

UNIVERSIDADE FEDERAL DE VIÇOSA

Estudo epidemiológico das arboviroses no Brasil e o uso de inteligência artificial na construção de um modelo preditivo de letalidade para febre amarela

Karen Oliveira Fracalossi
Magister Scientiae

**VIÇOSA - MINAS GERAIS
2021**

KAREN OLIVEIRA FRACALOSSI

Estudo epidemiológico das arboviroses no Brasil e o uso de inteligência artificial na construção de um modelo preditivo de letalidade para febre amarela

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Nutrição, para obtenção do título de *Magister Scientiae*.

Orientadora: Glauce Dias da Costa

Coorientadores: Rodrigo S. Batista
Tiago Ricardo Moreira
Rosângela M. M. Cotta

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

F814e
2021

Fracalossi, Karen Oliveira, 1994-
Estudo epidemiológico das arboviroses no Brasil e o uso de inteligência artificial na construção de um modelo preditivo de letalidade para febre amarela / Karen Oliveira Fracalossi. – Viçosa, MG, 2021.

1 dissertação eletrônica (133 f.): il. (algumas color.).

Inclui anexos.

Orientador: Glauce Dias da Costa.

Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Nutrição e Saúde, 2021.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2024.835>

Modo de acesso: World Wide Web.

1. Arboviroses - Prognóstico. 2. Febre amarela.
3. Mineração de dados (Computação). 4. Inteligência artificial - Aplicações médicas. I. Costa, Glauce Dias da, 1979-.
II. Universidade Federal de Viçosa. Departamento de Nutrição e Saúde. Programa de Pós-Graduação em Ciência da Nutrição.
III. Título.

CDD 22. ed. 616.9185075

KAREN OLIVEIRA FRACALLOSSI

Estudo epidemiológico das arboviroses no Brasil e o uso de inteligência artificial na construção de um modelo preditivo de letalidade para febre amarela

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Nutrição, para obtenção do título de *Magister Scientiae*.

APROVADA: 29 de janeiro de 2021.

Assentimento:

Karen Oliveira Fracalossi
Autora

Glauce Dias da Costa
Orientadora

Essa dissertação foi assinada digitalmente pela autora em 02/01/2025 às 20:31:49 e pela orientadora em 03/01/2025 às 11:34:07. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **1S8U.NTRF.Q3EB** e clique no botão 'Validar documento'.

AGRADECIMENTOS

Agradeço primeiramente à Deus por permitir que tudo pudesse acontecer, nas horas mais difíceis era a fé que me sustentava.

A minha família por estar comigo por todo o caminho, por me segurar nas horas em que eu fraquejava e por sempre estarem disponível nas horas mais escuras.

A Glauce, minha orientadora, por ser uma força que me motivou durante todo o processo, se hoje chegamos onde estamos é tanto por ela quanto por mim.

Aos meus amigos, os novos e os antigos, por serem um oásis no meio do deserto, por serem um ponto de descanso no meio da correria, e um ouvido atento para minhas queixas.

A Fran, minha parceira de casa, e amiga para a vida toda.

Ao Daniel, uma pessoa muito especial que eu conheci durante esse caminho, muito obrigada por toda a ajuda, sem você nada disso seria possível.

Aos professores, os que compõem a minha banca, os meus coorientadores e os outros professores com quem venho aprendendo por todo o meu caminho.

Ao meu grupo VigSUS e ao PRODUS, lugares e pessoas que me acolheram, me motivaram, pessoas carinhosas, amorosas e trabalhadoras, as vezes tão diferentes de mim, mas que me complementavam, me estimulavam, me alegravam, pessoas com quem discutir, pessoas para pedir ajuda, pessoas para rir junto, se desesperar junto e as vezes chorar junto, e tomar aquele cafezinho, porque não?

A todos do Departamento de Nutrição por sempre me acolherem tão bem, e me fazerem me sentir em casa, mesmo tão longe de casa.

Ao Programa de Pós-Graduação em Ciência da Nutrição, pela oportunidade e pelo cuidado.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

FRACALOSSI, Karen Oliveira, M.Sc., Universidade Federal de Viçosa, janeiro de 2021. **Estudo epidemiológico das arboviroses no Brasil e o uso de inteligência artificial na construção de um modelo preditivo de letalidade para febre amarela.** Orientadora: Glauce Dias da Costa. Coorientadores: Rodrigo Siqueira Batista, Tiago Ricardo Moreira e Rosangela Minardi Mitre Cotta.

As arboviroses são doenças causadas por arbovírus e representam uma crescente preocupação para a saúde pública no Brasil e no mundo, principalmente pelo potencial de dispersão, pela capacidade de adaptação a novos ambientes e hospedeiros, e possibilidade de causar epidemias extensas. O compartilhamento de um mesmo vetor de transmissão entre dengue, zika, chikungunya e febre amarela, a co-circulação dessas doenças e a grande propagação do *Aedes aegypti* pelo território brasileiro podem causar epidemias de caráter emergente. Ainda, a similaridade entre os sintomas apresentados para essas doenças dificulta o manejo clínico. Nesse sentido, os estudos epidemiológicos e a utilização de recursos computacionais na forma de modelos preditivos, podem aprimorar a pesquisa em saúde, avaliação da qualidade da atenção, prevenção, prognóstico resultando em tratamentos mais efetivos, bem como no controle de doenças, direcionamento de recursos e melhorias na gestão. Assim, o objetivo desse trabalho foi estimar as prevalências das arboviroses dengue, zika, chikungunya e febre amarela nos períodos de 2000 a 2018 e construir um modelo preditivo para óbito por febre amarela no Brasil utilizando inteligência artificial. No primeiro estudo foi realizada uma revisão sistemática com metanálise baseada no protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Foi realizada a busca nas bases de dados PubMed/MEDLINE, Scielo (Scientific Electronic Library Online), ScienceDirect e BVS (Biblioteca Virtual de Saúde), sob o termo de busca (dengue OR zika OR chikungunya OR "febre amarela") AND prevalência). A metanálise foi realizada por modelo de efeito randômico. A heterogeneidade foi avaliada pelo teste do qui-quadrado (χ^2) com significância de $P < 0,10$, sendo sua magnitude apurada pelo o I-quadrado (I^2). As análises foram realizadas com os comandos Metaprop e Metareg do programa Stata (versão 13.0). Dos 8052 artigos incluídos na busca 61 artigos foram selecionados. Os anos dos estudos publicados estavam entre os anos de 1992 e 2019, sendo 65,57% dos artigos ($n=40$) foram publicados nos últimos 10 anos. A maioria dos artigos selecionados descrevem principalmente dados epidemiológicos regionais derivados de pesquisas e estudos conduzidos em regiões específicas, sendo utilizados para apoiar os dados nacionais no que diz

respeito à prevalência e sorotipo. Não foi possível realizar a metanálise com os artigos de febre amarela, pois poucos estudos foram selecionados. O desfecho primário foi a prevalência de dengue, zika e chikungunya com intervalo de confiança de 95% (IC95%). A prevalência agrupada de dengue, zika e chikungunya no modelo randômico foi, respectivamente, 35%; 19% e 28%. Para as três arboviroses, dengue, zika e chikungunya, nenhuma das variáveis, na análise subgrupos e metaregressão, explicou a grande heterogeneidade encontrada. No artigo original, foram utilizados dados secundários do SINAN disponibilizados pelo Ministério da Saúde e solicitados por meio do Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC). Esses dados passaram por uma etapa de padronização e organização. Para realização do modelo, foram incluídos apenas as pessoas com Classificação final confirmada para febre amarela. Para realização das análises foi utilizado o Google Colaboratory. Para o problema encontrado foram testados 8 algoritmos: Naive Bayes, Decision Tree, k-Nearest Neighbour, Logistic Regression, Multilayer Perceptron, Support Vector Machine, Random Forest e CatBoost. Para nosso conjunto de dados, foi utilizada o percentual de 75% da base para treinamento e 25% para teste. Os parâmetros tiveram suas configurações modificadas. Para a validação do modelo utilizamos a Validação Cruzada 10 partes. Para visualização foi utilizada a biblioteca SHAP. O algoritmo selecionado para o modelo foi o CatBoost por apresentar métricas de desempenho melhores, por reduzir a necessidade de ajustes de hiperparâmetros, além de ser melhor com bancos categóricos e lidar com valores faltantes. Ao analisar o modelo, as variáveis que mais influenciaram no modelo foram resultado do exame imunológico de IgM, presença de distúrbio renal, presença de dor abdominal, critério de confirmação utilizado e presença de sinais hemorrágicos. O modelo proposto para previsão de óbito em pacientes com febre amarela teve um bom resultado com 81,1% de acurácia e 79,8% de precisão. O uso dessa ferramenta pode auxiliar os profissionais de saúde na prática clínica. A análise do comportamento do modelo permitiu aprofundar aspectos importantes do banco do SINAN confirmando alguns elementos de prognóstico já apontados pela literatura e indicando novas possibilidades de estudo a fim de aprofundar aspectos relacionados ao prognóstico da febre amarela e construir estratégias para um melhor manejo clínico desta enfermidade e prevenção do óbito.

Palavras-chave: prevalência. arbovirose. modelo preditivo. mineração de dados.

ABSTRACT

FRACALOSSO, Karen Oliveira, M.Sc., Universidade Federal de Viçosa, January, 2021. **Epidemiological study of arboviruses in Brazil and the use of artificial intelligence in the construction of a predictive model of lethality for yellow fever.** Adviser: Glauce Dias da Costa. Co-advisers: Rodrigo Siqueira Batista, Tiago Ricardo Moreira and Rosangela Minardi Mitre Cotta.

Arboviruses are diseases caused by arboviruses and represent a growing concern for public health in Brazil and around the world, mainly due to their potential for dispersal, their ability to adapt to new environments and hosts, and the possibility of causing extensive epidemics. The sharing of the same transmission vector between dengue, zika, chikungunya and yellow fever, the co-circulation of these diseases and the widespread spread of *Aedes aegypti* throughout Brazilian territory can cause emerging epidemics. Furthermore, the similarity between the symptoms presented for these diseases makes clinical management difficult. In this sense, epidemiological studies and the use of computational resources in the form of predictive models can improve health research, assessment of the quality of care, prevention, prognosis resulting in more effective treatments, as well as disease control, targeting of resources and improvements in management. Thus, the objective of this work was to estimate the prevalence of the arboviruses dengue, zika, chikungunya and yellow fever in the periods from 2000 to 2018 and build a predictive model for death from yellow fever in Brazil using artificial intelligence. In the first study, a systematic review with meta-analysis was carried out based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol. A search was carried out in the databases PubMed/MEDLINE, Scielo (Scientific Electronic Library Online), ScienceDirect and VHL (Virtual Health Library), under the search term (dengue OR zika OR chikungunya OR "yellow fever") AND prevalence). The meta-analysis was performed using a random effect model. Heterogeneity was assessed by the chi-square test (χ^2) with a significance of $P < 0.10$, with its magnitude determined by the I-square (I^2). The analyzes were carried out using the Metaprop and Metareg commands of the Stata program (version 13.0). Of the 8052 articles included in the search, 61 articles were selected. The years of published studies were between 1992 and 2019, with 65.57% of articles ($n=40$) being published in the last 10 years. Most of the selected articles mainly describe regional epidemiological data derived from surveys and studies conducted in specific regions, being used to support national data with regard to prevalence and serotype. It was not possible to carry out the meta-

analysis with yellow fever articles, as few studies were selected. The primary outcome was the prevalence of dengue, zika and chikungunya with a 95% confidence interval (95%CI). The pooled prevalence of dengue, zika and chikungunya in the random model was, respectively, 35%; 19% and 28%. For the three arboviruses, dengue, zika and chikungunya, none of the variables, in the subgroup and meta-regression analysis, explained the great heterogeneity found. In the original article, secondary data from SINAN made available by the Ministry of Health and requested through the Electronic System of the Citizen Information Service (e-SIC) were used. This data went through a standardization and organization stage. To create the model, only people with a confirmed final classification for yellow fever were included. Google Colaboratory was used to carry out the analyses. For the problem encountered, 8 algorithms were tested: Naive Bayes, Decision Tree, k-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Support Vector Machine, Random Forest and CatBoost. For our dataset, 75% of the base was used for training and 25% for testing. The parameters have had their settings modified. To validate the model, we used 10-part Cross Validation. The SHAP library was used for visualization. The algorithm selected for the model was CatBoost because it presents better performance metrics, reduces the need for hyperparameter adjustments, in addition to being better with categorical banks and dealing with missing values. When analyzing the model, the variables that most influenced the model were the result of the IgM immunological test, presence of renal disorder, presence of abdominal pain, confirmation criteria used and presence of hemorrhagic signs. The proposed model for predicting death in patients with yellow fever had a good result with 81.1% accuracy and 79.8% precision. The use of this tool can assist healthcare professionals in clinical practice. The analysis of the model's behavior allowed us to delve deeper into important aspects of the SINAN database, confirming some prognostic elements already highlighted in the literature and indicating new possibilities for studies in order to delve deeper into aspects related to the prognosis of yellow fever and construct strategies for better clinical management of this disease. and prevention of death.

Keywords: prevalence. arbovirus. predictive model. data mining.

SUMÁRIO

1. INTRODUÇÃO	10
2. REVISÃO DE LITERATURA	13
2.1. Arboviroses.....	13
2.1.1. Dengue.....	13
2.1.1.1. Epidemiologia	14
2.1.1.2. Aspectos Clínicos e Classificação.....	15
2.1.2. Zika.....	17
2.1.2.1. Epidemiologia	18
2.1.2.2. Aspectos Clínicos.....	19
2.1.3. Chikungunya.....	20
2.1.3.1. Epidemiologia	20
2.1.3.2. Aspectos clínicos.....	21
2.1.4. Febre Amarela.....	22
2.1.4.1. Epidemiologia	23
2.1.4.2. Patogenia	25
2.1.4.3. Aspectos Clínicos e Classificação.....	26
2.2. Sistema Nacional de Agravos de Notificação (SINAN).....	27
2.3. Mineração de Dados na Saúde.....	28
3. OBJETIVOS	30
3.1. Objetivo geral	30
3.2. Objetivos específicos.....	30
4. METODOLOGIA.....	31
4.1. Artigo de Revisão	31
4.2. Artigo Original	32
4.2.1. Knowledge Discovery in Databases – KDD	32
4.2.2. Dados.....	34

4.2.3.Análise Exploratória	35
4.2.4.Seleção de Atributos	36
4.2.5.Aprendizagem de Máquina.....	37
4.2.6.Algoritmos de Aprendizagem de Máquina.....	38
4.2.7.Processamento de Dados	40
4.2.8. Pós-processamento de Dados	41
4.2.9.Avaliação	42
5. RESULTADO E DISCUSSÕES	43
5.1. Artigo de Revisão - prevalência de dengue, zika, chikungunya e febre amarela no brasil: revisão sistemática e metanálise	44
5.1.1.Material Suplementar – Artigo de Revisão	71
5.2. Artigo Original – Febre Amarela no Brasil: estudo de prognóstico utilizando técnicas de Inteligência Artificial	89
CONSIDERAÇÕES FINAIS.....	122
REFERÊNCIAS	123
ANEXO A – Dispensa do comitê de ética	131
ANEXO B – Ficha de notificação de febre amarela	132

1. INTRODUÇÃO

As arboviroses, doenças transmitidas por artrópodes, têm causado um impacto significativo na saúde pública de todo o mundo (MAYER, TESH e VASILAKIS, 2017). As mudanças climáticas, a crescente urbanização, as adaptações e a disseminação dos vetores globalmente, principalmente do *Aedes aegypti* que é altamente antropofílico, são fatores relacionados com a extensa expansão geográfica dos arbovirus (KRAEMER, et al., 2015; KRAEMER, et al., 2019; WEAVER, REISEN, 2009). Os arbovirus da dengue, zika e chikungunya têm se expandido em todo o mundo de forma rápida e os mosquitos que transmitem essas doenças são grandes ameaças para a população humana em muitas regiões tropicais e subtropicais (MAYER, TESH e VASILAKIS, 2017; PATTERSON, SAMMON e GARG, 2016).

Na última década, diversos surtos de arboviroses transmitidas por *Aedes aegypti* ocorreram na Europa e América do Norte. Os surtos de febre amarela no Brasil, África e Europa, fizeram esse vírus voltar à visibilidade, e mostraram a fragilidade do sistema de vacinação atual (JÁCOME, et al., 2019). Casos de zika já eram relatados no pacífico Ocidental, Sudeste asiático e África antes da epidemia de 2015. O vírus chikungunya, que normalmente circulava na África e Ásia, em 2013 teve seus primeiros relatos registrados nas Américas, mantendo valores altos de casos confirmados até 2019. A febre amarela é uma doença endêmica das Américas Central e do Sul e da África, em 2016, teve casos exportados para outros países, e um aumento nos números de casos nas regiões consideradas endêmicas. A dengue é considerada endêmica em mais de 100 países nas regiões das Américas, Sudeste Asiático e Pacífico Ocidental. Os números de casos confirmados óbitos vem crescendo a números alarmantes em diversas regiões, com ocorrência recorrente de surtos e epidemias (WHO, 2022a; WHO, 2022b; WHO, 2022c; WHO, 2022d).

Os vetores dessas doenças, em especial o *A. aegypti*, estão difundidos em diversos territórios, por esse motivo são doenças que se espalham rápida e facilmente, trazendo aumento de custos para o sistema de saúde, bem como sequelas que podem comprometer as atividades ocupacionais dos pacientes (LIMA-CAMARA, 2016). Uma revisão realizada com estudos entre 2000 e 2018 apresenta valores de custos globais para dengue e febre amarela variando de 2,1 a 57,3 bilhões de dólares e custos dos programas com o controle de vetores variaram de 5,62 a 73,5 milhões de dólares (THOMPSON; MARTIN DEL CAMPO; CONSTENLA, 2020). No Brasil, em 2016, o custo total aproximado com o manejo das arboviroses foi 2,3 bilhões de reais, esse total inclui o investimento no combate do vetor, custos médicos diretos e custos indiretos, como a perda da produtividade e afastamento. O valor estimado do custo médico

direto para dengue, zika e chikungunya em 2019 excede os 3 bilhões de reais (PINHEIRO; DE BARROS ALVES; DA SILVA, 2020).

São muitos os desafios a serem superados para o controle dessas doenças. Para grande parte das arboviroses não existem vacinas como método profilático ou antivirais efetivos para o tratamento (LOPES, NOZAWA e LINHARES, 2014). Ainda se considerarmos que essas doenças não ocorrem de forma isolada e que a sintomatologia entre elas muitas vezes é parecida, a realização de um diagnóstico clínico preciso pode ser difícil (LOPES, NOZAWA e LINHARES, 2014; LIMA-CAMARA, 2016).

A imprecisão dos dados de incidência de doenças no Brasil é uma barreira importante no planejamento e execução de novos projetos (AVELINO-SILVA e RAMOS, 2017). A incompletude e duplicidade dos dados gerados pelo sistema de informações, aponta para uma necessidade de avaliação da qualidade da informação gerada pelos sistemas de informação em saúde. A má qualidade nesses dados prejudica a adoção de medidas de controle e o mapeamento das regiões de risco (SIQUEIRA, et al., 2020). Nesse sentido os estudos epidemiológicos possibilitam uma compreensão do cenário das doenças, a identificação de áreas de risco e as populações mais vulneráveis, o que é determinante para o desenvolvimento de políticas públicas (PORTO, et al., 2019).

Um bom sistema de informação, com dados consistentes e de qualidade são fundamentais para a tomada de decisão da gestão em saúde e pode contribuir para melhorias no manejo clínico e prognóstico dos pacientes. O Sistema de Informação de Agravos de Notificação (SINAN) é alimentado com as notificações de casos de doenças em todo o Brasil (LIMA, 2018). Esse grande volume de informações geradas anualmente, se bem geridas, podem ser utilizadas na redução de gastos com os sistemas de saúde e podem contribuir na resposta dos pacientes aos tratamentos (HILLESTAD, et al., 2005; ABOUZAHAR, ADJEI e KANCHANACHITRA, 2007; RUDD, et al., 2019).

Vale ressaltar, no entanto, que a utilização desses dados enfrenta pode apresentar problemas, entre eles: problemas com diferenças nas estruturas dos dados, dificuldades na padronização, problemas na segurança, no armazenamento e na transferência dos dados, além da falta de treinamento dos operadores no preenchimento das informações, o que pode levar a dados errados, duplicados ou faltantes, entre outros (KRUSE, et al., 2020; DATASUS, 2021).

Nesse sentido, o uso da análise preditiva, pode ser uma ferramenta que auxilie na análise e interpretação dos dados, visto que, essa metodologia usa algoritmos para compreender os dados existentes e gerar regras para as associações. Na área da saúde, os modelos preditivos podem ser usados de diversas formas como formas de auxiliar os profissionais de saúde na toma

de decisões (SANTOS, et al. 2019). Temos, por exemplo, o trabalho de Ranapurwala, et al., 2019 utilizando modelos preditivos para prever o risco de ferimentos ou morte em acidentes com veículos agrícolas. Já Mohamdlou, et al., 2018 desenvolveram um modelo de predição de lesão renal aguda usando dados de registros eletrônicos de saúde. Ainda, Senanayake, et al., 2019 fizeram uma revisão sobre modelos preditivos para previsão de falha do enxerto após o transplante renal, na qual 18 estudos foram incluídos. Por esse motivo, a construção de modelos de previsão podem ser uma importante ferramenta no processo de diagnóstico de pacientes, controle das doenças e direcionamento dos recursos. Os modelos de previsão vêm como uma opção para simplificar problemas complexos com o auxílio de algoritmos de aprendizagem de máquina. Dessa forma, novas informações podem ser geradas e consequentemente utilizadas na atuação profissional dos profissionais de saúde e gestão do serviço de saúde (PIROS, 2019; RANAPURWALA, et al., 2019).

2. REVISÃO DE LITERATURA

2.1. ARBOVIROSES

As arboviroses são doenças transmitidas por artrópodes. Os artrópodes hematófagos que são os vetores mais frequentes incluem mosquitos, carrapatos, flebotomíneos e, possivelmente, ácaros. Existe uma grande variedade de arbovírus; os mais comuns no Brasil são vírus Dengue, vírus Chikungunya, vírus Zika e vírus Febre Amarela. A co-circulação de dengue, febre amarela e outros flavivírus e a ocorrência de sucessivas infecções por flavivírus diferentes têm resultado em reatividade cruzada nos testes sorológicos disponíveis, isso impacta dificultando o diagnóstico específico dessas doenças e na vigilância das arboviroses (MEDEIROS e VASCONCELOS, 2019).

As arboviroses urbanas muitas vezes possuem sinais e sintomas semelhantes o que dificulta a suspeita inicial da doença, o que pode comprometer o manejo clínico adequado, levando o paciente ao agravamento e conseqüentemente à morte. Por esse motivo, a identificação dos casos suspeitos e o diagnóstico precoce são tão difíceis. A dengue, zika, chikungunya e febre amarela são doenças de notificação compulsória e suas informações vão ser usadas para alimentar o Sistema de Informação de Agravos de Notificação (SINAN) (BRASIL, 2019a).

2.1.1. Dengue

A dengue é uma doença febril aguda com um amplo espectro de sintomas. Ela pode ser assintomática ou em casos graves levar ao óbito. É a arbovirose urbana mais prevalente nas Américas, incluindo o Brasil. O cenário epidemiológico atual do Brasil conta com a circulação dos 4 sorotipos de dengue. O agente etiológico é o vírus da dengue (DENV), vírus de RNA do gênero *Flavivirus*, pertencente à família *Flaviviridae*, O vírus da dengue é transmitido por mosquitos fêmeas, principalmente da espécie *A. aegypti* e, em menor proporção, da espécie *Aedes albopictus*. (BRASIL, 2019a; OPAS Brasil, 2019a).

Um mosquito infectado, depois de um período de incubação do vírus de 4-10 dias, é capaz de transmitir o vírus por toda sua vida. Os humanos são os principais portadores e multiplicadores do vírus, quando infectados os pacientes podem transmitir a infecção pelo vetor por 4-12 dias (OPAS Brasil, 2019a).

O mosquito *A. aegypti* prefere se alimentar durante o dia, sendo os picos de atividade no início da manhã e antes de anoitecer. Se reproduzem em recipiente com água parada e vivem em habitats urbanos. Enquanto que o *Aedes albopictus*, principal vetor da Ásia, é altamente adaptável e pode sobreviver a temperaturas muito baixas. A tolerância a temperaturas abaixo de zero, hibernação e capacidade de se abrigar em micro-habitats, são características que explicam sua propagação (OPAS Brasil, 2019a).

Conforme dispõe a Portaria de Consolidação nº 4, de 28 de setembro de 2017, dengue, chikungunya e zika são doenças de notificação compulsória, ou seja, todo caso suspeito e/ou confirmado deve ser obrigatoriamente notificado ao Serviço de Vigilância Epidemiológica da Secretaria Municipal de Saúde (SMS) (BRASIL, 2019a; NUNES, et al., 2019).

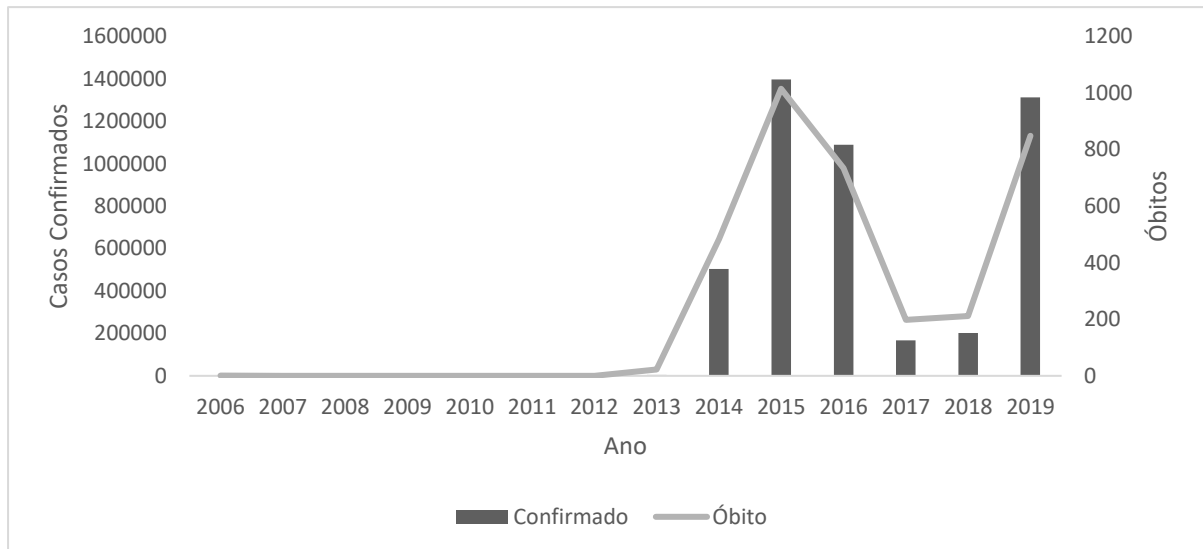
2.1.1.1.Epidemiologia

Existem mais de 100 países nos quais a dengue é uma doença endêmica. Aproximadamente 3,2 milhões de casos de dengue ocorreram em 2015 nas Américas, Sudeste Asiático e Região do Pacífico Ocidental. Na região do Pacífico Ocidental em 2016 foram registrados 375.000 casos suspeitos. A Índia, Indonésia, Mianmar, Sri Lanka e Tailândia estão entre os 30 países mais endêmicos do mundo para dengue. De 2015 a 2019, os casos de dengue na região do Sudeste Asiático aumentaram 46% (de 451.442 para 658.301), enquanto as mortes diminuíram 2% (de 1.584 para 1.555). Nas Américas entre 2014 a 2020 o ano de 2019 tem o maior registro com 3.190.778 casos relatados (OPS, 2022; WHO, 2022b).

Os primeiros sintomas compatíveis com a dengue no Brasil foram relatados em 1846, no início da década de 80 houve o primeiro surto da doença no país, após 4 anos sem casos da doença em 1986 houve um novo surto de dengue, e com o passar dos anos os casos foram aumentando e novos sorotipos da doença foram aparecendo no país (NUNES, et al., 2019).

Dados do Ministério da Saúde mostram que houve um aumento expressivo no número de casos confirmados e óbitos por dengue a partir de 2014, com queda em 2017 e 2018, e aumento em 2019 (Figura 1) (BRASIL, 2024).

Figura 1 - Número de casos confirmados e de óbitos por dengue de 2006 - 2019, Brasil.



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

2.1.1.2. Aspectos Clínicos e Classificação

A infecção pelo vírus dengue pode ser sintomática ou assintomática. Quando sintomática as manifestações podem variar de um quadro oligossintomático até quadros graves. A doença apresenta três fases clínicas: **Fase febril** - Os sintomas mais comuns são: febre acima de 38 °C, cefaleia, astenia, mialgia e dor retro orbitária. O exantema está presente em 50% dos casos. Anorexia, náuseas e vômitos podem estar presentes. **Fase crítica** - Tem início com a queda da febre. Pacientes nessa fase podem evoluir para as formas graves, acompanhada do surgimento dos sinais de alerta (dor abdominal intensa, vômito persistentes, acúmulo de líquido, sangramento de mucosa, letargia e/ou irritabilidade, hipotensão postural e/ou lipotimia, hepatomegalia, aumento progressivo do hematócrito); **Fase de recuperação** - Fase em que ocorre o retrocesso dos sintomas, melhora do estado hemodinâmico e do apetite. (BRASIL, 2019a; BRASIL 2016).

Atualmente os pacientes com dengue são divididos em grupos de acordo com o estadiamento clínico da doença (A, B, C e D). Os pacientes do **grupo A** são casos suspeitos de dengue com ausência de sinais de alarme, sem apresentar comorbidades, não são pessoas do grupo de risco e não apresentam condições clínicas especiais; os do **grupo B** são caso suspeito de dengue, com ausência de sinais de alarme, mas com sangramento espontâneo de pele (petéquias) ou induzido (prova do laço positiva); pessoas com comorbidades, são pessoas do grupo de risco e que apresentam condições clínicas especiais. As pessoas do **grupo C** são casos suspeitos que apresentem algum dos sinais de alarme; e o **grupo D** é formado por casos

suspeitos de dengue com presença de sinais de choque, sangramento grave ou disfunção grave de órgãos (BRASIL, 2016).

O tratamento é sintomático sendo a reposição volêmica adequada a principal preocupação para pacientes com dengue grave, considerando o estadiamento clínico da doença (BRASIL, 2019a). No diagnóstico clínico deve ser considerado caso suspeito o indivíduo que apresente, resida ou tenha frequentado a área afetada, que apresente febre entre 2 e 7 dias e duas ou mais das seguintes condições: náusea/vômitos, exantema, mialgia/artralgia, cefaleia/dor retro orbital, petéquias/prova do laço positiva, leucopenia (BRASIL, 2019a).

A sintomatologia de dengue é parecida com outras doenças. Por isso, é necessário o diagnóstico diferencial com as síndromes febris (enterovirose, influenza e outras viroses respiratórias, hepatites virais, malária, febre tifoide, chikungunya, zika e outras arboviroses); as síndromes exantemáticas febris (rubéola, sarampo, escarlatina, eritema infeccioso, etc.); as síndromes hemorrágicas febris (hantavirose, febre amarela, leptospirose, etc.); as síndromes dolorosas abdominais (apendicite, obstrução intestinal, etc.); a síndrome do choque (meningococemia, septicemia, etc.); e as síndrome meníngeas (meningites virais, meningite bacteriana e encefalite) (BRASIL, 2016).

Com o cenário brasileiro de concomitância de dengue, zika e chikungunya, algumas especificidades do diagnóstico diferencial entre essas doenças é explicado no Quadro 1 abaixo.

Quadro 1 - Diagnóstico diferencial dengue versus zika versus chikungunya

(continua)

Sinais/sintomas	Dengue	Zika	Chikungunya
Febre	Febre alta (>38 °C)	Sem febre ou febre baixa (<=38 °C)	Febre alta (>38,5 °C)
Duração	2-7 dias	1-2 dias subfebril	2-3 dias
Exantema	Surge do 3º ao 6º dia	Surge do 1º ao 2º dia	Surge do 2º ao 5º dia
Mialgias (frequência)	+++	++	++
Artralgia (frequência)	+	++	+++
Artralgia (intensidade)	Leve	Leve/moderada	Moderada/intensa
Edema da articulação (frequência)	Raro	Frequente	Frequente

Quadro 1 - Diagnóstico diferencial dengue versus zika versus chikungunya

(conclusão)

Sinais/sintomas	Dengue	Zika	Chikungunya
Edema da articulação (intensidade)	Leve	Leve	Moderado/intenso
Conjuntivite	Raro	50 - 90% dos casos	30%
Cefaleia	+++	++	++
Linfonomegalia	+	+++	++
Discrasia hemorrágica	++	Ausente	+
Acometimento neurológico	+	+++	++
Leucopenia	+++	++	++
Linfopenia	Incomum	Incomum	Frequente
Trombocitopenia	+++	+	++

Fonte: BRASIL, 2019a

Com o passar dos anos, as referências de classificação da dengue foram mudando. O Brasil já adotou 3 referências: a da Organização Mundial da Saúde (OMS) de 1997, que classificava a doença em febre hemorrágica da dengue e síndrome do choque da dengue; em 2000, o Ministério da Saúde do Brasil passou a adotar a classificação de dengue com complicações; e em 2009 a OMS passou a classificar a doença em dengue com sinais de alerta e dengue grave. O Brasil só passou a adotar essa classificação em 2014 (NUNES, et al., 2019).

2.1.2. Zika

A zika é uma arbovirose transmitida pelas fêmeas do *A. aegypti*, causada pelo vírus zika (ZIKV). O vírus zika foi isolado pela primeira vez em 1947 em sangue de macaco. Por muito tempo a infecção por zika vírus foi considerada benigna. Em relatos de casos humanos, os sintomas mais comuns são febre, mialgia, dor de cabeça, conjuntivite e erupção cutânea. No entanto, em casos raros, é possível que ocorra o agravamento da doença com a evolução ao óbito (BRASIL, 2019a; CARDONA-OSPINA, et al., 2019)

A zika é uma doença de notificação compulsória, ou seja, todo caso suspeito e/ou confirmado deve ser obrigatoriamente notificado ao Serviço de Vigilância Epidemiológica da Secretaria Municipal de Saúde (SMS) (BRASIL, 2019a).

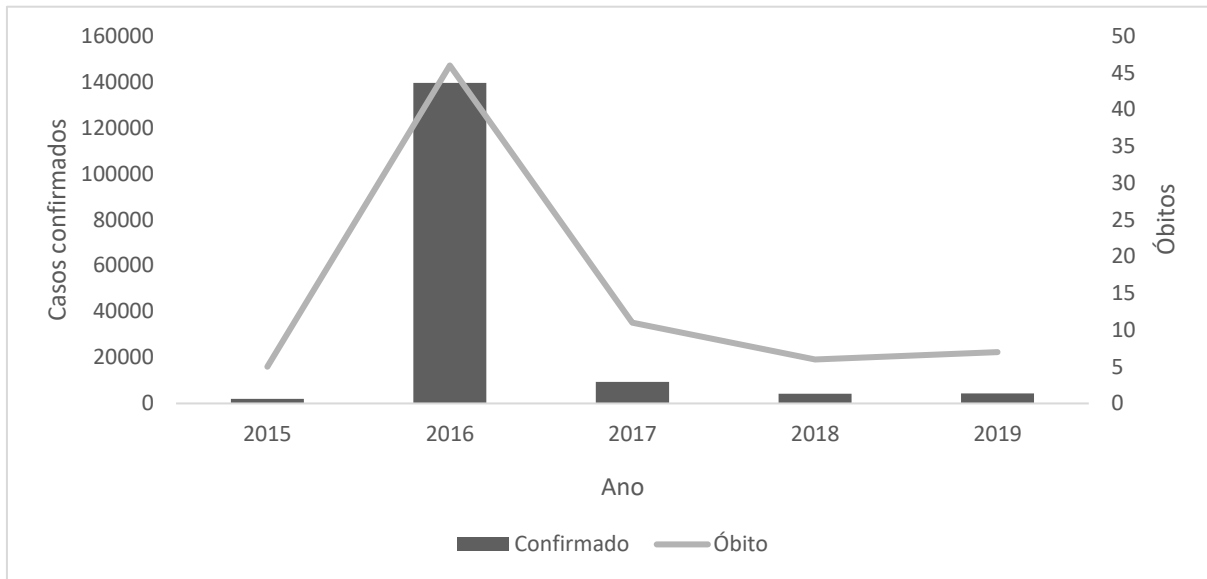
Além da forma vetorial, pela picada do mosquito, a doença apresenta outros meios de transmissão: sexual, pós-transfusional e vertical. O vírus zika como outros flavivírus é neurotrópico, também apresenta potencial teratogênico, por isso o vírus zika é causa de uma série de distúrbios neurológicos, anomalia em fetos e também má formações congênitas, incluindo a microcefalia (BRASIL, 2019a). Desde 2015 com o aumento nos casos de microcefalia devido a epidemia de zika, principalmente no nordeste brasileiro, uma grande rede de informação e ação foi realizada com o objetivo de melhor atender essa população. No entanto é uma preocupação que com a redução no número de casos de zika parte dessa capacidade operacional que foi estruturada em uma situação de crise vá se perdendo e a população fique menos assistida (PAIXÃO, et al., 2019).

2.1.2.1.Epidemiologia

Em relação ao zika vírus, até julho de 2019 casos autóctones de transmissão foram relatados em 87 países ou territórios. Na região do Pacífico Ocidental casos da doença são documentados desde 2007. Antes da epidemia de 2015, na Região Africana, 10 países já tinham casos relatados de zika, 4 países do Sudeste Asiático e 9 territórios do Pacífico Ocidental, durante a epidemia mais 8 territórios desta região relataram casos. Nenhum país do Mediterrâneo Oriental relatou casos autóctones. A incidência da infecção por zika nas Américas teve pico em 2016 (652.207 casos relatados) e foi diminuindo nos próximos anos, os únicos países do continente americano não afetados foram o Chile continental, Uruguai e Canadá (WHO, 2022d; OPS, 2024).

Os casos de zika no Brasil tiveram pico em 2016, com 139.642 casos e 46 óbitos confirmados por zika registrados (Figura 2). A Bahia é o estado que mais concentra casos de zika, com valores próximos a 50%. Em 2019 houve um pequeno aumento no número casos e óbitos confirmados pela doença (BRASIL, 2018; BRASIL,2019a; BRASIL,2019b; BRASIL, 2024).

Figura 2 – Número de casos confirmados e de óbitos por zika de 2015 - 2019, Brasil.



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

2.1.2.2. Aspectos Clínicos

Estima-se que apenas 50% das pessoas infectadas com zika apresentem sintomas. As pessoas que apresentam sintomas podem apresentar sintomas brandos, até complicações neurológicas e má formações congênitas em fetos. O período de incubação varia de 2-7 dias. A doença é autolimitada e dura de 4-7 dias. Os sintomas mais comuns são: febre baixa ($\leq 38,5^{\circ}\text{C}$) ou ausente, exantema (geralmente pruriginoso e maculopapular craniocaudal) de início precoce, conjuntivite não purulenta, artralgias. Pode ocorrer ainda edema periarticular, cefaleia, linfonomegalia, astenia, mialgia, dor abdominal, náuseas e diarreia (BRASIL, 2019a; FALCÃO, et al., 2016).

Gestantes infectadas podem transmitir o vírus ao feto, o que pode resultar em aborto espontâneo, óbito fetal e malformações congênitas. O aparecimento de quadro neurológicos em qualquer ponto é motivo de atenção (BRASIL, 2019a).

O tratamento da doença semelhante a outras arboviroses é sintomático. É recomendado que o paciente se hidrate e não faça uso de medicamentos que contenham ácido acetilsalicílico (BRASIL, 2019a). É preciso se atentar que a zika apresenta quadro clínico semelhante ao de dengue e outros arbovírus, incluindo trombocitopenia e sangramento, choque ou doença neurológica grave, incluindo encefalite, por isso é necessário exame diferencial (CARDONA-OSPINA, et al., 2019).

