-mínimos	13	21,0%	6	6,7%	9	7,0%
	Total	62	100,0%	90	100,0%	128	100,0%
ÁGUA PRA CONSUMO	Poço artesiano	13	21,0%	27	30,0%	50	39,1%
	Poço boca larga	2	3,2%	0	0,0%	3	2,3%
	Rede de água encanada	47	75,8%	63	70,0%	75	58,6%
	Total	62	100,0%	90	100,0%	128	100,0%
TRATAMENTO DA ÁGUA	Coadada ou nenhum outro método	17	27,4%	46	51,1%	53	41,4%
	Filtrada, fervida ou clorada	39	62,9%	34	37,8%	70	54,7%
	Mineral	6	9,7%	10	11,1%	5	3,9%

		C1		C2		C3	
		n	%	n	%	n	%
	Total	62	100,0%	90	100,0%	128	100,0%
CLASSIFICACAO OPERACIONAL – INDICE	Paucibacilar (PB)	5	8,1%	90	100,0%	0	0,0%
	Multibacilar (MB)	57	91,9%	0	0,0%	128	100,0%
	Total	62	100,0%	90	100,0%	128	100,0%
FORMA CLÍNICA – INDICE	Indeterminada (I)	1	1,6%	44	48,9%	0	0,0%
	Neural Pura (NP)	0	0,0%	7	7,8%	0	0,0%
	Tuberculóide (T)	4	6,5%	39	43,3%	0	0,0%
	Dimorfa (D)	7	11,3%	0	0,0%	8	6,3%
	Borderline Tuberculóide (BT)	19	30,6%	0	0,0%	71	55,5%
	Borderline Borderline (BB)	7	11,3%	0	0,0%	20	15,6%
	Borderline Virchowiana (BV)	10	16,1%	0	0,0%	5	3,9%
	Virchowiana (V)	14	22,6%	0	0,0%	24	18,8%
	Total	62	100,0%	90	100,0%	128	100,0%

Na Tabela 9 pode ser observada a distribuição dos comunicantes entres os 3 *clusters*: 22,15% (62/280) no C1, 32,15% (90/280) no C2 e 45,7% (128/280) no C3. A análise epidemiológica da Tabela 9 apresentou algumas características podem auxiliar a entender as diferenças dos *clusters*, como por exemplo o grau de parentesco, onde a maioria dos componentes dos *clusters* C1 e C2 foram classificados como filho(a), 30,6% (19/62) e 31,1% (28/90), respectivamente, contra a classificação de parente não consanguíneo no C3, 25,1% (33/128). A variável dormitório com mais de 2 pessoas também sinalizou uma diferença entre os *clusters*, com o C1 tendo a maioria para não com 61,3% (38/62) e os *clusters* C2 e C3 tendo a maioria para sim com 56,7% (51/90) e 57,8% (74/128), respectivamente. A variável renda familiar apontou uma condição mais favorável, na maioria dos casos, para os *clusters* C1 e C2, apontando uma renda de até dois salários mínimos com 58,1% (36/62) e 37,8% (34/90) respectivamente, do que para o *cluster* C3, com uma renda de um salário mínimo na maioria dos seus componentes com 48,4% (62/128). Outra característica interessante foi a classificação operacional do caso índice, onde a totalidade (100%) dos comunicantes alocados no *cluster* C2 tinham como caso índice pacientes classificados como PB, em contraste com os comunicantes alocados no *cluster* C3, cuja totalidade (100%) tinham casos índices classificados como MB.

Para complementar a análise epidemiológica mostrada na Tabela 9, foram avaliados os valores do anti-PGL-1 dos componentes de cada *cluster*, como pode ser observado no gráfico da Figura 27.

Figura 27 - Anti-PGL-1 por *Clusters* C1, C2 e C3.

Fonte: Base de dados.

Com a análise dos casos novos diagnosticados entre os comunicantes distribuídos entre os 3 *clusters*, obtivemos que dos 35 casos novos, 6 foram alocados no *cluster* C1, 10 no *cluster* C2 e 19 no *cluster* C3. A análise desta distribuição e as características epidemiológicas, classificamos os *clusters* em diferentes riscos de infecção para a hanseníase: C1 como “baixo risco”, C2 como “médio risco” e C3 como “alto risco”. A análise epidemiológica dos casos novos distribuídos entre os 3 *clusters* pode ser observada na Tabela 10.

Tabela 10. Características epidemiológicas dos casos novos (35) de hanseníase diagnosticados no município de Santarém-Pa de acordo com os 3 *clusters* C1, C2, 3 C3 encontrados e as variáveis: gênero, faixa etária, grau de parentesco, período de convivência, BCG, dormitórios, renda Familiar, água para consumo e tratamento, classificação operacional e forma clínica (casos índices).

