

**Do Registro ao Insight: Modelagem Preditiva do Risco de
Osteonecrose dos Maxilares a Partir de Dados Clínicos
Odontológicos**

ANTONIO JOSÉ RIBEIRO DE CASTRO

Trabalho de Conclusão de Curso
MBA em Inteligência Artificial e Big Data

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

**Do Registro ao *Insight*: Modelagem
Preditiva do Risco de Osteonecrose
dos Maxilares a Partir de Dados
Clínicos Odontológicos**

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Antonio José Ribeiro de Castro

**Do Registro ao *Insight*: Modelagem Preditiva do
Risco de Osteonecrose dos Maxilares a Partir de
Dados Clínicos Odontológicos**

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial.

Orientadora: Profa. Dra. Alessandra Alaniz Macedo

USP - São Carlos

2025

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

C355r Castro, Antonio José Ribeiro de
Do Registro ao Insight: Modelagem Preditiva do
Risco de Osteonecrose a Partir de Dados Clínicos
Odontológicos / Antonio José Ribeiro de Castro;
orientador Alessandra Alaniz Macedo. -- São Carlos,
2025.
72 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2025.

1. Osteonecrose dos maxilares. 2. Bisfosfonatos.
3. Modelagem preditiva. 4. Aprendizado de máquina.
I. Macedo, Alessandra Alaniz , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Antonio José Ribeiro de Castro

Do Registro ao *Insight*: Modelagem Preditiva do Risco de Osteonecrose dos Maxilares a Partir de Dados Clínicos Odontológicos

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial.

Orientadora: Profa. Dra. Alessandra Alaniz Macedo

USP - São Carlos

2025

DEDICATÓRIA

Para meus pais Nair e Rogério.

Para minha esposa Maria Isabel, minha maior incentivadora.

Para meus filhos Maria Antonia e João Pedro.

AGRADECIMENTOS

À minha orientadora, Profa. Dra. Alessandra Alaniz Macedo, pelos ensinamentos, pela paciência, por me acolher no grupo InRedd e por reacender em mim a alegria de pesquisar, que eu acreditava estar adormecida.

Ao grupo do InRedd por aceitar um dentista computeiro.

À 4ª turma do MBA em IA e Big Data do ICMC-USP, uma turma nota 10! Apesar do curso online impossível maior troca, união e ajuda.

Ao prof. Dr. Leandro Dorigan de Macedo e sua equipe (Hilton Marcos Alves Ricz, Lara Maria Alencar Ramos Innocentini, Victor Augusto Minari e Mariana Andrade Costa) pela criação e seção do banco de dados utilizados neste trabalho.

Ao corpo docente do MBA em IA e Big Data pelo curso excelente.

Ao meu pai que em 1983, com sua visão de futuro, plantou a semente da paixão por tecnologia e pela computação, ao me presentear com um “computador” TK85 (onde aprendi programação em Basic).

Foram utilizadas ferramentas de inteligência artificial, especificamente os modelos chat GPT-5, Manus IA 1.5 e Gemini Pro como apoio à revisão linguística do texto e ao aprimoramento de trechos de código em linguagem Python. Todas as decisões conceituais, analíticas e redacionais permaneceram sob responsabilidade integral do autor

RESUMO

CASTRO, A.J.R. **Do Registro ao *Insight*: Modelagem Preditiva do risco de Osteonecrose dos Maxilares a Partir de Dados Clínicos Odontológicos** 2025. 73 f. Monografia (MBA em Ciências de Dados) – Centro de Ciências Matemáticas Aplicadas à Indústria, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2025.