Devido a sintomatologia parecida com outras doenças é necessário o diagnóstico diferencial para dengue, chikungunya, eritema infeccioso, rubéola, sarampo, riquetsioses, malária, leptospirose, enterovirose, sífilis secundária, reação de hipersensibilidade, síndrome da mononucleose, reações alérgicas e doenças autoimunes. Para confirmação da zika é necessário a execução de exames específicos (BRASIL, 2019a).

2.1.3. Chikungunya

O vírus chikungunya (CHIKV) é um arbovírus reemergente que pertence ao gênero *Alphavirus* da família *Togaviridae*. A principal forma de transmissão da chikungunya é vetorial, mas em raras situações a transmissão vertical pode acontecer, e em casos de mães acometidas no período perinatal podem passar o vírus na hora do parto. Apesar disso o vírus não interfere no curso da gravidez e o vírus não é transmitido pelo leite materno (BRASIL, 2019a; WIMALASIRI-YAPA, et al., 2019).

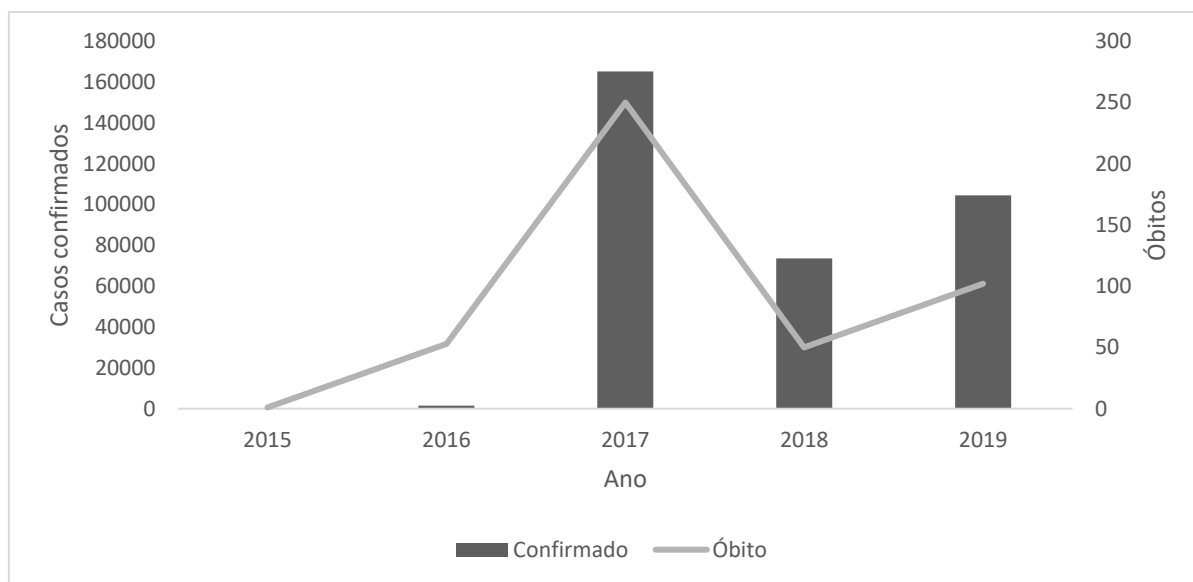
Foi introduzido no continente americano em 2013, o Brasil tem relatos da doença a partir de 2014, com casos autóctones no Amapá e na Bahia, desde então todos os estados da federação apresentam casos autóctones (BRASIL, 2019a; KIM, et al., 2019). Hoje a doença faz parte da lista de doenças de notificação compulsória (BRASIL, 2019a).

2.1.3.1. Epidemiologia

O vírus Chikungunya é conhecido por circular na África e Ásia. Os primeiros relatos foram registrados em 1824. Em 1999, foram registrados surtos na Indonésia. Em 2007 casos foram relatados na Itália. No Quênia em 2016, houve um surto de chikungunya tendo mais de 1700 casos registrados até o primeiro semestre deste ano. O Senegal registrou circulação do vírus chikungunya em 2015, os últimos registros na região foram relatados entre 2009 e 2010. Em 2019, o Congo declarou surto. Em 2013 o vírus foi detectado na Região das Américas, nas Ilhas de São Bartolomeu e São Martinho, Guadalupe e Martinica e se espalhou por vários territórios do continente. De 2013 a 2017 foram registrados nas américas 2.673.671 casos, o ano de 2014 representa 40,8% dos casos desse período. No período de 2018 a 2019 foram confirmados 275.770 casos de chikungunya, destes 36,7% em 2019 (PAHO, 2022a; WHO, 2022a).

No Brasil o pico no número de casos e óbitos ocorreu em 2017. Foram 165.292 casos e 250 óbitos confirmados pela doença. Em 2018 houve queda e em 2019 aumento nos casos e óbitos confirmados por chikungunya (Figura 3).

Figura 3 – Número de casos confirmados e de óbitos por chikungunya de 2015 - 2019, Brasil.



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

2.1.3.2. Aspectos clínicos

A maioria dos pacientes com chikungunya desenvolve sintomas, cerca de 70% das pessoas são sintomáticas. O período de incubação pode variar entre 1-12 dias. A doença possui um amplo espectro de manifestações clínicas e pode afetar todas as faixas etárias. No entanto, em crianças e idosos os casos costumam ser mais graves, podendo levar ao óbito mesmo na ausência de comorbidades (WIMALASIRI-YAPA, et al., 2019).

Como a sintomatologia da doença é similar à outras doenças febris são necessários exames específicos para o diagnóstico. Também deve ser feito o diagnóstico diferencial com outras doenças febris agudas associadas a artralgia. São testadas malária, leptospirose, febre reumática, artrite séptica, zika, mayaro, oropouche, sífilis, lúpus eritematosos sistêmicos, artrite reumatoide, entre outras, podem também cursar com febre e dores articulares (BRASIL, 2017b; BRASIL, 2019a).

Não existe tratamento específico para a doença, são recomendados cuidados com hidratação e repouso e analgesia e suporte. A fisioterapia é parte importante no tratamento da doença e deve começar ainda na fase aguda. Os anti-inflamatórios não esteroides,

corticosteroides e o ácido acetilsalicílico não devem ser utilizados na fase aguda da doença (BRASIL, 2017b; BRASIL, 2019a).

2.1.4. Febre Amarela

A febre amarela é uma doença infecciosa febril aguda e prevenível por meio da vacinação. O vírus da febre amarela (YFV) é um arbovírus do gênero *Flavivirus*, protótipo da família Flaviviridae. A febre amarela é uma doença com elevada importância epidemiológica devido à possibilidade de fácil disseminação da doença, principalmente quando a transmissão é urbana por *A. aegypti*, e também pela elevada letalidade das suas formas mais graves (BRASIL, 2019a).

Todo caso suspeito da doença deve ser informado às autoridades competentes visto que a doença é de notificação compulsória e imediata. Para isso, a Ficha de Investigação da Febre Amarela deve ser preenchida e suas informações adicionadas ao Sistema de Informação de Agravos de Notificação (SINAN) (BRASIL, 2019a).

A febre amarela possui dois ciclos de transmissão, um selvático ou silvestre e outro urbano. No ciclo silvestre os primatas não humanos são os principais hospedeiros, amplificadores do vírus e vítimas da doença. Os principais mosquitos silvestres envolvidos na transmissão da doença são *Haemagogus janthinomys* e *Haemagogus leucocelaenus*, além de diversas espécies do gênero *Sabethes*. Basicamente o ciclo selvagem envolve a transmissão entre macacos e mosquitos selvagens. Já quanto ao ciclo urbano o mosquito envolvido é o *A. aegypti*, ele é o responsável pela transmissão da doença, sendo o ser humano o único hospedeiro (BRASIL, 2019a; DO CARMO CUPERTINO, et al. 2019).

A maior frequência da doença ocorre entre dezembro e maio, quando a densidade vetorial é maior devido as chuvas, que coincidem com a época de maior atividade agrícola. A doença acomete com maior frequência pessoas do sexo masculino e faixa etária superior a 15 anos, pois está associada a prática laboral desse grupo ter maior exposição às áreas de risco da doença, locais com matas, cachoeiras, parques, rios e que possuam hospedeiros e vetores disponíveis. Outro grupo de risco são as pessoas não vacinadas que residem próxima de ambientes silvestres (CÂMARA; DE CARVALHO; GOMES, 2013).

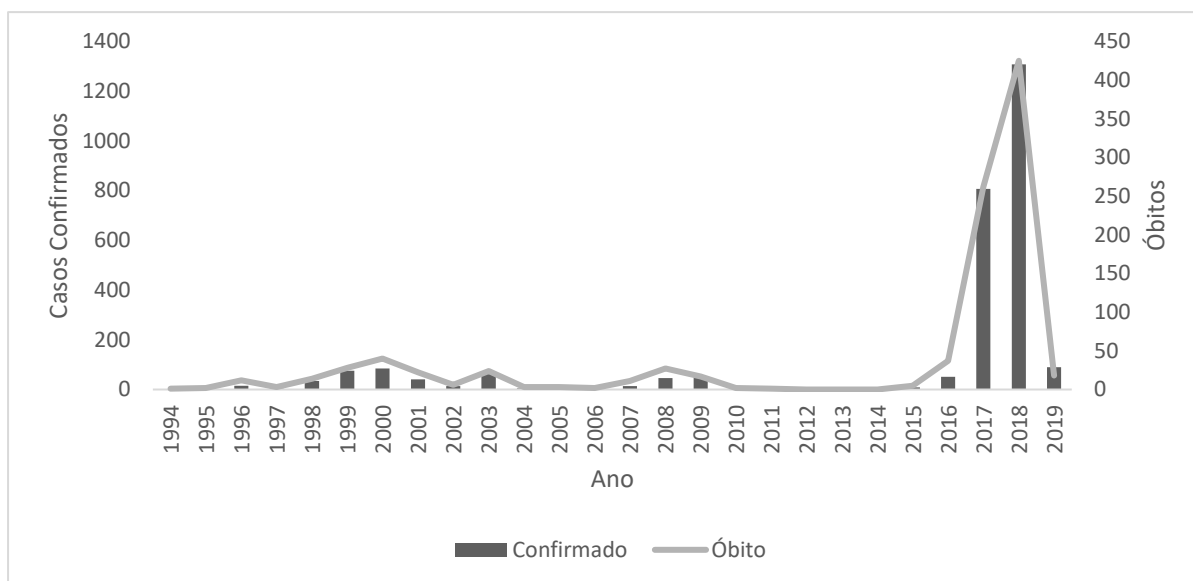
2.1.4.1.Epidemiologia

A febre amarela é endêmica em 34 países na África e 13 nas Américas Central e do Sul. A OMS estima 200.000 casos da doença e 30.000 mortes por ano. Em 2016, dois surtos, um em Luanda e outro na Angola, geraram casos que foram exportados para outros países, dentre os quais a China. Nas Américas, em 2018, 14 países possuíam condições para transmissão da febre amarela. Entre 1960 e 2019, foram relatados 19.360 casos da doença, o ano com maior registro foi 2018 com 1.401 casos. Em 2019, a África confirmou 216 casos de febre amarela e a América confirmou 89 casos (WHO, 2020; OPAS, 2022c; PAHO, 2022b; WHO, 2022).

Os primeiros casos supostos de febre amarela no Brasil são datados de 1640, por meio de documentos consulares de Pernambuco. Acredita-se que o vírus tenha chegado ao Brasil por meio das embarcações e comércio de escravos (FRANCO, 1969).

Entre 1960 e 2019 foram notificados no Brasil 3830 casos de febre amarela e 408 óbitos pela doença, destes, 1266 casos confirmados e 415 óbitos ocorreram no período de julho de 2017 à maio de 2018. Esse foi o evento mais expressivo da doença no país, no qual o vírus alcançou a costa leste brasileira, principalmente os estados de Minas Gerais, São Paulo e Rio de Janeiro. Outros processos de ocorrência da doença, menos expressivos, foram relatados em 1998/1999 no Pará; 2002/2003 em Minas Gerais e no Rio Grande do Sul; no Centro-Oeste em 2007/2008; e em 2008/2009 em São Paulo, no Paraná e Rio Grande do Sul e um surto em 2016/2017. Devido aos surtos recentes, foi observada a expansão do vírus para municípios que não eram considerados áreas de risco, por esse motivo uma série de medidas tiveram de ser tomadas, incluindo a expansão das áreas com recomendação para vacinação (Figura 4) (BRASIL, 2019a; CALADO e PAZ, 2020).

Figura 4 – Número de casos confirmados e óbitos por febre amarela no Brasil de 1994 – 2019



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

A vacinação é uma importante medida de prevenção e controle da doença. A vacina vem sendo produzida no Brasil desde 1937, pelo Instituto em Imunobiológicos Bio-Manguinhos. Utilizando-se da cepa 17DD e constituída por derivados de uma amostra atenuada do vírus amarelíco selvagem africano. Apresenta uma eficácia superior a 95%, e é reconhecidamente segura. Contudo, alguns eventos adversos podem ocorrer, devendo-se ter atenção aos indivíduos que apresentarem dor abdominal intensa pois esse é u sinal de agravo da doença (BRASIL, 2019a).

Em 2017, o número de doses de vacinas oferecidas, cresceu no Brasil, bem como a cobertura de vacinação, principalmente devido ao surto nesse ano. Contudo nos anos seguintes esses valores não se mantiveram, apesar da reincidência dos casos nos anos de 2018 e 2019 e dos frequentes alertas da Organização Mundial da Saúde, principalmente aos viajantes (OPAS Brasil, 2019b; BRASIL, 2021).

Desde 2016 o Brasil observa uma queda na cobertura vacinal devido a diversos fatores. Essa hesitação na vacinação pode levar a reintegração de diversas doenças já controladas no Brasil, ou frequentes epidemias de doenças imunopreveníveis. Isso além de acarretar um enorme risco para a saúde da população, gera um encargo financeiro enorme ao sistema de saúde (SATO, 2018).

2.1.4.2. Patogenia

Apesar de ser uma doença antiga, ainda se desconhece muitos aspectos da doença, inclusive de sua patogênese. Os recentes surtos no Brasil ocasionaram uma oportunidade ímpar de pesquisar algumas características clínicas da doença, mas ainda há muito a ser estudado (SACCHETTO, et al., 2018).

Após a picada do inseto transmissor, o vírus vai para a corrente sanguínea e rapidamente atinge os linfonodos regionais, saindo da circulação pelas próximas 24 horas. Nos linfonodos, o vírus vai infectar preferencialmente as células linfoides e macrófagos. Depois do processo de replicação viral, as partículas virais vão voltar para a corrente sanguínea pelos vasos linfáticos e então vão alcançar o fígado. O período de viremia varia bastante de acordo com as manifestações clínicas e coincide com o período prodrômico na doença (VASCONCELOS, 2002).

No fígado o vírus vai infectar as células de Küpffer causando degeneração acidófila em zonas focais ainda no início do período de replicação viral. Após esse evento, ocorre degeneração baloniforme e depois causa necrose do tipo hialina. Nos hepatócitos, células também infectadas, o vírus causa necrose em grandes extensões do parênquima hepático (VASCONCELOS, 2002).

As lesões são localizadas nos lóbulos hepáticos e são seletivas com as estruturas que afetam, visto que as áreas entre o espaço porta e a veia centrolobular são mais atingidas pela necrose. Esse tipo de lesão é muito associada com a infecção amarílica, apesar de poder estar presente em outras viroses. A infecção causa uma desorganização da arquitetura hepática normal, com destruição das traves de Remack. Contudo, por vezes a necrose é muito extensa o que compromete o diagnóstico histopatológico (VASCONCELOS, 2002).

Na área necrosada é visível o infiltrado inflamatório com predomínio de células mononucleares, debris celulares e vários tipos e graus de lesões degenerativas. A degeneração hialina, acidófila dos hepatócitos é um forte indicativo da febre amarela, apesar de não ser característico apenas dessa doença. Outro sinal comum é a ocorrência de degeneração gordurosa (esteatose), encontrada em células necrosadas e/ou preservadas. Os corpúsculos de Torres e Villela podem ser vistos em hepatócitos, células de Küpffer e macrófagos com maior raridade. Exames de necrópsia apontam que a histopatologia da febre amarela é semelhante nos humanos e macacos (VASCONCELOS, 2002).

2.1.4.3. Aspectos Clínicos e Classificação

Os principais sintomas da doença são febre alta, dor de cabeça forte e duradoura, inapetência, náuseas e mialgia, geralmente com icterícia. No entanto, as características clínicas podem variar podendo a doença ser assintomática ou levar ao óbito. Os sinais clínicos mais comuns na forma mais grave é uma insuficiência renal e hepática. O tratamento da doença é principalmente sintomático. No entanto, para as formas mais graves é recomendado a hospitalização com o intuito de reduzir as complicações e o risco de morte (BRASIL, 2019a; DO CARMO CUPERTINO, et al. 2019).

Mais de 90% dos casos são oligossintomáticos ou assintomáticos. As formas assintomáticas ou os casos subclínicos e formas leves da doença, costumam ocorrer em crianças de pouca idade a qual a mãe foi vacinada e transmitiu a imunidade por via transplacentária (VASCONCELOS, 2002).

Quando hospitalizados, o perfil dos indivíduos em maior risco muda. Estudos que avaliaram o perfil de mortalidade dos pacientes com febre amarela, observaram maior número de mortes em pacientes de maior idade; contagem elevada de neutrófilos, aumento da aspartato aminotransferase (AST); maior carga viral; alterações na excreção renal; alterações na alanina aminotransferase (ALT); Icterícia elevada³¹; alterações no tempo de protrombina; Encefalopatia hepática grau IV; Pior índice APACHE II (*Acute Physiology And Chronic Health Evaluation II*) (TUBOI, et al., 2007; KALLAS, et al., 2019; ESCOSTEGUY, et al., 2019; DE ÁVILA, et al., 2020).

No diagnóstico clínico, deve ser considerado caso suspeito o indivíduo que apresente exposição a área afetada, com quadro febril agudo de até sete dias, acompanhado de dois ou mais dos sintomas a seguir: cefaleia, mialgia, lombalgia, mal-estar, calafrios, náuseas, icterícia ou manifestações hemorrágicas, e seja residente de uma região de risco, que não possa apresentar comprovante de vacinação, ou que tenha sido vacinado nos últimos 30 dias (LITVOC, NOVAES e LOPES, 2018). O quadro clínico pode ser classificado como leve, moderado, grave ou maligno (Quadro 2)

Quadro 2 - Manifestações clínicas e laboratoriais comuns da febre amarela

Forma	Sinais e sintomas	Alterações laboratoriais
Leve/ moderada	Febre; cefaleia; mialgia; náuseas; icterícia ausente ou leve	Plaquetomia; elevação moderada de transaminases; bilirrubinas normais ou discretamente elevadas; trombocitopenia.
Grave	Todos os anteriores Icterícia intensa Manifestações hemorrágicas; oligúria; diminuição de consciência	Plaquetopenia intensa; aumento da creatinina; elevação importante de transaminases; trombocitopenia intensa; aumento das transaminases.
Maligna	Todos os sintomas da forma grave intensificados	Todos os anteriores Coagulação intravascular disseminada com consumo de fibrinogênio; acúmulo de dímero D

Fonte: BRASIL, 2017c; LITVOC, NOVAES e LOPES, 2018.

Em uma situação de epidemia geralmente é fácil diagnosticar clinicamente a febre amarela, mas quando os casos são esporádicos é possível confundir a febre amarela com outras doenças, principalmente as que evoluem com casos de febre hemorrágica, icterícia e com comprometimento hepato-renal. Logo nos testes diferenciais deve ser testado a malária, hepatites virais, febre tifoide, dengue hemorrágico, septicemias, púrpura trombocitopênica idiopática e algumas formas de envenenamento, inclusive a picada de cobras com peçonha que induz a presença de hemorragias, que podem apresentar sintomatologia semelhante (VASCONCELOS, 2002).

2.2.SISTEMA NACIONAL DE AGRAVOS DE NOTIFICAÇÃO (SINAN)

O SINAN é um sistema que foi implantado no Brasil no início da década de 1990; no entanto, não havia uma coordenação e acompanhamento do sistema por parte das três esferas do governo. É a partir de 1998, que o uso do SINAN foi regulamentado pela Portaria GM/MS nº 1882 de 18 de dezembro de 1997, sendo então obrigatória a alimentação regular no sistema em todo território nacional (BRASIL, 2006; BRASIL, 1997).

Deve ser alimentado principalmente pelas notificações e investigação de casos de doenças e agravos que constam da lista nacional de doenças de notificação compulsória (Portaria de Consolidação nº 4, de 28 de setembro de 2017, anexo V - Capítulo I), no entanto os estados e

municípios tem liberdade para incluir as enfermidades que sejam de importância para sua região, grande parte das arboviroses estão nesta lista (BRASIL, 2017; PORTAL SINAN, 2019).

O uso do sistema de forma descentralizada, permite que os profissionais de saúde de várias partes do país tenham acesso aos dados contribuindo para a democratização da informação. Logo, o sistema é um instrumento para auxiliar o planejamento da saúde, definir prioridades de intervenção, e permite que seja avaliado o impacto das intervenções (PORTAL SINAN, 2019).

A real utilização do sistema permite a realização do diagnóstico dinâmico da ocorrência de um evento na população; pode indicar os motivos para os agravos de notificação compulsória, além de contribuir no esclarecimento de riscos aos quais os cidadãos estão sujeitos. Dessa forma podem contribuir para a determinação da real situação epidemiológica de área ao qual se referem os dados (BRASIL, 2006).

Entretanto, o desenvolvimento e aplicação de um sistema de informação em saúde possui uma série de desafios como o desconhecimento das etapas do processo pelos alimentadores do sistema e a despreocupação com a correta alimentação desses sistemas. Outro problema é a ilusão que a utilização de um recurso informatizado vai solucionar todos os problemas, sendo que muitas vezes inexitem investimentos em recursos materiais e humanos, bem como a pouca ou nenhuma a disseminação das informações (LAGUARDIA, et al., 2004).

2.3. MINERAÇÃO DE DADOS NA SAÚDE

A mineração de dados é utilizada em diversas áreas, inclusive na área da saúde. É mais utilizada na área clínica de saúde, e é uma ferramenta importante utilizada para otimizar o tratamento, diagnóstico e prognóstico. Baseada em um grande volume de dados, a mineração de dados pode contribuir no ganho de eficiência e eficácia dos serviços e na tomada de decisões a médio e longo prazos, levando a um ganho na qualidade de vida dos pacientes (CARVALHO, ESCOBAR e TSUNODA, 2014)

Por exemplo, na área da saúde mental foi utilizada como suporte para decisão clínica, previsão de diagnósticos e melhoria da qualidade de vida do paciente de doenças como depressão, esquizofrenia e distúrbios bipolares, possibilitando a diferenciação entre pessoas deprimidas e saudáveis, e oferecendo sugestões para manutenção e bem-estar de pessoas que já tiveram depressão, bem como auxiliar na detecção de esquizofrenia e distúrbios bipolares (ALONSO, et al., 2004).

Ainda foi utilizada para analisar dados epidemiológicos de poluição do ar, por meio de diversas técnicas de mineração de dados, prevendo a distribuição de certos poluentes ou a qualidade geral do ar, identificando regiões ou exposições, também foram utilizadas técnicas de mineração com o objetivo de melhorar infraestrutura de coleta de dados, assim possibilitando gerar novas hipóteses sobre a contaminação do ar e problemas de saúde (BELLIGER, et al., 2017).

E também na busca de informações sobre o diabetes mellitus, principalmente na identificação de biomarcadores, diagnóstico e previsão da doença, mas também para prever gastos com assistência médica dos diabéticos, e o status de risco à saúde e financeiro e quais os fatores genéticos envolvido com a doença e para prever a sequência de medicamentos a ser utilizada, e diversas outras áreas (KAVAKIOTIS, et al., 2017).

Foi utilizada também a mineração para identificar fatores de risco para a deficiência de vitamina D em meninas adolescentes e identificar fatores associados à resposta a suplementação de vitamina D (GONOODI, et al., 2019a; GONOODI, et al., 2019b). Foi utilizada para prever a gravidade de dengue em crianças por meio das características clínicas e indicadores laboratoriais e prever cenários de Dengue como em Phitsanulok, Tailândia, em que parâmetros climáticos e físicos foram utilizados para mapear os habitats de criação do mosquito (PHAKHOUNTHONG, et al., 2018; SARFRAZ. Et al., 2014)

A rede neural artificial é uma das técnicas de mineração de dados que foram utilizadas em estudo em Guangdong na China para melhorar a triagem para risco de câncer gástrico precoce, no qual os modelos de mineração puderam avaliar rapidamente a progressão do câncer gástrico, permitindo uma intervenção precoce (LIU, et al., 2018). Em Ontário, Canadá, a rede neural foi utilizada para determinar o prognóstico de pacientes com insuficiência cardíaca, usando dados do Teste Cardiopulmonar de Exercício, foi observado que os usos dos dados completos do teste apresentavam uma melhor resposta de prognóstico se comparado a utilização dos índices resumidos comumente utilizados na prática clínica (HEARN, et al., 2018).

Em Mashhad, Irã, a rede neural foi utilizada para identificar fatores de risco associados ao Diabetes Mellitus tipo 2, utilizando-se variáveis demográficas, antropométricas e bioquímicas dos participantes (ESMAEILY, et al., 2015). Outro estudo utilizou a rede neural para prever doenças cardíacas a partir de um banco de dados de doenças cardíacas publicamente disponível para previsão foram utilizados 15 parâmetros médicos do paciente como idade, sexo, pressão arterial, colesterol, etc. Assim, seria possível auxiliar profissionais da área de saúde em um melhor diagnóstico e em uma intervenção precoce (SINGH, SINGH e PANDI-JAIN, 2018).

3. OBJETIVOS

3.1.OBJETIVO GERAL

Estimar as prevalências das arboviroses dengue, zika, chikungunya e febre amarela nos períodos de 2000 a 2018 e construir um modelo preditivo para óbito por febre amarela no Brasil utilizando inteligência artificial.

3.2.OBJETIVOS ESPECÍFICOS

Descrever a prevalência das arboviroses dengue, zika, chikungunya e febre amarela no Brasil no período de 2000 a 2018 (Artigo de Revisão);

Descrever características dos casos confirmados com febre amarela no Brasil no período de 2000 a 2018 (Artigo original);

Construir um modelo preditivo para óbito por febre amarela (Artigo original).

4. METODOLOGIA

O trabalho consiste na execução de dois artigos um de revisão sobre a prevalência de dengue, zika, chikungunya e febre amarela no Brasil e um artigo original sobre o desenvolvimento de um modelo preditivo de óbito para febre amarela no Brasil utilizando inteligência artificial.

4.1. ARTIGO DE REVISÃO

Foi realizada uma revisão sistemática com metanálise para determinar quais as prevalências de dengue, zika, chikungunya e febre amarela no Brasil, buscando responder a pergunta “Qual a prevalência das arboviroses no Brasil ao longo dos anos e os possíveis fatores associados a variação da prevalência?”. A revisão foi baseada no protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) e registrada na base *International Prospective Register of Systematic Reviews* (PROSPERO). Foi realizada a busca nas bases de dados PubMed/MEDLINE, Scielo (*Scientific Electronic Library Online*), ScienceDirect e BVS (Biblioteca Virtual de Saúde), sob o termo de busca (dengue[DeCS] OR zika[TIAB] OR chikungunya[TIAB] OR "febre amarela[DeCS] ") AND prevalência[DeCS]).

Foram incluídos estudos originais, realizados no Brasil e que apresentassem valores de prevalência ou incidência de, pelo menos, uma das quatro doenças (dengue, zika, chikungunya e febre amarela). Foram excluídos estudos com animais, estudos envolvendo apenas vacinas e medicamentos; estudos sobre mutação ou genotipagem do vírus ou que retratassem apenas formas de diagnóstico das doenças, estudos que estimassem os valores das prevalências das doenças ou que retratassem modelos de expansão da doença ou vetor, e os estudos que focassem no vetor de transmissão (*Aedes aegypti*), assim como aqueles que focassem em outra doença ou que trouxesse dados apenas de casos confirmados, estudos com apenas dados dos Sistemas de Informação.

A avaliação dos trabalhos foi realizada pelas autoras KOF e MMC de forma independente. O programa StArt (*State of the Art through Systematic Review*) foi utilizado como ferramenta de auxílio na seleção dos artigos. Na extração dos dados foi utilizado o Microsoft® Office Excel. Foram extraídos região, ano das coletas de material para análise, número de pessoas testadas, número de pessoas confirmadas, método de detecção da doença, qualidade dos estudos, se o estudo era um inquérito sorológico, se foi realizado em uma área de epidemia, se os casos considerados eram casos suspeitos, se as pessoas investigadas eram apenas crianças e adolescentes, se eram apenas mulheres, ou se eram casos de população geral (ambos os sexos

e qualquer idade). Para dengue especificamente ainda foram analisados os subgrupos a estratificação do ano (antes de 2000, entre 2000 e 2010 e depois de 2010), e a classificação da doença utilizada.

A avaliação da qualidade foi realizada pela lista de verificação para estudos de prevalência do Instituto Joanna Briggs (MUNN, et al., 2015). O desfecho primário foi a prevalência das doenças com IC95%. Os estudos de febre amarela não passaram pela metanálise, pois apenas três estudos foram incluídos, o que impossibilita a análise. As análises foram realizadas utilizando modelo randômico com os comandos **Metaprop** e **Metareg** do programa Stata (versão 13.0). Foram realizadas metarregressões para identificar as causas da heterogeneidade, utilizando o teste de Knapp e Hartung. A heterogeneidade foi avaliada pelo teste do qui-quadrado (χ^2) com significância de $P < 0,10$, sendo sua magnitude apurada pelo o I-quadrado (I^2). O efeito de estudos pequenos foi avaliado pela inspeção visual do gráfico de funil e pelo teste de Egger (MOREIRA, et al., 2019).

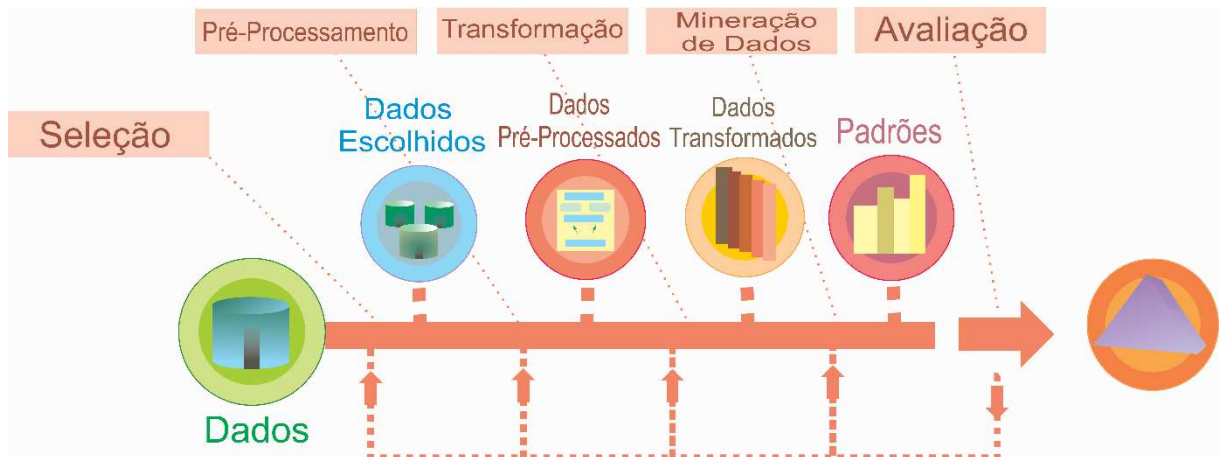
4.2.ARTIGO ORIGINAL

4.2.1. Knowledge Discovery in Databases – KDD

O artigo original foi executado conforme o processo de extração do conhecimento também conhecido como Extração do Conhecimento em Bancos de Dados (*Knowledge Discovery in Databases – KDD*). O KDD se refere a todo um processo complexo, iterativo, iterativo e não trivial para a geração de informações úteis por meio de dados brutos. Esse processo vai contar com a aplicação de métodos computacionais para encontrar dentro de um conjunto de dados, padrões de comportamento, tendências, predições ou anomalias (FERNANDES, 2017).

O processo de KDD é baseado em seis etapas: seleção (1), pré-processamento (2), transformação (3), mineração de dados (4), avaliação (5) e interpretação (6) das informações adquiridas (Figura 5). Podendo ainda ser divididas em três fases que reúnem as etapas do processo, são elas: pré-processamento, processamento e pós-processamento dos dados (PINHEIRO, 2008; FERNANDES, 2017).

Figura 5 – Etapas do Processo de KDD



Fonte: A autora

Na fase de pré-processamento os dados são preparados para a mineração, fazem parte dessa fase as etapas 1, 2 e 3. Nessa fase há a entrada dos dados brutos, desses serão selecionados os dados relevantes, logo os dados serão manipulados para remover possíveis problemas (dados faltantes, duplicados, incompletos, remover ruídos ou qualquer dado que possa comprometer a detecção de informações relevantes). Assim os dados serão transformados em dados em formatos adequados para as análises posteriores. Um trabalho ruim nessa etapa pode comprometer toda a qualidade do trabalho futuro (PINHEIRO, 2008; WITTEN, FRANK e HALL, 2011; FERNANDES, 2017).

A fase de processamento de dados é a mineração de dados em si. Nesse momento ocorre a escolha dos algoritmos de aprendizagem que serão utilizados no conjunto de treinamento para analisar os dados e obter padrões úteis (WITTEN, FRANK e HALL, 2011; FERNANDES, 2017).

Já a fase de pós-processamento envolve as etapas 5 e 6. E tem como objetivo garantir que a mineração tenha funcionado e que os resultados obtidos são de fato os mais interessantes e adequados para a tomada de decisão do problema estudado. Em 5 a performance dos modelos obtidos é testada e validada, por métodos e métricas estatísticas. Em 6 o analista interage com os resultados, verificando se os padrões encontrados respondem aos objetivos ou perguntas realizadas e se é necessária uma análise mais profunda. Após todo esse processo é possível tomar uma decisão (WITTEN, FRANK e HALL, 2011; FERNANDES, 2017).

O KDD é um processo que funciona com feedback, ou seja, a partir de qualquer momento é permitido voltar nas etapas anteriores, de forma a sempre tentar alcançar o melhor resultado (WITTEN, FRANK e HALL, 2011).

A transformação é a última etapa do KDD antes da mineração em si. Ela é responsável pela alteração nos dados, sendo que essas mudanças podem afetar a performance de aprendizagem. Logo, deve-se escolher métodos adequados que melhorem a precisão dos algoritmos e que mantenha a qualidade da informação (FERNANDES, 2017).

4.2.2. Dados

Foram utilizados, no trabalho, dados secundários do SINAN. Esses dados foram disponibilizados pelo Ministério da Saúde e solicitados por meio do Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC). Foram solicitados dados referentes a febre amarela desde o ano 2000 até o ano de 2018. Os dados solicitados advêm das informações contidas na ficha de notificação para febre amarela (ANEXO B). Abaixo estão sendo mencionadas as variáveis solicitadas (Quadro 3).

Quadro 3 - Variáveis solicitadas segundo ficha de notificação febre amarela.

Doença	Variáveis Descritivas	Variáveis Clínicas
Febre Amarela	Agravo/doença, data da notificação, município de notificação, estado de notificação, código (IBGE), unidade de notificação, data dos primeiros sintomas, data de nascimento, idade, sexo, gestante, raça/cor, escolaridade., município de residência, código IBGE, zona, país, data da investigação, ocupação, se o caso é autóctone do município de residência; Município, evolução do caso, data do óbito, data do encerramento	Sinais e Sintomas; Vacinação contra febre amarela; Exames inespecíficos; Sorologia IgM, data da coleta, resultado. Exame NS1, resultado. Isolamento, resultado. RT-PCR, data da coleta, resultado. Histopatologia, imunohistoquímica, se ocorreu hospitalização Data da internação, município do hospital, Classificação Final, Critério de confirmação ou descarte.

Fonte: dados da pesquisa

Segundo o artigo 1º, parágrafo único da Resolução nº 510 de 07 de Abril de 2016 do Conselho Nacional de Saúde, não deverão ser registradas nem avaliadas pelo sistema CEP (Comitê de Ética e Pesquisa)/CONEP (Comissão Nacional de Ética em Pesquisa) as pesquisas que utilizam informações de acesso público, nos termos da Lei no 12.527, de 18 de novembro de 2011, logo esse trabalho dispensa a apresentação e submissão ao Comitê de Ética. Essa não

necessidade de registro e avaliação do projeto pelo CEP foi confirmada pelo órgão responsável na Universidade Federal de Viçosa via e-mail, no mês de outubro de 2019 (ANEXO A).

Realizou-se toda a padronização do banco de dados advindo do SINAN com a correção de grafia, bem como, a exclusão de variáveis redundantes, aquelas que não agregam informação; fusão ou agregação das variáveis que apresentem respostas similares ou altamente correlacionadas, limpeza dos dados, tratamento dos dados ausentes, inconsistentes ou fora dos padrões. De forma que os dados brutos estejam em formato apropriado para as análises. O banco original possuía 17942 casos de notificações de febre amarela. Contudo após uma primeira organização e padronização dos dados foram excluídos 2208 casos sendo esses referentes a dados inconsistentes, faltantes, ou que não eram de interesse para essa pesquisa, como casos de doença em estrangeiros. Das 32 variáveis iniciais foram removidas as variáveis DATA_NOTIF (data da notificação), DATA_SINTOMA (data do início dos sintomas), DATA_INTERNACAO (data da internação do paciente), DATA_OBITO (data do óbito do paciente), DATA_ENCERRAMENTO (data do registro de encerramento do caso), BT (bilirrubina total), BD (bilirrubina direta), AST_TGO (aspartato aminotransferase), ALT_TGP (alanina aminotransferase) devido a dados faltantes e informações equivocadas.

4.2.3. Análise Exploratória

Após o processo de organização foi realizada a análise exploratória que foi realizada com o intuito de se conhecer as características dos dados e procurar algumas relações entre as variáveis. Foi utilizado do *Jupyter Notebook* em linguagem de programação Python. Inicialmente realizamos uma análise estatística com valores de média, desvio padrão, valores mínimos e máximos, para as variáveis quantitativas e valores de frequência absoluta e relativa para variáveis qualitativas. Foi avaliada também a normalidade das variáveis por meio da análise visual de boxplot e histogramas e confirmadas com o teste de Shapiro-Wilk. Ainda realizamos uma análise bivariada pela Correlação de Pearson para variáveis categóricas, buscando conhecer a correlação entre as variáveis.

Por meio da análise dos boxplot foi possível verificar que não existe uma normalidade na distribuição das variáveis. Essa informação foi confirmada pela análise de histograma e pelo teste de Shapiro-Wilk. No resultado da correlação de Pearson as variáveis que foram altamente correlacionadas positivamente (coef. de Pearson ≥ 0.75) foram: região de notificação e estado de notificação; período sazonal e mês dos sintomas; idade e idade categorizada; histopatologia e imunohistoquímica. Essas variáveis são redundantes, pois uma traz a mesma ideia que a outra,

nesse sentido uma delas foi excluída antes da mineração de dados. Para a histopatologia e imunohistoquímica foram excluídas visto que esses testes são executados depois do óbito do paciente, contudo o objetivo do modelo preditivo é prever fatores relacionados ao óbito ou cura.

4.2.4. Seleção de Atributos

A seleção de atributos tem a função de simplificar o modelo, tornar a construção dele mais rápida, sem perder informações importantes e produzindo os melhores resultados. É uma forma de evitar a maldição de dimensionalidade, que diz que quanto maior a dimensionalidade, mais os dados se dispersam no espaço que ocupam, em outras palavras, quanto maior a dimensionalidade mais difícil é criar um modelo (SANTORO, 2005; TAN, STEINBACH e KUMAR, 2009).

A seleção computacional de atributos pode ser:

- **Tipo filtro** – aplicados antes do processo de aprendizagem, e vão definir a relevância de cada atributo a partir de um critério pré-definido; esse método é recomendado para grandes volumes de dados porque são rápidas, já que não depende do classificador (SANTORO, 2005).