		C1		C2		C3	
		n	%	n	%	n	%
SEXO	Feminino	4	66,7%	8	80,0%	11	57,9%
	Masculino	2	33,3%	2	20,0%	8	42,1%
	Total	6	100,0%	10	100,0%	19	100,0%
FAIXA ETARIA	0 - 14 ANOS	2	33,3%	2	20,0%	5	26,3%
	15 - 39 ANOS	3	50,0%	3	30,0%	9	47,4%
	40 - 59 ANOS	0	0,0%	4	40,0%	3	15,8%
	ACIMA 60 ANOS	1	16,7%	1	10,0%	2	10,5%
	Total	6	100,0%	10	100,0%	19	100,0%
GRAU DE PARENTESCO	Cônjuge	1	16,70%	2	20,00%	1	5,30%
	Não parente	0	0,00%	1	10,00%	2	10,50%
	Parente não consanguíneo	0	0,00%	3	30,00%	4	21,10%
	Filho(a)	4	66,70%	1	10,00%	3	15,80%
	Irmão	0	0,00%	1	10,00%	1	5,30%

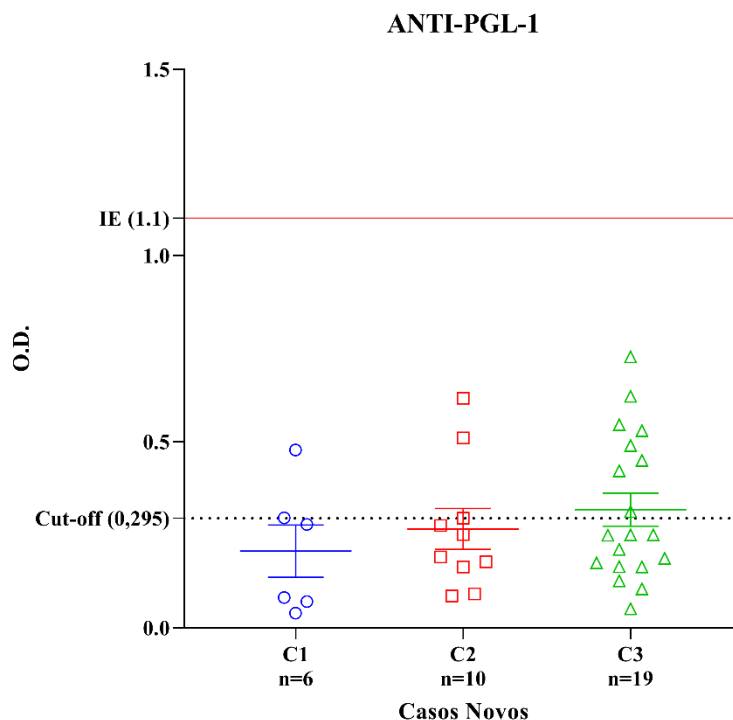
		C1		C2		C3	
		n	%	n	%	n	%
	Pai/Mãe	0	0,00%	0	0,00%	2	10,50%
	Parente consanguíneo	1	16,70%	2	20,00%	6	31,60%
	Total	6	100,0%	10	100,0%	19	100,0%
PERIODO DE CONVIVENCIA	Mais que um ano	0	0,0%	1	10,0%	2	10,5%
	Entre um a dois anos	0	0,0%	0	0,0%	2	10,5%
	Entre dois a cinco anos	0	0,0%	1	10,0%	0	0,0%
	Entre cinco a dez anos	0	0,0%	0	0,0%	2	10,5%
	Mais que dez anos	6	100,0%	8	80,0%	13	68,4%
	Total	6	100,0%	10	100,0%	19	100,0%
BCG	0	0	0,0%	1	10,0%	5	26,3%
	1,0	6	100,0%	9	90,0%	13	68,4%
	2,0	0	0,0%	0	0,0%	1	5,3%
	Total	6	100,0%	10	100,0%	19	100,0%
DORMINTORIOS > 2 PESSOAS	Não	5	83,3%	2	20,0%	10	52,6%
	Sim	1	16,7%	8	80,0%	9	47,4%
	Total	6	100,0%	10	100,0%	19	100,0%
RENDIA FAMILIAR	Menor que um salário-mínimo	0	0,0%	1	10,0%	0	0,0%
	Um salário-mínimo	0	0,0%	3	30,0%	12	63,2%
	Até dois salários-mínimos	5	83,3%	4	40,0%	4	21,1%
	Até três salários-mínimos	1	16,7%	1	10,0%	2	10,5%
	Maior que três salários-mínimos	0	0,0%	1	10,0%	1	5,3%
	Total	6	100,0%	10	100,0%	19	100,0%
ÁGUA PRA CONSUMO	Poço artesiano	2	33,3%	0	0,0%	5	26,3%
	Rede de água encanada	4	66,7%	10	100,0%	14	73,7%
	Total	6	100,0%	10	100,0%	19	100,0%
TRATAMENTO DA ÁGUA	Coadá ou nenhum outro método	1	16,7%	4	40,0%	11	57,9%
	Filtrada, fervida ou clorada	5	83,3%	5	50,0%	7	36,8%
	Mineral	0	0,0%	1	10,0%	1	5,3%
	Total	6	100,0%	10	100,0%	19	100,0%
CLASSIFICACAO OPERACIONAL - INDICE	Paucibacilar (PB)	1	16,7%	10	100,0%	0	0,0%
	Multibacilar (MB)	5	83,3%	0	0,0%	19	100,0%
	Total	6	100,0%	10	100,0%	19	100,0%
FORMA CLINICA – INDICE	Indeterminada (i)	0	0,0%	1	10,0%	0	0,0%
	Neural Pura (NP)	0	0,0%	2	20,0%	0	0,0%
	Tuberculóide (T)	1	16,7%	7	70,0%	0	0,0%
	Dimorfa (D)	1	16,7%	0	0,0%	2	10,5%
	Borderline Tuberculóide (BT)	1	16,7%	0	0,0%	9	47,4%
	Borderline Borderline (BB)	0	0,0%	0	0,0%	4	21,1%
	Borderline Virchowiana (BV)	1	16,7%	0	0,0%	1	5,3%
	Virchowiana (V)	2	33,3%	0	0,0%	3	15,8%
	Total	6	100,0%	10	100,0%	19	100,0%