A osteonecrose dos maxilares (ONM) representa uma complicação na clínica de odontologia relevante em pacientes que utilizam bisfosfonatos. Este trabalho teve como objetivo desenvolver e validar modelos preditivos para estimar o risco de ONM, aplicando técnicas de mineração de dados e modelagem preditiva a um conjunto de dados clínicos estruturados de 354 pacientes oncológicos provenientes do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da USP (HCFMRP-USP). O processo incluiu o pré-processamento dos dados clínicos, com tratamento de inconsistências, codificação de variáveis e normalização das informações, seguido tratamento do forte desbalanceamento da amostra (5,4% de prevalência de ONM) por meio da técnica SMOTE (Synthetic Minority Over-sampling Technique), Análise de Componentes Principais (PCA), e clusterização (K-Means). Modelos de Regressão Logística e Random Forest (RF) foram comparados, com foco na interpretabilidade através de ferramentas de IA Explicável (XAI), como SHAP. Os resultados da mineração de dados apontaram que a ausência de avaliação odontológica prévia, o número de doses acumuladas e o uso concomitante de corticoides e imunossupressores são fatores críticos associados à necrose óssea. O modelo Random Forest demonstrou desempenho superior, alcançando uma Área Sob a Curva (AUC-ROC) de 0,783 no conjunto de teste. O modelo final obteve alta especificidade (98,0%) e um valor preditivo negativo (VPN) de 96,1%. Embora a sensibilidade tenha sido limitada (33,3%), um resultado esperado em eventos de baixa prevalência, conclui-se que o modelo é uma ferramenta de grande utilidade para identificar pacientes de baixo risco com alta confiança, permitindo o direcionamento de recursos preventivos e apoiando a odontologia de precisão.

Palavras-chave: Osteonecrose dos Maxilares. Bisfosfonatos. Modelagem Preditiva.

Aprendizado de Máquina.

ABSTRACT

CASTRO, A.J.R. *From Records to Insights: Predictive Modeling of the risk of osteonecrosis of the jaws based on Clinical Dental Data*. 2025. 73 f. Monograph (MBA in Data Science) – Center for Applied Mathematical Sciences to Industry, Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, 2025.

Medication-Related Osteonecrosis of the Jaws (MRONJ) represents a relevant clinical complication in dentistry, particularly among patients using bisphosphonates. This study aimed to develop and validate predictive models to estimate the risk of MRONJ by applying data mining and predictive modeling techniques to a structured clinical dataset of 354 oncologic patients from the Hospital das Clínicas, Ribeirão Preto Medical School, University of São Paulo (HCFMRP-USP). The process included data preprocessing—addressing inconsistencies, encoding categorical variables, and normalizing clinical information—followed by handling the strong class imbalance (5.4% MRONJ prevalence) using the SMOTE (Synthetic Minority Over-sampling Technique) method, Principal Component Analysis (PCA), and clustering (K-Means). Logistic Regression and Random Forest (RF) models were compared, emphasizing interpretability through Explainable Artificial Intelligence (XAI) tools such as SHAP. Data mining results indicated that the absence of prior dental evaluation, higher cumulative doses, and concomitant use of corticosteroids and immunosuppressants are critical factors associated with bone necrosis. The Random Forest model showed superior performance, achieving an Area Under the Curve (AUC-ROC) of 0.783 on the test set. The final model reached high specificity (98.0%) and a negative predictive value (NPV) of 96.1%. Although sensitivity was limited (33.3%)—as expected for low-prevalence events—the model proved highly useful for reliably identifying low-risk patients, enabling targeted preventive strategies and supporting precision dentistry.

Keywords: Medication-Related Osteonecrosis of the Jaws. Bisphosphonates. Predictive Modeling. Machine Learning.

LISTA DE FIGURAS

Figura 1 — Fluxograma de trabalho	18.
Figura 2 — Distribuição da idade e relação entre idade e ocorrência dos pacientes.	21.
Figura 3 — Associação entre uso de medicamentos concomitantes e desenvolvimento de osteonecrose dos maxilares	22.
Figura 4 — Relação entre número de doses acumuladas de bisfosfonatos e ocorrência de osteonecrose dos maxilares.	24.
Figura 5 — Relação do tempo e da frequência de uso de bisfosfonatos na ocorrência de osteonecrose dos maxilares.	24.
Figura 6 — Matriz de correlação	26.
Figura 7 — Principais variáveis correlacionadas com a ocorrência de ONM	27.
Figura 8 — Variância explicada por cada componente principal.	29.
Figura 9 — Método do cotovelo para determinação do número ótimo de clusters	30.
Figura 10 — Comparação das características numéricas dos clusters 0 (alto risco) e 1 (baixo risco).	32.
Figura 11 — Visualização dos clusters nos dois primeiros componentes principais. ...	33.
Figura 12 — Comparação entre quatro métodos de análise (correlações, Random Forest, SHAP e ranking consolidado) mostrando as 10 variáveis mais importantes.	34.
Figura 13 — Gráfico de barras agrupadas comparando a importância das 10 principais variáveis segundo Random Forest, SHAP e correlações.	37.
Figura 14 — Comparação de modelos usando a validação cruzada (boxplots do AUC-ROC)	41.
Figura 15 — Curva ROC do modelo Random Forest	42.
Figura 16 — Matriz de confusão detalhada	43.
Figura 17 — Importância relativa das variáveis no modelo Random Forest)	44.