Os filtros se dividem em dois:

- **Rankers:** Que calculam individualmente a pontuação de cada atributo. Gera uma lista decrescente de pontuação. É eficiente em tempo de execução, contudo ignora correlações importantes entre os atributos (TAN, STEINBACH e KUMAR, 2009).

- **Aqueles que avaliam subconjuntos:** considera relações entre conjuntos. Seleciona subconjuntos nos quais os atributos estão fortemente correlacionados com a classe e fracamente entre eles (TAN, STEINBACH e KUMAR, 2009).

- **Tipo wrapper** – Abordagem de alto custo computacional que usa critérios e pontos chave para explorar o espaço de busca dos subconjuntos de atributos. Depende de algoritmos de classificação, o que melhora a precisão de classificação, mas limita o método pelo alto custo computacional gerado (TAN, STEINBACH e KUMAR, 2009).

Nesse trabalho não foram utilizadas formas computacionais para a seleção dos atributos. A seleção foi realizada de forma manual por meio da análise dos pesquisadores da análise exploratória. Foram 15734 notificações de febre amarela e 21 variáveis para a análise exploratória. Com base no critério de redundância ou por ter outra variável que representasse melhor os dados foram excluídas as variáveis ANO_NOT (ano de notificação), REG_NOTIF (região de notificação), AUTOCTONE (se o caso era ou não autóctone do município), MES_SINTOMA (mês de apresentação de sintomas), IDADE_CAT (categorização de idade),

HISTOPATOLOGIA (resultado da histopatologia), IMUNOHISTOQUIMICA (resultado da imunohistoquímica), perfazendo um total de 14 variáveis consideradas no modelo final.

Considerando que o objetivo do trabalho é baseado no óbito ou cura para febre amarela, as pessoas sem informação e as mortes por outras causas foram excluídas do modelo final dessa forma perfazendo um total de 2405 indivíduos confirmados com febre amarela. Os casos confirmados para febre amarela, foram baseados na variável CLAS_FINAL (Classificação final) que dividia o banco em casos confirmados e descartados.

4.2.5. Aprendizagem de Máquina

A mineração de dados, 4º passo do processo de KDD, é o processo de descoberta de padrões e tendências, de forma automática ou semiautomática, ocultos em um grande banco de dados por meio de algoritmos que permitem a exploração, análise e modelagem dos dados (PINHEIRO, 2008; TAN, STEINBACH e KUMAR, 2009; WITTEN, FRANK e HALL, 2011; FERNANDES, 2017).

Na prática a mineração é categorizada em duas formas de análise, e a escolha de cada categoria pode variar conforme o objetivo da pesquisa, o problema a ser estudado e as bases de dados pesquisadas (TAN, STEINBACH e KUMAR, 2009; FERNANDES, 2017).

A primeira categoria é de previsão, o objetivo dessa é prever um valor de um determinado atributo (variável) baseado nos valores de outros atributos. Ou seja, tem-se variáveis alvo ou dependente (atributos a serem previsto) e variáveis explicativas ou independentes (usadas para fazer a previsão). No modelo de previsão existem dois tipos de tarefa a classificação para variáveis alvo discretas e a regressão para variáveis alvo continua. O objetivo de ambas as tarefas é aprender um modelo que erre menos os valores previsto e real da variável alvo (TAN, STEINBACH e KUMAR, 2009).

A segunda categoria é a descritiva, nessa o objetivo é derivar padrões (correlações, tendências, grupos, trajetórias e anomalias) para que formem um resumo da relação subjacente dos dados. As tarefas dessa categoria são geralmente exploratórias e é comum necessitar de técnicas de pós-processamento para validar e explicar os resultados (TAN, STEINBACH e KUMAR, 2009).

Devido as características do banco de dados e o objetivo do trabalho esse é um estudo de previsão com tarefas de classificação.

A aprendizagem de máquina é ferramental matemático e computacional que normalmente é utilizado para se realizar as tarefas de mineração. A aprendizagem de máquina é o estudo e o

desenvolvimento de algoritmos que permitam que sistemas computacionais melhores sua performance em determinada tarefa, por meio de aprendizagem automatizada (FERNANDES, 2017). A aprendizagem supervisionada é a aplicação de um algoritmo em um conjunto de treinamento, com instâncias e registros já classificados cujos valores e rótulos são conhecidos, dessa forma permitindo a construção de um modelo para classificar novos componentes (TAN, STEINBACH e KUMAR, 2009).

O principal objetivo de um algoritmo de aprendizagem de máquina é produzir um modelo que identifique bem a relação entre o vetor de atributos e o rótulo da classe, prevendo a classe de componentes futuros baseados em seus atributos. Existem diversos algoritmos e cada um funciona de formas diferentes, são mais eficientes para determinadas tarefas (TAN, STEINBACH e KUMAR, 2009).

4.2.6. Algoritmos de Aprendizagem de Máquina

Árvore de Decisão (*Decision Tree* - DT): Baseado no conceito de uma árvore, a técnica de árvore de decisão cada ramo representa um tópico de classificação. Essa técnica mantém a qualidade original dos dados. Suas análises são de mais rápidas e de fácil entendimento (PINHEIRO, 2008). A precisão de suas previsões, em geral, possui taxas de acertos competitivas em relação a de outros modelos (HAN, KAMBER e KAUFMANN, 2012). Uma árvore de decisão possui um nodo raiz, não possui arestas chegando, apenas arestas saindo; nodos internos, cada um com uma aresta chegando e duas ou mais arestas saindo; e os nodos folhas ou terminais, que possuem uma aresta chegando e nenhuma aresta saindo (TAN, STEINBACH e KUMAR, 2009).

Redes Neurais Artificiais: Essa técnica foi bem mais desenvolvida na comunidade de inteligência artificial e implementa padrões de detecção e algoritmos de aprendizagem de máquina. Essa técnica possui altos valores de predição e foi inspirada na tentativa de simular sistemas neurais biológicos (TAN, STEINBACH e KUMAR, 2009; PINHEIRO, 2008). As principais razões para a utilização da técnica de redes neurais é a capacidade intrínseca de aprendizagem a partir de dados de entrada, permitindo posteriormente a generalização. Elas tornam as suposições mais precisas, pois são não-paramétricas, independente da distribuição dos dados de entrada; são altamente não-lineares no espaço de características e podem criar limites de decisão; como são modelos não-lineares não necessitam de suposições ou premissas sobre os dados base do modelo (PINHEIRO, 2008).

Redes bayesianas (*Naive Bayes* – NB): É um algoritmo baseado no Teorema de Bayes, que tentava provar a existência de Deus. Esse determina como conceitos feitos antes dos experimentos podem ser modificados depois da coleta de evidências, na análise dos resultados. Quando os eventos estão conectados em forma de rede, ele pode ajudar a prever o resultado de um determinado evento a partir de outros eventos (PINHEIRO, 2008). A principal característica do algoritmo é desconsiderar completamente a correlação entre as variáveis. O cálculo para estimar a probabilidade do vetor de atributos condicionados à classe é simples, visto que o algoritmo assume que os atributos são independentes. Logo, sem dependência ou com a dependência fraca o algoritmo tem um melhor resultado, com um tempo de resposta mais curto (TAN, STEINBACH e KUMAR, 2009).

***k*-Nearest Neighbour (kNN):** Faz a classificação utilizando os vizinhos mais próximos (MITCHELL, MICHALSKI e CARBONELL, 2013). Depende da métrica escolhida para definição da vizinhança, de forma que quando mais próximo o vizinho maior o peso dele. (MITCHELL, MICHALSKI e CARBONELL, 2013; TAN, STEINBACH e KUMAR, 2009)

***Logistic Regression* (LR):** Método estatístico parecido com a regressão linear, contudo as variáveis respostas da LR podem ser categóricas ou contínuas, diferente da regressão linear. Utiliza o método de máxima verossimilhança e *Odds Ratio* em vez de probabilidade para prever a associação ao grupo. Pressupõe independência entre as variáveis (DIGANGI e HEFNER, 2013).

***Multilayer Perceptron* (MP):** São camadas de elementos de processamento sigmoide simples, de dois estados (nós) ou neurônios que interagem usando conexões ponderadas. Entre a camada de entrada e a de saída, existem diversas camadas intermediárias ou ocultas. Quando todos os neurônios de uma camada estão ligados em camadas adjacentes, as conexões dentro de uma mesma camada são impossíveis. Os pesos vão medir o grau de correlação entre a atividade de neurônios conectados (PAL e MITRA, 1992)

***Support Vector Machine* (SVM):** Tem fundamento na teoria de aprendizagem estatística. Funciona bem com dados de alta dimensionalidade o que evita o problema da dimensionalidade. Esse algoritmo utiliza vetores de suporte que são subconjuntos dos exemplos de treinamento. Esses são utilizados em um limite de decisão (TAN, STEINBACH e KUMAR, 2009).

***Random Forest* (RF):** Utiliza o resultado de várias árvores de decisão, sendo um método de classificação em conjunto. O resultado final é obtido a partir da classe majoritária das várias árvores em um esquema de votação. Cada árvore é construída a partir de uma amostra do conjunto de treinamento, e as amostras são construídas a partir de um subconjunto aleatório dos atributos originais (TAN, STEINBACH e KUMAR, 2009).

CatBoost (CB): Esse algoritmo é uma implementação de *gradient boosting* que usa árvores de decisão binárias como preditores base. É um método *ensemble*, ou seja, cria várias árvores de decisão, sendo que cada nova árvore é construída em comparativo com as outras. O resultado agrega a resposta de todas as árvores construídas. Pode usar variáveis categóricas diretamente e é escalonável por natureza. Retorna bons resultados sem um treinamento extensivo de dados e fornece suporte avançado para os formatos de dados mais descritivos (HANCOCK e KHOSHGOFTAAR, 2020).

Nesse trabalho foram testados 8 algoritmos. Naive Bayes (NB), Decision Tree (DT), k-Nearest Neighbour (kNN), Logistic Regression (LR), Multilayer Perceptron (MP), Support Vector Machine (SVM), Random Forest (RF) e CatBoost.

4.2.7. Processamento de Dados

Para a realização do processo de mineração de dados foi utilizado *Google Colaboratory* ou *Google Colab* em linguagem de programação Python. Nos scripts do Python foram aplicadas como principais bibliotecas: *numpy* e *pandas* para manipulação de dados, *matplotlib* e *shap* para visualização de dados, *catboost* e *sklearn* para os algoritmos de aprendizagem de máquina.

- **Numpy:** é utilizada principalmente para realizar cálculo em Arrays Multidimensionais. Possui um grande conjunto de funções e operações para facilitar cálculos numéricos (NUMPY, 2020);
- **Pandas:** é uma ferramenta de código aberto para análise e manipulação dos dados (PANDAS, 2021);
- **Matplotlib:** é uma biblioteca abrangente para criação de visualizações estatísticas, figuras interativas e animadas, e que podem ser customizadas (MATPLTLIB, 2021);
- **Shap:** é uma biblioteca para explicar o impacto dos atributos na saída de um modelo qualquer de aprendizagem de dados (SHAP, 2021).
- **Catboost:** é uma biblioteca para *gradient boosting* em árvores de decisão (CATBOOST, 2021).
- **Sklearn:** uma ferramenta para a análise preditiva de dados. Pode ser usada para problemas de classificação, regressão, clustering, redução de dimensionalidade, seleção de modelo e pré-processamento (SCIKIT-LEARN, 2021).

O treinamento do modelo foi realizado pelo método *Hold-Out*. Para nosso conjunto de dados foi utilizada o percentual de 75% da base para treinamento e 25% para teste. O

Underfitting e o *Overfitting* foram verificados, e não ocorreram. Para seleção de hiperparâmetros utilizamos o método *GridSearchCV* com validação cruzada 5 partes para encontrar os valores otimizados para um conjunto de importantes hiperparâmetros para cada um dos algoritmos testados. Os parâmetros tiveram suas configurações modificadas.

4.2.8. Pós-processamento de Dados

A avaliação do modelo de dados é a 5ª etapa do processo KDD, nada mais é que a utilização de métodos e métricas para estimar a precisão e robustez dos modelos na prática. Seguem métodos utilizados para analisar a performance dos modelos criados (TAN, STEINBACH e KUMAR, 2009).

A validação cruzada é muito utilizada para avaliação da capacidade de generalização. Inicialmente todo o conjunto é particionado. Cada parte vai ser testada com o resto do grupo individualmente, ao final se escolhe outra parte e repete o processo até todos serem testados. O resultado final é a acurácia, o que indica como o modelo generaliza (TAN, STEINBACH e KUMAR, 2009).

Para a validação do modelo utilizamos a Validação Cruzada 10 partes. Para selecionar o modelo preditivo final, foi utilizada a validação cruzada 10 partes em todos os algoritmos e foram executados 30 vezes com 30 sementes geradas aleatoriamente (valores inteiros aleatórios entre 0 e 232-1) para obter o desempenho médio dos classificadores.

Uma outra abordagem comum para avaliação de modelos de aprendizagem é a análise de matriz de confusão. A matriz de confusão é similar a uma matriz matemática, formada por linhas e colunas. Os valores previstos corretamente são verdadeiros positivos e verdadeiros negativos e se encontram na diagonal principal da matriz. Já os registros previstos erroneamente são os falsos positivos e falsos negativos, e são representados fora da diagonal principal. Essa matriz permite a determinação da performance do modelo. São utilizadas algumas métricas para sumarizar o resultado da matriz de confusão (TAN, STEINBACH e KUMAR, 2009). Algumas delas são:

- **Acurácia:** proporção entre predições corretas em relação ao número total de predições (corretas e erradas) (TAN, STEINBACH e KUMAR, 2009; SILVA, 2018);
- **Taxa de Erro:** proporção entre as predições erradas em relação ao número total de predições (TAN, STEINBACH e KUMAR, 2009);

- **Precisão:** proporção entre previsões verdadeiramente positivas em relação as previsões ditas positivas (falso positivo e verdadeiro positivo) (TAN, STEINBACH e KUMAR, 2009; SILVA, 2018);
- **Sensibilidade:** Proporção entre previsões positivas corretas em relação as previsões positivas (verdadeiro positivo e falso negativo) (TAN, STEINBACH e KUMAR, 2009);
- **Especificidade:** Proporção entre previsões negativas corretas em relação as previsões negativas (falso positivo e verdadeiro negativo) (TAN, STEINBACH e KUMAR, 2009);
- **Medida F:** É a média harmônica entre a precisão e a sensibilidade (SILVA, 2018).
- **Área sobre a curva ROC:** Representa a relação entre a sensibilidade (y) e especificidade (x) de um modelo. Um bom modelo de classificação tem a curva mais aberta. Quanto maior a área sob a curva maior o poder de previsão do modelo (PRATI, BATISTA e MONARD, 2008).

Para esse trabalho foram utilizadas a acurácia, precisão, sensibilidade, medida F e área sob a curva ROC.

4.2.9.Avaliação

Após a validação do modelo e a obtenção das métricas, realizou-se as comparações estatísticas dos modelos por meio de testes não paramétricos. Utilizou-se o teste de Friedman, seguido do teste Post-HOC de Nemenyi para identificar diferenças entre os modelos de acordo com a média da Curva ROC com nível de confiança de 95%. E para visualização dos resultados utilizamos a biblioteca SHAP (*SHapely Additive exPlanations*).

5. RESULTADO E DISCUSSÕES

Os resultados e discussões deste trabalho serão apresentados em dois artigos, um de revisão e um original, buscando atender aos objetivos propostos.

Os artigos que compõem esta dissertação são intitulados:

Artigo de revisão: Prevalência de dengue, zika, chikungunya e febre amarela no brasil:
Revisão Sistemática e Metanálise

Artigo Original: Febre amarela no Brasil: estudo de prognóstico utilizando técnicas de Inteligência Artificial

5.1. ARTIGO DE REVISÃO - PREVALÊNCIA DE DENGUE, ZIKA, CHIKUNGUNYA E FEBRE AMARELA NO BRASIL: REVISÃO SISTEMÁTICA E METANÁLISE

RESUMO

Introdução: As arboviroses são doenças causadas por arbovírus e representam uma crescente preocupação para a saúde pública no Brasil e no mundo, principalmente pelo potencial de dispersão, pela capacidade de adaptação a novos ambientes e hospedeiros, pela possibilidade de causar epidemias extensas. Pela similaridade de sintomas e de sinais, as notificações de casos suspeitos de dengue, zika e chikungunya podem se confundir e além disso, a cocirculação de infecção por esses vírus no Brasil dificulta o manejo clínico. **Objetivo:** descrever a prevalência das arboviroses dengue, zika, chikungunya e febre amarela no Brasil. **Metodologia:** Foi realizada uma revisão sistemática com metanálise baseada no protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) e registrada na base International Prospective Register of Systematic Reviews (PROSPERO). Foi realizada a busca nas bases de dados PubMed/MEDLINE, Scielo (Scientific Electronic Library Online), ScienceDirect e BVS (Biblioteca Virtual de Saúde), sob o termo de busca (dengue OR zika OR chikungunya OR "febre amarela") AND prevalência) nos idiomas português, inglês ou espanhol. A avaliação dos trabalhos foi realizada por dois autores de forma independente. A avaliação da qualidade foi realizada pela lista de verificação para estudos de prevalência do Instituto Joanna Briggs. A metanálise foi realizada por modelo de efeito randômico. A heterogeneidade foi avaliada pelo teste do qui-quadrado (χ^2) com significância de $P < 0,10$, sendo sua magnitude apurada pelo o I-quadrado (I^2). As análises foram realizadas com os comandos Metaprop e Metareg do programa Stata (versão 13.0). **Resultados:** Foram selecionados 61 estudos publicados entre os anos de 1992 e 2019, sendo 65,57% dos artigos ($n=40$) foram publicados nos últimos 10 anos. A maioria dos artigos selecionados descrevem principalmente dados epidemiológicos regionais derivados de pesquisas e estudos conduzidos em regiões específicas, sendo utilizados para apoiar os dados nacionais no que diz respeito à prevalência e sorotipo. Não foi possível realizar a metanálise com os artigos de febre amarela, pois poucos estudos foram selecionados. A prevalência agrupada entre os indivíduos testados, no modelo randômico, de zika foi de 19% (IC95%: 0,14 a 0,25), de chikungunya foi de 28% (IC95%: 0,24 a 0,31) e de dengue foi de 35% (IC95%: 0,30 a 0,39). Para as três arboviroses, dengue, zika e chikungunya, nenhuma das variáveis, na análise subgrupos e metaregressão, explicou a grande heterogeneidade encontrada. **Conclusão:** As prevalências de zika (19%) e chikungunya (28%) refletem valores de uma

pesquisa pontual, focada apenas em epidemias ou casos suspeitos, sendo necessário o avanço dos estudos para um melhor entendimento dessas doenças no país, para melhor preparar os serviços de saúde, garantindo um melhor atendimento da população. Sobre a dengue vemos que a prevalência da doença é alta no Brasil.

Palavras—chave: Prevalência. Arbovirose. Epidemiologia. Metanálise. Revisão de literatura.

INTRODUÇÃO

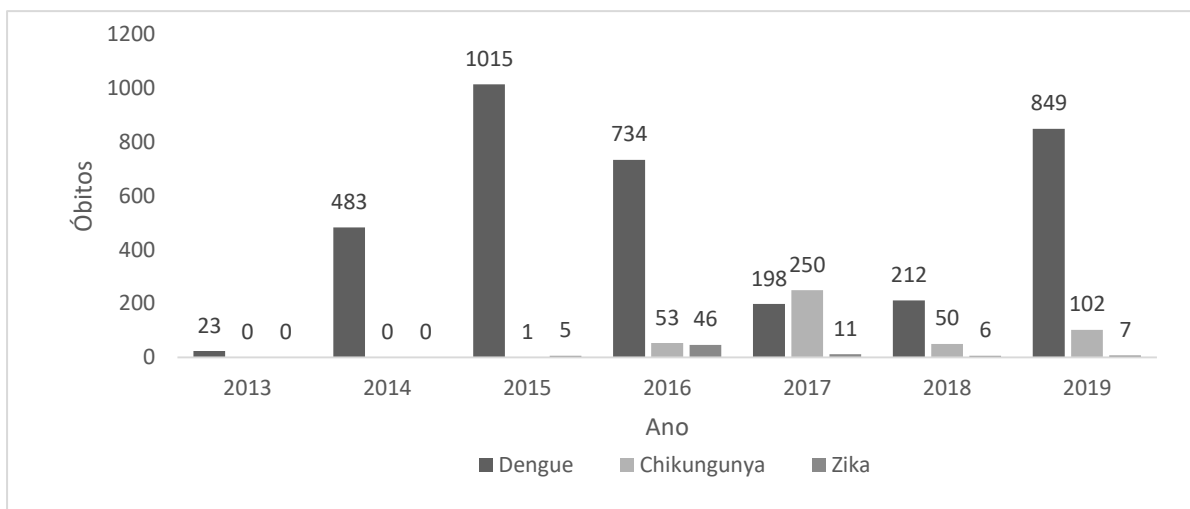
As arboviroses são doenças causadas por arbovírus, patógenos que compartilham a característica de serem transmitidos por artrópodes, em sua maioria mosquitos hematófagos. Entre os agentes infecciosos deste grupo, os mais importantes para a saúde humana são aqueles transmitidos por culicídeos, principalmente dos gêneros *Culex* e *Aedes*, e pertencem a cinco famílias: Bunyaviridae, Togaviridae (que inclui o vírus Chikungunya), Flaviviridae (que inclui os vírus da dengue, zika e febre amarela), Reoviridae e Rhabdoviridae (AVELINO-SILVA e RAMOS, 2017; DONALISIO et al., 2017; LOPES, NOZAWA e LINHARES, 2014; WEAVER e REISEN, 2010).

No atual contexto epidemiológico brasileiro, dengue, chikungunya, zika e febre amarela, constituem problemas graves em termos de saúde coletiva. Cada arbovirose com seus processos históricos, sociais e epidemiológicos distintos, impõem grandes desafios sociais e de saúde nos territórios onde estão presentes. Representam uma crescente preocupação para a saúde pública no Brasil e no mundo, principalmente pelo potencial de dispersão, pela capacidade de adaptação a novos ambientes e hospedeiros, pela possibilidade de causar epidemias extensas, pela susceptibilidade universal e pela gravidade clínica, com acometimento neurológico, articular e hemorrágico (AVELINO-SILVA e RAMOS, 2017; DONALISIO et al., 2017; EVANGELISTA et al., 2018).

Entre 1958 a 1973, o *A. aegypti* chegou a ser erradicado do país por duas vezes. Entretanto, em 1976, surgiram os primeiros registros da reintrodução do vetor no Brasil. Desde então, o *A. aegypti* está presente em todas as Unidades da Federação, distribuído em, aproximadamente, 4.523 municípios. A dengue é a doença de maior importância epidemiológica, observada de forma praticamente ininterrupta no país desde a década de 1980 com aumento importante de prevalência e número de óbitos a partir do ano de 2004. (AVELINO-SILVA e RAMOS, 2017; VALLE, PIMENTA e AGUIAR, 2016; ZARA et al., 2016). Dados do Ministério da Saúde mostram que houve um aumento expressivo no número de casos confirmados e óbitos por dengue a partir de 2014, com queda em 2017 e 2018, e aumento em 2019 (Figuras 1 e 2) (BRASIL, 2024).

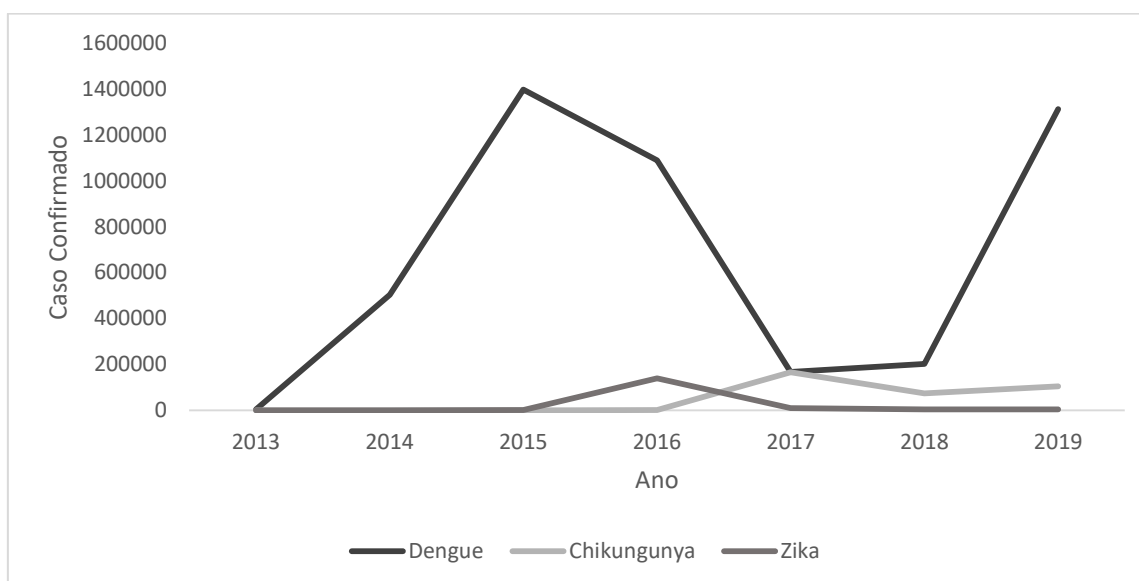
Os casos de zika no Brasil tiveram pico em 2016, com 139.642 casos e 46 óbitos confirmados por zika registrados (Figuras 1 e 2). A Bahia é o estado que mais concentra casos de zika, com valores próximos a 50%. Em 2019 houve um pequeno aumento no número casos e óbitos confirmados pela doença (BRASIL, 2018; BRASIL, 2019a; BRASIL, 2019b; BRASIL, 2024).

Figura 1 – Número de óbitos por dengue, chikungunya e zika, de 2013 – 2019, Brasil



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

Figura 2 – Número de casos confirmados por dengue, chikungunya e zika, de 2013 – 2019, Brasil

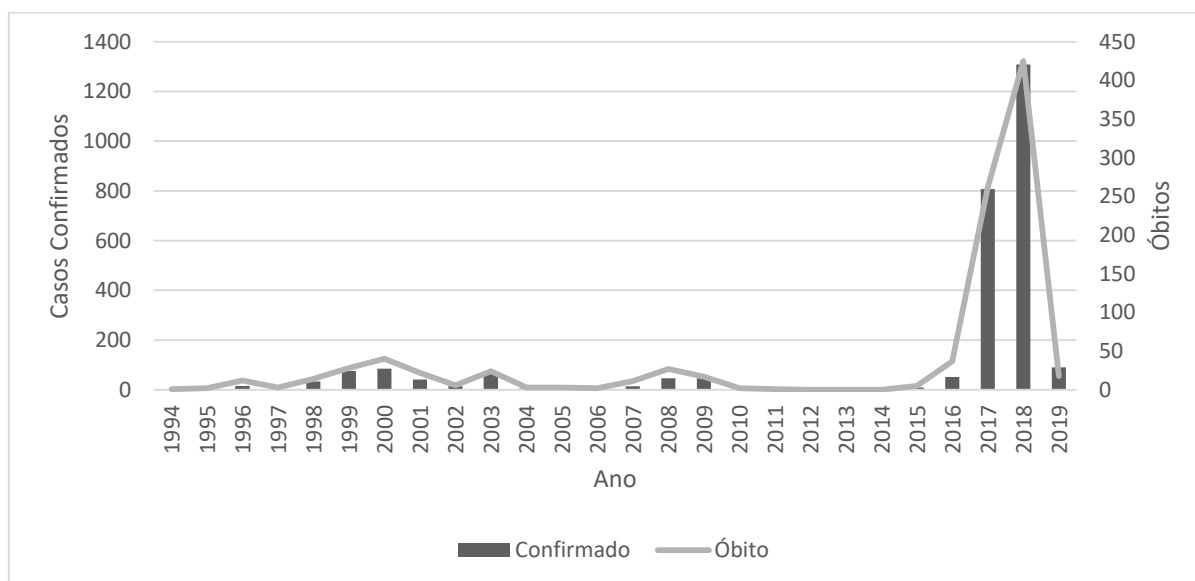


Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

Quanto a chikungunya o pico no número de casos e óbitos ocorreu em 2017. Foram 165.292 casos e 250 óbitos confirmados pela doença. Em 2018 houve queda e em 2019 aumento nos casos e óbitos confirmados por chikungunya (Figuras 1 e 2).

Quanto aos números de casos humanos de febre amarela, ele vem crescendo desde 2014/2015, se tornando uma preocupação devido aos surtos ocorridos em 2016/2017 e 2017/2018, sendo o último o maior surto ocorrido no Brasil que afetou principalmente a região Sudeste, registrando 1.376 casos humanos e 483 óbitos (Figura 3) (BRASIL, 2018; BRASIL,2019a; BRASIL,2019b; BRASIL, 2020).

Figura 3 – Número de casos confirmados e óbitos por febre amarela no Brasil de 1994 – 2019



Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

Diante aumento dos casos e da visibilidade dessas doenças e dos surtos ocorridos ampliou-se a discussão a respeito dos determinantes das arboviroses, relacionados ao vetor (mosquito), ao hospedeiro (geralmente primatas, humanos ou não) e ao ambiente. Além disso, abordagens pertinentes no âmbito de saúde pública devem ser mais amplamente discutidas (AVELINO-SILVA e RAMOS, 2017; VALLE, PIMENTA e AGUIAR, 2016).

Pela similaridade de sintomas e de sinais, as notificações de casos suspeitos de dengue, zika e chikungunya podem muitas vezes apresentar erros. Além disso, a cocirculação de infecção por esses vírus Brasil dificulta o manejo clínico em razão de similaridades (AVELINO-SILVA e RAMOS, 2017; LOPES, NOZAWA e LINHARES, 2014).

Diante dos desafios de controle do vetor e de um quadro grave e preocupante em relação às arboviroses delineado pela expansão destes vírus em todo o mundo, torna-se imprescindível a adoção de estratégias específicas, com maiores investimentos em métodos adequados, que forneçam sustentabilidade às ações estabelecidas pelas redes de vigilância, além de ensejarem a análise de sua efetividade (ZARA et al., 2016). Além disso com a recente pandemia de Covid

– 19 a mobilização das equipes de vigilância epidemiológica gerou uma situação de atraso ou subnotificações das arboviroses. Outro fator que pode influir na redução do número de notificações nesse período é o receio da população em procurar o serviço de saúde no contexto da pandemia (BRASIL, 2020a).

Diante dessas preocupações o objetivo desse trabalho é descrever a prevalência das arboviroses dengue, zika, chikungunya e febre amarela no Brasil.

METODOLOGIA

O protocolo da revisão foi registrado na base *International Prospective Register of Systematic Reviews* (PROSPERO) sob o número CRD42019134631. Essa revisão sistemática foi baseada no protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*). O trabalho buscou responder a pergunta “Qual a prevalência das arboviroses no Brasil ao longo dos anos e os possíveis fatores associados a variação da prevalência?”

Fonte de Informações e estratégia de busca

A busca foi realizada nas bases de dados PubMed/MEDLINE, Scielo (*Scientific Electronic Library Online*), ScienceDirect e BVS (Biblioteca Virtual de Saúde), sob o termo de busca (dengue[DeCS] OR zika[TIAB] OR chikungunya[TIAB] OR "febre amarela[DeCS] ") AND prevalência[DeCS]) nos idiomas português, inglês ou espanhol. No Pubmed também foi acrescentado o filtro de humanos, no ScienceDirect o filtro utilizado foi o de artigos originais, nas demais não foram utilizados outros filtros. A busca foi realizada em setembro 2020.

Critérios de elegibilidade

Foram incluídos estudos originais, realizados no Brasil e que apresentassem valores de prevalência de, pelo menos, uma das quatro doenças (dengue, zika, chikungunya e febre amarela). Foram excluídos estudos com animais, estudos envolvendo apenas vacinas e medicamentos; estudos sobre mutação ou genotipagem do vírus ou que retratassem apenas formas de diagnóstico das doenças. Ainda excluímos estudos que estimassem os valores das prevalências das doenças ou que retratassem modelos de expansão da doença ou vetor, e os estudos que focassem no vetor de transmissão (*Aedes aegypti*), assim como aqueles que

focassem em outra doença ou que trouxesse dados apenas de casos confirmados. Os estudos que trouxessem apenas dados dos Sistemas de Informação também foram excluídos.

Seleção dos estudos e extração dos dados

A avaliação dos trabalhos foi realizada pelas autoras KOF e MMC de forma independente. As comparações foram realizadas após cada etapa do processo de triagem, sendo elas leitura de título, leitura de resumo e leitura de textos completos. Os conflitos foram resolvidos em comum acordo entre as autoras.

O programa StArt (*State of the Art through Systematic Review*) foi utilizado como ferramenta de auxílio na seleção dos artigos. Na extração dos dados foi utilizado o Microsoft® Office Excel. Foram extraídos região, ano das coletas de material para análise, número de pessoas testadas, número de pessoas confirmadas, método de detecção da doença, qualidade dos estudos, se o estudo era um inquérito sorológico, se foi realizado em uma área de epidemia, se os casos considerados eram casos suspeitos, se as pessoas investigadas eram apenas crianças e adolescentes, se eram apenas mulheres, ou se eram casos de população geral (ambos os sexos e qualquer idade). Para dengue especificamente ainda foram analisados os subgrupos a estratificação do ano (antes de 2000, entre 2000 e 2010 e depois de 2010), considerando as três classificações já adotadas pelo Brasil o ano que foram criadas e que começaram a ser implementadas no país, e a classificação da doença utilizada.

Avaliação da qualidade metodológica dos estudos incluídos

Para avaliação da qualidade dos trabalhos utilizamos a lista de verificação para estudos de prevalência do Instituto Joanna Briggs (MUNN, et. al., 2015). A lista é composta por 9 perguntas: 1) O quadro de amostra foi apropriado para abordar a população-alvo?; 2) Os participantes do estudo foram amostrados de maneira apropriada?; 3) O tamanho da amostra foi adequado? 4) Os sujeitos do estudo e o ambiente foram descritos em detalhes?; 5) A análise dos dados foi realizada com cobertura suficiente da amostra identificada?; 6) Foram usados métodos válidos para a identificação da doença?; 7) A condição foi medida de maneira padrão e confiável para todos os participantes?; 8) Houve análise estatística apropriada?; 9) A taxa de resposta foi adequada e, em caso negativo, a baixa taxa de resposta foi gerenciada de maneira adequada? Cada critério atendido pontua um ponto, logo a pontuação máxima é 9 pontos (MUNN, et. al., 2015). Nesse trabalho consideramos os estudos com pontuação 7 a 9 como de

alta qualidade; moderada qualidade, 4 a 6 pontos; e de baixa qualidade, 0 a 3 pontos. A qualidade dos artigos não foi utilizada como motivo de exclusão dos artigos, sendo parâmetro para estudo da heterogeneidade e da análise de subgrupos.

Análise dos dados

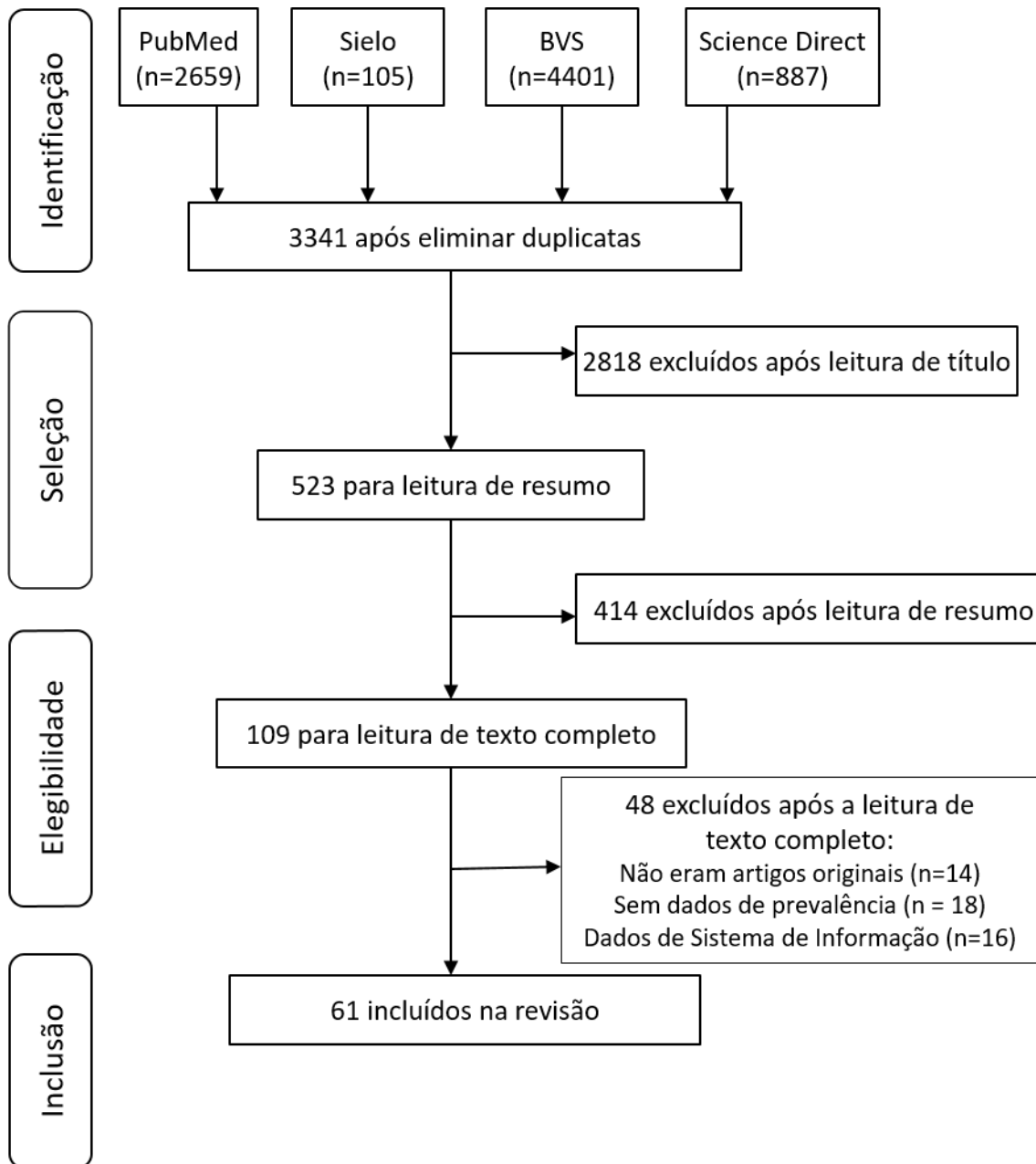
O desfecho primário foi a prevalência das doenças com IC95%. Dado os critérios de seleção dos artigos, foram sumarizadas as prevalências das doenças observadas nos artigos, sendo que os artigos não podiam trazer informação secundária da realização do exame e nem trabalharem com casos já confirmados da doença.

Os estudos de febre amarela não passaram pela metanálise pois, apenas três estudos foram incluídos, o que impossibilita a análise. Os estudos incluídos apresentavam por vezes mais de uma das doenças buscadas em um mesmo trabalho, análise de anos diferentes ou resultados de métodos diagnósticos diferentes, para cada uma dessas situações de interesse o estudo foi acrescentado como um novo estudo na metanálise. Para a chikungunya devido a um dos estudos apresentar prevalência igual a zero, foi imputado um valor de caso igual a 0,1 para esse estudo. A metanálise foi realizada para dengue, zika e chikungunya de forma isolada, usando um modelo de efeito randômico. A heterogeneidade foi avaliada pelo teste do qui-quadrado (χ^2) com significância de $P < 0,10$, sendo sua magnitude apurada pelo o I-quadrado (I^2). As análises foram realizadas com os comandos Metaprop e Metareg do programa Stata (versão 13.0). Foram realizadas metarregressões para identificar as causas da heterogeneidade, utilizando o teste de Knapp e Hartung. Primeiro, realizamos uma análise univariada testando todas as variáveis associadas com a prevalência das doenças, aquelas que apresentaram $p \leq 0,20$ foram incluídas no modelo multivariado. Na análise multivariada o nível de significância adotado foi de 5%. O efeito de estudos pequenos foi avaliado pela inspeção visual do gráfico de funil e pelo teste de Egger (MOREIRA, et al., 2019).