Fonte: Base de dados.

Após classificar os *clusters* em diferentes níveis de risco de infecção, foi realizado uma análise dos valores do anti-PGL-1 referente aos casos novos alocados em cada *cluster*, como pode ser observado no gráfico da Figura 28. É interessante observar que a carga bacilar do conjunto de casos novos do *clusters* C3 foi maior do que os do *cluster* C2, que por conseguinte a carga bacilar do conjunto do *cluster* C2 foi maior do que os do *cluster* C1. Essa informação veio corroborar a classificação do risco de infecção encontrado com o processo de clusterização. Também é importante notar que a variável anti-PGL-1 não foi utilizada no processo de composição dos *clusters*, o que torna o resultado ainda mais significativo de

definição de que os componentes (comunicantes) do *clusters* C3 possuem maior risco de contaminação para a hanseníase, seguidos pelo C2 e C1 respectivamente.

Figura 28 - Densidade óptica (O.D.) dos casos novos por *cluster*.



Fonte: Base de dados.

Vários estudos demonstraram que a sorologia poderia ser utilizada para detectar anticorpos anti-PGL-1 e a sua titulação poderia ser útil na classificação dos pacientes, monitoramento dos casos, identificação do risco de recidivas e identificação de contatos intradomiciliares com maior risco de desenvolver a hanseníase (MOURA et al., 2008b; OSKAM; SLIM; BÜHRER-SÉKULA, 2003; SPENCER et al., 2012).

A sorologia anti-PGL-1 apresenta de moderada a boa correlação com o índice baciloscópico do paciente (NAVES et al., 2013; ROCHE et al., 1990), entretanto não possui a sensibilidade necessária para ser utilizada como um teste diagnóstico de forma isolada. Neste sentido, nossos achados mostram que com o auxílio de técnicas de mineração de dados baseadas em clusterização foi possível potencializar a utilização da sorologia anti-PGL-1 juntamente com os dados socioeconômicos e clínicos na identificação de diferentes riscos de infecção para a hanseníase entre os comunicantes de pacientes diagnosticados e tratados do SINAN. Esse achado representa um conhecimento importante para o suporte à decisão das equipes de vigilância de contatos, que podem priorizar seus esforços nos grupos de contatos com maiores riscos na tentativa de melhorar o desempenho do combate da hanseníase.

6 CONCLUSÃO

Na busca ativa realizada no ano de 2014, foram detectados 35 casos novos de hanseníase em Santarém, revelando que a maioria dos casos diagnosticados tanto na série histórica analisada como na semana de busca ativa estavam na forma clínica MB, o que caracteriza o diagnóstico tardio e por conseguinte um estado de disseminação ativa da doença.

Nesse estudo, identificou-se uma possível infecção-subclínica e a necessidade permanente do acompanhamento dos comunicantes com elevada carga bacilar, principalmente os que tem casos MB diagnosticados na família. Nossos achados expressam que há ocorrência de diagnóstico tardio tendo em vista que o período de incubação é longo, ou seja, o indivíduo já vem transmitindo a doença há muitos anos.

Evidencia-se a importância da vigilância de contatos de pacientes diagnosticados com hanseníase periodicamente durante, no mínimo, o período médio de incubação da doença. Nossos resultados reforçam a necessidade da efetiva avaliação clínico-laboratorial desses comunicantes para a detecção de novos casos de hanseníase na forma inicial da doença.