LISTA DE TABELAS

Tabela 1 — Os valores de Inércia e do Coeficiente de Silhueta médio, calculados para cada valor de K.....	30
Tabela 2 — Detalhamento das características dos Clusters 0 (alto risco de ONM) e Cluster 1 (baixo risco)	32
Tabela 3 — Top 10 Correlações com ONM	35
Tabela 4 — Resultados dos testes estatísticos, confirmando os achados da análise das correlações com osteonecrose dos maxilares	36
Tabela 5 — Ranking consolidado das variáveis TOP 10.....	38

LISTA DE ABREVIATURAS E SIGLAS

ANMWS: Variação da técnica SMOTE

AUC: *Area Under the Curve* (Área Sob a Curva)

AUROC: *Area Under the Receiver Operating Characteristic curve*

BDA: *Big Data Analytics*

CAAE: Certificado de Apresentação para Apreciação Ética

CEP: Comitê de Ética em Pesquisa

CNN: Redes Neurais Convolucionais

DMFT: Índice de dentes cariados, perdidos e obturados

FORP: Faculdade de Odontologia de Ribeirão Preto

GBM: *Gradient Boosting Machine*

GBDT: *Gradient Boosting Decision Tree*

GIGO: *Garbage In, Garbage Out*

HCFMRP-USP: Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da USP

IA: Inteligência Artificial

ICA: Análise de Componentes Independentes

ICMC: Instituto de Ciências Matemáticas e de Computação

INREDD: *Interdisciplinary Research Group in Digital Dentistry*

KDD: *Knowledge Discovery in Databases* (Descoberta de Conhecimento em Bases de Dados)

LGPD: Lei Geral de Proteção de Dados

LIME: *Local Interpretable Model-agnostic Explanations*

LLMs: *Large Language Models* (Grandes Modelos de Linguagem)

MBA: *Master of Business Administration*

ML: *Machine Learning* (Aprendizado de Máquina)

NLP: *Natural Language Processing* (Processamento de Linguagem Natural)

ONM: Osteonecrose dos Maxilares

PCA: Análise de Componentes Principais

RF: *Random Forest*

RL: Regressão Logística

SHAP: *SHapley Additive exPlanations*

SMOTE: *Synthetic Minority Oversampling Technique*

SUS: Sistema Único de Saúde

SVM: *Support Vector Machines* (Máquinas de Vetor de Suporte)

TCC: Trabalho de Conclusão de Curso

TCLE: Termo de Consentimento Livre e Esclarecido

TIC: Tecnologias de Informação e Comunicação

USP: Universidade de São Paulo

WCSS: *Within-Cluster Sum of Squares* (Soma dos Quadrados Intra-Cluster)

XAI: *Explainable Artificial Intelligence* (Inteligência Artificial Explicável)

XGBoost: *Extreme Gradient Boosting*

SUMÁRIO

1. Introdução	1
2. Fundamentação Teórica	4
2.1. Mineração de dados e análise preditiva	4
2.2. Mineração de dados e análise preditiva na saúde	5
2.3. Bisfosfonatos e osteonecrose dos maxilares	7
2.4. Dados desbalanceados	8
3. Trabalhos Relacionados	10
4. Desenvolvimento	13
4.1. Materiais	13
4.1.1. Conjunto de dados	13
4.2. Métodos	15
4.2.1. Ambiente computacional	15
4.2.2. Fluxo de trabalho	15
4.2.3. Tratamento do desbalanceamento	16
4.2.4. Avaliação e validação	17
4.2.5. Algoritmo	18
5. Resultados e Discussão	20
5.1. Características da população estudada	20
5.2. Fatores associados à osteonecrose dos maxilares	21
5.2.1. Avaliação odontológica prévia	21
5.2.2. uso de medicamentos concomitantes	22
5.2.3. Quimioterapia prévia e indicação do uso	23
5.2.4. Características do tratamento com bisfosfonato	23
5.3. Mineração de dados	25
5.3.1. Análise de correlação	25
5.3.2. Análise de componentes principais (PCA)	28
5.3.3. Análise de clusters	29
5.4. Análise preditiva	39
5.4.1. Seleção de variáveis e preparação dos dados	39
5.4.2. Random Forest	41