RESULTADOS

O processo de seleção baseado nas diretrizes do PRISMA está representado na Figura 4. A busca nas bases de dados resultou em 8052 artigos, dos quais 61 foram elegíveis para inclusão após o processo de triagem, onde 4711 foram excluídos por duplicata, 2818 pela leitura de títulos, 414 pelos resumos e 48 após a leitura de texto completo.

Figura 4 - Fluxograma de seleção de artigos sobre prevalência no Brasil de dengue, zika, chikungunya e febre amarela.



Fonte: Os autores

A seguir temos as características dos artigos incluídos de acordo com as informações das Tabelas suplementares 1, 2, 3, 4.

Os estudos incluídos nessa revisão foram publicados entre os anos de 1992 e 2019, sendo 65,57% dos artigos (n=40) foram publicados nos últimos 10 anos. Os estudos consideraram dados de coleta de amostra sorológica de surtos das arboviroses ocorridos em anos anteriores

aos anos de publicação, ou de estudos sorológicos independente da ocorrência de surto, dessa forma apresentam dados de 1986 a 2016. Observa-se que os dados de 2015 e 2016 se repetem em uma maior quantidade de estudos (n=13).

A dengue foi a arbovirose mais estudada nos artigos incluídos, estando presente em 49 deles (80,32%), seguida pela zika, em 11 estudos (18,03%), chikungunya, abordada em 10 estudos (16,39%) e febre amarela, em apenas 3 estudos (4,91%). Os sorotipos da dengue não foram definidos em todos os artigos, mas dentre os que definiram, os sorotipos 1 e 2 foram os mais estudados.

Dentre os 61 artigos selecionados, 73,77% utilizaram ensaio de imunoabsorção enzimática (ELISA) e / ou reação em cadeia de polimerase (RT-PCR ou RT-qPCR). Outros métodos como imunoaglutinação, isolamento viral, soroneutralização e outros foram utilizados nos estudos.

A maioria dos artigos descrevem principalmente dados epidemiológicos regionais derivados de pesquisas e estudos conduzidos em regiões específicas, sendo utilizados para apoiar os dados nacionais no que diz respeito à prevalência e sorotipo. Estudos de dezenove dos vinte e seis estados brasileiros foram incluídos na revisão, com exceção de Roraima, Piauí, Rio Grande do Norte, Alagoas, Espírito Santo, Santa Catarina e Paraná. Rio de Janeiro e São Paulo relataram o maior número de estudos (14 e 13 respectivamente). O estudo que abrangeu a maior amostra foi o de Macedo, et al., 2013, realizado no Rio de Janeiro, que contou com a análise do soro de 14408 indivíduos, o menor é o estudo de Faria, et al., 2016 com 61 indivíduos. A média de pessoas estudadas entre os estudos é de 1302,97 pessoas.

Resumo das características dos estudos de dengue, zika e chikungunya (Tabela 1).

Tabela 1 - Distribuição das principais características dos estudos de dengue, zika e chikungunya

(continua)

Região	Dengue	Zika	Chikungunya
	n (%)	n (%)	n (%)
Nordeste	23 (27,4)	2 (18,2)	8 (53,5)
Norte	8(9,5)	0 (0,0)	2 (13,3)
Centro-Oeste	13 (15,5)	1 (9,1)	1 (6,7)
Sudeste	39 (46,4)	6 (54,5)	4 (26,7)
Sul	1 (1,2)	0 (0,0)	0 (0,0)
Brasil	0 (0,0)	2 (18,2)	0 (0,0)

Tabela 1 - Distribuição das principais características dos estudos de dengue, zika e chikungunya
(conclusão)

	Dengue	Zika	Chikungunya
	n (%)	n (%)	n (%)
Método de Detecção			
Detecção de anticorpos específicos	30 (35,7)	1 (9,1)	9 (60,0)
Isolamento Viral	14 (16,7)	0 (0,0)	4 (26,7)
Detecção de ácido nucléico	16 (19,0)	8 (72,7)	0 (0,0)
Detecção de anticorpos virais	3 (3,6)	0 (0,0)	0 (0,0)
Múltiplos	21 (25,0)	2 (18,2)	2 (13,3)
Classificação de Dengue *			
1997	19 (22,6)	-	-
2000	35 (41,7)	-	-
2009	30 (35,7)	-	-
Inquérito sorológico			
Não	54 (64,3)	11 (100,0)	11 (73,3)
Sim	30 (35,7)	0 (0,0)	4 (26,7)
Área de Epidemia			
Não	57 (67,9)	6 (54,5)	6 (40,0)
Sim	27 (32,1)	5 (45,5)	9 (60,0)
Caso suspeito			
Não	38 (45,2)	5 (54,5)	7 (46,7)
Sim	46 (54,8)	6 (45,5)	8 (53,3)
Apenas crianças e adolescentes			
Não	70 (83,3)	11 (100,0)	15 (100,0)
Sim	14 (16,7)	0 (0,0)	0 (0,0)
Apenas mulheres			
Não	82 (97,6)	9 (81,8)	15 (100,0)
Sim	2 (2,4)	2 (18,2)	0 (0,0)
Casos gerais			
Não	41 (48,8)	7 (63,6)	8 (53,3)
Sim	43 (51,2)	4 (36,4)	7 (46,7)
Qualidade categorizada			
Baixa	0 (0,0)	0 (0,0)	0 (0,0)
Moderada	66 (78,6)	11 (100,0)	11 (73,3)
Alta	18 (21,4)	0 (0,0)	4 (26,7)

*classificação referente apenas à dengue, conforme referencias adotadas pela Organização Mundial da Saúde.

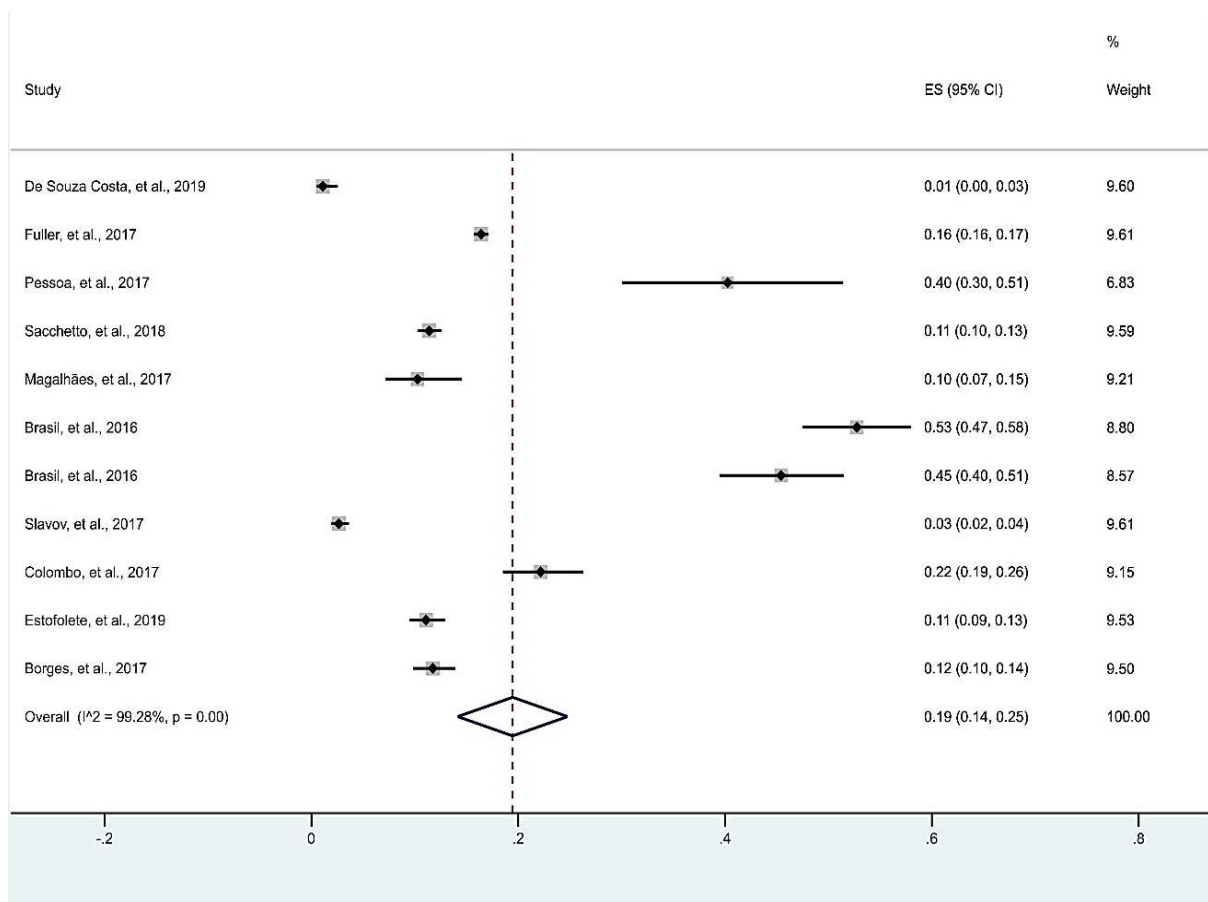
Fonte: dados dos artigos selecionados

Em relação à prevalência das arboviroses encontradas é possível observar uma variação de 0,27 a 94% de dengue em relação aos casos estudados, 0,0% a 75,35% de chikungunya, 0,66 a 68,42% de febre amarela e 1,1% a 52,75% de zika. Esses dados mostram que o número de casos flutuou durante todo o período da revisão.

A prevalência agrupada de zika entre indivíduos testados foi de 19% (IC95%: 0,14 a 0,25) no modelo randômico. A prevalência agrupada de chikungunya entre indivíduos testados foi de 28% (IC95%: 0,24 a 0,31) no modelo randômico. A prevalência agrupada de dengue entre indivíduos testados foi de 35% (IC95%: 0,30 a 0,39) no modelo randômico (Figura 5, Figura 6 e Figura 7).

Os estudos incluídos para zika, chikungunya e dengue foram avaliados quanto à heterogeneidade e ao viés de publicação. A heterogeneidade entre os estudos foi significativa ($P < 0,001$) para todas as doenças. Os resultados de teste Q ($P < 0,00$) e estatística I² ($I^2 = 99,28\%$) para zika. Teste Q ($P < 0,00$) e estatística I² ($I^2 = 99,71\%$) para chikungunya. E para dengue teste Q ($P < 0,00$) e estatística I² ($I^2 = 99,90\%$) (Figura 5, Figura 6 e Figura 7).

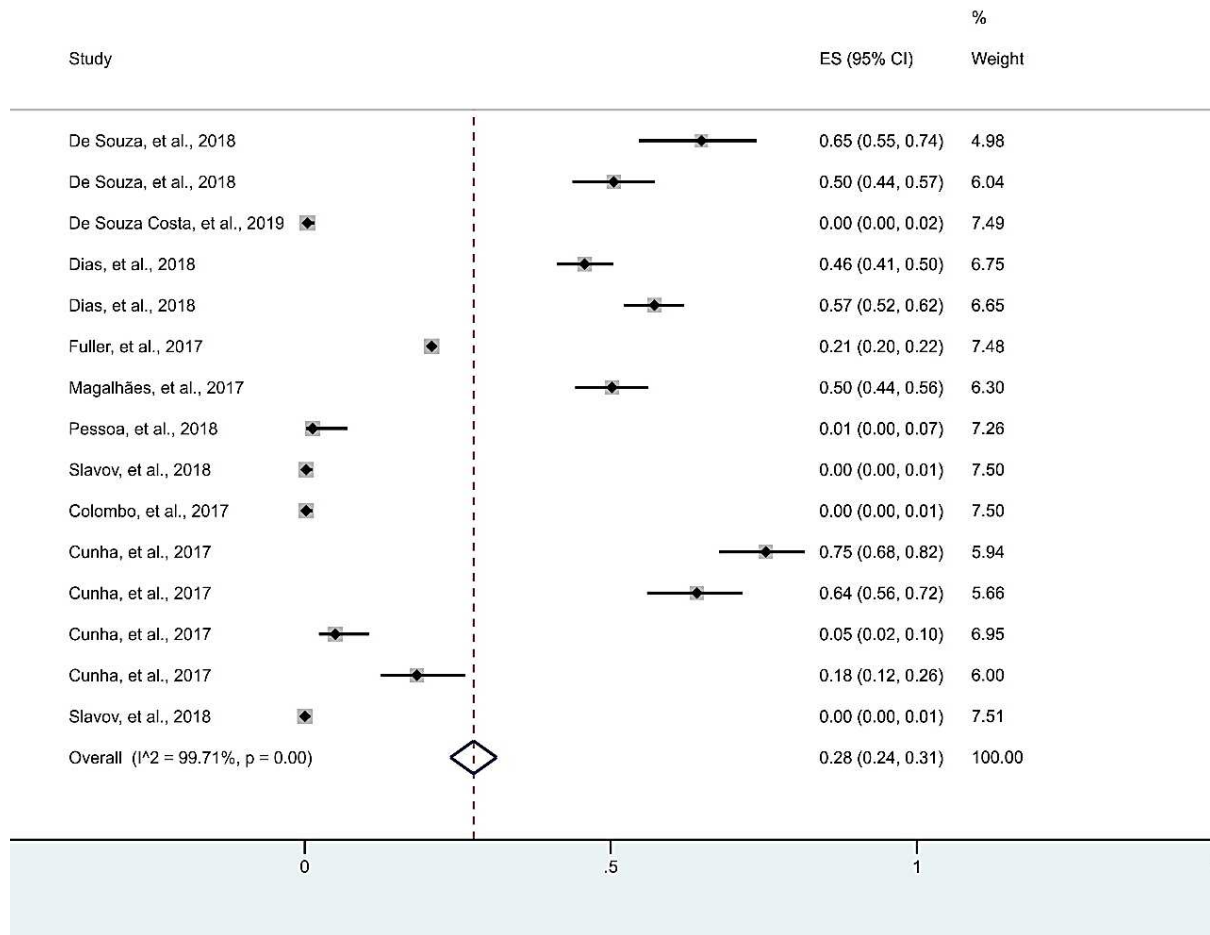
Figura 5 - Prevalência agrupada de zika



EP: estimativa de prevalência; 95%IC: intervalo de confiança de 95%.

Fonte: dados dos artigos selecionados

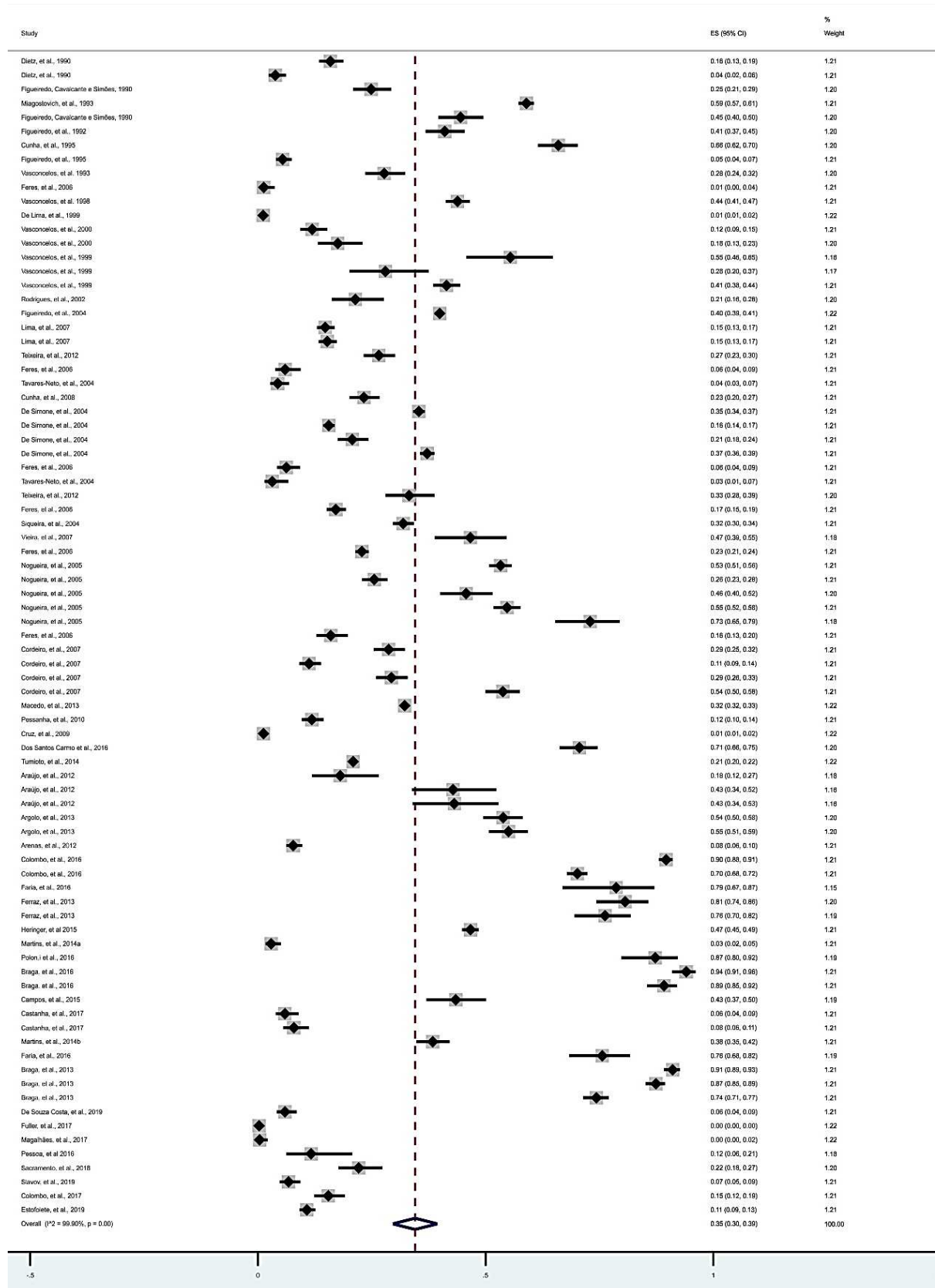
Figura 6 - Prevalência agrupada de chikungunya



EP: estimativa de prevalência; 95%IC: intervalo de confiança de 95%.

Fonte: dados dos artigos selecionados

Figura 7 - Prevalência agrupada de dengue

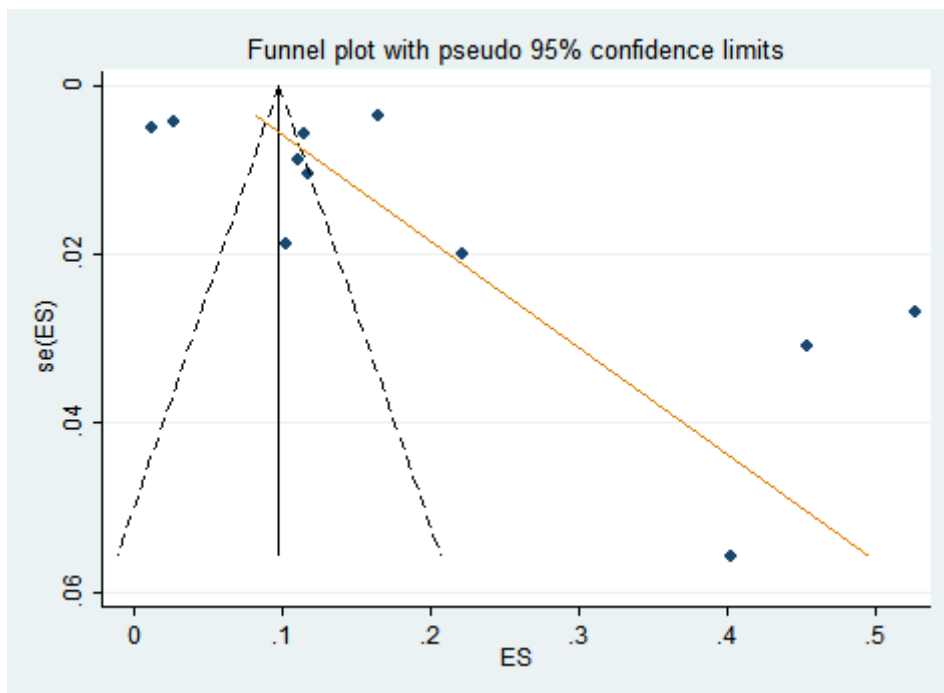


EP: estimativa de prevalência; 95%IC: intervalo de confiança de 95%.

Fonte: dados dos artigos selecionados

Para as três arboviroses, dengue, zika e chikungunya, nenhuma das variáveis, na análise subgrupos e metaregressão, explicou a grande heterogeneidade encontrada. Para zika, na análise visual do gráfico de funil, foi observado assimetria entre as investigações, mas não foi confirmado pelo teste de Egger ($P = 0,202$), apontando que é provável estudos pequenos, que possivelmente encontraram prevalências baixas, não terem sido publicados, contudo como poucos estudos entraram na revisão essa informação não pode ser confirmada (Figura 8).

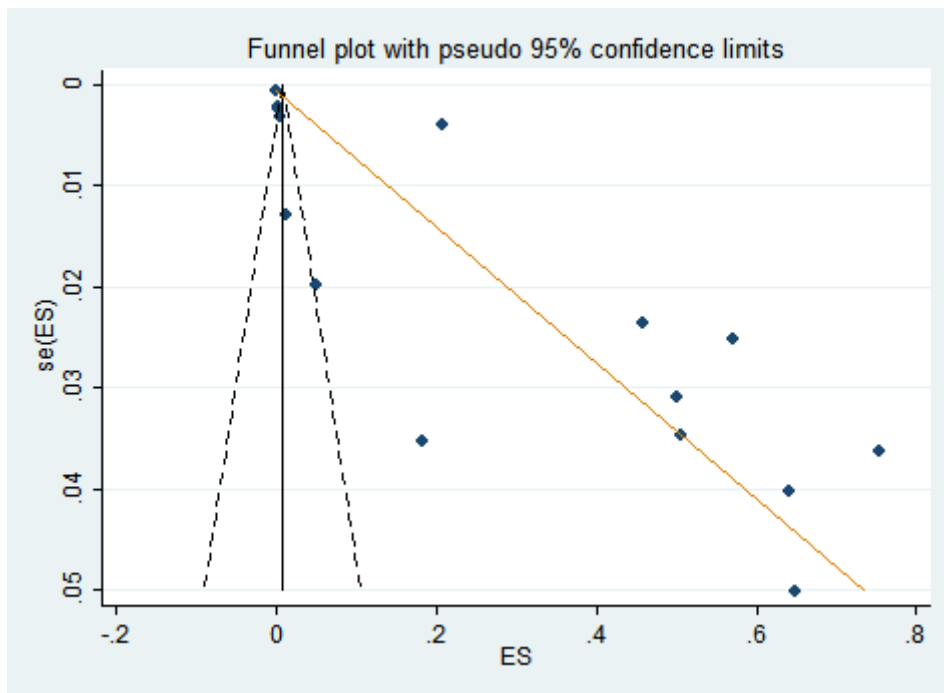
Figura 8 - Gráfico de funil: prevalência de zika por estudo pelo desvio padrão dos estudos



Fonte: os autores

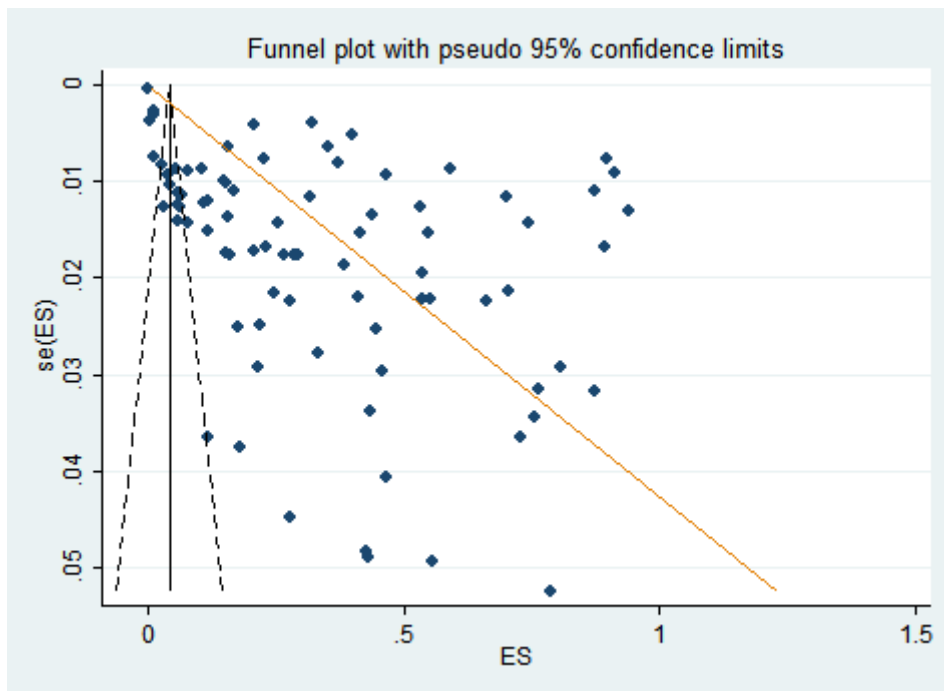
Para chikungunya e dengue, na análise visual do gráfico de funil, foi observado assimetria entre as investigações, confirmado pelo teste de Egger ($P = 0,003$ e $P = 0,000$, respectivamente), apontando que é provável estudos pequenos, que possivelmente encontraram prevalências baixas, não terem sido publicados. (Figura 9 e Figura 10).

Figura 9 - Gráfico de funil: prevalência de chikungunya por estudo pelo desvio padrão dos estudos



Fonte: os autores

Figura 10 - Gráfico de funil: prevalência de dengue por estudo pelo desvio padrão dos estudos



Fonte: os autores

DISCUSSÃO

Foram selecionados 61 estudos publicados entre os anos de 1992 e 2019, sendo 65,57% dos artigos (n=40) foram publicados nos últimos 10 anos. A maioria dos artigos selecionados descrevem principalmente dados epidemiológicos regionais derivados de pesquisas e estudos conduzidos em regiões específicas, sendo utilizados para apoiar os dados nacionais no que diz respeito à prevalência e sorotipo das doenças.

Não foi possível realizar a metanálise com os artigos de febre amarela, pois poucos estudos foram selecionados. A prevalência agrupada entre os indivíduos testados, no modelo randômico, de zika foi de 19% (IC95%: 0,14 a 0,25), de chikungunya foi de 28% (IC95%: 0,24 a 0,31) e de dengue foi de 35% (IC95%: 0,30 a 0,39).

Considerando que 2020 foi um ano atípico, devido a pandemia de COVID 19, que a manutenção de profissionais, equipamentos e suprimentos, foi comprometida, e que o contato social e as indicações para buscar o sistema de saúde estavam limitadas e as demais mudanças de comportamento ocasionadas pela pandemia, é possível que no ano de 2020 a notificação de outras doenças, como as arboviroses, tenham sido subnotificadas (LEANDRO et al., 2020; OPAS, 2020).

Nesse sentido, para comparação dos resultados encontrados, serão utilizados os dados fornecidos pelo Ministério da Saúde para o ano de 2019. E considerando que a estimativa de população brasileira para o ano de 2019 foi de 210.147.125 habitantes segundo o IBGE (Instituto Brasileiro de Geografia e Estatística), e que o número de casos prováveis de Zika, chikungunya e dengue, registrados nas Informações de Saúde (TABNET) do Sistema de Informação de Agravos de Notificação, foram estimadas as prevalências das doenças (Tabela 2).

Tabela 2 – Número de casos prováveis e estimativas de prevalência

Doença	Casos Prováveis* (n)	Prevalência Estimada (%)	Prevalência Agrupada (%)
Zika	30324	0,01	19
Chikungunya	177845	0,08	28
Dengue	1543665	0,73	35

Fonte: As autoras

*Fonte: BRASIL, Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net

A prevalência agrupada entre os indivíduos testados, no modelo randômico, de zika foi de 19% (IC95%: 0,14 a 0,25), muito superior a prevalência estimada de 0,01% para zika no ano de 2019. Já para chikungunya, a prevalência agrupada foi de 28% (IC95%: 0,24 a 0,31), também valor muito superior a prevalência estimada de 0,08% de chikungunya em 2019. Quanto a dengue, a prevalências agrupadas dos estudos foi de 35% (IC95%: 0,30 a 0,39), enquanto que a estimada para o ano de 2019 foi de 0,73%. Para as três arboviroses, dengue, zika e chikungunya, nenhuma das variáveis, na análise subgrupos e metaregressão, explicou a grande heterogeneidade encontrada.

Muitos dos estudos avaliados são inquéritos soropidemiológicos. Vasconcelos et al. 1998 descrevem a importância epidemiológica de inquéritos soro epidemiológicos que avaliam a população em geral, como forma de rastrear as infecções dado que muitos dos casos se apresentam de forma assintomática, considerando que esses indivíduos podem funcionar como fonte silenciosa da transmissão da doença.

A alta heterogeneidade dos estudos se manteve nas análises de subgrupos, o que impossibilita a extrapolação dos resultados. As grandes diferenças dos estudos quanto as principais características consideradas na caracterização dos estudos são fatores que interferem diretamente na prevalência de cada investigação. Além desses pontos, desigualdades sociais e acesso ao serviço de saúde também podem ter interferido nas prevalências observadas (BARRETO, 2017).

Outros fatores que influenciam a prevalência dessas doenças são as variações no número de casos das doenças no Brasil, de um ano para o outro, e dentro de um mesmo ano, dependendo da estação. E as particularidades de cada região, cada sub-região, cada estado e casa cidade, o que aumenta ainda mais a variação.

Os fatores ambientais, como precipitação, temperatura, humidade, localização geográfica e as mudanças climáticas podem ocasionar variação na transmissão das doenças, podendo, em alguns casos, ser preditor de potenciais surtos (NOGUEIRA et al., 2005; LIMA et al., 2007; FULLER et al., 2017; COLÓN-GONZALEZ et al., 2018). É visível que as características ambientais são determinantes nessas doenças, contudo alguns aspectos sobre os vetores e sua relação com o meio ambiente não estão claros, e considerando a vastidão territorial do Brasil e sua variedade de ambiental, novos dados precisam ser soletados.

O *A. aegypti* nas Américas, está bem estabelecido, e os aspectos macroestruturais, socioeconômicos e ambientais, são fatores que definem os caminhos das doenças, influenciam na mutação viral e nas adaptações genéticas dos vírus a hospedeiros, vetores e novos ambientes. Esses aspectos extrapolam o setor saúde e são historicamente ignorados em prol de intervenções

meramente biomédicas ou tecnológicas. Esse processo faz o controle do *Aedes* (vetor) ser um importante desafio, especialmente nos países em desenvolvimento (DONALISIO et al., 2017; VALLE, PIMENTA e AGUIAR, 2016).

Nesse sentido, o nível de urbanização das cidades e comunidades, o tipo de moradia, a ocupação do espaço, e o acesso a água municipal, são características que influem nas doenças. A dengue, por exemplo apresenta um perfil socioeconômico claro nas pessoas contaminadas, pessoas de baixa renda e estilo de vida em condições mais precárias (CUNHA et al. 2008; BRAGA et al., 2010; FULLER et al., 2017).

A co-circulação de diferentes tipos de arbovírus, cenário atual do Brasil, é um sério problema e desafio de saúde pública, reflete em termos de dinâmica de transmissão, competência de vetores, espectro clínico, resultados de saúde, diagnóstico e complicações das doenças. (FARIA et al., 2016; COLOMBO et al., 2017; ESTOFOLETE et al., 2019)

A prevenção dessas doenças necessita mais que a atuação individual, é preciso a intervenção do setor público no controle de endemias e nas medidas de prevenção. As prioridades precisam ser a educação, o saneamento básico e a melhoria das condições de moradia da população (LEFÈVRE, et al., 2007). A educação, comunicação e mobilização social são repetidamente citadas como fatores importantes no controle vetorial, mas pouco se discute sobre como fazer. As ações são escassas e desconectadas. As políticas públicas voltadas para esse assunto muitas vezes são incertas e os documentos norteadores são superficiais. As ações precisam ser, na maioria dos casos, intersetoriais, o que é pouco visto e muito menos executado (FARIA, et al 2023).

A verba para financiamento é outro assunto contraditório, visto que na maioria das vezes as secretarias não possuem verba para manter as ações das políticas públicas, e não conseguem ou não podem articular com instituições do segundo e/ou terceiro setor. A continuidade de serviços com a troca de mandatos políticos é outra preocupação, visto que, em muitas cidades brasileiras as ações realizadas pelo Estado podem ser facilmente interrompidas com a troca de gestão (AVELINO-SILVA e RAMOS, 2017; FARIA, et al. 2023).

A elaboração de estudos e pesquisas, se torna essencial, para encontrar formas alternativas para melhorar o planejamento e execução dessas políticas públicas. Contudo o processo com agências regulatórias pode ser burocrático e demorado, o que limita não apenas a execução de novas pesquisas, mas também, o usufruto da sociedade de melhorias (AVELINO-SILVA e RAMOS, 2017; FARIA, et al. 2023).

Como limitações entendemos que as características dos estudos eram bastante distintas o que pode ter dificultado a obtenção de resultados após os agrupamentos. E que os estudos publicados vieram de apenas quatro bases de dados, limitando a amplitude da busca.

CONCLUSÃO

O estudo demonstra uma alta prevalência de zika (19%) e chikungunya (28%) relacionadas a uma pesquisa pontual, focada apenas em de epidemias ou casos suspeitos, logo sendo necessário o avanço dos estudos para um melhor entendimento dessas doenças no país, para assim melhor preparar os serviços de saúde, garantindo um melhor atendimento da população. Sobre a dengue vemos que a prevalência da doença é alta (35%) contudo o novo cenário de COVID-19 associado a dengue e outras arboviroses requer uma atitude de vigilância e controle efetiva no país. Quanto a febre amarela, os poucos estudos encontrados indicam uma necessidade investigar os casos de febre amarela no país para assim melhor entender a realidade da doença no Brasil e possivelmente preencher as lacunas acerca da patologia da doença.

Todas essas arboviroses são extremamente complexas e não se relacionam apenas a patologia específica de cada uma. São doenças também de caráter socioambiental que requerem comprometimento dos governantes para adequação e efetividade do sistema de saúde ampliando e qualificando as ações de prevenção, controle, monitoramento e tratamento. De maneira geral é necessário que mais estudos foquem em estudar a prevalência para mais que um número, mas para entender fatores que levam a perpetuação dessas doenças.

AGRADECIMENTOS

Este trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 através de bolsas de estudo para estudantes da pós graduação.

REFERÊNCIAS

ARENAS, C.J. V.et al . The burden of dengue: Jundiaí, Brazil - January 2010. **Revista da Associação Médica Brasileira**. São Paulo , v. 58, n. 4, p. 477-480, ago. 2012. Disponível em: from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-42302012000400020&lng=en&nrm=iso>. Acesso em: 11 jan. 2021.

ARGOLO, A. F. L. T. et al. Prevalence and incidence of dengue virus and antibody placental transfer during late pregnancy in central Brazil. **BMC infectious diseases**. v. 13, n. 1, p. 254, 2013.

AVELINO-SILVA, V. I.; RAMOS, J. F. Arboviruses and public policies in Brazil. *Health Sciences Journal*, v. 7, n. 3, p. 1-2. 2017. Disponível em: <<https://doi.org/10.21876/rcsfmit.v7i3.675>>. Acesso em: 19 jan. 2021

BARRETO, M. L. Desigualdades en salud: una perspectiva global. **Ciência & Saúde Coletiva**. v. 22, p. 2097-2108, 2017.

BORGES J. R. E. et al. Zika Virus Outbreak-Should assisted reproduction patients avoid pregnancy?. **JBRA Assisted Reproduction**. v. 21, n. 3, p. 208, 2017.

BRAGA, C. et al. Seroprevalence and risk factors for dengue infection in socio-economically distinct areas of Recife, Brazil. **Acta tropica**, v. 113, n. 3, p. 234-240, 2010.

BRAGA, C. et al. Prospective birth cohort in a hyperendemic dengue area in Northeast Brazil: methods and preliminary results. **Cadernos de saude publica**, v. 32, p. e00095815, 2016.

BRASIL, P. et al. Zika virus outbreak in Rio de Janeiro, Brazil: clinical characterization, epidemiological and virological aspects. **PLoS neglected tropical diseases**, v. 10, n. 4, p. e0004636, 2016.

BRASIL. Ministério da Saúde. DATASUS. Tabnet. Sistema de Informação de Agravos de Notificação (SINAN) net. Brasília, DF: Ministério da Saúde, 2024 Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>. Acesso em: 15 set. 2024

BRASIL. Ministério da Saúde. Informe no 02 | 2018/2019. Monitoramento do Período Sazonal da Febre Amarela Brasil – 2018/2019. 2019a. Disponível em: <https://portalarquivos2.saude.gov.br/images/pdf/2018/dezembro/17/informe-FA-n.2_12dez18.pdf>. Acesso em: 19 jan. 2021.

BRASIL. Ministério da Saúde. Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes (dengue, chikungunya e Zika), Semanas Epidemiológicas 01 a 52. Secretaria de Vigilância em Saúde. v. 51. Jan. 2020b. Disponível em: <<https://antigo.saude.gov.br/images/pdf/2020/janeiro/20/Boletim-epidemiologico-SVS-02-1-.pdf>>. Acesso em: 19 jan. 2021

BRASIL. Ministério da Saúde. Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes Aegypti (dengue, chikungunya e zika), semanas epidemiológicas 1 a 50, 2020 Secretaria de Vigilância em Saúde. v. 51. Dez. 2020c. Disponível em: <https://www.gov.br/saude/pt-br/media/pdf/2020/dezembro/28/boletim_epidemiologico_svs_51.pdf>

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Monitoramento do Período Sazonal da Febre Amarela Brasil-2017/2018 MONITORAMENTO DA SITUAÇÃO EPIDEMIOLÓGICA DA FEBRE AMARELA NO BRASIL. 12. 2018.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. SVS. Monitoramento de Febre Amarela Brasil 2019 - informe no 18 – 9, Jun 2019. Secretaria de Vigilância em Saúde 1–8. 2019b.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Monitoramento dos casos de arboviroses urbanas transmitidas pelo *Aedes Aegypti* (dengue, chikungunya e zika), semanas epidemiológicas 1 a 38, 2020. 2020a

CAMPOS, G. S. et al. Dengue Virus 4 (DENV-4) re-emerges after 30 years in Brazil: Co-circulation of DENV 2, DENV-3 and DENV-4 in Bahia. **Japanese journal of infectious diseases**, p. JJID. 2014.063, 2014.

CASTANHA, P. M. S. et al. Incidence and risk factors for dengue virus (DENV) infection in the first 2 years of life in a Brazilian prospective birth cohort. **Epidemiology & Infection**, v. 145, n. 14, p. 2971-2979, 2017.

COLOMBO, T. E. et al. Clinical, laboratory and virological data from suspected ZIKV patients in an endemic arbovirus area. **Journal of Clinical Virology**, v. 96, p. 20-25, 2017.

COLOMBO, T. E. et al. Dengue virus surveillance: Detection of DENV-4 in the city of São José do Rio Preto, SP, Brazil. **Acta tropica**, v. 164, p. 84-89, 2016.

COLÓN-GONZÁLEZ, F. J. et al. Limiting global-mean temperature increase to 1.5–2° C could reduce the incidence and spatial spread of dengue fever in Latin America. **Proceedings of the National Academy of Sciences**, v. 115, n. 24, p. 6243-6248, 2018.

CORDEIRO, M. T. et al. Characterization of a dengue patient cohort in Recife, Brazil. **The American journal of tropical medicine and hygiene**, v. 77, n. 6, p. 1128-1134, 2007.

CRUZ, A. C. R. et al. Vigilância sorológica para arbovírus em Juruti, Pará, Brasil. **Cadernos de Saúde Pública**, v. 25, p. 2517-2523, 2009.

CUNHA, M. P. et al. Outbreak of chikungunya virus in a vulnerable population of Sergipe, Brazil—A molecular and serological survey. **Journal of Clinical Virology**, v. 97, p. 44-49, 2017a.