A sorologia anti-PGL-1 apresenta de moderada a boa correlação com o índice baciloscópico do paciente, entretanto não possui a sensibilidade necessária para ser utilizada como um teste diagnóstico, sendo significativamente mais elevada entre os pacientes MB, com progressivo e consistente aumento até o polo virchowiano. Sua titulação pode ser útil na classificação dos pacientes, monitoramento dos casos, identificação do risco de recidivas, infecção subclínica e identificação de comunicantes intradomiciliares com maior risco de desenvolver a hanseníase.

Por fim, com a utilização de técnicas de mineração de dados baseadas em clusterização foi possível potencializar a utilização da sorologia anti-PGL-1 juntamente com os dados socioeconômicos e clínicos na identificação de diferentes níveis de risco, baixo, médio e alto risco, de desenvolver a doença. Esse achado representa um conhecimento importante para o suporte à decisão das equipes de vigilância de contatos, que podem priorizar seus esforços nos grupos de contatos com maiores riscos na tentativa de melhorar o desempenho do combate da hanseníase, além de contribuir para a identificação de fatores de vulnerabilidades sociais e econômicas dos indivíduos infectados pelo *M. leprae*.

REFERÊNCIAS

ALVES, E. D.; FERREIRA, T. L.; FERREIRA, I. N. **Hanseníase : avanços e desafios**. Brasília/DF: NESPROM, 2014.

BARRETO, J. G. et al. High rates of undiagnosed leprosy and subclinical infection amongst school children in the Amazon region. **Memorias do Instituto Oswaldo Cruz**, v. 107, n. SUPPL.1, p. 60–67, 2012.

BASSO, M. E. DE M.; ANDRADE, R. F. DE; SILVA, R. L. F. DA. Trend of epidemiological indicators of leprosy in an endemic state of the Amazon region. **Revista gaucha de enfermagem**, v. 42, p. e20190520, 2021.

BOIGNY, R. N. et al. Persistência da hanseníase em redes de convívio domiciliar: sobreposição de casos e vulnerabilidade em regiões endêmicas no Brasil. **Cadernos de Saúde Pública**, v. 35, n. 2, 18 fev. 2019.

BONITA, R.; BEAGLEHOLE, R.; KJELLSTROM, T. **Epidemiologia básica**. 2ª ed. São Paulo, Santos.: Livraria Santos Editora Com. Imp. Ltda., 2010.

BRASIL. MINISTÉRIO DA SAÚDE. **Diretrizes para vigilância, atenção e eliminação da Hanseníase como problema de saúde pública : manual técnico-operacional**. Brasília: Ministerio da Saúde, 2016.

BRASIL. MINISTÉRIO DA SAÚDE. Caracterização da situação epidemiológica da hanseníase e diferenças por sexo, Brasil, 2012-2016. **Boletim epidemiológico**, v. 49, n. 4, p. 1–12, 2018.

BRASIL. MINISTÉRIO DA SAÚDE. **Protocolo Clínico e Diretrizes Terapêuticas da Hanseníase**. Brasília/DF: Ministério da Saúde, 2021a.

BRASIL. MINISTÉRIO DA SAÚDE. **Indicadores e Dados Básicos de Hanseníase nos Municípios Brasileiros**. Disponível em: <http://indicadoreshanseniaze.aids.gov.br/>. Acesso em: 3 abr. 2022b.

BRASIL. MINISTÉRIO DA SAÚDE. COORDENAÇÃO GERAL DO PROGRAMA NACIONAL DE IMUNIZAÇÕES/DEPARTAMENTO DE IMUNIZAÇÃO E DOENÇAS TRANSMISSÍVEIS/SECRETARIA DE VIGILÂNCIA EM SAÚDE. **Instrução Normativa Referente ao Calendário de Vacinação**. Disponível em: <https://www.saude.go.gov.br/files/imunizacao/legislacao/InstrucaoNormativaCalendarioVacinao.2019.pdf>. Acesso em: 2 ago. 2022.

BRASIL. MINISTÉRIO DA SAÚDE. SECRETARIA DE POLÍTICAS DE SAÚDE. DEPARTAMENTO DE ATENÇÃO BÁSICA. **Guia para o controle da hanseníase**. 3ª ed. Brasília/DF: Ministerio da Saúde, 2002.

BRASIL. MINISTÉRIO DA SAÚDE. SECRETARIA DE VIGILÂNCIA EM SAÚDE. DEPARTAMENTO DE DOENÇAS DE CONDIÇÕES CRÔNICAS E INFECÇÕES SEXUALMENTE TRANSMISSÍVEIS. **PROTOCOLO CLÍNICO E DIRETRIZES TERAPÊUTICAS DA HANSENÍASE**. Brasília/DF: Ministério da Saúde, 2022.