CUNHA, R. V. et al. Seroprevalence of chikungunya virus in a rural community in Brazil. **PLoS neglected tropical diseases**, v. 11, n. 1, p. e0005319, 2017b.

CUNHA, M. da C. M. et al. Fatores associados à infecção pelo vírus do dengue no Município de Belo Horizonte, Estado de Minas Gerais, Brasil: características individuais e diferenças intra-urbanas. **Epidemiologia e serviços de saúde**, v. 17, n. 3, p. 217-230, 2008. Disponível em: <http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742008000300007&lng=pt&nrm=iso>. Acesso em: 11 jan. 2021. <http://dx.doi.org/10.5123/S1679-49742008000300007>.

CUNHA, R. V. da et al. Secondary dengue infection in schoolchildren in a dengue endemic area in the State of Rio de Janeiro, Brazil. **Revista do Instituto de Medicina Tropical de São Paulo**, São Paulo, v. 37, n. 6, p. 517-521, dez. 1995. Disponível em:

<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0036-46651995000600008&lng=en&nrm=iso>. Acesso em: 11 Jan. 2021.

DE ARAÚJO, J. M. G. et al. Evaluation of a generic RT-nested-PCR for detection of flaviviruses in suspected fatal cases of dengue infection, Rio de Janeiro, Brazil. **Journal of virological methods**, v. 186, n. 1-2, p. 167-170, 2012. Disponível em: doi: 10.1016/j.jviromet.2012.08.010.

DE LIMA, V. LC et al. Dengue: inquérito sorológico pós-epidêmico em zona urbana do Estado de São Paulo (Brasil). **Revista de saúde pública**, v. 33, n. 6, p. 566-574, 1999.

DE SIMONE, T. S. et al. Dengue virus surveillance: the co-circulation of DENV-1, DENV-2 and DENV-3 in the State of Rio de Janeiro, Brazil. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 98, n. 9, p. 553-562, 2004.

DE SOUZA COSTA, M. C. et al. Arbovirus investigation in patients from Mato Grosso during Zika and Chikungunya virus introduction in Brazil, 2015–2016. **Acta tropica**, v. 190, p. 395-402, 2019.

DE SOUZA, T. M. A. et al. Following in the footsteps of the chikungunya virus in Brazil: the first autochthonous cases in Amapá in 2014 and its emergence in Rio de Janeiro during 2016. **Viruses**, v. 10, n. 11, p. 623, 2018.

DIAS, J. P. et al. Seroprevalence of chikungunya virus after its emergence in Brazil. **Emerging infectious diseases**, v. 24, n. 4, p. 617, 2018.

DLETZ, V. J. et al. Epidemic dengue 1 in Brazil, 1986: evaluation of a clinically based dengue surveillance system. **American journal of epidemiology**, v. 131, n. 4, p. 693-701, 1990.

DO ROSÁRIO M.S. e DE SIQUEIRA I.C. Preocupações sobre surtos simultâneos de COVID-19 e arbovirais (chikungunya, dengue, zika). **The Brazilian Journal of Infectious Diseases**. v. 24, n. 6, p. 583-584, 2020. Disponível em: <doi: 10.1016 / j.bjid.2020.08.008>

DONALISIO, M. R.; FREITAS, A. Ri. R.; ZUBEN, A. P. B. Von. Arboviroses emergentes no Brasil: desafios para a clínica e implicações para a saúde pública. **Revista de Saúde Pública**, São Paulo , v. 51, 30, 2017. Disponível em:

<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102017000100606&lng=en&nrm=iso>. Acesso em: 11 Jan. 2021.

DOS SANTOS CARMO, A. M. et al. Maintenance of demographic and hematological profiles in a long-lasting dengue fever outbreak: implications for management. **Infectious diseases of poverty**, v. 5, n. 1, p. 84, 2016.

ESTOFOLETE, C. F. et al. Co-infection between Zika and different Dengue serotypes during DENV outbreak in Brazil. **Journal of Infection and Public Health**, v. 12, n. 2, p. 178-181, 2019.

EVANGELISTA, J. G. et al. Agentes de combate às endemias: construção de identidades profissionais no controle da dengue. **Trabalho, educação e saúde**, Rio de Janeiro, v. 17, n. 1, e0017303, 2019.

FARIA, M. T. da S. et al. Saúde e saneamento: uma avaliação das políticas públicas de prevenção, controle e contingência das arboviroses no Brasil. *Ciência & Saúde Coletiva*, v. 28, p. 1767-1776, 2023. Disponível em: <https://www.scielo.org/article/csc/2023.v28n6/1767-1776/#>. Acesso em: 17 nov. 2024

FARIA, N. R. da C. et al. Dengue epidemics in two distinct periods reveal distinct epidemiological, laboratorial and clinical aspects in a same scenario: analysis of the 2010 and 2013 epidemics in Mato Grosso do Sul, Brazil. **Transactions of The Royal Society of Tropical Medicine and Hygiene**, v. 110, n. 4, p. 228-236, 2016.

FERES, V. C. R. et al. Laboratory surveillance of dengue virus in Central Brazil, 1994–2003. **Journal of clinical virology**, v. 37, n. 3, p. 179-183, 2006.

FERRAZ, F. O. et al. Evaluation of laboratory tests for dengue diagnosis in clinical specimens from consecutive patients with suspected dengue in Belo Horizonte, Brazil. **Journal of Clinical Virology**, v. 58, n. 1, p. 41-46, 2013.

FIGUEIREDO, L. T. M. et al. Dengue serologic survey in Ribeirão Preto, São Paulo, Brazil. **Bulletin of the Pan American Health Organization (PAHO)**. v. 29, n. 1, mar. 1995.

FIGUEIREDO, L. T. M. et al. Laboratory diagnosis and symptoms of dengue studied during an outbreak in the Ribeirão Preto Region, SP, Brazil. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 34, n. 2, p. 121-130, 1992.

FIGUEIREDO, L. T. M.; CAVALCANTE, S. M. B.; SIMOES, M. C. Dengue serologic survey of schoolchildren in Rio de Janeiro, Brazil, in 1986 and 1987. **Bulletin of the Pan American Health Organization (PAHO)**. v. 24, n. 2, 1990.

FIGUEIREDO, R. M. P. de et al. Doenças exantemáticas e primeira epidemia de dengue ocorrida em Manaus, Amazonas, no período de 1998-1999. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 37, n. 6, p. 476-9, 2004.

FULLER, T. L. et al. Behavioral, climatic, and environmental risk factors for Zika and Chikungunya virus infections in Rio de Janeiro, Brazil, 2015-16. **PloS one**, v. 12, n. 11, p. e0188002, 2017.

HERINGER, M. et al. Dengue type 4 in Rio de Janeiro, Brazil: case characterization following its introduction in an endemic region. **BMC infectious diseases**, v. 17, n. 1, p. 410, 2017.

HERINGER, M. et al. Impact of the emergence and re-emergence of different dengue viruses' serotypes in Rio de Janeiro, Brazil, 2010 to 2012. **Transactions of The Royal Society of Tropical Medicine and Hygiene**, v. 109, n. 4, p. 268-274, 2015.

IBGE. SIDRA. Estimativas da população. Disponível em:

<https://sidra.ibge.gov.br/tabela/6579>

LEANDRO, C. dos S. et al. Redução da incidência de dengue no Brasil em 2020: controle ou subnotificação de casos por COVID-19?. **Research, Society and Development**, v. 9, n. 11, p. e76891110442-e76891110442, 2020.

LEFÈVRE, A. M.C. et al. Representações sobre dengue, seu vetor e ações de controle por moradores do Município de São Sebastião, Litoral Norte do Estado de São Paulo, Brasil. **Cadernos de Saúde Pública**, v. 23, p. 1696-1706, 2007.

LIMA, V. L. C. de et al. Dengue: inquérito populacional para pesquisa de anticorpos e vigilância virológica no Município de Campinas, São Paulo, Brasil. **Caderno de Saúde Pública**, Rio de Janeiro, v. 23, n. 3, p. 669-680, mar. 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2007000300025&lng=en&nrm=iso>. Acesso em: 11 jan. 2021.

LOPES, N.; NOZAWA, C.; LINHARES, R. E. C. Características gerais e epidemiologia dos arbovírus emergentes no Brasil. **Revista Pan-Amazônica de Saúde**, v. 5, n. 3, p. 10-10, 2014. Disponível em <http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S2176-62232014000300007&lng=pt&nrm=iso>. Acesso em: 11 jan. 2021.

MACEDO, Gleicy Amorim et al. Virological surveillance for early warning of dengue epidemics in the state of Rio de Janeiro, Brazil. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 107, n. 3, p. 141-146, 2013.

MAGALHAES, T. et al. Zika virus displacement by a chikungunya outbreak in Recife, Brazil. **PLoS neglected tropical diseases**, v. 11, n. 11, p. e0006055, 2017.

MARTINS, A. C. et al. Seroprevalence and seroconversion of dengue and implications for clinical diagnosis in Amazonian children. **Interdisciplinary Perspectives on Infectious Diseases**, v. 2014, 2014b.

MARTINS, V. do Carmo Alves et al. Clinical and virological descriptive study in the 2011 outbreak of dengue in the Amazonas, Brazil. **PLoS One**, v. 9, n. 6, p. e100535, 2014a.

MASCARENHAS, M. D. M. et al. Ocorrência simultânea de COVID-19 e dengue: o que os dados revelam?. **Cadernos de Saúde Pública**, v. 36, p. e00126520, 2020.

MIAGOSTOVICH, M. P. et al. Dengue epidemic in the state of Rio de Janeiro, Brazil: virological and epidemiological aspects. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 35, n. 2, p. 149-154, Mar./Apr. 1993.

MOREIRA, T. R. et al. Prevalência de tuberculose na população privada de liberdade: revisão sistemática e metanálise. **Revista Panamericana de Salud Pública**, v. 43, p. e16, 2019.

MUNN, Z. et al. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. **International journal of evidence-based healthcare**, v. 13, n. 3, p. 147-153, 2015.

NOGUEIRA, R. M. R. et al. Dengue virus type 3, Brazil, 2002. **Emerging infectious diseases**, v. 11, n. 9, p. 1376, 2005.

OPAS Brasil. Casos de dengue nas Américas chegam a 1,6 milhão, o que destaca a necessidade do controle de mosquitos durante a pandemia. 2020. Disponível em: <https://www.paho.org/bra/index.php?option=com_content&view=article&id=6205:casos-de-dengue-nas-americas-chegam-a-1-6-milhao-o-que-destaca-a-necessidade-do-controle-de-mosquitos-durante-a-pandemia&Itemid=812#:~:text=23%20de%20junho%20de%202020,%C3%A0%20pandemia%20de%20COVID%2D19>.

PESSANHA, J. E. M. et al. Dengue em três distritos sanitários de Belo Horizonte, Brasil: inquérito soropidemiológico de base populacional, 2006 a 2007. **Revista Panamericana de Salud Pública**, v. 27, p. 252-258, 2010.

PESSÔA, R. et al. Investigation into an outbreak of dengue-like illness in Pernambuco, Brazil, revealed a cocirculation of Zika, Chikungunya, and dengue virus type 1. **Medicine**, v. 95, n. 12, 2016.

POLONI, T. R. et al. High prevalence of clinically unsuspected dengue disease among children in Ribeirão Preto City, Brazil. **Journal of medical virology**, v. 88, n. 10, p. 1711-1719, 2016.

RODRIGUES, E. M. S. et al. Epidemiologia da infecção pela dengue em Ribeirão Preto, SP, Brasil. **Revista de Saúde Pública**, v. 36, n. 2, p. 160-165, 2002.

SACCHETTO, L. et al. Screening for Zika virus RNA in sera of suspected cases: a retrospective cross-sectional study. **Virology journal**, v. 15, n. 1, p. 155, 2018.

SACRAMENTO, R. H. M. et al. Dengue Fever and *Aedes aegypti* in indigenous Brazilians: seroprevalence, risk factors, knowledge and practices. **Tropical Medicine & International Health**, v. 23, n. 6, p. 596-604, 2018.

SIQUEIRA, J. B. et al. Household survey of dengue infection in central Brazil: spatial point pattern analysis and risk factors assessment. **The American journal of tropical medicine and hygiene**, v. 71, n. 5, p. 646-651, 2004.

SLAVOV, S. N. et al. Zika virus RNA detection in asymptomatic blood donors during an outbreak in the northeast region of São Paulo State, Brazil, 2016. **Transfusion**, v. 57, n. 12, p. 2897-2901, 2017.

SLAVOV, S. N. et al. Dengue seroprevalence among asymptomatic blood donors during an epidemic outbreak in Central-West Brazil. **Plos one**, v. 14, n. 3, p. e0213793, 2019.

SLAVOV, S. N. et al. Seroprevalence of Chikungunya virus in blood donors from Northern and Southeastern Brazil. **Hematology, transfusion and cell therapy**, v. 40, n. 4, p. 358-362, 2018.

TAVARES-NETO, J. et al. Pesquisa de anticorpos contra arbovírus e o vírus vacinal da febre amarela em uma amostra da população de Rio Branco, antes e três meses após a vacina 17D. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 37, n. 1, p. 1-6, 2004.

TEIXEIRA, M. G. et al. Risk factors for the incidence of dengue virus infection in preschool children. **Tropical Medicine & International Health**, v. 17, n. 11, p. 1391-1395, 2012.

TUMIOTO, G. L. et al. Laboratory surveillance of dengue in Rio Grande do Sul, Brazil, from 2007 to 2013. **PLoS One**, v. 9, n. 8, p. e104394, 2014.

VALLE, D.; PIMENTA, D. N.; AGUIAR, R. Zika, dengue e chikungunya: desafios e questões. **Epidemiologia e Serviços de Saúde**, v. 25, p. 419-422, 2016. Disponível em: <http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742016000200419&lng=pt&nrm=iso>. Acesso em: 11 jan. 2021.

VASCONCELOS, P. F. C. et al. Dengue epidemic in a Northeastern Brazil: random epidemiological serum survey. **Revista de Saúde Pública**, v. 32, n. 5, p. 447-454, 1998.

VASCONCELOS, P. F. C. et al. Epidemia de febre clássica de dengue causada pelo sorotipo 2 em Araguaiana, Tocantins, Brasil. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 35, n. 2, p. 141-148, 1993.

VASCONCELOS, P. F. C. et al. Inquérito soro-epidemiológico na Ilha de São Luís durante epidemia de dengue no Maranhão. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 32, n. 2, p. 171-179, 1999.

VASCONCELOS, P. F. C. et al. Epidemia de dengue em Ipujiara e Prado, Bahia. Inquérito soro-epidemiológico. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 33, n. 1, p. 61-67, 2000.

VIEIRA, D. Souza et al. Characterization of dengue virus serotype 1 in epidemics in Porto Velho, Rondônia, in 2001-2003. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 40, n. 3, p. 268-271, 2007.

WEAVER, S. C.; REISEN, W. K. Present and future arboviral threats. **Antiviral research**, v. 85, n. 2, p. 328-345, 2010.

ZARA, A. L. de S. A. et al. Estratégias de controle do *Aedes aegypti*: uma revisão. **Epidemiologia e Serviços de Saúde**, v. 25, p. 391-404, 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222016000200391&lng=en&nrm=iso>. Acesso em: 11 Jan. 2021.

5.1.1. Material Suplementar – Artigo de Revisão

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continua)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
De Araújo, et al., 2012	Rio de Janeiro	DENV 2	2008	105	18	Casos Fatais Suspeitos de dengue	RT-PCR genérico	17,1	4
	Rio de Janeiro	DENV 3	2008	105	1	Casos Fatais Suspeitos de dengue	RT-PCR genérico	1,0	4
	Rio de Janeiro	DENV 2	2008	105	40	Casos Fatais Suspeitos de dengue	RT-PCR convencional	38,1	4
	Rio de Janeiro	DENV 3	2008	105	5	Casos Fatais Suspeitos de dengue	RT-PCR convencional	4,8	4
	Rio de Janeiro	DENV	2008	102	44	Casos Fatais Suspeitos de dengue	IGM	43,1	4
Arenas, et al., 2012	Jundiaí/SP	DENV	2010	876	68	Moradores dos Bairros de novo horizonte e São Camilo	ELISA	7,8	7
Argolo, et al., 2013	Goiânia/GO	DENV	2009-2010	505		Parturientes	RT-PCR; antígeno NS; ELISA; IgM	53,9	6
	Goiânia/GO	DENV	2009-2010	505		Recém-nascidos	RT-PCR; antígeno NS; ELISA; IgM	55,0	6
Braga, et al., 2013	Recife/PE	DENV	2015-2016	976		Área de privação	ELISA	91,1	8
	Recife/PE	DENV	2015-2016	923		Área intermediária	ELISA	87,4	8
	Recife/PE	DENV	2015-2016	920		Área rica	ELISA	74,3	8
Braga, et al., 2016	Recife/PE	DENV	2011-2012	321	302	Recém-nascidos	IgM ou RT-PCR	94,1	5
	Recife/PE	DENV	2011-2014	342	305	Parturientes	IgM ou RT-PCR	89,2	5

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Campos, et al., 2015	Bahia	DENV	2011-2012	214	93	Geral	(RT-PCR)/nested PCR	43,5	5
	Bahia	DENV 4	2011-2012	214	77	Geral	(RT-PCR)/nested PCR	36,0	5
	Bahia	DENV 3	2011-2012	214	3	Geral	(RT-PCR)/nested PCR	1,4	5
	Bahia	DENV 2	2011-2012	214	13	Geral	(RT-PCR)/nested PCR	6,1	5
Dos Santos Carmo, et al., 2016	Marília/SP	DENV	2007	456	322	Pacientes com diagnóstico clínico de dengue	Ensaio imunoenzimático de antígeno não estrutural 1 (NS1)	70,6	5
Castanha, et al., 2017	Recife/PE	DENV	2011-2014			crianças menores de 1 ano	IgM/IgG; RT-PCR	107,6/1000 pessoas-ano	7
	Recife/PE	DENV	2011-2014			crianças com idade entre 1-2 anos	IgM/IgG; RT-PCR	93,3/1000 pessoas-ano	7
Colombo, et al., 2016	São José do Rio Preto/SP	DENV	2010 – 2014	1549	1.389	Geral	NS1	89,7	5
	São José do Rio Preto/SP	DENV	2010 – 2014	1549	1087	Geral	Multiplex-Nested-PCR	70,2	5
	São José do Rio Preto/SP	DENV 4	2010 – 2014	1087	528	Geral	Multiplex-Nested-PCR NS1	48,6	5
	São José do Rio Preto/SP	DENV 1	2010 – 2014	1087	449	Geral	Multiplex-Nested-PCR NS1	41,3	5
	São José do Rio Preto/SP	DENV 2	2010 – 2014	1087	104	Geral	Multiplex-Nested-PCR NS1	9,6	5

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Colombo, et al., 2017	São José do Rio Preto/SP	DENV 1	2016	433	9	Geral	RT - PCR	2,1	6
	São José do Rio Preto/SP	DENV 2	2016	433	56	Geral	RT - PCR	12,9	6
	São José do Rio Preto/SP	DENV 4	2016	433	2	Geral	RT - PCR	0,5	6
Cordeiro, et al., 2007	Recife/PE	DENV	2004-2006	658	189	Geral, maior ou igual a cinco anos	RT-PCR	28,7	5
	Recife/PE	DENV	2004-2006	658	74	Geral, maior ou igual a cinco anos	Isolamento do vírus	11,2	5
	Recife/PE	DENV	2004-2006	658	193	Geral, maior ou igual a cinco anos	ELISA IgG	29,3	5
	Recife/PE	DENV	2004-2006	658	354	Geral, maior ou igual a cinco anos	ELISA IgM	53,8	5
Cruz, et al., 2009	Jurite/PA	DENV	2007-2008	1597	20	Demanda passiva de pessoas com ou sem sintomas febris	MAC-ELISA	1,3	5
Cunha, et al., 1995	Niterói/RJ	DENV	1991-1992	450	297	Escolares entre 5 e 19 anos de escolas públicas	Inibição da hemaglutinação (IH)	66,0	5
Cunha, et al., 2008	Belo Horizonte/MG	DENV	2000	627	146	Geral	Teste de soroneutralização.	23,3	6

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
De Simone, et al., 2004	Rio de Janeiro/RJ	DENV 1	2000—2001	606	380	Geral	Isolamento viral e/ou RT—PCR	62,7	6
	Rio de Janeiro/RJ	DENV 2	2000—2001	606	147	Geral	Isolamento viral e/ou RT—PCR	24,3	6
	Rio de Janeiro/RJ	DENV 3	2000—2001	606	79	Geral	Isolamento viral e/ou RT—PCR	13,0	6
Dietz, et al., 1990	Nova Iguaçu/RJ	DENV	1986	709	113	Geral	MAC-ELISA	15,9	7
	Niterói/RJ	DENV	1986	416	16	Geral	MAC-ELISA	3,8	7
Estofolete, et al., 2019	São José do Rio Preto/SP	DENV 1	2016	1254	41	Geral	Multiplex-Nested-PCR (MN-PCR)	3,3	6
	São José do Rio Preto/SP	DENV 2	2016	1254	89	Geral	Multiplex-Nested-PCR (MN-PCR)	7,1	6
	São José do Rio Preto/SP	DENV 4	2016	1254	5	Geral	Multiplex-Nested-PCR (MN-PCR)	0,4	6
Faria, et al., 2016	Campo Grande/MS	DENV	2010	61	48	Geral	RT-PCR, MACELISA, NS1 ELISA e IgG-ELISA.	78,7	6
	Campo Grande/MS	DENV	2010	48	41	Geral	MAC ELISA	85,4	6
	Campo Grande/MS	DENV	2010	48	14	Geral	NS1 ELISA	29,2	6
	Campo Grande/MS	DENV	2010	48	12	Geral	RT-PCR	25,0	6

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Faria, et al., 2016	Campo Grande/MS	DENV 1	2010	12	6	Geral	RT-PCR	50,0	6
	Campo Grande/MS	DENV 2	2010	12	6	Geral	RT-PCR	50,0	6
	Campo Grande/MS	DENV	2013	156	118	Geral	RT-PCR, MACE LISA, NS1 ELISA e IgG-ELISA.	75,6	6
	Campo Grande/MS	DENV	2013	118	55	Geral	MAC ELISA	46,6	6
	Campo Grande/MS	DENV	2013	118	84	Geral	NS1 ELISA	71,2	6
	Campo Grande/MS	DENV	2013	118	36	Geral	RT-PCR	30,5	6
	Campo Grande/MS	DENV 4	2013	36	33	Geral	RT-PCR	91,7	6
	Campo Grande/MS	DENV 1	2013	36	2	Geral	RT-PCR	5,6	6
	Campo Grande/MS	DENV 2	2013	36	1	Geral	RT-PCR	2,8	6
Ferraz, et al., 2013	Belo Horizonte/MG	DENV	2010	181	146	Maiores de 18 anos com suspeita de Dengue	NS1/IgM	80,7	5
	Belo Horizonte/MG	DENV	2010	181	138	Maiores de 18 anos com suspeita de Dengue	PCR em tempo real	76,2	5

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Figueiredo, et al., 1992	Ribeirão Preto/SP	DENV	1990-1991	502	206	Geral	Inibição da hemaglutinação (IH)	41,0	5
Figueiredo, et al., 1995	Ribeirão Preto/SP	DENV 1	1992	662	36	5 - 15 anos de idade	Imunoensaio Enzimático em Células Cultivadas Infectadas (EI-ICC)	5,4	7
Figueiredo, et al., 2004	Manaus/AM	DENV	1998 – 1999	8557	3416	Usuários de unidade de saúde	MAC-ELISA	39,9	6
Figueiredo, Cavalcante e Simões, 1990	Rio de Janeiro/RJ	DENV 1	1986	402		CRIANÇAS - 4 A 14 ANOS	Inibição da hemaglutinação (IAH)	24,9	6
	Rio de Janeiro/RJ	DENV 1	1987	384		CRIANÇAS - 4 A 14 ANOS	Inibição da hemaglutinação (IAH)	44,5	6
Fuller, et al., 2017	Rio de Janeiro	DENV	2015-2016	10459	29	Geral	CDC RT-PCR	0,3	6
Heringer, et al., 2015	Rio de Janeiro	DENV	2010-2012	2833	1323	Casos suspeitos	MAC-ELISA ; (IgG-ELISA;	46,7	5
Heringer, et al., 2017	Rio de Janeiro/RJ	DENV- 4	2011-2013	3727	1593	Geral	isolamento viral (teste de anticorpo fluorescente indireto), RT-PCR e ELISA de captura de antígeno NS1, MAC-ELISA e IgG-ELISA	42,7	6

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Heringer, et al., 2017	Rio de Janeiro/RJ	DENV 4	2011-2013	3727	298	Geral	isolamento viral (teste de anticorpo fluorescente indireto)	8,0	6
	Rio de Janeiro/RJ	DENV 4	2011-2013	705	704	Geral	RT-PCR	99,9	6
	Rio de Janeiro/RJ	DENV 4	2011-2013	264	104	Geral	MAC-ELISA	39,4	6
	Rio de Janeiro/RJ	DENV 4	2011-2013	557	231	Geral	ELISA de captura de antígeno NS1	41,5	6
	Rio de Janeiro/RJ	DENV 4	2011-2013	557	488	Geral	NS1 Capture ELISA after dissociation	87,6	6
Lima, et al., 2007	Campinas/SP	DENV	1998	1.257	186	Geral	EIA-ICC, um teste imunoenzimático	14,8	8
	Campinas/SP	DENV	1998	1.260	192	Geral	IgG	15,2	8
Macedo, et al., 2013	Região Metropolitana do Rio de Janeiro	DENV	2004-2008	14408	4648	Pacientes que apresentavam doença febril aguda com duas ou mais das seguintes manifestações clínicas: cefaleia, dor retrobulbar, mialgia, artralgia, erupção cutânea, epistaxe e / ou hemorragia.	RT-PCR, Isolação Viral, Testes sorológicos	32,3	5

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Magalhães, et al., 2017	Paulista/PE	DENV	2015-2016	263	1	Pacientes com sugestão de doença febril aguda atendidos	DENV NS1 Ag;	0,4	5
Martins, et al., 2014a	Assis Brasil/AC	DENV	2010	411	12	Crianças de 6 meses a 12 anos	ELISA IgG	2,9	5
Martins, et al., 2014b	Manaus/AM	DENV	2011	677	260	Geral, tirando pacientes com malária	RT - PCR	38,4	5
	Manaus/AM	DENV 1	2011	260	26	Geral, tirando pacientes com malária	RT - PCR	10,0	5
	Manaus/AM	DENV 2	2011	260	120	Geral, tirando pacientes com malária	RT - PCR	46,2	5
	Manaus/AM	DENV 3	2011	260	24	Geral, tirando pacientes com malária	RT - PCR	9,2	5
	Manaus/AM	DENV 4	2011	260	77	Geral, tirando pacientes com malária	RT - PCR	29,6	5
Martins et al., 2014b	Manaus/AM	Co-infecções (DENV-1/2, DENV-1/4, DENV-2/4 e DENV-3/4)	2011	260	13	Geral, tirando pacientes com malária	RT - PCR	5,0	5
Miagostovich, et al., 1993	Rio de Janeiro/RJ	DENV	1986-1987	3178	1874	Geral	Isolamento viral/ MAC ELISA	59,0	6

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Nogueira, et al., 2005	Rio de Janeiro/RJ	DENV	2002	1559	831	Geral	Isolamento viral, RT-PCR, MAC ELISA, IgG ELISA	53,3	6
	Rio de Janeiro/RJ	DENV	2002	927	237	Geral	isolamento viral	25,6	6
	Rio de Janeiro/RJ	DENV 3	2002	282	129	Geral	RT - PCR	45,7	6
	Rio de Janeiro/RJ	DENV 2	2002	282	1	Geral	RT - PCR	0,4	6
	Rio de Janeiro/RJ	DENV 1	2002	282	2	Geral	RT - PCR	0,7	6
	Rio de Janeiro/RJ	DENV	2002	1060	580	Geral	MAC ELISA	54,7	6
	Rio de Janeiro/RJ	DENV	2002	148	108	Geral	IgG ELISA	73,0	6
Pessanha, et al., 2010	Belo Horizonte/MG	DENV 1	2006-2007	709	15	Geral, maior ou igual a um ano	soroneutralização (SN)	2,1	8
	Belo Horizonte/MG	DENV 1 - 2	2006-2007	709	8	Geral, maior ou igual a um ano	soroneutralização (SN)	1,1	8
	Belo Horizonte/MG	DENV 1 - 2 - 3	2006-2007	709	28	Geral, maior ou igual a um ano	soroneutralização (SN)	3,9	8
	Belo Horizonte/MG	DENV 1 - 3	2006-2007	709	8	Geral, maior ou igual a um ano	soroneutralização (SN)	1,1	8

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Pessanha, et al., 2010	Belo Horizonte/MG	DENV 2	2006-2007	709	9	Geral, maior ou igual a um ano	soroneutralização (SN)	1,3	8
	Belo Horizonte/MG	DENV 2 - 3	2006-2007	709	6	Geral, maior ou igual a um ano	soroneutralização (SN)	0,8	8
	Belo Horizonte/MG	DENV 3	2006-2007	709	10	Geral, maior ou igual a um ano	soroneutralização (SN)	1,4	8
Pessoa, et al 2016	Tuparetama/PE	DENV	2015	77	9	Pacientes ambulatório do Hospital Severino de Souto Siqueira	qRT-PCR, IgM	11,7	5
Poloni, et al., 2016	Ribeirão Preto/SP	DENV	2010-2011	110	96	Crianças maiores de 1 ano com doença febril	ELISA	87,3	5
Rodrigues, et al., 2002	Ribeirão Preto/SP	DENV	1997	196	42	Internos e funcionários de uma instituição correcional de adolescentes infratores	MAC-ELISA	21,4	5
Sacramento, et al., 2018	Ceará	DENV	2015	290	62	Índios da etnia Tremembé da aldeia de Tapera	ELISA	21,4	6

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(continuação)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Sacramento, et al., 2018	Ceará	DENV 1	2015	27	21	Índios da etnia Tremembé da aldeia de Tapera	microneutralização	77,8	6
	Ceará	DENV 2	2015	27	19	Índios da etnia Tremembé da aldeia de Tapera	microneutralização	70,4	6
	Ceará	DENV 3	2015	27	4	Índios da etnia Tremembé da aldeia de Tapera	microneutralização	14,8	6
	Ceará	DENV 4	2015	27	3	Índios da etnia Tremembé da aldeia de Tapera	microneutralização	11,1	6
Siqueira, et al 2004	Goiania/GO	DENV	2001	1585		Geral	IgM/IgG	29,5	8
Slavov, et al., 2019	Brasília/DF	DENV	2015 – 2016	475	32	DOADORES DE SANGUE	ELISA IgM	6,7	8
	Brasília/DF	DENV	2015 – 2016	32	20	DOADORES DE SANGUE	ELISA IgG	62,5	8
Tavares-Neto, et al., 2004	Rio Branco/AC	DENV 1	1999	390	1	Maiores de 2 anos	anticorpos inibidores da hemaglutinação (IH)	0,3	5
	Rio Branco/AC	DENV 2	1999	390	16	Maiores de 2 anos	anticorpos inibidores da hemaglutinação (IH)	4,1	5
	Rio Branco/AC	DENV 3	2000	190	6	Maiores de 2 anos após vacinação para febre amarela	anticorpos inibidores da hemaglutinação (IH)	3,2	5
Teixeira, et al., 2012	Salvador/BA	DENV	1998	625	166	Crianças	ELISA	26,6	4
	Salvador/BA	DENV	2000	289	96	Crianças de deram resultado negativo no teste anterior	ELISA	33,2	4

Tabela Suplementar 1 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre dengue.

(conclusão)

Autor/Ano de Publicação	Local	Sorotipo*	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Tumioto, et al., 2014	Rio Grande do Sul	DENV	2007-2013	9779	2046	Pacientes com suspeita clínica de Dengue	QRT-PCR; IgM; NS1	20,9	5
Vasconcelos, et al. 1993	Araguaina/TO	DENV	1993	400	111	Geral	IgM anti-Dengue	27,8	6
Vasconcelos, et al. 1998	Fortaleza/CE	DENV 2 Primária	1994	1341	93	Geral	inibição da hemaglutinação (IH)	6,9	7
	Fortaleza/CE	DENV 2 Secundária	1994	1341	495	Geral	inibição da hemaglutinação (IH)	36,9	7
Vasconcelos, et al., 1999	Paço do Lumiar/MA	DENV	1996	101	56	Geral	inibição da hemaglutinação (IH)	55,4	7
	São José de Ribamar/MA	DENV	1996	100	28	Geral	inibição da hemaglutinação (IH)	28,0	7
	São Luís/MA	DENV	1996	1.016	421	Geral	inibição da hemaglutinação (IH)	41,4	7
Vasconcelos, et al. 2000	Ipupiara/BA	DENV 1	1995	461	55	Geral	inibição da hemaglutinação (IH)	11,9	6
	Prado/BA	DENV 2	1995	228	40	Geral	inibição da hemaglutinação (IH)	17,5	6
Vieira, et al., 2007	Porto Velho/RO	DENV	2001-2003	150	70	Geral	Imunofluorescência	46,7	5
	Porto Velho/RO	DENV 1	2001-2003	70	70	Geral	RT/PCR	100,0	5

*Para aqueles descritos como DENV o artigo não especificou o sorotipo nestes casos.

CI – Casos investigados; CP – Casos positivos; EQ – Escore de Qualidade.

Tabela Suplementar 2 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre zika.

(continua)

Autor/Ano de Publicação	Local	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Borges, et al., 2017	Brasil	2016	954	112	Mulheres que se submeteram a teste para fertilização in vitro	Imunoensaio de cromatografia	11,7	5
Brasil, et al., 2016	Rio de Janeiro/RJ	2015-2016	345	182	Mulheres grávidas que se apresentaram ao ambulatório de doenças febris agudas da Fundação Oswaldo Cruz com uma erupção cutânea nos 5 dias anteriores	RTPCR	52,8	5
	Rio de Janeiro/RJ	2015	262	119	Pacientes com Doença febril aguda atendidos pelo Centro Regional de Pesquisa em Dengue da Fundação Oswaldo Cruz	qRT-PCR	45,4	5
Colombo, et al., 2017	São José do Rio Preto/SP	2016	433	96	Geral	qRT-PCR	22,2	6
De Souza Costa et al., 2019	Mato Grosso	2015 - 2016	453	5	Geral	Protocolos de RT-PCR, sequenciamento nucleotídico e isolamento viral, testes colorimétricos rápidos, como recomendado pelo fabricante para anti-ZIKV IgM e IgG e anti-CHIKV IgM.	1,1	5
Estofolete, et al., 2019	São José do Rio Preto/SP	2016	1254	139	Geral	qRT-PCR	11,1	6

Tabela Suplementar 2 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre zika.

(conclusão)

Autor/Ano de Publicação	Local	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Fuller, et al., 2017	Rio de Janeiro	2015-2016	10459	1717	Geral	O soro e a urina foram testados por qRT-PCR (em tempo real)	16,4	6
Magalhães, et al., 2017	Paullista/PE	2015-2016	263	27	Pacientes com sugestão de doença febril aguda atendidos	qRTPCR	10,3	5
Pessoa, et al 2016	Tuparetama/PE	2015	77	31	Pacientes ambulatório do Hospital Severino de Souto Siqueira	qRT-PCR	40,3	5
Sacchetto, et al., 2018	Brasil	2015-2016	3001	342	Amostras séricas adquiridas passivamente por meio de uma rede de laboratórios privados no Brasil.	RT-qPCR	11,4	6
Slavov, et al., 2017	Ribeirão Preto/SP	2015-2016	1393	37	Doadores de sangue	RNA do ZIKV; reação em cadeia da polimerase em tempo real	2,7	6

CI – Casos investigados; CP – Casos positivos; EQ – Escore de Qualidade.

Tabela Suplementar 3 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre chikungunya.

(continua)

Autor/Ano de Publicação	Local	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
De Souza Costa, et al., 2019	Mato Grosso	2015 - 2016	453	2	GERAL	Protocolos de RT-PCR, sequenciamento nucleotídico e isolamento viral, testes colorimétricos rápidos, como recomendado pelo fabricante para anti-ZIKV IgM e IgG e anti-CHIKV IgM.	0,4	5
Cunha, et al., 2017 ^a	Sergipe	2016	142	107	Pacientes com sintomas de arboviroses	qRT-PCR	75,4	5
	Sergipe	2016	142	91	Pacientes com sintomas de arboviroses	IgM-ELISA	64,1	5
Cunha, et al., 2017 ^b	Riachão do Jacuípe/BA	2016	120	6	Geral	Anticorpos IgM CHIKV	5,0	7
	Riachão do Jacuípe/BA	2016	120	22	Geral	Anticorpos IgG CHIKV	18,3	7
Dias, et al., 2018	Riachão do Jacuípe/BA	2015	446	204	Geral	ELISA		7
	Feira de Santana/BA	2015	385	220	Geral	ELISA	57,1	7

Tabela Suplementar 3 - Identificação dos artigos selecionados para revisão sistemática e metanálise sobre chikungunya.

(conclusão)

Autor/Ano de Publicação	Local	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
Fuller, et al., 2017	Rio de Janeiro	2015-2016	10459	2170	Geral	Soro e a urina foram testados para o CHIKV por qRT-PCR ou por ensaio imunoenzimático (ELISA) para detectar anticorpos IgM	20,7	6
Magalhães, et al., 2017	Paullista/PE	2015-2016	263	132	Pacientes com sugestão de doença febril aguda atendidos pelo	qRTPCR	50,2	5
Pessoa, et al., 2016	Tuparetama/PE	2015	77	1	Pacientes ambulatório do Hospital Severino de Souto Siqueira	CHIKV IgM	1,3	5
Slavov, et al., 2018	Macapá/AP	2015	442	1	Doadores de sangue	ELISA IgG	0,2	5
	Ribeirão Preto/SP	2016	455	0	Doadores de sangue	ELISA IgG	0,0	5
De Souza, et al., 2018	Rio de Janeiro	2014-2015	91	59	Geral	qRT-PCR	64,8	5
	Amapá	2014-2015	208	105	Geral	Anti-CHIKV IgM	50,5	5
Colombo, et al., 2017	São José do Rio Preto/SP	2016	433	1	Geral	qRT-PCR	0,2	6

CI – Casos investigados; CP – Casos positivos; EQ – Escore de Qualidade.

Tabela Suplementar 4 - Identificação dos artigos selecionados para revisão sistemática sobre febre amarela.

Autor/Ano de Publicação	Local	Ano de Referência	CI	CP	População de Referência	Método de Detecção	P%	EQ
De Souza Costa, et al., 2019	Mato Grosso	2015 - 2016	453	3	Geral	Protocolos de RT-PCR, sequenciamento nucleotídico e isolamento viral, testes colorimétricos rápidos, como recomendado pelo fabricante para anti-ZIKV IgM e IgG e anti-CHIKV IgM.	0,7	5
Moreno et al., 2011	São Carlos, Rinção, Ribeirão Preto/SP	2008	577	5	Suspeitas de infecção por febre amarela	MacElisa; RT-PCR; Isolado de células; Isolado em ratos; Histopatologia/Imunohistologia	0,9	5
Tavares-Neto, et al., 2004	Rio Branco/AC	1999	390	106	Maiores de 2 anos	Anticorpos inibidores da hemaglutinação (IH)	27,2	5
	Rio Branco/AC	2000	190	130	Maiores de 2 anos após vacinação para febre amarela	Anticorpos inibidores da hemaglutinação (IH)	68,4	5

CI – Casos investigados; CP – Casos positivos; EQ – Escore de Qualidade.

5.2.ARTIGO ORIGINAL – FEBRE AMARELA NO BRASIL: ESTUDO DE PROGNÓSTICO UTILIZANDO TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL

RESUMO

Introdução: A febre amarela é uma doença infecciosa hemorrágica febril aguda prevenível por meio da vacinação. A presença do vírus, o vetor em grande quantidade e o contato com o hospedeiro fazem da febre amarela na atualidade um problema de Saúde Pública com importância no Brasil. Entre 2016/2017 e 2018/2019 a taxa de letalidade da doença variou de 33,63% à 41,66%, valores altos para uma doença imunizável. Logo, identificar métodos que possam auxiliar seu controle são de extrema importância. Assim, a utilização de recursos computacionais na forma de modelos preditivos podem aprimorar a pesquisa em saúde, avaliação da qualidade da atenção e prevenção e controle de doenças, bem como prognóstico e manejo clínico da população. **Objetivo:** Aplicar técnicas de aprendizagem de máquina supervisionado para construção de um modelo preditivo de mortalidade de febre amarela utilizando dados relativos a casos confirmados no Sistema de Informação de Agravos de Notificação do Brasil (SINAN). **Metodologia:** Foram utilizados dados secundários de febre amarela do período 2000 a 2018 do SINAN disponibilizados pelo Ministério da Saúde e solicitados por meio do Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC). Esses dados passaram por uma etapa de padronização e organização. Para realização do modelo foram incluídos apenas as pessoas com Classificação final confirmada para febre amarela, perfazendo 2405 casos. Para realização das análises foi utilizado o *Google Colaboratory*. Foram testados 8 algoritmos: *Naive Bayes*, *Decision Tree*, *k-Nearest Neighbour*, *Logistic Regression*, *Multilayer Perceptron*, *Support Vector Machine*, *Random Forest* e *CatBoost*, para a solução da problemática do artigo. No treinamento do conjunto de dados foi utilizada o percentual de 75% da base para treinamento e 25% para teste. Após a calibração dos hiperparâmetros os parâmetros tiveram suas configurações modificadas. Para a validação do modelo utilizamos a Validação Cruzada 10 partes. E para a visualização do impacto dos atributos foi utilizada a biblioteca SHAP (*SHapley Additive exPlanations*). **Resultados:** Foi observado um maior número de casos nos anos de 2017 e 2018. A maioria dos casos confirmados ocorreu na Região Sudeste, no período sazonal da doença, entre homens e adultos. O algoritmo selecionado para o modelo foi o *CatBoost* por apresentar métricas de desempenho melhores, por reduzir a necessidade de ajustes de hiperparâmetros, além de ser melhor com bancos categóricos e não precisar ter valores faltantes imputados. A acurácia do modelo foi de 81,1% e a precisão de 79,8%. As

variáveis que mais influenciaram no modelo foram: resultado do exame imunológico de IgM, presença de distúrbio renal, presença de dor abdominal, critério de confirmação utilizado e presença de sinais hemorrágicos. O resultado de IgM reagente associou-se com cura, o resultado positivo de PCR se relacionou com o óbito. A presença de distúrbio renal e sinais hemorrágicos foram associados com o óbito, o inverso ocorreu com o sinal de Faget. As variáveis dor abdominal e resultado de isolamento viral, quando respondidos associaram-se com a cura e os valores ausentes com o óbito, o inverso ocorreu com o critério de confirmação. Regiões Sul e Centro-Oeste se associaram ao óbito. Enquanto idade inferiores a 40 anos estão relacionadas com a cura do paciente. A ocorrência da hospitalização associou-se com o óbito. O período sazonal da doença (jan.-abr.) foi associado a cura. A vacinação quando realizada se associou com a cura e o sexo feminino também. **Considerações finais:** O modelo proposto para previsão de óbito em pacientes com febre amarela teve um bom resultado com 81,1% de acurácia e 79,8% de precisão. O uso dessa ferramenta pode auxiliar os profissionais de saúde na prática clínica. A análise do comportamento do modelo permitiu aumentar o conhecimento sobre o grau de impacto de importantes variáveis sobre a doença, reforçando pontos de atenção que podem contribuir para a construção do estadiamento da doença e consequentemente proporcionando um melhor e manejo clínico da febre amarela.

Palavras-chave: Aprendizagem de máquina. Mineração de dados. Modelo Preditivo. Febre amarela. Arbovirose.

INTRODUÇÃO

A febre amarela é uma doença infecciosa hemorrágica febril aguda prevenível por meio da vacinação. É causada pelo vírus da febre amarela (YFV) um arbovírus protótipo do gênero *Flavivirus*, da família *Flaviviridae*. Apresenta importância epidemiológica e clínica por ser uma doença transmitida por vetores existentes em nosso meio e pela possível reintrodução na sua forma urbana. É mais comum que a doença se apresente de forma leve ou moderada, quando a febre amarela evolui para a forma grave a letalidade é alta, chegando a 50% dos casos (BRASIL, 2019a; DO CARMO CUPERTINO, et al. 2019).

A febre amarela possui dois ciclos de transmissão um silvestre e outro urbano. O ciclo silvestre envolve a transmissão entre macacos e mosquitos selvagens. Já no ciclo urbano o mosquito envolvido é o *Aedes aegypti*, sendo ele o responsável pela transmissão da doença, e o ser humano o único hospedeiro (BRASIL, 2019a; DO CARMO CUPERTINO, et al. 2019). Existem 47 países no mundo onde a enfermidade é endêmica ou que possuem áreas endêmicas. Desses, 34 são na África e 13 na América Central ou do Sul, onde afeta mais de 200.000 pessoas anualmente (OPAS BRASIL, 2019).

No Brasil, a forma de transmissão é silvestre, tendo o último caso de febre amarela urbana sido descrito em 1942. Os casos registrados desde então ocorriam sobretudo na Região da Amazônia Legal, mas não eram casos muito abundantes (BRASIL, 2004). Os registros de casos humanos e epizootias em primatas não humanos começaram a se espalhar pelo território brasileiro em 2014/2015, iniciando-se na Região Norte, expandindo-se no sentido leste e sul do país, afetando em 2015/2016 principalmente a Região Centro-Oeste. Os casos entre 2016/2017 ocorreram principalmente a Região Sudeste, registrando-se 779 casos humanos e 262 óbitos. Entre 2017/2018, foi registrado o surto mais expressivo no Brasil, que afetou principalmente a Região Sudeste, sendo registrados 1.376 casos humanos e 483 óbitos. No período epidemiológico 2018/2019 foram confirmados 12 casos humanos e 5 óbitos (BRASIL, 2018; BRASIL, 2019b; BRASIL, 2019c). A taxa de letalidade de febre amarela foi superior a 30% os períodos de 2016/2017 (33,63%) e 2017/2018 (35,10%). Em 2018/2019, a taxa de letalidade foi 41,66%.

Os principais sintomas e sinais da doença são: febre alta, dor de cabeça forte e duradoura, inapetência, náuseas e mialgia, geralmente com icterícia. Contudo 40% a 60% dos casos são oligossintomáticos ou assintomáticos. Os casos mais graves podem aparecer em 20% a 40%: desses até 50% podem evoluir para óbito. Os sinais clínicos mais comuns na forma mais grave são a insuficiência renal e hepática. É comum nesta forma, após um período prodromico,

uma aparente remissão, que é seguida de um o agravamento da icterícia, insuficiência renal e fenômenos hemorrágicos mais intensos, que geralmente leva o paciente a óbito em uma semana (BRASIL, 2019a; DO CARMO CUPERTINO, et al. 2019;). O diagnóstico precoce pode resultar em um melhor manejo da doença, visto que as manifestações clínicas podem ser inespecíficas e evoluir em pouco tempo para um curso fulminante (KALLAS e WILDER-SMITH, 2019).

É imprescindível identificar fatores que se associem com uma maior mortalidade. A idade avançada, sexo masculino, maior contagem de leucócitos e neutrófilos, maior alanina aminotransferase (ALT), aspartato aminotransferase (AST), bilirrubina e creatinina, tempo de protrombina prolongado, e maior carga viral plasmática de RNA do vírus da febre amarela são fatores descritos por Kallas, et al., 2019 como associados a maior mortalidade de febre amarela. Nesse sentido Ho, et. al. 2019 trazem que, pacientes com diabetes mellitus apresentam maior taxa de letalidade e que as principais causas de morte nos pacientes com febre amarela grave são sangramento gastrointestinal grave, estado epiléptico, acidose metabólica grave, pancreatite necro-hemorrágica e falência múltipla de órgãos.

Como a febre amarela é na atualidade um problema de Saúde Pública com importância em nosso país, identificar métodos que possam auxiliar seu controle é de extrema importância. Logo chega-se aos modelos de previsão, que podem ser utilizados para aprimorar a pesquisa em saúde e avaliar a qualidade da atenção e agir na prevenção e controle de doenças. Para isso têm sido utilizados recursos computacionais com a intenção de solucionar problemas complexos por meio da utilização de algoritmos de aprendizagem de máquina. Com essa ação, novas informações podem ser geradas e conseqüentemente utilizadas pelos profissionais de saúde no atendimento, gerenciamento e diagnóstico (TAYLOR, et al., 2016; LUO, et al. 2017; MANSOOR, et. al., 2017; WANG, et. al. 2018; LUND, et al., 2019; PIROS, et al., 2019).

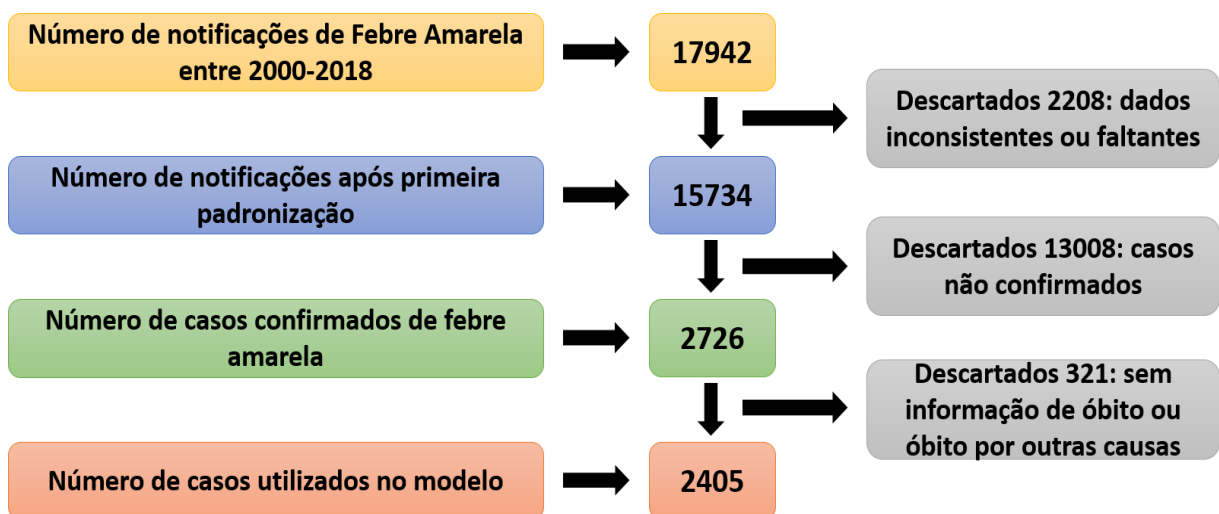
Assim o objetivo desse trabalho é aplicar técnicas de aprendizagem de máquina supervisionado para construção de um modelo preditivo de mortalidade de febre amarela utilizando dados relativos a casos confirmados no Sistema de Informação de Agravos de Notificação do Brasil.

MÉTODOS

Banco de dados

Foram utilizados dados secundários do SINAN disponibilizados pelo Ministério da Saúde e solicitados por meio do Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC). Os dados solicitados são referentes às notificações de febre amarela do ano 2000 até o ano de 2018. O banco original possuía 17942 casos de notificações de febre amarela. Contudo após uma primeira organização e padronização dos dados foram excluídos 2208 casos sendo esses referentes a dados inconsistentes, faltantes, ou que não eram de interesse para essa pesquisa, como casos de doença em estrangeiros. Do total de 15734 notificações de febre amarela, 13008 (82,67%) foram descartados, enquanto 2726 (17,37%) foram casos confirmados para febre amarela, baseados na variável CLAS_FINAL (Classificação final) que dividia o banco em casos confirmados e descartados. Dos 2726 casos confirmados trezentos e dezenove não apresentavam informação de cura ou óbito e duas mortes foram determinadas por outras causas. Assim, considerando que o objetivo do trabalho é baseado no óbito ou cura para febre amarela, as pessoas sem informação e as mortes por outras causas foram excluídas do modelo final dessa forma perfazendo um total de 2405 indivíduos confirmados com febre amarela. Essas informações estão apresentadas de forma resumida na Figura 1.

Figura 1 - Resumo do processo de preparação do banco de dados



Fonte: Dados da pesquisa

Atributos

A seleção de atributos é uma forma de reduzir a complexidade do conjunto de dados e consequentemente tornar a construção do modelo mais rápida, podendo produzir melhores resultados (CERQUEIRA, et al. 2014). Tal procedimento pode ser realizado de três formas: por meio de métodos computacionais; por ação humana; e em uma combinação dos dois métodos anteriores. Nesse trabalho a seleção de atributos foi feita apenas por meio da análise humana das variáveis.

Inicialmente o banco de dados possuía 32 variáveis; foram removidas as variáveis DATA_NOTIF (data da notificação), DATA_SINTOMA (data do início dos sintomas), DATA_INTERNACAO (data da internação do paciente), DATA_OBITO (data do óbito do paciente), DATA_ENCERRAMENTO (data do registro de encerramento do caso), BT (bilirrubina total), BD (bilirrubina direta), AST_TGO (aspartato aminotransferase), ALT_TGP (alanina aminotransferase) devido a dados faltantes e informações equivocadas. Assim apenas 21 variáveis foram para a análise exploratória. Com base no critério de redundância ou por ter outra variável que representasse melhor os dados foram excluídas as variáveis: ANO_NOT (ano de notificação), REG_NOTIF (região de notificação), AUTOCTONE (se o caso era ou não autóctone do município), MES_SINTOMA (mês de apresentação de sintomas), IDADE_CAT (categorização de idade), HISTOPATOLOGIA (resultado da histopatologia), IMUNOHISTOQUIMICA (resultado da imunohistoquímica), perfazendo um total de 14 variáveis consideradas no modelo final.

Ferramenta de aprendizagem de máquina

Utilizou-se o *Google Colaboratory* ou *Google Colab* que é um projeto de pesquisa para prototipar modelos de aprendizagem de máquina em opções de hardware poderosas, como *Graphics Processing Unit* (GPU) e *Tensor Processing Unit* (TPU). É um ambiente de *Jupyter Notebook* que não requer configuração e é executado na nuvem, além de ser gratuito (BISONG, 2019). O *Google Colab* é um ambiente de desenvolvimento em linguagem de programação Python. Nos scripts do Python foram aplicadas como principais bibliotecas: numpy e pandas para manipulação de dados, matplotlib e shap para visualização de dados, catboost e sklearn para os algoritmos de aprendizagem de máquina.

Algoritmos de aprendizagem

Diversos algoritmos podem ser utilizados para a tarefa de classificação com o objetivo de prever em qual classe um determinado registro pertence (PINHEIRO, 2008). Após pesquisa

e análise em artigos publicados e na literatura cinzenta foram selecionados, para serem testados, os algoritmos mais utilizados para o problema de classificação dentro do campo da biomedicina: *Naive Bayes* (NB), *Decision Tree* (DT), *k-Nearest Neighbour* (kNN), *Logistic Regression* (LR), *Multilayer Perceptron* (MLP), *Support Vector Machine* (SVM), *Random Forest* (RF) e *CatBoost* (CB) (OZER, 2008; NANCY e RAMANI, 2011; MAHINDRAKAR e HANUMANTHAPPA, 2013; ALONSO, et. al. 2018; SIQUEIRA-BATISTA e DA SILVA, 2020).

Treinamento e validação do modelo

Para treinar qualquer tipo de modelo de aprendizagem de máquina é preciso que haja dados de treinamento e dados de teste. Quando não há um conjunto específico para teste é preciso dividir o conjunto de dados em dados de treinamento e dados de teste. Os algoritmos utilizam os dados de treinamento para encontrar um modelo preditivo, que em linhas gerais é o estabelecimento de uma relação do vetor de atributos com a classe, de modo que futuros registros (vetor de atributos sem classe associada) possam ter sua classe predita com alta acuraria. Para testar os modelos construídos, aplicam-se os mesmos dados de testes para avaliar o percentual de acerto, dentro outras estatísticas importantes. Todo esse processo é conhecido como *Hold-Out*. Para nosso conjunto de dados foi utilizado o percentual de 75% da base para treinamento e 25% para teste (WITTEN, FRANK e HALL, 2011; TAN, STEINBACH e KUMAR, 2009). Dois problemas devem ser evitados na construção do modelo: o *Underfitting* – quando o algoritmo não consegue aprender o padrão, que é muito comum de ocorrer quando se tem dados insuficientes, o que impossibilita a continuação do trabalho e o *Overfitting* – quando o modelo apresenta uma sobre adequação ao conjunto de treinamento, não conseguindo generalizar para novos registros (TAN, STEINBACH e KUMAR, 2009).

Os valores dos hiperparâmetros, que são as variáveis de configuração do algoritmo, foram definidos a partir da verificação na literatura. Após algumas análises, verificou-se que a otimização dos valores poderia trazer um aumento de desempenho no modelo. Como foram poucos os hiperparâmetros testados, utilizamos o método *GridSearchCV* com validação cruzada 5 partes para encontrar os valores otimizados para um conjunto de importantes hiperparâmetros para cada um dos algoritmos testados. A Tabela 1 mostra os valores dos hiperparâmetros após o processo de calibração.

Tabela 1 - Hiperparâmetros considerados na otimização e utilizados no modelo, segundo algoritmo.

Algoritmo	Hiperparâmetros*	Considerados na otimização	Utilizados no modelo
<i>Naives Bayes</i>	Padrão	Padrão	Padrão
<i>Decision Tree</i>	Criterion max_features min_samples_split min_samples_leaf class_weight	gini, entropy auto, sqrt, log2 2,3,4,5,6,7,8,9,10,11,12,13,14,15 1,2,3,4,5,6,7,8,9,10,11 balanced, None	gini Auto 12 5 None
<i>Random Forest</i>	Criterion n_estimators min_samples_leaf min_samples_split max_features	gini, entropy 10,15,20,25,30,50,75,100,150,20, 500,1000 1,2,3,4,5 1,2,3,4,5,6,7 auto, sqrt, log2	Gini 150 5 2 Auto
<i>kNN</i>	n_neighbors leaf_size Weights Algorithm	1,2,3,4,5,6,7,8,9,10,11 1,2,3,4,5 uniform, distance auto, ball_tree, kd_tree, brute	10 2 Uniform ball_tree
<i>Logistic Regression</i>	Penalty C Solver max_iter	l1, l2 0.001, 0.01, 0.1, 1, 10, 100, 1000 liblinear, lbfgs 100, 200, 300	l2 10 Liblinear 100
<i>MLP</i>	hidden_layer_sizes Activation Solver learning_rate_init learning_rate	(10,30,10), (20,), (100), (64,64), (50,50,50), (50,100,50) tanh, relu, logistic sgd, adam, lbfgs 0.1, 0.01, 0.001 constant, adaptive	(64,64) Tanh Sgd 0.01 Constant
<i>SVM</i>	C Kernel class_weight	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 20 linear, rbf, poly, sigmoid balanced', None	9 Rbf None
<i>Catboost</i>	Iterations learning_rate Depth l2 leaf reg	10, 25, 50, 100, 111, 200, 300, 500, 1000 0.001, 0.01, 0.02, 0.03, 0.05, 0.1, 0.11, 0.15, 0.2, 0.3, 0.5 2, 4, 6, 8, 10 1, 3, 5, 7, 9	200 0.05 6 5

*Os hiperparâmetros não mencionados foram mantidos segundo configuração padrão/default das bibliotecas do Sklearn e do Catboost

Fonte: Dados da pesquisa

Para a validação do modelo utilizamos a Validação Cruzada 10 partes, que diferente do *Hold-Out*, divide o conjunto de dados em 10, usa 9 para treinar e um para testar, esse processo

é repetido com cada uma das partes como teste e o resultado do desempenho é a média. Esse processo verifica a capacidade de generalização do modelo pois expõe o modelo a uma variância maior dos dados (TAN, STEINBACH e KUMAR, 2009).

Para selecionar o modelo preditivo final, foi utilizada a validação cruzada 10 partes em todos os algoritmos e foram executados 30 vezes com 30 sementes geradas aleatoriamente (valores inteiros aleatórios entre 0 e $2^{32}-1$) para obter o desempenho médio dos classificadores. Em seguida, as comparações estatísticas dos modelos foram realizadas por meio de testes não paramétricos. Utilizou-se o teste de Friedman, seguido do teste Post-HOC de Nemenyi para identificar diferenças entre os modelos de acordo com a média da Curva ROC com nível de confiança de 95%.

O gráfico de área sob a curva ROC permite a visualização e avaliação de modelos de classificação e a análise independe do limiar de classificação. A análise ROC possibilita uma avaliação mais rica se comparado com a avaliação de um modelo de classificação com apenas uma medida (PRATI, BATISTA e MONARD, 2008).

Ferramenta de Visualização

SHapely Additive exPlanations (SHAP)

SHAP é uma biblioteca que provê uma ferramenta de visualização para interpretação da saída de modelos de caixa-preta, nos quais a complexidade de análise é alta, e a estrutura interna de processamento é desconhecida ou não é considerada para a análise dos resultados. Este é o caso do *CatBoost*, que por ser composto de diversas árvores de decisão tem o entendimento de análise dificultado. SHAP é baseada na *Coalitional Game Theory*, que proporciona por meio de um conjunto de ferramentas matemáticas a análise de interações complexas entre jogadores racionais, com objetivo de elevar o ganho dos participantes considerando seus interesses pessoais (SAAD, 2009; VAKILINIA e SENGUPTA, 2017; MOLNAR, 2019). Permite uma visualização em nível global, lista os atributos em ordem decrescente de importância e local como cada atributo impacta positivamente/negativamente a saída do modelo para cada classe (PRAKHAR, 2020).

O uso do SHAP possibilitou, neste trabalho, a visualização global dos atributos, gerando gráficos que listam os atributos em ordem decrescente de contribuição para o modelo e o relacionamento de cada atributo com a variável destino.

Aspectos Éticos

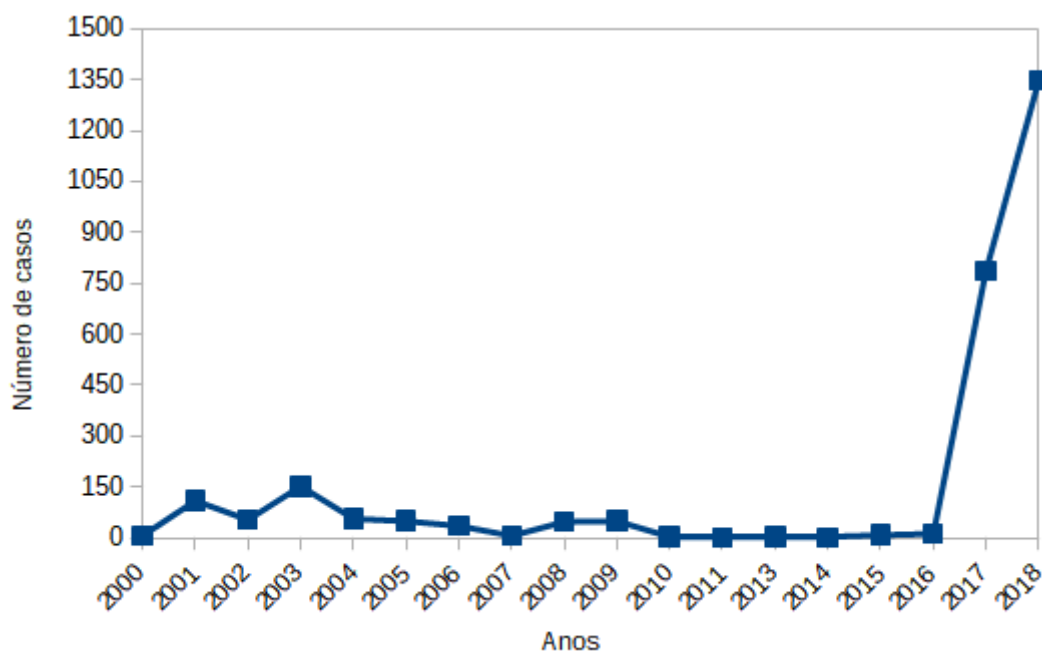
Esse trabalho dispensa a apresentação e submissão ao Comitê de Ética segundo o artigo 1º, parágrafo único da Resolução nº 510 de 07 de Abril de 2016 do Conselho Nacional de Saúde, que informa que não deverão ser registradas nem avaliadas pelo sistema CEP (Comitê de Ética e Pesquisa)/CONEP (Comissão Nacional de Ética em Pesquisa) as pesquisas que utilizam informações de acesso público, nos termos da Lei nº 12.527, de 18 de novembro de 2011.

RESULTADOS

Características do banco de dados de febre amarela

É possível observar que o número de casos confirmados desde o ano de 2000 sofria poucas alterações, com concentração principal nas áreas endêmicas. Todavia, a partir do ano de 2017 a doença passou a ocorrer por uma grande parte do país aumentando muito o número de casos confirmados, tornando-se uma das grandes preocupações de saúde dos anos de 2017 e 2018 (Figura 2).

Figura 2 - Número de casos confirmados de febre amarela no Brasil entre 2000-2018



Fonte: Dados da pesquisa

Na Tabela 2 podemos visualizar algumas das características das pessoas que tiveram a classificação final na ficha de notificação como caso confirmado para febre amarela. É possível perceber que algumas características apresentam dados ausentes e/ou ignorados.

Tabela 2 - Caracterização das pessoas confirmados com febre amarela – Cura ou óbito por febre amarela – Brasil - 2000-2018.

(continua)

	Total	Óbito por febre amarela	Cura
		<i>n₁</i> (%)	<i>n₂</i> (%)
Região			
Norte	99	83 (8,7)	16 (1,1)
Nordeste	8	3 (0,3)	5 (0,4)
Sudeste	2170	771 (80,8)	1399 (96,4)
Sul	41	22 (2,3)	19 (1,3)
Centro-Oeste	87	75 (7,9)	12 (0,8)
Período Sazonal			
Período sazonal (janeiro - abril)	2040	659 (69,1)	1381 (95,2)
Período de baixa ocorrência (maio - agosto)	122	104 (10,9)	18 (1,2)
Período pré-sazonal (setembro - dezembro)	243	191 (20,0)	52 (3,6)
Sexo			
Feminino	479	183 (19,2)	296 (20,4)
Masculino	1926	771 (80,8)	1155 (79,6)
Faixa etária			
0-12 anos incompletos	57	28 (2,9)	29 (2,0)
12-18 anos incompletos	126	49 (5,1)	77 (5,3)
18-60 anos incompletos	1797	676 (70,9)	1121 (77,3)
Acima de 60 anos	425	201 (21,1)	224 (15,4)
Situação Vacinal			
Não vacinado	1654	545 (57,1)	1109 (76,4)
Vacinado	393	148 (15,5)	245 (16,9)
Dados ausentes/ignorados	358	261 (27,4)	97 (6,7)

Tabela 2 - Caracterização das pessoas confirmados com febre amarela – Cura ou óbito por febre amarela – Brasil - 2000-2018.

(continuação)

	Total	Óbito por febre amarela	Cura
		<i>n₁</i> (%)	<i>n₂</i> (%)
Sinais Hemorrágicos			
Ausente	1156	316 (33,1)	840 (57,9)
Presente	551	364 (38,2)	187 (12,9)
Dados ausentes/ignorados	698	274 (28,7)	424 (29,2)
Distúrbio Renal			
Ausente	1137	306 (32,1)	831 (57,3)
Presente	508	351 (36,8)	157 (10,8)
Dados ausentes/ignorados	760	297 (31,1)	463 (31,9)
Sinal de Faget			
Ausente	1118	435 (45,6)	683 (47,1)
Presente	465	163 (17,1)	302 (20,8)
Dados ausentes/ignorados	822	356 (37,3)	466 (32,1)
Dor Abdominal			
Ausente	454	86 (9,0)	368 (25,4)
Presente	919	276 (28,9)	643 (44,3)
Dados ausentes/ignorados	1032	592 (62,1)	440 (30,3)
Hospitalização			
Não	158	17 (1,8)	141 (9,7)
Sim	1326	639 (67,0)	687 (47,3)
Dados ausentes/ignorados	921	298 (31,2)	623 (43,0)
Resultado IgM			
Não reagente	403	272 (28,5)	131 (9,0)
Reagente	1090	287 (30,1)	803 (55,3)
Inconclusivo	15	11 (1,1)	4 (0,3)
Dados ausentes/ignorados	897	384 (40,3)	513 (35,3)
Resultado PCR			
Negativo	139	15 (1,6)	124 (8,5)
Positivo	1331	459 (48,1)	872 (60,1)
Inconclusivo	2	1 (0,1)	1 (0,1)
Dados ausentes/ignorados	933	479 (52,2)	454 (31,3)

Tabela 2 - Caracterização das pessoas confirmados com febre amarela – Cura ou óbito por febre amarela – Brasil - 2000-2018.

(conclusão)

Histopatologia			
Negativo	13	12 (1,3)	1 (0,1)
Positivo	24	24 (2,5)	0 (0,0)
Inconclusivo	180	166 (17,4)	14 (0,9)
Dados ausentes/ignorados	2188	752 (78,8)	1436 (99,0)
Imuno-histoquímica			
Negativo	15	14 (1,5)	1 (0,1)
Positivo	57	56 (5,9)	1 (0,1)
Inconclusivo	161	146 (15,3)	15 (1,0)
Dados ausentes/ignorados	2172	738 (77,3)	1434 (98,8)
Resultado Isolamento viral			
Não reagente	114	31 (3,2)	83 (5,7)
Reagente	69	11 (1,2)	58 (4,0)
Inconclusivo	5	4 (0,4)	1 (0,1)
Dados ausentes/ignorados	2217	908 (95,2)	1309 (90,2)
Critério de Confirmação			
Laboratorial	513	392 (41,1)	121 (8,3)
Clínico-Epidemiológico	42	32 (3,3)	10 (0,7)
Laboratorial e Clínico-Epidemiológico	16	12 (1,3)	4 (0,3)
Dados ausentes/ignorados	1834	518 (54,3)	1316 (90,7)
Total	2405	954	1451

Fonte: Dados da pesquisa

O início dos sintomas é mais comum entre janeiro a abril, sendo o ápice em janeiro. O maior número de mortes ocorre no mesmo período, entre janeiro e abril. A doença é mais comum nos homens e adultos, seguido por idosos para ambos os sexos. O número de crianças curadas e que morrem pela doença é próximo. Para os adolescentes e adultos aproximadamente 60% das pessoas são curadas. Os valores de cura e óbito para idosos são próximos (Tabela 2).

A região sudeste tem o maior número de casos sendo que o estado de Minas gerais lidera com o maior número de casos confirmados, seguidos pelo estado de SP. A grande maioria dos casos é autóctone do município de residência. Na região Norte, 84% dos infectados por febre amarela foram a óbito. No Nordeste 38% foram a óbito. No Sudeste, região que teve aproximadamente 90% do total de confirmados na base de dados utilizada, 64% das pessoas

infectadas foram curadas e 36% foram a óbito. No Sul, 46% foram curados e 54% foram a óbito. No Centro-Oeste 86% das pessoas confirmadas foram a óbito e 14% foram curadas. As regiões Norte e Centro-Oeste se assemelham por terem uma alta porcentagem de mortes e baixa de cura. A região Sudeste mesmo um alto percentual de morte, foi a região com maior percentual de cura (Tabela 2).

Dos casos confirmados para febre amarela, 80% das pessoas não foram vacinadas. Entre as pessoas vacinadas, 60% é curada. Em todas as categorias de idade, a quantidade de infectados que não tomaram a vacina é superior aos que tomaram (principalmente em adultos e idosos). A maioria dos sintomas está ausente nos pacientes com exceção da dor abdominal, contudo dentre as pessoas que foram a óbito, mais da metade apresentaram sinais hemorrágicos, distúrbios renais e dor abdominal. A grande maioria das pessoas confirmadas com febre amarela foi hospitalizada, dentre os que faleceram, 97% foi hospitalizada (Tabela 2).

Os resultados reagentes ou positivos foram comuns entre os três testes listados, PCR, IgM e isolamento viral. Dentre os que morreram com febre amarela, a maioria tinha resultado positivo ou reagente para os testes. Com exceção da região Nordeste, o exame laboratorial é o método mais utilizado para confirmação da doença (Tabela 2). A quantidade de ignorados referente ao critério de confirmação da doença cresceu substancialmente a partir de 2017.

De forma geral, a maioria das pessoas é curada da doença com letalidade de 35%. O ano de 2018 é o com maior número de casos e de óbitos.

Algoritmos testados

Conforme os valores nas diversas métricas de desempenho do teste de validação cruzada 10-partes, o *CastBoost* obteve o melhor resultado com exceção da precisão (Tabela 3).

Tabela 3 - Métricas de desempenho do Teste de Validação de 10-partes segundo classificador
(continua)

Classificador	Acc	Pr	Rc	F1	AUC
NB	76,7	78,7	56,4	65,5	0,816
DT	77,0	75,5	62,8	68,3	0,850
RF	80,9	81,3*	67,4	73,6	0,888
kNN	79,3	80,0	63,7	70,8	0,854
LR	78,2	76,9	64,4	70,0	0,853

Tabela 3 - Métricas de desempenho do Teste de Validação de 10-partes segundo classificador
(conclusão)

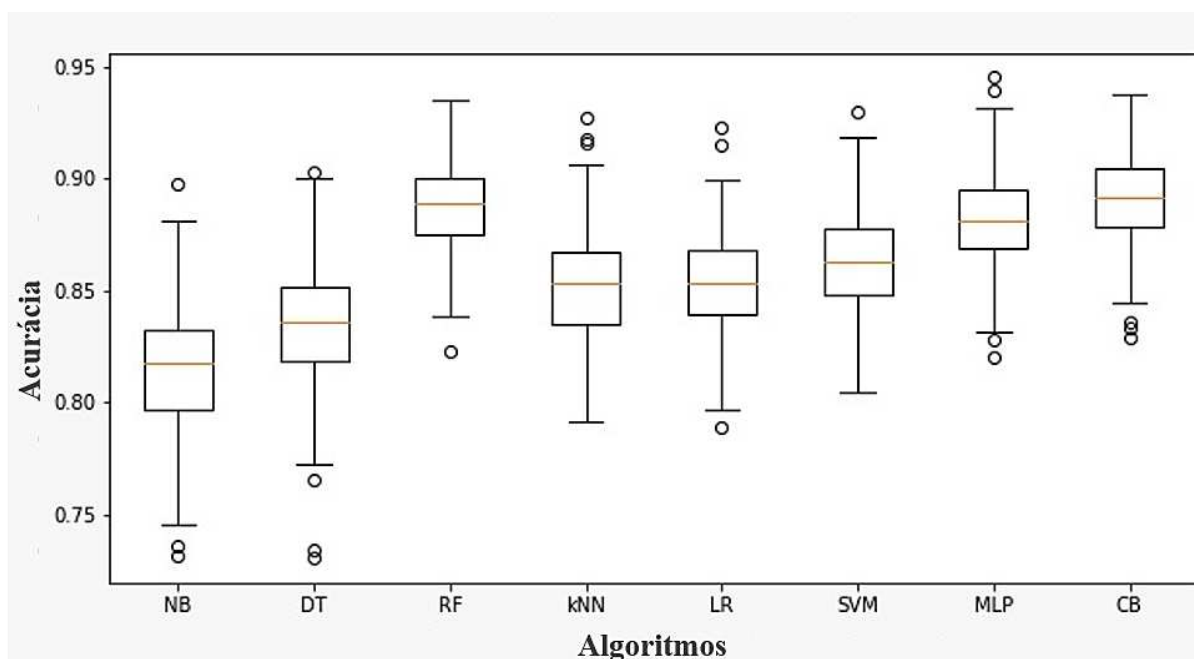
Classificador	Acc	Pr	Rc	F1	AUC
MLP	80,8	77,5	70,2*	73,5	0,883
SVM	80,2	80,0	66,9	72,7	0,863
CB	81,1*	79,8	70,2*	74,6*	0,892*

*Valores em negritos são os maiores valores de cada métrica. Acc = Acurácia; Pr = Precisão; Rc = Sensibilidade; F1 = Medida F; AUC = Área sob a curva Roc

Fonte: Dados da pesquisa

O teste de Friedman indicou que houve uma diferença estatística entre os algoritmos ($p < 0,05$). Considerando a classificação apresentada na Curva ROC, o *CastBoost* e o *Random Forest* têm as melhores performances (Figura 3). E de acordo com o Nemenyi, o o *CastBoost* e o *Random Forest*, não diferem estatisticamente entre si (Figura 4).

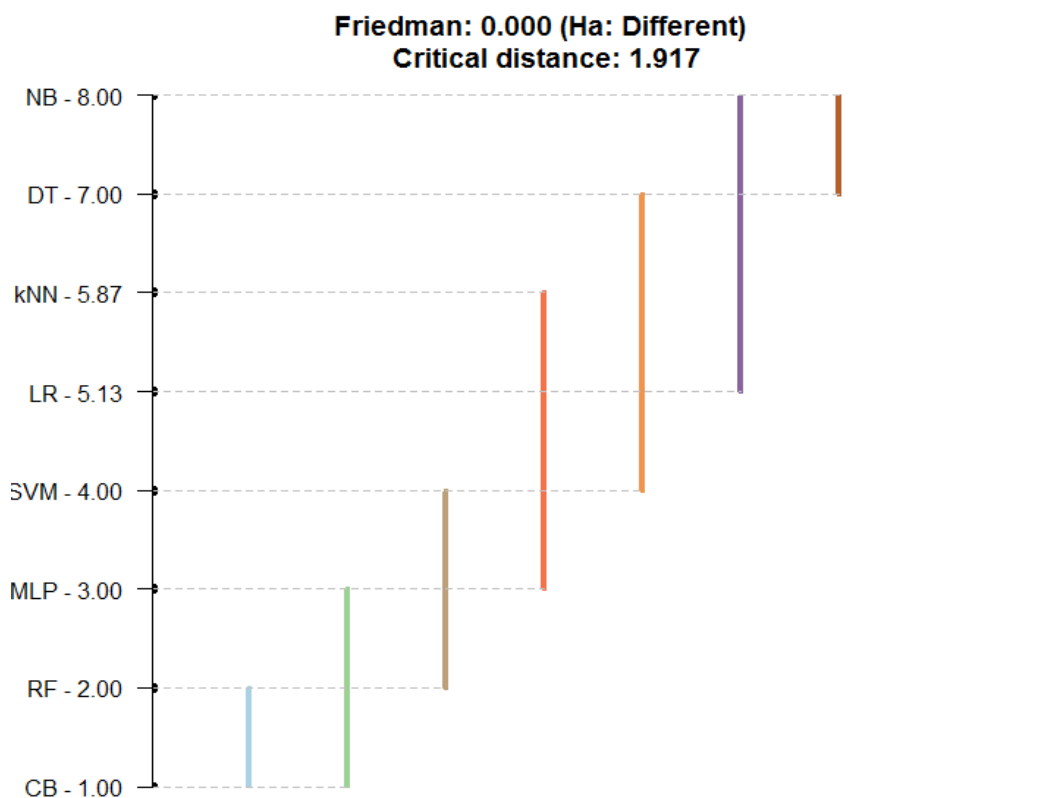
Figura 3 - Comparação dos modelos de aprendizagem de máquina



NB, Naive Bayes; DT, Decision Tree; kNN, k-Nearest Neighbour; LR, Logistic Regression; MLP, Multilayer Perceptron; SVM, Support Vector Machine; RF, Random Forest; CB, CatBoost.

Fonte: Dados da pesquisa

Figura 4 - Comparação dos modelos de aprendizagem de máquina pelo Teste de Nemenyi



NB, Naive Bayes; DT, Decision Tree; kNN, k-Nearest Neighbour; LR, Logistic Regression; MLP, Multilayer Perceptron; SVM, Support Vector Machine; RF, Random Forest; CB, CatBoost. No Teste de Nemenyi distância crítica (CD-Critical Distance) foi definida como 1.917, o que indica que as distâncias entre os algoritmos devem ser maiores do que este valor para caracterizar como diferença significativa.