BRASIL. MINISTÉRIO DA SAÚDE. SECRETARIA DE VIGILÂNCIA EM SAÚDE. DEPARTAMENTO DE DOENÇAS DE CONDIÇÕES CRÔNICAS E INFECÇÕES SEXUALMENTE TRANSMISSÍVEIS. **Hanseníase no Brasil: caracterização das incapacidades físicas**. 1ª ed. Brasília/DF: Ministério da Saúde, 2020.

BRASIL. MINISTÉRIO DA SAÚDE. SECRETARIA DE VIGILÂNCIA EM SAÚDE. DEPARTAMENTO DE VIGILÂNCIA EPIDEMIOLÓGICA. **Guia de procedimentos técnicos: baciloscopia em hanseníase**. Brasília/DF: Ministério da Saúde, 2010.

BRASIL. MINISTÉRIO DA SAÚDE. SECRETARIA DE VIGILÂNCIA EM SAÚDE. GUIA DE VIGILÂNCIA EPIDEMIOLÓGICA. **Guia de Vigilância Epidemiológica**. 6ª ed. Brasília/DF: Ministério da Saúde, 2005.

BRASIL. MINISTÉRIO DA SAÚDE. **Hanseníase – Informações Técnicas**. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/h/hanseníase>. Acesso em: 4 maio. 2021.

BRASIL. **MEDIDA PROVISÓRIA Nº 373 DE 24 DE MAIO DE 2007**. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/mpv/373.htm. Acesso em: 10 ago. 2022a.

BRASIL. **LEI Nº 11.520, DE 18 DE SETEMBRO DE 2007**. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/11520.htm. Acesso em: 10 ago. 2022b.

BRASIL. **Acompanhamento dos dados de Hanseníase - Pará. Sistema de Informação de Agravos de Notificação - SINAN Net**. Brasília: Tecnologia da Informação a Serviço do SUS - DATASUS, 2022. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/hanswpa.def>. Acesso em: 7 ago. 2022.

BRENNAN, P. J.; SPENCER, J. S. The Physiology of Mycobacterium leprae Morphology and Cell Wall. **The International Textbook of leprosy**, n. 6, p. 1–51, 2019.

BRITO, K. K. RYSTIN. G. DE et al. Epidemiological analysis of leprosy in an endemic state of northeastern Brazil. **Revista gaúcha de enfermagem / EENFUFGRS**, v. 36, p. 24–30, 2015.

BROOKS, G. et al. **Microbiologia médica de Jawetz, Melnick e Adelberg**. Porto Alegre: AMGH EDITORA LTDA, 2014.

CAMILO, C.; SILVA, J. Mineração de Dados: Conceitos, tarefas, métodos e ferramentas. **Universidade Federal de Goiás (UFG)**, p. 29, 2009.

CASTRO, S. S. DE et al. Leprosy incidence, characterization of cases and correlation with household and cases variables of the Brazilian states in 2010. **Anais Brasileiros de Dermatologia**, v. 91, n. 1, p. 28–33, 2016.

CHAVES, E. C. et al. Índice de carência social e hanseníase no estado do Pará em 2013: análise espacial. **Epidemiologia e serviços de saúde : revista do Sistema Único de Saúde do Brasil**, v. 26, n. 4, p. 807–816, 2017.

DESSUNTI, E. M. et al. Hanseníase: o controle dos contatos no município de Londrina-PR em um período de dez anos. **Revista Brasileira de Enfermagem**, v. 61, n. spe, p. 689–693, 2008.

DUARTE, M. T. C.; AYRES, J. A.; SIMONETTI, J. P. Perfil socioeconômico e demográfico de portadores de hanseníase atendidos em consulta de enfermagem. **Revista Latino-Americana de Enfermagem**, v. 15, n. SPEC. ISS., p. 774–779, 2007.

EIDT, L. M. Breve história da hanseníase: sua expansão do mundo para as Américas, o Brasil e o Rio Grande do Sul e sua trajetória na saúde pública brasileira. **Saúde e Sociedade**, v. 13, n. 2, p. 76–88, ago. 2004.

EMERSON, L. E. et al. Poor wash (Water, sanitation, and hygiene) conditions are associated with leprosy in North Gondar, Ethiopia. **International Journal of Environmental Research and Public Health**, v. 17, n. 17, p. 1–10, 2020.

FABRI, A. DA C. O. C. **Análise comparativa da reatividade anti-LID-1, NDO-LID, NDO-HSA e PGL-1 em hanseníase**. Belo Horizonte - MG: UFMG, 2015.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. Mining association rules in graphs based on frequent cohesive itemsets. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, 1996.

FIGINI, S.; GIUDICI, P. **Applied data mining for business and industry**. 2. ed. Nova Jersey, EUA: John Wiley & Sons Ltd, 2009.