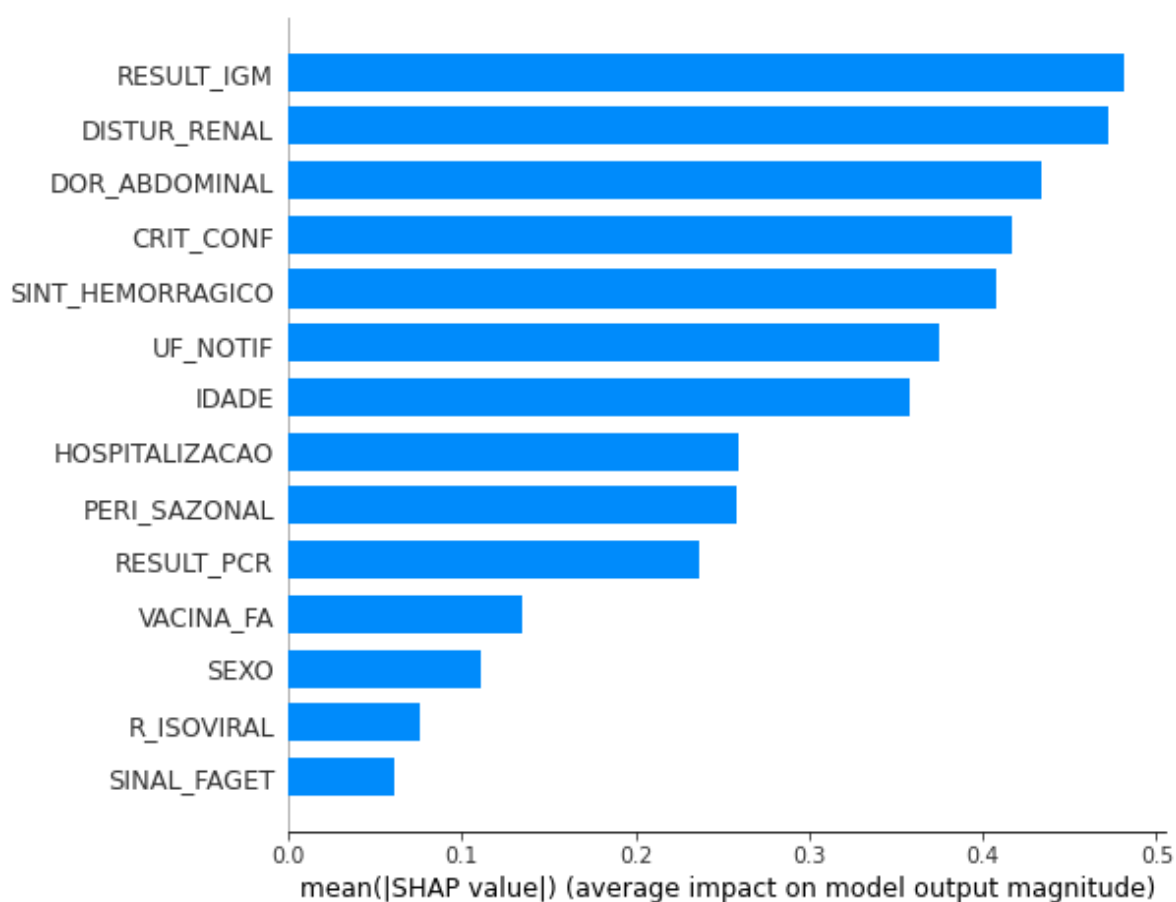
Fonte: Dados da pesquisa

Considerando que o *CatBoost* é um algoritmo recente desenvolvido em 2017 por pesquisadores e engenheiros da *Yandex*. Foi criado inicialmente para competições, mas já é usado para pesquisa, sistemas de recomendação, assistentes pessoais, carros autônomos, previsão do tempo, entre outras áreas. É adequado para base de dados categóricas, pois pode ser usado sem necessidade de converter previamente categorias em números. É um algoritmo robusto, o que reduz a necessidade de ajustes de hiperparâmetros, tem o código aberto, que pode ser usado em linha de comando em uma interface de programação de aplicações para *Python* ou R (PROKHORENKOVA, et al. 2018; BANERJEE, 2020; CATBOOST, 2020), optou-se por atotar o *CatBoost* como algoritmo para a criação do modelo final.

Variáveis preditivas de letalidade por febre amarela

As variáveis que mais influenciaram no modelo foram: resultado do exame imunológico de IgM, presença de distúrbio renal, presença de dor abdominal, critério de confirmação utilizado e presença de sinais hemorrágicos (Figura 5).

Figura 5 - Gráfico de importância das variáveis do modelo.



RESULT_IGM, resultado do teste IgM; DISTUR_RENAL, presença de distúrbio renal; DOR_ABDOMINAL, presença de dor abdominal; CRIT_CONF, critério de confirmação da doença; SINT_HEMORRAGICO, presença de sinais hemorrágicos; UF_NOTIF, estado de notificação da doença; IDADE, idade da pessoa afetada; HOSPITALIZACAO, ocorrência da hospitalização; PERI_SAZONAL, período sazonal da febre amarela; RESULT_PCR, resultado do teste PCR; VACINA_FA, situação vacinal do paciente; SEXO, sexo do paciente; R_ISO VIRAL, resultado do isolamento viral; SINAL_FAGET, presença de sinal de Faget. Um gráfico de importância de variável lista as variáveis mais significativas em ordem decrescente. As variáveis superiores contribuem mais para o modelo do que as inferiores e, portanto, têm alto poder preditivo.

Fonte: Dados da pesquisa

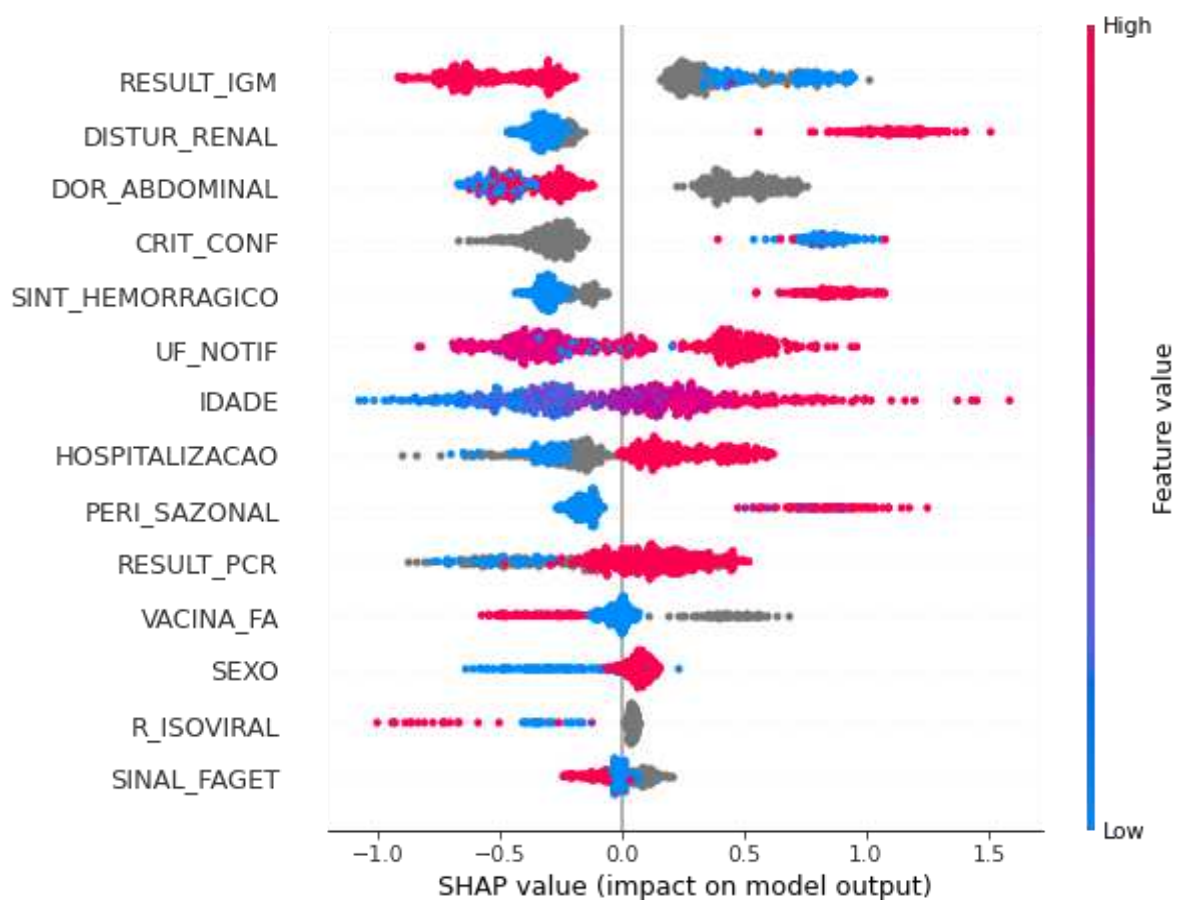
Na Tabela 4 as variáveis foram compiladas em quatro dimensões: características sociodemográficas, sintomatologia, diagnósticos laboratoriais e estrutura. A partir das Figuras 5 e 6, verifica-se a associação dos preditores de óbitos por febre amarela.

Tabela 4 - Dimensões e magnitude dos preditores de óbito.

DIMENSÕES	VARIÁVEIS	MAGNITUDE DOS PREDITORES DE ÓBITO
Características sócio-demográfica	Idade	0,375072
	Sexo	0,110619
	Estado	0,375072
Sintomas	Dor abdominal	0,434294
	Sinal Faget	0,061167
	Distúrbio renal	0,471834
	Hemorrágicos	0,407354
Diagnósticos laboratoriais	Resultado IGM	0,481636
	Resultado PCR	0,236988
	Isolamento viral	0,075644
Estrutura	Critério de confirmação	0,416866
	Hospitalização	0,259365
	Vacinação	0,135144
	Período Sazonal	0,259365

Fonte: Dados da pesquisa

Figura 6 - Gráfico resumo SHAP para preditores de óbito por febre amarela



Legenda:

- | | | | | | |
|----------|--|-----------|--|-----------|--|
| 1 | <ul style="list-style-type: none"> ■ Negativo ■ Positivo ■ Não realizado ou inconclusivo | 6 | <ul style="list-style-type: none"> ■ Norte e Nordeste ■ Sudeste ■ Sul e Centro-Oeste | 11 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado |
| 2 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado | 7 | <ul style="list-style-type: none"> ■ < 40 anos ■ 40 – 50 anos ■ 50 – 70 anos ■ > 70 anos | 12 | <ul style="list-style-type: none"> ■ Feminino ■ Masculino |
| 3 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado | 8 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado | 13 | <ul style="list-style-type: none"> ■ Não reagente ■ Reagente ■ Não realizado ou inconclusivo |
| 4 | <ul style="list-style-type: none"> ■ Laboratorial ■ Clínico-Epidemiológico ■ Laboratorial e Clínico-Epidemiológico ■ Ignorado | 9 | <ul style="list-style-type: none"> ■ Período Sazonal ■ Período de baixa ocorrência ■ Período pré-sazonal | 14 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado |
| 5 | <ul style="list-style-type: none"> ■ Não ■ Sim ■ Ignorado | 10 | <ul style="list-style-type: none"> ■ Negativo ■ Positivo ■ Não realizado ou inconclusivo | | |

RESULT_IGM (1), resultado do teste IgM; DISTUR_RENAL (2), presença de distúrbio renal; DOR_ABDOMINAL (3), presença de dor abdominal; CRIT_CONF (4), critério de confirmação da doença; SINT_HEMORRAGICO (5), presença de sinais hemorrágicos; UF_NOTIF (6), estado de notificação da doença; IDADE (7), idade da pessoa afetada; HOSPITALIZACAO (8), ocorrência da hospitalização; PERI_SAZONAL (9), período sazonal da febre amarela; RESULT_PCR (10), resultado do teste PCR; VACINA_FA (11), situação

vacinal do paciente; SEXO (12), sexo do paciente; R_ISOVIKAL (13), resultado do isolamento viral; SINAL_FAGET (14), presença de sinal de Faget. O gráfico de valor SHAP pode mostrar os relacionamentos positivos e negativos dos preditores com a variável de destino. Pontos vermelhos, valores altos; Pontos azuis, valores baixos; Pontos cinzas, valores faltantes.

Fonte: Dados da pesquisa

Analisando as características sociodemográficas, o óbito foi associado ao sexo masculino, residentes das Regiões Sul e Centro-Oeste e indivíduos com idade superior a 60 anos. Já a cura foi associada ao sexo feminino, residentes das regiões Norte, Nordeste e Sudeste e indivíduos com idade inferior a 40 anos (Figura 6).

Com relação aos sintomas, a presença de distúrbios hemorrágicos e/ou renais e a presença do sinal de Faget (dissociação pulso-temperatura) foram fatores associados a um pior prognóstico e evolução para o óbito, enquanto a presença ou não dor abdominal se associa a cura. (Figura 6).

Com relação ao diagnóstico laboratorial da febre amarela, o óbito foi associado ao resultado não reagente ou inconclusivo no IgM e o resultado positivo no PCR. A cura foi associada ao resultado reagente no IgM, o resultado negativo no PCR e o resultado do isolamento viral positivo ou negativo (Figura 6).

Quanto às variáveis da dimensão de estrutura, o óbito foi associado com o uso de algum critério de confirmação da febre amarela, com a hospitalização do paciente, com o período de baixa ocorrência e período pré-sazonal e com a falta de vacinação do indivíduo. A cura foi associada à ausência de dados no critério de confirmação, à não hospitalização, ao período sazonal (entre janeiro e abril) e à vacinação do indivíduo (Figura 6).

DISCUSSÃO

O modelo utilizado associou o óbito com sexo masculino, residentes das Regiões Sul e Centro-Oeste, indivíduos com idade superior a 60 anos, a presença de distúrbios hemorrágicos e/ou renais, a presença do sinal de Faget, resultado não reagente ou inconclusivo no IgM, o resultado positivo no PCR, com o uso de algum critério de confirmação da febre amarela, com a hospitalização do paciente, com o período de baixa ocorrência e período pré-sazonal e com a falta de vacinação do indivíduo.

Os achados deste estudo utilizando métodos de inteligência artificial podem contribuir de forma promissora para inovações de pesquisa no campo das doenças infecciosas. A utilização dessa ferramenta no campo biomédico, além de auxiliar na previsão da evolução e do desfecho

doença, pode facilitar o diagnóstico (FATHIMA, MANIMEGALAI e HUNDEWALE, 2011; SHAKIL, ANIS e ALAN, 2015; SHAUKAT, et. al., 2015)

Considerando todos os algoritmos testados, o *CatBoost* foi o executado em menos tempo, alcançou bons resultados nos testes já realizados, aceita valores faltantes e sem normalização, o que justifica sua utilização neste trabalho.

Em outros estudos como na criação de um modelo de previsão de risco em paciente com osteoporose, o *Catboost* teve melhor desempenho que o FRAX, SMV e *Logist regression* foi ainda capaz de propor novos preditores (KONG, et. al., 2020). Enquanto que em um estudo para prever reserva de fluxo fracionário em lesões de artéria coronária intermediária o *Adaboost* e o *Random Forest* tiveram resultado superior ao *Catboost* (LEE, et. al., 2020).

De maneira geral o *Catboost* é um bom algoritmo para conjunto de dados heterogêneos e com atributos categóricos. Contudo não se deve descartar o uso de outros algoritmos. Por outro lado, como o *Catboot* possui um bom desempenho e um tratamento automático de valores categóricos, o que facilita o seu uso, é esperado que ele apareça cada vez mais em estudos da área da saúde (HANCOCK e KHOSHGOFTAAR, 2020).

Os achados do estudo podem ajudar a identificar as áreas para melhor estruturar o sistema de saúde, em relação a organização e administração dos recursos, e melhor acolher os pacientes, além de fornecer informações de auxiliar no diagnóstico e tratamento da doença.

A idade a partir dos 40 anos esteve associada com a evolução para óbito, quanto maior a idade mais forte a associação com o óbito. No envelhecimento, é bem sabido que as alterações decorrentes desse processo natural levam a vulnerabilidade do idoso, inclusive a fisiológica (MOREIRA, et. al., 2007; GIRONDI, et. al., 2013). A imunosenescência conjunta com a vacinação para febre amarela é um assunto delicado, isso porque a imunodepressão pode não controlar a replicação do vírus vacinal, resultando no desenvolvimento da doença com quadro próximo à febre amarela silvestre ou doença viscerotrópica aguda, que é um efeito adverso grave com alta letalidade para idosos. Logo, a vacinação de pacientes com 60 anos ou mais deverá ser realizada após uma séria avaliação do risco/benefício dessa ação (MARINHO, et al., 2017; LINHEIRA-BISETTO, et al., 2016).

Sobretudo, a presença de comorbidades determina a gravidade da doença em paciente com arboviroses, visto que essas outras doenças que podem favorecer o surgimento de complicações e os idosos geralmente apresentam muitas comorbidades, associadas a uma polifarmácia, que resultam em uma série de efeitos adversos nos idosos (GÓIS e VERA, 2010; VIANA, et al., 2018). Com relação a doenças por flavivirus como dengue e febre do nilo ocidental, Badawi, et al., 2018 encontraram uma maior prevalência de comorbidades crônicas

em casos graves da doença se comparado aos não graves. Logo os idosos são um grupo que requerem atenção.

O estudo aponta que homens são os que mais evoluem a óbito por febre amarela. Os homens adultos jovens com ocupação rural, ou que vivem e trabalham próximos a regiões de matas são um dos grupos vulneráveis. Isso porque o ciclo silvestre do vírus envolve primatas não humanos e mosquitos (BRASIL, 2015). Além disso, os homens são resistentes a procurar o serviço de saúde para atendimentos, inclusive em busca de vacinação, o que além de comprometer a prevenção, pode atrasar o tratamento (FIGUEIREDO, 2005; MOURA, et. al., 2014; PINHEIRO, et al., 2002). As pessoas que adentram áreas de mata em ambientes rurais e silvestres onde o vírus ocorre naturalmente, principalmente aquelas que não são imunizadas, podem se infectar e ser um perigo para a reurbanização da doença, já que o vetor da doença (*Aedes aegypti*) está presente em grande quantidade nas cidades brasileiras. (CASALI, et. al., 2019).

Pessoas que residem nas Regiões Sul e Centro-Oeste têm maior associação ao óbito conforme indica o modelo. A região Sul não estava preparada para o aumento de casos de febre amarela devido ao surto de 2017/2018. A expansão das áreas de circulação do vírus e consequentemente o aparecimento da doença em locais não esperados, resultou na necessidade de intervenção do Ministério da Saúde. Nesse sentido foram ampliadas as áreas com recomendação da vacina e a implantação de ações de vacinação com o objetivo de alcançar 95% da população, sendo que na Região Sul a cobertura vacinal sequer chegava a 60% em 2017. Essas ações mostraram resultados na disseminação e controle da doença com a redução no número de casos e óbitos em 2019. Contudo no ano de 2020 esses casos votaram a crescer, reforçando a necessidade de acompanhamento dessas medidas implantadas para um efeito mais duradouro (FANTINI, et al. 2021). A vacina foi incorporada no calendário de vacinação na Regiões Centro-Oeste em 1999, em 2001 com a expansão da circulação do vírus, as áreas de vacinação também foram expandidas, nas Regiões Centro-Oeste, Sul e Sudeste, algo similar aconteceu entre 2008 e 2009. Com a reemergência do vírus em 2008/2009 nas Regiões Centro-Oeste, Sul e Sudeste, as ações de vigilância e monitoramento da doença foram intensificadas. Entre 2012 e 2013, 125 casos de epizootias foram registrados no Brasil com maior frequência na Região Centro-Oeste, seguida das regiões Sudeste e Sul. A partir de 2014 ocorreu um aumento de epizootias e casos humanos culminando no surto de 2017/2018 (NORONHA e CAMACHO, 2017). As informações são contraditórias, o questionamento realmente é se as regiões não estavam verdadeiramente preparadas para o surto da doença ou se foi uma questão de descaso aos avisos que foram sendo dados no decorrer dos anos.

O sinal de Faget é um achado comum no período toxêmico da doença. É o período no qual o vírus deixa a corrente sanguínea e passa a ser encontrado no baço e fígado, e em outros órgãos. Isso determina uma piora do quadro clínico do paciente (XAVIER, et. al., 2018). A dissociação pulso-temperatura, febre alta com bradicardia ou sinal de Faget é um sinal característico da febre amarela, mas pode ocorrer ainda na febre tifoide, leishmaniose, brucelose, doença do Legionário e psitacose (FROES, 1943; SPENCER, 2015). Nesse estudo a presença do sinal de Faget se associou com o óbito do paciente.

Os sintomas hemorrágicos são sinais de complicação da febre amarela, ou seja, quando presente indica que o paciente se encontra em uma situação de maior gravidade, estando relacionado a imunidade ineficaz e altas cargas virais (PAESLER e WALKER, 2012). Assim, a existência de distúrbio hemorrágico estará provavelmente associada a quadros mais graves, onde a lesão hepática é maior, com diminuição da síntese dos fatores de coagulação, deposição de fibrina e fibrinólise e presença das manifestações como a hematêmese, petéquias, sangramento uterino, epistaxe, melena, hemorragias do tegumento, gengiva e orelha e outras sufusões hemorrágicas (XAVIER, et al., 2018). Sua presença indica gravidade e como foi detectado pelo algoritmo é fator de associação com o óbito.

A forma mais grave da doença caracteriza-se pela presença de manifestações de insuficiência hepática e renal. Nestes casos a letalidade global da doença pode chegar a 50% (VASCONCELOS, 2002; CAVALCANTE e TAUIL, 2017). O distúrbio renal se inicia com insuficiência pré-renal, nas fases iniciais da doença, devido ao aumento do catabolismo e da hipovolemia. Com o agravamento do quadro, o vírus migra para os rins causando alterações glomerulares e tubulares, o que compromete a capacidade de filtração aumentam os resíduos a níveis perigosos, ainda com a coagulação intravascular disseminada e fibronólise secundária, podendo desencadear um quadro de hemorrágico sistêmico (VASCONCELOS, 2003; DE ALMEIDA NETTO, 1991). Dessa forma, a presença desse sinal indica um agravamento da doença e aumento do risco de morte, o que foi detectado pelo algoritmo, visto que a presença de distúrbios renais foi associada com o óbito do paciente.

Na cavidade torácica e abdominal são observados acúmulo de líquido pleural e ascítico. Além disso, é possível observar pequenas erosões e petéquias no tubo digestivo, distensão da vesícula biliar, sufusões hemorrágicas da mucosa da bexiga, e focos hemorrágicos no fígado (BRASIL, 2014). Esses fatores podem estar relacionados com a dor abdominal nesse sentido é esperado que a presença de dor abdominal seja um fator associado ao óbito. No entanto, no modelo tanto a presença quanto a ausência da dor abdominal são associados com a cura. É

importante ressaltar que os dados não informados, valores cinza, sobrescreveram os resultados e indicam uma associação que é questionável.

Assim como a dor abdominal, o isolamento viral é um atributo que é possível visualizar um comportamento claro dos dados, contudo a resposta desse teste quando positiva ou negativa foi associada com a cura e a ausência de informação com o óbito. Nesse sentido, entendemos que essa resposta gerada pelo modelo não apresenta uma informação clara, visto que o não preenchimento de um campo não pode ser considerado um valor de óbito ou cura do paciente, a interpretação de forma isolada acaba comprometida.

O método de PCR é uma técnica eficiente para determinação da carga viral do vírus da febre amarela, tanto *in vivo* quanto *in vitro* (GOMES, 2012). A carga viral é um fator de melhor ou pior resposta terapêutica ou prognóstico da doença, como é o caso do HIV, em que maiores cargas virais são um pior prognóstico do paciente, do citomegalovírus, que a carga viral está relacionado com o prognóstico da doença e para a hepatite C em que quanto menor a carga viral maiores são as perspectivas de bons resultados (STRAUSS, 2001; BRAVO, et al., 2006; CRUZ, 2017). Nesse sentido seria possível que a carga viral da febre amarela esteja interferindo nesse resultado, de forma que as pessoas que tiveram um resultado PCR positivo estivessem com maior carga viral, e, portanto, com pior prognóstico levando a óbito estes pacientes.

Quanto ao IgM, esse é um dos testes mais utilizados para confirmação da febre amarela, e nossos achados apresentam que o resultado positivo de IGM está associado a cura. Uma reação positiva desse teste pode indicar uma infecção recente ocorrida nos últimos 2 ou 3 meses (BRASIL, 2019a). A coleta das amostras para testes é realizada na primeira semana de sintomas. Numa amostra até 5º dia de infecção o IgM é capaz de detectar resultado nesse período, no entanto, apenas 80% dos casos são detectáveis. Entre o 6º e 10º dias, 93% a 99% dos casos são detectáveis nesse sentido para maior segurança é esperado uma segunda coleta realizada com 2 semana de sintomas (BRASIL, 2004). Ainda sabemos que os casos graves da doença costumam levar o paciente ao óbito entre sete a dez dias (OPAS, 2019). Desta forma nossos achados nos levam a sugerir que esse resultado pode ocorrer porque foram realizadas ações para melhor acolher esse paciente no serviço de saúde ou porque dentro dos 15 dias necessário para a realização do exame e da repetição, aqueles pacientes que complicaram foram a óbito e os pacientes que não complicaram tiveram tempo de desenvolver uma resposta imunológica e sobreviveram.

O método de confirmação quando respondido se associa ao óbito, ou seja, tanto faz se o critério de confirmação for laboratorial, clínico-epidemiológico ou os dois, tudo é associado ao óbito, enquanto que os valores ignorados são relacionados a cura. Assumir que o não

preenchimento de uma informação na ficha de notificação leva ao óbito ou cura do paciente é arriscado e imprudente analisar essa variável de forma isolada é impossível. O modelo não está errado, na verdade ele encontrou uma forma de contornar um problema, que no caso é a falta de dados, e diante das suas possibilidades apresentar uma resposta ao problema solicitado que é a capacidade de previsão de óbito. Para uma melhor análise dessa e das outras variáveis em que o óbito ou cura ficou condicionada à complementação do banco, seria necessário um melhor preenchimento das fichas de notificação, possibilitando visualizar a relação real dessas variáveis.

A hospitalização de pacientes com febre amarela é condicionada à gravidade da doença. Pacientes com casos moderados ou graves, que se apresentem sem sinais de hemorragias ativas e com consciência, e com os exames laboratoriais com alterações discretas no hemograma, são recomendados para a hospitalização em enfermaria. Nos casos com manifestações clínicas graves ou malignas como, por exemplo, icterícia, hemorragias, dor abdominal intensa, diminuição do nível de consciência, oligúria, etc., para esses pacientes é recomendada a hospitalização em unidade de terapia intensiva (BRASIL, 2017). Nesse sentido é esperado que os pacientes hospitalizados venham a óbito, porque eles são hospitalizados por apresentarem formas mais graves da doença. Afirmção que concorda com o resultado do trabalho. A ficha de notificação, no entanto não permite a separação entre tipos de hospitalização esse nível de informação poderia ser uma informação interessante no que se refere a previsão de óbito ou cura desse paciente.

A vacinação quando não realizada se associa com o óbito. Essa associação é visível, mas não apresenta muita força, em parte isso pode ocorrer devido ao grande número de valores ignorados dessa variável. A vacinação continua sendo o ponto principal da resposta ao surto, mas um estoque baixo de vacina forçou uma estratégia de dose poupada, que até agora foi implementada nos países africanos afetados e agora no Brasil (DOMINGO, et al., 2018). A vacina é constituída do vírus atenuado e confere imunidade à doença por pelo menos 10 anos. Essa é uma importante forma de prevenção por interromper o ciclo de transmissão da doença gerando uma barreira de imunidade coletiva (BRASIL, 2004; FERREIRA, et. al. 2011). A forma da vacina, entretanto está em debate, foi comprovado que a dose fracionada da vacina contra a febre amarela 17DD foi eficaz na indução da soroconversão, vários estudos apoiam o uso da vacina fracionada principalmente porque a vacina de rotina da febre amarela, na maioria das vezes, possui um excesso de partículas virais (ROUKENS, et al., 2018; CASEY, et al., 2019; BURKHARD, et al., 2020; WEC, et al., 2020). Visser, 2019, traz que a necessidade de revacinação vai depender da situação epidemiológica local, no caso de um surto urbano, no

caso de uma transmissão entre humanos, 80% da população com os anticorpos já garante a proteção da população mesmo com a queda da imunidade com o passar dos anos. Contudo em um quadro de transmissão silvestre, que o vetor está em contato com os seres humanos, a revacinação pode ser uma forma de aumentar a proteção. A revacinação é um ponto importante, não se pode ser feita muito cedo nem muito tarde e enquanto ainda não se tem uma resposta de quanto tempo a aplicação a dose fracionada permite de proteção, a recomendação é a vacinação de 10 em 10 anos (CAMPI-AZEVEDO, et al., 2019).

Devido ao alarme gerado pela veiculação na mídia de uma emergência epidemiológica sobre a febre amarela, de maneira errônea em 2008, que levou a população a uma busca desnecessária pela vacina em áreas por vezes que não estavam em estado de alerta, causando uma sobrecarga no sistema. Por isso, em 2008, para melhor vigilância das áreas, foram definidas prioridades de vigilância, prevenção e controle, e as áreas de risco passaram a ser classificadas como Área Com Recomendação de Vacina (ACRV) e Área Sem Recomendação de Vacina (ASRV). Para as ACRV as ações são bem mais pontuais e focadas, sendo a vacinação de 100% da população, investigação em menos de 24H de casos humanos e epizootias e captação precoce e tratamento de casos de febre amarela (ROMANO, et. al., 2011). Assim espera-se que no período sazonal da doença o serviço de saúde esteja melhor preparado para lidar com essa doença, facilitando a suspeita médica e possivelmente o diagnóstico. O modelo sugere que pacientes que notificam a doença no período sazonal estão associados a cura, o que é esperado diante do apontado que no período sazonal da doença os pacientes sejam curados.

Esse trabalho tem como limitação a qualidade do banco utilizado. As informações do banco do SINAN muitas vezes são incompletas, sejam por falta de preenchimento, por preenchimento incorreto (DE SOUZA MELO, et al. 2018). Um melhor preenchimento das fichas de notificação, e conseqüentemente uma alimentação do sistema com dados mais completos, poderia ter deixado mais claro algumas relações e permitido a análise de diversos outros pontos que não puderam ser trabalhados porque as informações não constavam no banco de dados. Outro ponto é que várias das variáveis apresentam informações gerais, por vezes um detalhamento maior de algumas informações poderia ter enriquecido os nossos resultados como é o caso da hospitalização.

CONSIDERAÇÕES FINAIS

O modelo proposto para previsão de óbito em pacientes com febre amarela teve um bom resultado com 81,1% de acurácia e 79,8% de precisão. Esses resultados permitiram aumentar o

conhecimento sobre o grau de impacto de importantes variáveis sobre a febre amarela. Possivelmente valores mais expressivos seriam obtidos se os bancos de dados iniciais tivessem uma melhor qualidade. Por isso é preciso reforçar a importância de valorização dos dados gerados nos sistemas de informação nacionais, bem como a necessidade de manutenção e capacitação constante dos profissionais que fazem a alimentação desses sistemas. O entendimento por parte desses profissionais da importância e do valor que esses dados têm é um caminho para a melhora do sistema em si, visto que são esses dados que são utilizados para a criação e direcionamento de políticas públicas. Ainda assim, o uso dessa ferramenta pode auxiliar os profissionais de saúde na prática clínica.

Além disso, o maior entendimento do comportamento do modelo permitiu uma análise da doença, reforçando pontos de atenção, indicando a possibilidade de construção de um estadiamento da febre amarela como temos para a dengue e conseqüentemente trazer melhorias importantes no manejo clínico desta enfermidade.

AGRADECIMENTOS

Este trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 através de bolsas de estudo para estudantes da pós graduação.

REFERÊNCIAS

ALONSO, S. G. et al. Data mining algorithms and techniques in mental health: A systematic review. **Journal of medical systems**. v. 42, n. 9, p. 161, 2018.

APS, L. R. de M. M. et al. Adverse events of vaccines and the consequences of non-vaccination: a critical review. **Revista de saúde pública**. v. 52, p. 40, 2018. Disponível em: <https://doi.org/10.11606/S1518-8787.2018052000384>. Acessado 18 dez. 2020.

BADAWI, A. et al. “Prevalence of chronic comorbidities in dengue fever and West Nile virus: A systematic review and meta-analysis.” **PloS one**. v. 13, n. 7, Jul. 2018. Disponível em: doi:10.1371/journal.pone.0200200

BANERJEE, P. CatBoost Classifier in Python. 2020. Disponível em: <https://www.kaggle.com/prashant111/catboost-classifier-in-python>. Acesso em: 18 dez. 2020

BISONG E. Google Colaboratory. In: Building Machine Learning and Deep Learning Models on Google Cloud Platform. Apress, Berkeley, CA, 2019. Disponível em: https://doi.org/10.1007/978-1-4842-4470-8_7

BRASIL. Ministério da Saúde. **Informe no 02 | 2018/2019. Monitoramento do Período Sazonal da Febre Amarela Brasil – 2018/2019**. 2019b.

BRASIL. Ministério da Saúde. Fundação Nacional da Saúde. **Manual de Vigilância Epidemiológica da Febre Amarela**. Brasília: Ministério da Saúde, 2004.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. **Febre amarela: guia para profissionais de saúde**. Ministério da Saúde, Secretaria de Atenção à Saúde. – Brasília : Ministério da Saúde, 2017. 67 p

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. **Guia de Vigilância em Saúde** : volume único [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. – 3ª. ed. – Brasília : Ministério da Saúde, 2019a.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. **Reemergência da febre amarela silvestre no Brasil, 2014/2015: situação epidemiológica e a importância da vacinação preventiva e da vigilância intensificada no período sazonal**. Boletim Epidemiológico. n. 29, v. 46, 2015. p. 1-10. Disponível em: <http://portalarquivos2.saude.gov.br/images/pdf/2015/outubro/19/2015-032---FA-ok.pdf>

BRASIL. **Ministério da Saúde. Secretaria de Vigilância em Saúde. Monitoramento do Período Sazonal da Febre Amarela Brasil-2017/2018**. Informe n. 26, mai, 2018.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Monitoramento do Período Sazonal da Febre Amarela Brasil – 2018/2019**. Informe n.3, jan. 2019c.

BRAVO, I. M. et al. Prevalence of oral lesions in HIV patients related to CD4 cell count and viral load in a Venezuelan population. **Medicina oral, patología oral y cirugía bucal**. Ed. inglesa, v. 11, n. 1, p. 8, 2006.

BURKHARD, J. et al. Long-term immunogenicity after yellow fever vaccination in immunosuppressed and healthy individuals. **Vaccine**. v. 38, n. 19, p. 3610-3617, 2020. Disponível em:10.1016/j.vaccine.2019.12.042

CAMPI-AZEVEDO, A. C. et al. 17DD Yellow Fever Revaccination and Heightened Long-Term Immunity in Populations of Disease-Endemic Areas, Brazil. **Emerging infectious diseases**. v. 25, n. 8, p. 1511-1521, 2019. Disponível em:10.3201/eid2508.181432

CASALI, F. T. et al. Análise das características epidemiológicas da febre amarela em um estado da Região Sudeste do Brasil. **Enfermería Actual de Costa Rica**, n. 37, p. 50-65, 2019.

CASEY, R. M. et al. Immunogenicity of Fractional-Dose Vaccine during a Yellow Fever Outbreak - Final Report. **The New England journal of medicine**. v. 381, n. 5, p. 444-454, 2019. Disponível em:10.1056/NEJMoa1710430.

CATBOOST. 2020. Disponível em: <https://catboost.ai/>. Acesso em: 18 dez. 2020

CAVALCANTE, K. R. L. J.; TAUILL, P. L. Risco de reintrodução da febre amarela urbana no Brasil. **Epidemiologia e Serviços de Saúde** [online]. v. 26, n. 3, p. 617-620, 2017.. Disponível em: <https://doi.org/10.5123/S1679-49742017000300018>. Acesso em: 12 dez. 2020.

CRUZ, F. B. C. P. da. Desenvolvimento de protocolo de PCR em tempo real para monitoramento da carga viral de cmv em pacientes pós-transplante renal. 2017. 69 f. Dissertação (Mestrado em Pesquisa Clínica em Doenças Infecciosas)-Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, 2017.

DE ALMEIDA NETTO, J. C. Aspectos clínicos e fisiopatológicos da febre amarela. **Revista de Patologia Tropical/Journal of Tropical Pathology**. v. 20, n. 1, 1991.

DE SOUZA MELO, M. A. et al. Percepção dos profissionais de saúde sobre os fatores associados à subnotificação no Sistema Nacional de Agravos de Notificação (Sinan). **Revista de Administração em Saúde**. v. 18, n. 71, 2018.

DIGANGI, E. A. e HEFNER, J. T. **Research Methods in Human Skeletal Biology**. Elsevier. 2013.

DOMINGO C, et al. Yellow fever in the diagnostics laboratory. **Emerging Microbes & Infections**. v.7, n. 1, p. 129, 2018. Disponível em: 10.1038/s41426-018-0128-8

FATHIMA, A. S.; MANIMEGALAI, D.; HUNDEWALE, N. A review of data mining classification techniques applied for diagnosis and prognosis of the arbovirus-dengue. **International Journal of Computer Science Issues (IJCSI)**. v. 8, n. 6, p. 322, 2011.

FERREIRA, K. V. et al. Histórico da febre amarela no Brasil e a importância da vacinação antiamarílica. **Arquivos Brasileiros de Ciências da Saúde**. n. 1, v. 36, 2011.

FIGUEIREDO, W. Assistência à saúde dos homens: um desafio para os serviços de atenção primária. **Ciência e Saúde coletiva**. v. 10, p. 105-109, 2005.

FROES, H. P. Estudo clínico da febre amarela. **Boletín de la Oficina Sanitaria Panamericana (OSP)**. v. 22, n. 6, 1943.

GARNELO, L. et al. Acesso e cobertura da Atenção Primária à Saúde para populações rurais e urbanas na região norte do Brasil. **Saúde em Debate**. v. 42, p. 81-99, 2018.

GIRONDI, J. B. R. et al. Estudo do perfil de morbimortalidade entre idosos. **Revista de Enfermagem da UFSM**. v. 3, n. 2, p. 197-204, 2013.

GÓIS, A. L. B. de; VERAS, Renato Peixoto. Informações sobre a morbidade hospitalar em idosos nas internações do Sistema Único de Saúde do Brasil. **Ciência & Saúde Coletiva**. v. 15, p. 2859-2869, 2010.

HAN, J.; KAMBER, M.; KAUFMANN, M. **Data Mining: Concepts and Techniques**. Elsevier: 2012.

HANCOCK, J.T.; KHOSHGOFTAAR, T. M. CatBoost for big data: an interdisciplinary review. **Journal of Big Data**. v. 7, n. 1, p. 94. 2020. Disponível em: 10.1186/s40537-020-00369-8.

Ho, Y. L. et al. Severe yellow fever in Brazil: clinical characteristics and management. **Journal of Travel Medicine**. v. 26, n. 5, 2019. Disponível em: 10.1093/jtm/taz040. PMID: 31150098.

JOHN, G. H.; LANGLEY, P. Estimating continuous distributions in Bayesian classifiers. In Proceedings of the Eleventh conference on Uncertainty in artificial intelligence (UAI'95). **Morgan Kaufmann Publishers Inc.**, San Francisco, CA, USA, p. 338–345, 1995.

KALLAS, E. G.; WILDER-SMITH, A. Managing severe yellow fever in the intensive care: lessons learnt from Brazil. **Journal of Travel Medicine**. v. 26, n. 5, jun. 2019. Disponível em: 10.1093/jtm/taz043.

KALLAS, E. G. et al. Predictors of mortality in patients with yellow fever: an observational cohort study. **The Lancet Infectious Diseases**. v. 19, n. 11, jul. 2019. Disponível em: 10.1016/S1473-3099(19)30125-2.

KONG, S. H., et al. A Novel Fracture Prediction Model Using Machine Learning in a Community-Based Cohort. **JBMR Plus**. v. 4, n. 3. fev. 2020. Disponível em: 10.1002/jbm4.10337.

LABOISSIÈRE, P. OMS alerta para possível 3ª onda de surto de febre amarela no Brasil. Agência Brasil. 2019.

LEE, J. G. et al. Intravascular ultrasound-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. **Atherosclerosis**. v. 292, p. 171-177, jan. 2020. Disponível em: 10.1016/j.atherosclerosis.2019.10.022.

LINHEIRA-BISETTO, Lúcia Helena et al. Ocorrência de eventos adversos pós-vacinação em idosos. **Cogitare Enfermagem**. v. 21, n. 4, p. 01-10, 2016.

LUCENA, A. R. F. P. et al. Fatores associados à gravidade dos eventos adversos pós-vacinação contra a febre amarela durante o maior surto da doença registrado no Brasil, 2016-2017. **Epidemiologia e Serviços de Saúde** [online]. Brasília, v. 29, n. 1, 2020. Disponível em: <<https://doi.org/10.5123/S1679-49742020000100017>>. Acesso em: 18 dez 2020

LUND, J. L. et al. Development and validation of a 5-year mortality prediction model using regularized regression and Medicare data. **Pharmacoepidemiology and Drug Safety**. v. 28, n. 5, p. 584-592, mai. 2019. Disponível em: 10.1002/pds.4769.

LUO, M. et al. A new scoring model for the prediction of mortality in patients with acute kidney injury. **Scientific Reports**. v.7, n. 1, p.7862, ago. 2017. Disponível em: 10.1038/s41598-017-08440-w.

MAHINDRAKAR, P.; HANUMANTHAPPA, M. Data mining in healthcare: A survey of techniques and algorithms with its limitations and challenges. **International Journal of Engineering Research and Applications**. v. 3, n. 6, p. 937-941, 2013.

- MALTA, D. C. et al. A cobertura da Estratégia de Saúde da Família (ESF) no Brasil, segundo a Pesquisa Nacional de Saúde, 2013. **Ciência & Saúde Coletiva**. v. 21, p. 327-338, 2016.
- MANSOOR, H. et al. Risk prediction model for in-hospital mortality in women with ST-elevation myocardial infarction: A machine learning approach. **Heart & Lung**. v. 46, n. 6, p. 405-411, nov.-dez. 2017. Disponível em: [10.1016/j.hrtlng.2017.09.003](https://doi.org/10.1016/j.hrtlng.2017.09.003).
- MARINHO, A. K. B. B. et al. Vacina contra a febre amarela: reações adversas e populações de risco. **Arquivos de Asma, Alergia e Imunologia**. v. 1, p. 245-256, 2017.
- MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. Machine learning: An artificial intelligence approach. **Springer Science & Business Media**. 2013
- MOLNAR, C. Interpretable machine learning. A Guide for Making Black Box Models Explainable. 2019. Disponível em: <https://christophm.github.io/interpretable-ml-book/>.
- MOREIRA, I. P. B. et al. Principais aspectos do tratamento das infecções no idoso. **Ciência, Cuidado e Saúde**. v. 6, p. 488-495, 2007.
- MOURA, E. C. de et al. Atenção à saúde dos homens no âmbito da Estratégia Saúde da Família. **Ciência & Saúde Coletiva** [online]. v. 19, n. 02, p. 429-438, 2014. Disponível em: <https://doi.org/10.1590/1413-81232014192.05802013>. Acesso em: 18 dez. 2020.
- NANCY, P.; RAMANI, R. G. A comparison on performance of data mining algorithms in classification of social network data. **International Journal of Computer Applications**. v. 32, n. 8, p. 47-54, 2011.
- NEVES, R. G. et al. Tendência temporal da cobertura da Estratégia Saúde da Família no Brasil, regiões e Unidades da Federação, 2006-2016. **Epidemiologia e Serviços de Saúde**. v. 27, p. e2017170, 2018.
- NORONHA, T. G. de e CAMACHO, L. A. B. Controvérsias sobre a ampliação das áreas com vacinação de rotina contra a febre amarela no Brasil. **Cadernos de Saúde Pública** [online]. v. 33, n. 10, 2017. Disponível em: <https://doi.org/10.1590/0102-311X00060917>. Acesso em: 18 jan. 2021.
- OPAS BRASIL. Folha informativa - Febre amarela. Disponível em: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5578:folha-informativa-febre-amarela&Itemid=875.
- OPAS. Folha informativa - Febre amarela. OPAS (2019). Available at: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5578:folha-informativa-febre-amarela&Itemid=875.
- OZER, P. Data Mining Algorithms for Classification. 2008.
- PAL, S. K.; MITRA, S. Multilayer perceptron, fuzzy sets, classification. 1992.
- PINHEIRO, R. S. et al. Gênero, morbidade, acesso e utilização de serviços de saúde no Brasil. **Ciência e Saúde coletiva**. v. 7, p. 687-707, 2002.

PINHEIRO, C. A. R. **Inteligência analítica: mineração de dados e descoberta do conhecimento**. Ciência Moderna, 2008.

PIROS, P. et al. Comparing machine learning and regression models for mortality prediction based on the Hungarian Myocardial Infarction Registry. **Knowledge-Based Systems**. v. 179, p. 1-7, 2019. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0950705119302060>)

PRAKHAR, R. A Novel Approach to Feature Importance — Shapley Additive Explanations. 2020. Disponível em: <https://towardsdatascience.com/a-novel-approach-to-feature-importance-shapley-additive-explanations-d18af30fc21b>

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. Curvas ROC para avaliação de classificadores. **IEEE América Latina**. v. 6, n. 2, 2008.

PROKHORENKOVA, L. et al. CatBoost: unbiased boosting with categorical features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). **Curran Associates Inc.**, Red Hook, NY, USA, p. 6639–6649. 2018.

ROMANO, A. P. M. et al. Febre amarela no Brasil: recomendações para a vigilância, prevenção e controle. *Epidemiologia e Serviços de Saúde*, Brasília, v. 20, n. 1, p. 101-106, mar. 2011. Disponível em: http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742011000100011&lng=pt&nrm=iso. Acessos em 18 dez. 2020.

ROUKENS, A. H. E. et al. Long-Term Protection After Fractional-Dose Yellow Fever Vaccination: Follow-up Study of a Randomized, Controlled, Noninferiority Trial. **Annals of internal medicine**. v. 169, n.11, p. 761-765, 2018. Disponível em: doi:10.7326/M18-1529

SAAD, W. et al. Coalitional game theory for communication networks. **Ieee signal processing magazine**, v. 26, n. 5, p. 77-97, 2009.

SHAKIL, K. A.; ANIS, S.; ALAM, M. Dengue disease prediction using weka data mining tool. **arXiv preprint arXiv:1502.05167**, 2015.

SHAUKAT, K. et al. Dengue fever prediction: A data mining problem. **Journal of Data Mining in Genomics & Proteomics**, v. 6, n. 3, 2015.

SIQUEIRA-BATISTA, R.; DA SILVA, E. Notas Sobre os Fundamentos Matemáticos Da Inteligência Artificial. **Revista de Ciência, Tecnologia e Inovação**. v. 4, n. 6, 2020.

SPENCER, I. M. **Febre: padrões de febre e o seu impacto na patologia**. Dissertação (mestrado integrado em medicina) - Faculdade de Medicina e Universidade de Coimbra, Coimbra, 2015.

STRAUSS, E. Hepatite C. *Rev. Soc. Bras. Med. Trop.*, Uberaba, v. 34, n. 1, p. 69-82, Feb. 2001. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0037-86822001000100011&lng=en&nrm=iso. Acesso em: 17 Jan. 2021. <https://doi.org/10.1590/S0037-86822001000100011>.

TAN, P. N.; STEINBACH, M.; KUMAR, V. **Introduction to data mining**. Pearson Education, São Carlos. 2009

TAUIL, P. L. Aspectos críticos do controle da febre amarela no Brasil. **Revista de Saúde Pública**. São Paulo (São Paulo), v. 44, p. 555-558, 2010. Disponível em: <https://www.scielo.org/article/rsp/2010.v44n3/555-558/#ModalArticles>.

TAYLOR, R. A. et al. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. **Academic Emergency Medicine**. v. 23, n. 3, p. 269-278, 2016. Disponível em: 10.1111/acem.12876

VAKILINIA, I.; SENGUPTA, S. A coalitional game theory approach for cybersecurity information sharing. In: MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM). Institute of Electrical and Electronics Engineers. 2017. p. 237-242.

VASCONCELOS, P. F. da C. Febre amarela: reflexões sobre a doença, as perspectivas para o século XXI e o risco da reurbanização. **Revista Brasileira de. Epidemiologia**. v. 5, p. 244–258, 2002.

VASCONCELOS, P. F. da C. Febre amarela. **Revista da Sociedade Brasileira de Medicina Tropical**. v. 36, n. 2, p. 275-293, 2003.

VIANA, L. R.de C. et al. Arboviroses reemergentes: perfil clínico-epidemiológico de idosos hospitalizados. **Revista da Escola de Enfermagem da USP**. v. 52, 2018.

VISSER, L. G. Fractional-dose yellow fever vaccination: how much more can we do with less?. **Current opinion in infectious diseases**. v. 32, n. 5, p. 390-393, 2019. Disponível em: 10.1097/QCO.0000000000000576

WAGGONER, J. J.; ROJAS, A.; PINSKY, B. A. Yellow Fever Virus: Diagnostics for a Persistent Arboviral Threat. **Journal of Clinical Microbiology**. v. 56, n. 10, 2018. Disponível em: 10.1128/JCM.00827-18.

WANG, Z. et al. Mortality prediction system for heart failure with orthogonal relief and dynamic radius means. **International Journal of Medical Informatics**. v. 115, . p.10-17, 2018. Disponível em: 10.1016/j.ijmedinf.2018.04.003

WEC, A. Z. et al. Longitudinal dynamics of the human B cell response to the yellow fever 17D vaccine. **Proceedings of the National Academy of Sciences of the United States of America** vo. 117. n.12, p. 6675-6685, 2020. Disponível em:10.1073/pnas.1921388117

WEI YEE LEONG, M. P. H. Novas ferramentas de diagnóstico para febre amarela, **Journal of Travel Medicine**. v. 25, n. 1, 2018. Disponível em: <https://doi.org/10.1093/jtm/tay079>

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data Mining: Pratical Machine learning tools and techniques**. Elsevier, 2011.

XAVIER, A. R. et al. Febre amarela: diagnóstico laboratorial e manifestações clínicas. **Jornal Brasileiro de Patologia e Medicina Laboratorial**. v. 54, n. 5, p. 296-305, 2018.

CONSIDERAÇÕES FINAIS

Observamos uma alta prevalência de zika (19%) e chikungunya (28%) relacionadas a uma pesquisa pontual, focada apenas em de epidemias ou casos suspeitos, logo sendo necessário o avanço dos estudos para um melhor entendimento dessas doenças no país, esse entendimento vai auxiliar em um melhor acompanhamento e direcionamento das ações de saúde.

A alta prevalência de dengue encontrada (35%) é preocupante, mas observamos que no Brasil essa é uma doença muito estudada no decorrer dos anos, contudo o novo cenário de COVID-19 associado a dengue e outras arboviroses e requer uma atitude de vigilância e controle efetiva no país. Quanto a febre amarela, os poucos estudos encontrados indicam uma necessidade investigar os casos de febre amarela no país para assim melhor entender a realidade da doença no Brasil e possivelmente preencher as lacunas acerca da patologia da doença.

Foi possível prever com acurácia de 81,1% e precisão de 79,8% se um paciente com febre amarela viria a óbito ou seria curado. Esse resultado demonstra que a inteligência artificial pode ser utilizada na área da saúde como um reforço em determinados assuntos. Ainda observamos que é necessário reforçar a importância de valorização dos dados gerados nos sistemas de informação nacionais, bem como a necessidade de manutenção e capacitação constante dos profissionais que fazem a alimentação desses sistemas. Esses profissionais precisam compreender a importância de seu trabalho na criação e direcionamento de políticas públicas e gerenciamento dos recursos. Possivelmente valores mais expressivos seriam obtidos se os bancos de dados iniciais tivessem uma melhor qualidade e fossem mais completos.

Além disso, o maior entendimento do comportamento do modelo preditivo permitiu uma análise da doença, reforçando pontos de atenção, indicando a possibilidade de construção de um estadiamento da febre amarela como temos para a dengue e conseqüentemente trazer melhorias importantes no manejo clínico desta enfermidade.

REFERÊNCIAS

ABOUZahr, C.; ADJEL, S.; KANCHANACHITRA, C. From data to policy: good practices and cautionary tales. **The Lancet**, v. 369, n. 9566, p. 1039-1046, 2007.

ALONSO, S.G. et al. Data mining algorithms and techniques in mental health: A systematic review. **Journal of medical systems**, v. 42, n. 9, p. 161, 2018.

AVELINO-SILVA, V. I.; RAMOS, J. F. Arboviroses e políticas públicas no Brasil / Arboviruses and public policies in Brazil. **Revista Ciências em Saúde**, v. 7, n. 3, p. 1-2, 13 set. 2017.

BELLINGER, C. et al. A systematic review of data mining and machine learning for air pollution epidemiology. **BMC public health**, v. 17, n. 1, p. 1-19, 2017.

BRASIL. Ministério da Saúde. Banco de dados do Sistema Único de Saúde-DATASUS.2021. Disponível em:
<<http://www2.datasus.gov.br/DATASUS/index.php?area=0203&id=29878153>>. Acesso em: 19 jan. 2021.

BRASIL. Ministério da Saúde. Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes Aegypti (dengue, chikungunya e zika), semanas epidemiológicas 1 a 50, 2020. Secretaria de Vigilância em Saúde. v. 51. Dez. 2020c. Disponível em:
<https://www.gov.br/saude/pt-br/media/pdf/2020/dezembro/28/boletim_epidemiologico_svs_51.pdf>

BRASIL. Ministério da saúde. Portaria de consolidação nº 4, de 28 de setembro de 2017a. Consolidação das normas sobre os sistemas e os subsistemas do Sistema Único de Saúde. Disponível em:
<http://bvsmms.saude.gov.br/bvs/saudelegis/gm/2017/prc0004_03_10_2017.html> Acesso em: 18 jan. 2021.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Febre amarela: guia para profissionais de saúde / Ministério da Saúde, Secretaria de Atenção à Saúde. – Brasília: Ministério da Saúde, 2017c. p. 67.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Boletim Epidemiológico. Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes (dengue, chikungunya e Zika): Semanas Epidemiológicas 1 a 36. v. 50, n.28, out. 2019b.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Boletim Epidemiológico. Monitoramento dos casos de arboviroses urbanas transmitidas pelo Aedes Aegypti (dengue, chikungunya e zika), semanas epidemiológicas 1 a 36, 2020. v. 51, n.38, set. 2020a.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Boletim Epidemiológico. Situação Epidemiológica da Febre Amarela no Monitoramento 2019/2020. v. 51, n.24, jun. 2020b.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. Guia de Vigilância em Saúde : volume único [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. 3ª ed. Brasília: Ministério da Saúde, 2019a. p. 740.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. Chikungunya : manejo clínico / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância das Doenças Transmissíveis. – Brasília: Ministério da Saúde, 2017b. p. 65.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Sistema de Informação de Agravos de Notificação–Sinan: normas e rotinas / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância Epidemiológica. – Brasília : Editora do Ministério da Saúde, 2006.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. Dengue: diagnóstico e manejo clínico: adulto e criança [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância das Doenças Transmissíveis. – 5. ed. – Brasília: Ministério da Saúde, 2016. p.58.

BRASIL. Portaria no 1882/GM, de 18 de dezembro de 1997. Estabelece o Piso da Atenção Básica - PAB e sua composição. Disponível em: <https://www.saude.pr.gov.br/sites/default/arquivos_restritos/files/documento/2020-05/portgm1882_18dezembro_1997.pdf> Acesso em: 18 jan. 2021.

CALADO, A. F. S.; PAZ, F. A. do N. Análise de perfil epidemiológico e incidência de febre amarela no Brasil. **Research, Society and Development**, v. 9, n. 3, p. 40, 2020. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=7340988>. Acesso em: 11 fev. 2022.

CÂMARA, F. P.; DE CARVALHO, L. M.; GOMES, A. L. B. Demographic profile of sylvatic yellow fever in Brazil from 1973 to 2008. **Transactions of The Royal Society of Tropical Medicine and Hygiene**, v. 107, n. 5, p. 324-327, 2013.

CARDONA-OSPINA, J. A. et al. Fatal Zika virus infection in the Americas: a systematic review. **International Journal of Infectious Diseases**, v. 88, p. 49-59, 2019.

CARVALHO, D. R.; ESCOBAR, Leandro Fabian Almeida; TSUNODA, Denise. Pontos de Atenção para o Uso da Mineração de Dados da Saúde. **Informação & Informação**, v. 19, n. 1, p. 249-272, 2014.

CATBOOST. Yandex. 2021 Disponível em: <https://catboost.ai/>. Acesso em: 20. jan. 2021

DE ÁVILA, R. E. et al. Clinical profiles and factors associated with mortality in adults with yellow fever admitted to an intensive care unit in Minas Gerais, Brazil. **International Journal of Infectious Diseases**, v. 93, p. 90-97, 2020.

DIGANGI, E. A.; HEFNER, J. T. Ancestry estimation. **Research methods in human skeletal biology**, p. 117-149, 2013.

DO CARMO CUPERTINO, M. et al. Epidemiological, prevention and control updates of yellow fever outbreak in Brazil. **Asian Pacific Journal of Tropical Medicine**, v. 12, n. 2, p. 49, 2019.

ESCOSTEGUY, C. C. et al. Yellow fever: profile of cases and factors associated with death in a hospital in the State of Rio de Janeiro, 2017–2018. **Revista de Saúde Pública**, v. 53, p. 89, 2019.

ESMAEILY, H. et al. Comparing three data mining algorithms for identifying the associated risk factors of type 2 diabetes. **Iranian biomedical journal**, v. 22, n. 5, p. 303, 2018.

FALCÃO, et al. Guia de manejo da infecção pelo vírus zika. **Sociedade Brasileira de Infectologia**, 2016.

FERNANDES, D. L. **Framework de mineração de dados para análise de experimentos com eye tracking**: uma aplicação em bioética. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Viçosa. Viçosa-MG, p.104. 2017.

FRANCO, O. **História da febre amarela no Brasil**. Ministério da Saúde. Departamento Nacional de Endemias Rurais. Rio de Janeiro- Brasil. 1969.

GONOODI, K. et al. An assessment of the risk factors for vitamin D deficiency using a decision tree model. **Diabetes & Metabolic Syndrome: Clinical Research & Reviews**, v. 13, n. 3, p. 1773-1777, 2019a.

GONOODI, K. et al. Determinants of the magnitude of response to vitamin D supplementation in adolescent girls identified using a decision tree algorithm. **BioFactors**, v. 45, n. 5, p. 795-802, 2019b.

HAN, J., KAMBER, M. e KAUFMANN, M. **Data Mining: Concepts and Techniques**. Elsevier, 2012.

HANCOCK, J.T.; KHOSHGOFTAAR, T. M. CatBoost for big data: an interdisciplinary review. **Journal of Big Data**. v. 7, n. 1, p. 94. 2020. Disponível em: 10.1186/s40537-020-00369-8.

HEARN, J. et al. Neural Networks for Prognostication of Patients With Heart Failure: Improving Performance Through the Incorporation of Breath-by-Breath Data From Cardiopulmonary Exercise Testing. **Circulation: Heart Failure**, v. 11, n. 8, p. e005193, 2018.

HILLESTAD, R. et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. **Health affairs**, v. 24, n. 5, p. 1103-1117, 2005.

JÁCOME, R. et al. A yellow flag on the horizon: The looming threat of yellow fever to North America. **International Journal of Infectious Diseases**, v. 87, p. 143-150, 2019.

KALLAS, E. G. et al. Predictors of mortality in patients with yellow fever: an observational cohort study. **The Lancet Infectious Diseases**, v. 19, n. 7, p. 750-758, 2019.

KAVAKIOTIS, I. et al. Machine learning and data mining methods in diabetes research. **Computational and structural biotechnology journal**, v. 15, p. 104-116, 2017.

KIM, Y. C. et al. Development of an e2 elisa methodology to assess chikungunya seroprevalence in patients from an endemic region of mexico. **Viruses**, v. 11, n. 5, p. 407, 2019.

KRAEMER, M. U. G. et al. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. **Nature microbiology**, v. 4, n. 5, p. 854-863, 2019.

KRAEMER, M. U. G. et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. **elife**, v. 4, p. e08347, 2015.

LAGUARDIA, J. et al. Sistema de informação de agravos de notificação em saúde (Sinan): desafios no desenvolvimento de um sistema de informação em saúde. **Epidemiologia e Serviços de Saúde**, v. 13, n. 3, p. 135-146, 2004.

LIMA, K. W. S. de. **Análise das condições de operação do Sistema de Informação de Agravos de Notificação (SINAN) nos municípios paulistas**. 2018. Tese (Doutorado em Epidemiologia) - Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, 2018. doi:10.11606/T.6.2018.tde-04052018-114645. Acesso em: 20 jan. 2021.

LIMA-CAMARA, T. N. Arboviroses emergentes e novos desafios para a saúde pública no Brasil. **Revista de Saúde Pública**, v. 50, p. 36, 2016.

LITVOC, M. N.; NOVAES, C. T. G.; LOPES, M. I. B. F. Yellow fever. **Revista da Associação Médica Brasileira**, v. 64, n. 2, p. 106-113, 2018.

LIU, Mi-Mi et al. Application of data mining methods to improve screening for the risk of early gastric cancer. **BMC Medical Informatics and Decision Making**, v. 18, n. 5, p. 121, 2018.

LOPES, N.; NOZAWA, C.; LINHARES, R. E. C. Características gerais e epidemiologia dos arbovírus emergentes no Brasil. **Revista Pan-Amazônica de Saúde**, Ananindeua, v. 5, n. 3, p. 55-64, set. 2014. Disponível em <http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S2176-62232014000300007&lng=pt&nrm=iso>. Acesso em 20 jan. 2021.

MATPLOTLIB. Matplotlib: Visualization with Python. 2021. Disponível em: <<https://matplotlib.org/>>. Acesso em: 20. jan. 2021.

MAYER, S.V.; TESH, R. B.; VASILAKIS, N. The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. **Acta tropica**, v. 166, p. 155-163, 2017.

MEDEIROS, D.; VASCONCELOS, P. F. C. Is the brazilian diverse environment is a crib for the emergence and maintenance of exotic arboviruses? **Anais da Academia Brasileira de Ciências**, v. 91, 2019.

MITCHELL, R. S.; MICHALSKI, J. G.; CARBONELL, T. M. **An artificial intelligence approach**. Berlin: Springer, 2013.

MOHAMADLOU, H. et al. Prediction of Acute Kidney Injury With a Machine Learning Algorithm Using Electronic Health Record Data. **Canadian journal of kidney health and disease**. vol. 5. 8 Jun. 2018, doi:10.1177/2054358118776326.

NUMPY. NumPy. 2020. Disponível em: <https://numpy.org/>. Acesso em: 20. jan. 2021.

NUNES, P. C. G. et al. 30 years of fatal dengue cases in Brazil: a review. **BMC public health**, v. 19, n. 1, p. 1-11, 2019.

OPAS Brasil. Casos de dengue nas Américas chegam a 1,6 milhão, o que destaca a necessidade do controle de mosquitos durante a pandemia. 2020. Disponível em: <https://www.paho.org/bra/index.php?option=com_content&view=article&id=6205:casos-de-dengue-nas-americas-chegam-a-1-6-milhao-o-que-destaca-a-necessidade-do-controle-de-mosquitos-durante-a-pandemia&Itemid=812#:~:text=23%20de%20junho%20de%202020,%C3%A0%20pandemia%20de%20COVID%2D19>.

OPAS Brasil. Folha informativa–Dengue e dengue grave. PAHO - Pan American Health Organization. 2019a. Disponível em: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5963:folha-informativa-dengue-e-dengue-grave&Itemid=812>. Acesso em: 18 jan. 2021.

OPAS Brasil. Folha informativa-Febre amarela. PAHO - Pan American Health Organization. 2019b. Disponível em: <https://www.paho.org/bra/index.php?option=com_content&view=article&id=5578:folha-informativa-febre-amarela&Itemid=875>. Acesso em: 19 jan. 2021.

Organização Pan-Americana da Saúde (OPAS). 2022. Febre amarela. Disponível em: <https://www.paho.org/pt/node/40>. Acesso em: 11 fev. 2022.

Organización Pañamericana de la Salud (OPS). PLISA (Plataforma de Información en SaludparalasAméricas). Zika. 2024. Disponível em: <https://www3.paho.org/data/index.php/es/temas/indicadores-zika.html>. Acesso em: 17 nov. 2024.

Organización Pañamericana de la Salud (OPS). PLISA (Plataforma de Información en SaludparalasAméricas). Dengue. 2022. Disponível em: <https://www3.paho.org/data/index.php/es/temas/indicadores-dengue/dengue-nacional/9-dengue-pais-ano.html>. Acesso em: 10 fev. 2022.

PAIXÃO, E. S. et al. Impact evaluation of Zika epidemic on congenital anomalies registration in Brazil: An interrupted time series analysis. **PLoS neglected tropical diseases**, v. 13, n. 9, p. e0007721, 2019.

PAL, S. K.; MITRA, S. Multilayer perceptron, fuzzy sets, classification. **IEEE Transactions on Neural Networks**. v. 3, n.5, 1992. Disponível em: <<http://library.isical.ac.in:8080/jspui/bitstream/10263/4569/1/308.pdf>>. Acesso em: 19 jan. 2021.

Pan American health organization (PAHO). 2022a. PLISA (Health Information Platform for the Americas). Chikungunya Cases by country or territory Disponível em: <https://www3.paho.org/data/index.php/en/mnu-topics/chikv-en/550-chikv-weekly-en.html>. Acesso em: 17 nov. 2024.

Pan American health organization (PAHO). 2022b. Yellow Fever: number of confirmed human cases and deaths in endemic countries of the Americas, since 1960. Disponível em: https://ais.paho.org/phis/viz/ed_yellowfever.asp. Acesso em: 11 fev. 2022.

PANDAS. Pandas. NumFOCUS. 2021. Disponível em: <https://pandas.pydata.org/#:~:text=pandas%20is%20a%20fast%2C%20powerful,of%20the%20Python%20programming%20language..> Acesso em: 20. jan. 2021.

PATTERSON, J.; SAMMON, M.a; GARG, M. Dengue, Zika and chikungunya: emerging arboviruses in the New World. **Western Journal of Emergency Medicine**, v. 17, n. 6, p. 671, 2016.

PHAKHOUNTHONG, K. et al. Predicting the severity of dengue fever in children on admission based on clinical features and laboratory indicators: application of classification tree analysis. **BMC pediatrics**, v. 18, n. 1, p. 109, 2018.

PINHEIRO, C. A. R. **Inteligência analítica: mineração de dados e descoberta do conhecimento**. Ciência Moderna, 2008.

PINHEIRO, T. F.; DE BARROS ALVES, J.; DA SILVA, Y. R.N. O impacto financeiro das arboviroses oriundas do Aedes Aegypti no Brasil: uma projeção para 2019. **Brazilian Journal of Development**, v. 6, n. 5, p. 30757-30767, 2020.

PIROS, P. et al. Comparing machine learning and regression models for mortality prediction based on the Hungarian Myocardial Infarction Registry. **Knowledge-Based Systems**. v. 179, p.1-7. 2019. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0950705119302060>

PORTAL SINAN. O SINAN. Portal SINAN, 2019. Disponível em: <http://portalsinan.saude.gov.br/o-sinan> Acesso em: 18 jan. 2021.

PORTO, W. L. et al. Cenário epidemiológico das arboviroses no Piauí. **Revista Eletrônica Acervo Saúde**, v. 11, n. 14, p. e1054-e1054, 2019.

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. Curvas ROC para avaliação de classificadores. **IEEE América Latina**. v. 6, n. 2, 2008.

RANAPURWALA, S. I et al. Public health application of predictive modeling: an example from farm vehicle crashes. **Injury epidemiology**. v. 6, n.31. 17 Jun. 2019, doi:10.1186/s40621-019-0208-9

RUDD, K. E. et al. Building workforce capacity for effective use of health information systems: Evaluation of a blended eLearning course in Namibia and Tanzania. **International Journal of Medical Informatics**, v. 131, p. 103945, 2019.

SACCHETTO, L. et al. Screening for Zika virus RNA in sera of suspected cases: a retrospective cross-sectional study. **Virology journal**, v. 15, n. 1, p. 155, 2018.

SANTORO, D. M. **Sobre o processo de seleção de subconjuntos de atributos** - As abordagens filtro e wrapper. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de São Carlos. São Carlos-SP, p.141. 2005.

SANTOS, H. G. et al. Machine learning para análises preditivas em saúde: exemplo de aplicação para predizer óbito em idosos de São Paulo, Brasil. *Cadernos de saúde pública*, v. 35, p. e00050818, 2019.

SARFRAZ, M. S. et al. Mapping urban and peri-urban breeding habitats of Aedes mosquitoes using a fuzzy analytical hierarchical process based on climatic and physical parameters. **Geospatial health**, p. S685-S697, 2014.

SATO, A. P. S. What is the importance of vaccine hesitancy in the drop of vaccination coverage in Brazil?. **Revista de saúde pública**, v. 52, p. 96, 2018.

SCIKIT-LEARN. Machine Learning in Python. 2021. Disponível em: < <https://scikit-learn.org/stable/>>. Acesso em: 20. jan. 2021.

SENANAYAKE, S. et al. Machine learning in predicting graft failure following kidney transplantation: A systematic review of published predictive models. **International journal of medical informatics** v. 130, 2019. doi:10.1016/j.ijmedinf.2019.103957

SHAP. shap 0.38.1. Python Software Foundation. 2021. Disponível em: < <https://pypi.org/project/shap/>>. Acesso em: 20. jan. 2021.

SILVA, T. A. Como interpretar as métricas Precisão, Revocação, Acurácia e Medida-F. 2018. Disponível em: <https://tiago.blog.br/precisao-revocacao-acuracia-e-medida-f/>>. Acesso em: 20. jan. 2021.

SINGH, P.; SINGH, S.; PANDI-JAIN, G. S. Effective heart disease prediction system using data mining techniques. **International journal of nanomedicine**, v. 13, n. T-NANO 2014 Abstracts, p. 121, 2018.

SIQUEIRA, P. C. et al. Completude das fichas de notificação de febre amarela no estado do Espírito Santo, 2017. **Epidemiologia e Serviços de Saúde**, v. 29, p. e2019402, 2020.

TAN, P.-N., STEINBACH, M. & KUMAR, V. **Introdução ao Data Mining. Mineração de Dados**. Ciência Moderna, 2009.

THOMPSON, R.; MARTIN DEL CAMPO, J.; CONSTENLA, D. A review of the economic evidence of Aedes-borne arboviruses and Aedes-borne arboviral disease prevention and control strategies. *Expert review of vaccines*, v. 19, n. 2, p. 143-162, 2020.

TUBOI, S. H. et al. Clinical and epidemiological characteristics of yellow fever in Brazil: analysis of reported cases 1998–2002. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 101, n. 2, p. 169-175, 2007.

VASCONCELOS, P. F. da C. Febre amarela. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 36, n. 2, p. 275-293, 2003.

VASCONCELOS, P. F. da C. Febre amarela: reflexões sobre a doença, as perspectivas para o século XXI e o risco da reurbanização. **Revista Brasileira de Epidemiologia**, v. 5, n. 3, p. 244-258, 2002.

WEAVER, S. C.; REISEN, W. K. Present and future arboviral threats. *Antiviral research*, v. 85, n. 2, p. 328-345, 2010.

WIMALASIRI-YAPA, BMC R. et al. Chikungunya virus in Asia–Pacific: a systematic review. **Emerging microbes & infections**, v. 8, n. 1, p. 70-79, 2019.

WITTEN, I. H., FRANK, E. e HALL, M. A. **Data Mining: Practical Machine learning tools and techniques**. 3 ed. Elsevier, 2011.

World Health Organization (WHO). 2022a. Chikungunya. Disponível em: https://www.who.int/health-topics/chikungunya#tab=tab_1. Acesso em: 09 fev. 2022.

World Health Organization (WHO). 2022b. Dengue cases include clinically suspected and laboratory confirmed cases. Disponível em: <https://ntdhq.shinyapps.io/dengue5/>. Acesso em: 10 fev. 2022.

World Health Organization (WHO). 2022c. Yellow fever. Disponível em: https://www.who.int/health-topics/yellow-fever#tab=tab_1. Acesso em: 11 fev. 2022.

World Health Organization (WHO). 2022d. Zika virus disease. Disponível em: https://www.who.int/health-topics/zika-virus-disease#tab=tab_1. Acesso em: 09 fev. 2022.

World Health Organization (WHO). Global yellow fever update, 2019. **Weekly Epidemiological Record**. N. 34, p. 393–408, 2020. Disponível em: <https://iris.who.int/bitstream/handle/10665/333916/WER9534-393-407-eng-fre.pdf?sequence=1>. Acesso em: 17 nov. 2024.

ANEXO A – DISPENSA DO COMITÊ DE ÉTICA

12/08/2020

Gmail - Solicitação de informação



Karen Fracalossi <karen.fracalossi@gmail.com>

Solicitação de informação

4 mensagens

Karen Fracalossi <karen.fracalossi@gmail.com>
Para: cep@ufv.br

17 de outubro de 2019 11:39

Bom Dia!!

Gostaria de uma informação sobre um projeto que tenho a proposta de desenvolver. Esse projeto tem o foco em analisar dados secundários das arboviroses Dengue, Zika, Chikungunya e Febre Amarela, fornecidos pelo Ministério da Saúde, por meio do Sistema Eletrônico do Serviço de Informação ao Cidadão. Os dados fornecidos não contem qualquer forma de identificação dos usuários do sistema. Dessa forma gostaria de saber se é necessário submeter esse projeto ao Comitê de Ética e se sim quais as ações a serem tomadas. Caso não seja necessário a submissão, qual o procedimento para receber uma declaração do Comitê para essa liberação?
Aguardo resposta. Atenciosamente,

-

Karen Oliveira Fracalossi
Nutricionista
Mestranda em Ciência da Nutrição - UFV
Tel.: (31) 9 75340581

CEP UFV <cep@ufv.br>
Para: Karen Fracalossi <karen.fracalossi@gmail.com>

18 de outubro de 2019 08:28

Bom dia, Karen.
O dados são de domínio público?
Atenciosamente,
[Texto das mensagens anteriores oculto]

Juliana Aparecida Ferreira
Secretária do Comitê de Ética em Pesquisa com Seres Humanos
Universidade Federal de Viçosa
www.cep.ufv.br
(31)3612-2316

Karen Fracalossi <karen.fracalossi@gmail.com>
Para: CEP UFV <cep@ufv.br>

20 de outubro de 2019 22:19

Boa noite!!
Sim os dados podem ser solicitados por qualquer cidadão por meio da Lei de Acesso à Informação. Sendo a "disponibilização no sistema e-SIC tem status de resposta oficial, dispensando solenidades, uma vez que os atos administrativos em geral devam se pautar pela celeridade, informalidade e eficiência". (Fonte: Aplicação da Lei de Acesso à Informação na Administração Pública, 2ª edição, CGU, 2016).
Atenciosamente,

[Texto das mensagens anteriores oculto]

CEP UFV <cep@ufv.br>
Para: Karen Fracalossi <karen.fracalossi@gmail.com>

21 de outubro de 2019 08:27

Bom dia, Karen.
De acordo com o Art.1º, II da Resolução CNS nº510/2016, as pesquisas que utilizem dados de acesso público não serão registradas nem avaliadas pelo sistema CEP/CONEP, nos termos da lei de acesso à informação.
Atenciosamente,
[Texto das mensagens anteriores oculto]

ANEXO B – FICHA DE NOTIFICAÇÃO DE FEBRE AMARELA

República Federativa do Brasil
Ministério da SaúdeSINAN
SISTEMA DE INFORMAÇÃO DE AGRAVOS DE NOTIFICAÇÃO
FICHA DE INVESTIGAÇÃO DE FEBRE AMARELA

N°

CASO SUSPEITO: Paciente com febre aguda (de até sete dias), de início súbito, com icterícia, procedente de área de risco para febre amarela ou de locais com ocorrência de epizootias em primatas não-humanos ou isolamento de vírus em vetores, nos últimos 15 dias, sem comprovação de ser vacinado contra febre amarela (apresentação do cartão de vacina).

Dados Gerais	1	Tipo de Notificação		2 - Individual				
	2	Agravado/enferma		FEBRE AMARELA				
	3	Código (CID-10)	A 9 5.9		8	Data da Notificação		
	4	UF	6	Município de Notificação	Código (IBGE)			
Notificação Individual	8	Unidade de Saúde (ou outra fonte notificadora)		Código				
	7	Data dos Primeiros Sintomas						
	8	Nome do Paciente		9		Data de Nascimento		
	10	(ou) Idade	1 - Hora 2 - Dia 3 - Mês 4 - Ano	11	Sexo M - Masculino F - Feminino I - Ignorado	12	Gestante 1 - 1º trimestre 2 - 2º trimestre 3 - 3º trimestre 4 - Idade gestacional ignorada 5 - Não 6 - Não se aplica 7 - Ignorado	13
14	Escolaridade 0 - Analfabeto 1 - 1ª a 4ª série incompleta do EF (antigo primário ou 1º grau) 2 - 4ª série completa do EF (antigo primário ou 1º grau) 3 - 5ª a 8ª série incompleta do EF (antigo ginásio ou 1º grau) 4 - Ensino fundamental completo (antigo ginásio ou 1º grau) 5 - Ensino médio incompleto (antigo colegial ou 2º grau) 6 - Ensino médio completo (antigo colegial ou 2º grau) 7 - Educação superior incompleta 8 - Educação superior completa 9 - Ignorado 10 - Não se aplica							
15	Número do Cartão SUS		18			Nome da mãe		
Dados de Residência	17	UF	18	Município de Residência	Código (IBGE)			
	19	Distrito						
	20	Bairro		21	Logradouro (rua, avenida, ...)			
	22	Número	23	Complemento (apto., casa, ...)		24	Geo campo 1	
	26	Geo campo 2		28	Ponto de Referência			
	27	CEP						
	28	(DDD) Telefone		29	Zona 1 - Urbana 2 - Rural 3 - Periurbana 9 - Ignorado			
	30	País (se residente fora do Brasil)						
Dados Complementares do Caso								
Ocupação	31	Data da Investigação		32			Ocupação	
	33	Informar os dados da investigação entomológica (mosquitos) e de epizootias 1 - Sim 2 - Não 9 - Ignorado <input type="checkbox"/> Ocorrência de Epizootias (Mortandade de macacos: conhecidos como guariba, bugio, saguis, micos, macaco aranha, macaco prego, gulgô, solm, etc.) <input type="checkbox"/> Isolamento de vírus em mosquitos <input type="checkbox"/> Presença de mosquito <i>Aedes aegypti</i> em área urbana (Observar período de viremia do paciente)						
Antecedentes Epidemiológicos	34	Vacinado Contra Febre Amarela		1 - Sim 2 - Não 9 - Ignorado	<input type="checkbox"/>	36	Caso Afirmativo, Data	
	37	Município		Código (IBGE)		38	Unidade de Saúde	
Dados Clínicos	38	Sinais e Sintomas		1 - Sim 2 - Não 9 - Ignorado	<input type="checkbox"/>	Sinais hemorrágicos (hematêmese, melena, epistaxe, gengivorragia, etc.)		
	<input type="checkbox"/>	Dor abdominal		<input type="checkbox"/>	Distúrbios de excreção renal (oligúria e/ou anúria)			
Atendimento	40	Ocorreu Hospitalização?		1 - Sim 2 - Não 9 - Ignorado	<input type="checkbox"/>	41	Data da Internação	
	43	Município		Código (IBGE)		44	Unidade de Saúde	
Dados do Laboratório	46	Exames Inespecíficos (anotar o maior valor encontrado, independente da data de coleta)						
	Bilirrubina Total _____ mg/dl		AST (TGO) _____ UI					
Bilirrubina Direta _____ mg/dl		ALT (TGP) _____ UI						