FRADE, M. C. A. et al. Unexpectedly high leprosy seroprevalence detected using a random surveillance strategy in midwestern Brazil: A comparison of ELISA and a rapid diagnostic test. 2017.

GOULART, I. M. B.; PENNA, G. O.; CUNHA, G. Immunopathology of leprosy: The complexity of the mechanisms of host immune response to Mycobacterium leprae. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 35, n. 4, p. 365–375, 2002.

HAN, J.; KAMBER, M.; PEI, J. **Data mining: Data mining concepts and techniques**. 3. ed. 225 Wyman Street, Waltham, MA 02451, USA: Elsevier Inc., 2014.

HAND, D.; MANNILA, H.; SMYTH, P. **Principles of Data Mining**. Cambridge: Bradford Book, 2001. v. 2001

KUMAR, B.; UPRETY, S.; DOGRA, S. Clinical diagnosis of leprosy cases. **Journal of the Indian Medical Association**, v. 104, n. 12, p. 676–679, 2006.

LASTÓRIA, J.; ABREU, M. Hanseníase: diagnóstico e tratamento. **Currículo Lattes**, v. 17, n. 4, p. 5–8, 2012.

LEANO, H. A. DE M. et al. Socioeconomic factors related to leprosy: an integrative literature review. **Revista brasileira de enfermagem**, v. 72, n. 5, p. 1474–1485, 16 set. 2019.

LEVINSON, W. **Microbiologia médica e imunologia**. São Paulo: ARTMED, 2010.

LOBATO, J. et al. Detection of Immunoglobulin G Antibodies to *Neospora caninum* in Humans: High Seropositivity Rates in Patients Who Are Infected by Human Immunodeficiency Virus or Have Neurological Disorders. **Clinical and Vaccine Immunology**, v. 13, n. 1, p. 84–89, jan. 2006.

LOCKWOOD, D. N. J. Editor's choice March 2012: the different challenges in leprosy around the world. **Leprosy review**, v. 83, n. 1, p. 1, mar. 2012.

LONGO, J. D. D. M. Contribuição ao conhecimento da transmissão da hanseníase em Campo Grande – Mato Grosso do Sul, 2006. p. 93, 2006.

LYON, S.; GROSSI, M. A. DE F. Diagnóstico e tratamento da Hanseníase. **Hanseníase: avanços e desafios**, p. 141–170, 2014.

MANDIGAN, M. T. et al. **Microbiologia de Brock**. 14. ed. Porto Alegre: Artemed, 2016.

MARIOTTI, V. et al. Probable early presence of leprosy in Europe in a Celtic skeleton of the 4th-3rd century BC (Casalecchio di Reno, Bologna, Italy). **International Journal of Osteoarchaeology**, v. 15, n. 5, p. 311–325, 2005.

MAURANO, F. **Tratado de Leprologia**. 2ª Edição ed. Rio de Janeiro, Brazil: Serviço Nacional de Lepra, 1950.

MONOT, M. et al. On the origin of leprosy. **Science**, v. 308, n. 5724, p. 1040–1042, 2005.

MONOT, M. et al. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. **Nature Genetics**, v. 41, n. 12, p. 1282–1289, 2009.

MONTEIRO, L. D. et al. Padrões espaciais da hanseníase em um estado hiperendêmico no Norte do Brasil, 2001-2012. **Revista de Saude Publica**, v. 49, p. 49–84, 2015.

MOURA, R. S. DE et al. Sorologia da hanseníase utilizando PGL-I: Revisão sistemática. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 41, n. SUPPL. 2, p. 11–18, 2008a.

MOURA, R. S. DE et al. Leprosy serology using PGL-I: a systematic review. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 41, n. Suplemento II, p. 1, 2008b.

NAVES, M. DE M. et al. Bacterial load in the nose and its correlation to the immune response in leprosy patients. **Leprosy review**, v. 84, n. 1, p. 85–91, mar. 2013.

NETO, J. M. et al. Analysis of control household contacts of people affected by leprosy in Brazil and the state of São Paulo de 1991 a 2012. **Hansenologia Internationalis**, v. 38, p. 68–78, 2014.

NIITSUMA, E. N. A. et al. Fatores associados ao adoecimento por hanseníase em contatos: revisão sistemática e metanálise. **Revista Brasileira de Epidemiologia**, v. 24, 2021.

NOGUEIRA, A. et al. Implementation of a Distributed System to Collect, Search and Manage Data of Patients with Leprosy in Hiperendemic Municipalities. In: 2019, G. H. (Ed.). . **The Eighth International Conference on Global Health Challenges**. Porto, Portugal: IARIA, 2019. v. 1.p. 28–31.

OLIVEIRA, K. S. DE et al. Avaliação dos indicadores epidemiológicos e operacionais para a hanseníase em municípios prioritários no estado do Paraná, 2001 a 2010. **Epidemiologia e Serviços de Saúde**, v. 24, n. 3, p. 507–516, 2015.

OMS. Rumo à zero hanseníase Rumo à zero hanseníase Estratégia Global de Hanseníase. p. 1–30, 2021.

OPROMOLLA, D. V. A. **Noções de hansenologia**. Bauru, SP: Centro de Estudos Dr. Reynaldo Quagliato, 1981.

OPROMOLLA, D. V. A. **Noções de hansenologia**. Bauru, SP: Centro de Estudos Dr. Reynaldo Quagliato, 2000.

OSKAM, L.; SLIM, E.; BÜHRER-SÉKULA, S. Serology: recent developments, strengths, limitations and prospects: a state of the art overview. **Leprosy review**, v. 74, n. 3, p. 196–205, set. 2003.

PARDILLO, F. E. F. et al. Methods for the classification of leprosy for treatment purposes. **Clinical Infectious Diseases**, v. 44, n. 8, p. 1096–1099, 2007.

PATNAIK, S. K. et al. Automated Skin Disease Identification using Deep Learning Algorithm. **Biomedical and Pharmacology Journal**, v. 11, n. 3, p. 1429–1436, 28 set. 2018.

PEIXOTO, B. K. DE S. et al. Aspectos epidemiológicos dos contatos de hanseníase no Município de São Luís-MA. **Hansen. int**, v. 36, n. 1, p. 23–30, 2011.

PESCARINI, J. M. et al. Epidemiological characteristics and temporal trends of new leprosy cases in Brazil. **Cad. Saúde Pública**, v. 37, n. 7, p. 130020, 2021.

PIERNEEF, L. et al. **Detection of anti-m. Leptrae antibodies in children in leprosy-endemic areas: A systematic review** PLoS Neglected Tropical Diseases, 2021.

PIRES, C. A. A. et al. Hanseníase em menores de 15 anos: A importância do exame de contato. **Revista Paulista de Pediatria**, v. 30, n. 2, p. 292–295, 2012.

RASTOGI, N.; LEGRAND, E.; SOLA, C. The mycobacteria: an introduction to nomenclature and pathogenesis. **Revue scientifique et technique (International Office of Epizootics)**, v. 20, n. 1, p. 21–54, abr. 2001.

REIS, A. DA S. DOS et al. Overlapping of new leprosy cases in household contact networks in two municipalities in North and Northeast Brazil, 2001-2014. **Cadernos de Saude Publica**, v. 35, n. 10, 2019.

RICARDO LINDEN. Técnicas de Agrupamento. **Revista de Sistemas de Informação**, v. 1, n. 4, p. 18–36, 2009.

RIDLEY, D. S.; JOPLING, W. H. A classification of leprosy for research purposes. **Leprosy review**, v. 33, p. 119–128, 1962.

RIDLEY, D. S.; JOPLING, W. H. Classification of leprosy according to immunity. A five-group system. **International Journal of Leprosy and Other Mycobacterial Diseases**, v. 34, n. 3, p. 255–273, 1966.

ROCHA, C. A. Caracterização Dos Comunicantes Em Um Ambulatório De Referência Para Hanseníase Na Cidade De Salvador-Bahia [Characterization of the Household Contacts in a Reference Outpatient Clinic for Hanseníase in the City of Salvador-Bahia]. 2016.

ROCHE, P. W. et al. Operational value of serological measurements in multibacillary leprosy patients: clinical and bacteriological correlates of antibody responses. **International journal of leprosy and other mycobacterial diseases : official organ of the International Leprosy Association**, v. 58, n. 3, p. 480—490, 1990.

RODRIGUES, L. C.; LOCKWOOD, D. N. J. Leprosy now: Epidemiology, progress, challenges, and research gaps. **The Lancet Infectious Diseases**, v. 11, n. 6, p. 464–470, 2011.

RODRIGUES, M. M. et al. UVB suscetibilidade como fator de risco para o desenvolvimento da hanseníase virchowiana. **Anais Brasileiros de Dermatologia**, v. 85, n. 6, p. 839–842, 2010.

ROUQUAYROL, M. Z.; GURGEL, M. **Rouquayrol: Epidemiologia & Saúde**. 8ª ed. Rio de Janeiro: Medbook, 2018.

SALES, A. M. et al. Leprosy among patient contacts: A multilevel study of risk factors. **PLoS Neglected Tropical Diseases**, v. 5, n. 3, p. 1–6, 2011.

SANTOS, K. C. B. DOS et al. Estratégias de controle e vigilância de contatos de hanseníase: revisão integrativa. **Saúde em Debate**, v. 43, n. 121, p. 576–591, abr. 2019.

SANTOS, L. A. DE C.; FARIA, L.; MENEZES, R. F. DE. Contrapontos da história da hanseníase no Brasil: cenários de estigma e confinamento. **Revista Brasileira de Estudos de População**, v. 25, n. 1, p. 167–190, jun. 2008.

SETIA, M. S. et al. The role of BCG in prevention of leprosy: a meta-analysis. **The Lancet Infectious Diseases**, v. 6, n. 3, p. 162–170, mar. 2006.

SILVA, R. E. DA et al. Modeling a Vulnerability Index for Leprosy Using Spatial Analysis and Artificial Intelligence Techniques in a Hyperendemic Municipality in the Amazon. In: BI, Y.; BHATIA, R.; KAPOOR, S. (Eds.). . **Intelligent Systems and Applications - Proceedings of the 2019 Intelligent Systems Conference, IntelliSys 2019, London, UK, September 5-6, 2019, Volume 2**. [s.l.] Springer, Cham, 2020. v. 1038p. 802–823.

SILVA, Y. E. D. DA et al. Application of Clustering Technique with Kohonen Self-organizing Maps for the Epidemiological Analysis of Leprosy. In: ARAI, K.; KAPOOR, S.; BHATIA, R. (Eds.). . **Intelligent Systems and Applications - Proceedings of the 2018 Intelligent Systems Conference, IntelliSys 2018, London, UK, September 6-7, 2018, Volume 2**. [s.l.] Springer, Cham, 2019. v. 869p. 295–309.

SILVA GAMA, R. et al. Revista da Sociedade Brasileira de Medicina Tropical Prospects for new leprosy diagnostic tools, a narrative review considering ELISA and PCR assays. **Journal of the Brazilian Society of Tropical Medicine**, p. 2020, 2020.

SOUZA-ARAÚJO, H. C. DE. Contribuição á Epidemiologia e Prophylaxia da Lepra no Norte do Brasil. **Memórias do Instituto Oswaldo Cruz**, v. 27, n. 3, p. 165–337, set. 1933.

SOUZA, L. R. DE. **Condicionantes sociais na delimitação de espaços endêmicos de**

hanseníase. São Paulo: Universidade de São Paulo - USP, 2012.

SPENCER, J. S. et al. Identification of serological biomarkers of infection, disease progression and treatment efficacy for leprosy. **Memorias do Instituto Oswaldo Cruz**, v. 107, n. SUPPL.1, p. 79–89, 2012.

TEIXEIRA, C. S. S. et al. Incidence of and Factors Associated With Leprosy Among Household Contacts of Patients With Leprosy in Brazil. **JAMA Dermatology**, v. 156, n. 6, p. 640, 1 jun. 2020.

TORTORA, G. J.; FUNKE, B. R.; CASE, C. L. **Microbiologia (12 a Edição).** [s.l: s.n.].

VALE, M. N. DO. **AGRUPAMENTOS DE DADOS: AVALIAÇÃO DE MÉTODOS E DESENVOLVIMENTO DE APLICATIVO PARA ANÁLISE DE GRUPOS.** Rio de Janeiro, Brazil: PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO, 29 ago. 2005.

VAN HOOIJ, A.; GELUK, A. Immunodiagnostics for Leprosy. **International Text Book of Leprosy**, n. 5, p. 1–25, 2016.

VIDAL, S. L. et al. Testes sorológicos anti-NDO-HSA, anti-LID-1 e anti-NDO-LID em contatos domiciliares de área não endêmica de hanseníase. **HU Revista**, v. 44, n. 3, p. 325–331, 2019.

WEBBER, C. G.; ZAT, D.; LIMA, M. D. F. W. DO P. Utilização De Algoritmos De Agrupamento Na Mineração De Dados Educacionais. **Renote**, v. 11, n. 1, p. 1–10, 2013.

WHO. **A GUIDE TO ELIMINATING LEPROSY AS A PUBLIC HEALTH PROBLEM.** 1. ed. Genova: World Health Organization, 1995.

WHO. WHO Expert Committee on Leprosy. **World Health Organization technical report series**, n. 968, p. 1–61, 1 p following 61, 2012.

WHO. **Global leprosy strategy: accelerating towards a leprosy-free world.** New Delhi-110002, India: World Health Organization, 2016.

WHO. Global leprosy update, 2018: moving towards a leprosy. **Weekly epidemiological record**, v. 94, n. August 2019, p. 389–412, 2019.

WHO. Global leprosy (Hansen disease) update, 2019: time to step-up prevention initiatives. **Weekly epidemiological record**, v. 95, n. 36, p. 417–440, 2020.

WHO. Leprosy. **World Health Organization**, 2022.

YAMANOUCHI, A. A. et al. Hanseníase e sociedade: um problema sempre atual. **Anais brasileiros de dermatologia**, 1993.