5.4.3. Interpretabilidade e Importância das Variáveis	43
5.5. Performance do modelo	45
5.6. Interpretação clínica dos principais resultados	45
5.7. Contribuições da inteligência artificial na saúde	46
6. Conclusão	47
6.1. Limitações e considerações	47
6.2. Trabalhos futuros	48
7. Referências Bibliográficas	49

1 INTRODUÇÃO

A Inteligência Artificial (IA) tem revolucionado a maneira como os diversos setores e áreas organizam e executam suas atividades; e a área da saúde não é exceção (Dash; Shakyawar; Sharma, 2019). O uso de Tecnologias de Informação e Comunicação (TIC) tem transformado a organização, a oferta de serviços de saúde e as políticas públicas. No Brasil, a estratégia de Saúde Digital, conhecida como digiSUS, busca integrar essas tecnologias ao Sistema Único de Saúde (SUS), estabelecendo a digitalização como dimensão fundamental para promover maior eficiência e transparência nos processos de saúde, ofertando plataformas digitais para gestores e cidadãos, (Futuro da Saúde, 2025).

Na área da Saúde, encontramos vários tipos de dados além dos dados clínicos compostos de histórias médicas, de resultados de exames e de anotações de consultas. Contudo, eles são apenas uma parte da informação que tenta descrever a saúde de um paciente. A saúde combina informações administrativas, laboratoriais, de imagem, genéticas e comportamentais, que buscam formar uma visão total da saúde de um paciente para auxiliar os profissionais na tomada de decisões, no diagnóstico e tratamento, bem como para monitorar o estado de saúde do indivíduo e melhorar o atendimento médico. A IA se destaca pela capacidade de processar e analisar grandes volumes de dados, identificando padrões e oferecendo informações, reduzindo erros e melhorando a qualidade do atendimento (SADIKU; EZE; MUSA, 2018; EKWONWUNE; UBOCHI; DUROHA, 2022). Sistemas computacionais, que manipulam dados clínicos de pacientes, exames de imagem, dados genômicos, dispositivos vestíveis, sensores biomédicos e plataformas de telemedicina, geram diariamente grandes volumes de dados heterogêneos, complexos e dinâmicos.

O termo *Big Data* designa o vasto volume de informações estruturadas e não estruturadas, de difícil gerenciamento por métodos convencionais, mas que, se adequadamente tratados, permitem descobrir correlações e gerar *insights* com aplicações clínicas, administrativas e epidemiológicas (DUCLOS, 2023; KUMARI; CHAKRABORTY, 2023). A mineração de dados e a análise preditiva constituem pilares centrais para o entendimento do real significado dos dados: a primeira voltada à identificação de padrões e associações relevantes, e a segunda dedicada à construção de modelos capazes de prever eventos futuros com base em informações históricas.

Na Odontologia Digital, técnicas de mineração de dados e análise preditiva já vêm sendo aplicadas em diferentes áreas, como a análise de imagens dentárias por redes neurais convolucionais (CNN), a modelagem preditiva de risco para doenças periodontais, a estratificação de pacientes para atendimento personalizado e até a integração de redes complexas no campo da *Network Medicine* (Lee et al., 2019; Ricci, 2018; Gili et al., 2021; Schwendicke; Krois, 2021). Além de aprimorar diagnósticos e tratamentos, esses recursos possibilitam melhor gestão de materiais e de recursos em clínicas e hospitais, aumentando a eficiência e a personalização no cuidado em saúde bucal (Herbst et al., 2022).

Entre os softwares de gestão hospitalar para registro clínico de pacientes, destaca-se o banco de dados do ROMEU, desenvolvido e utilizado pela Faculdade de Odontologia de Ribeirão Preto (FORP) da USP há mais de duas décadas. Esse sistema acumula um grande volume de informações referentes a agendamentos, prontuários e procedimentos odontológicos, configurando uma base de dados clínicos singular para investigação científica. Entre as condições clínicas registradas, uma que se pode destacar é o acompanhamento de pacientes em uso de bisfosfonatos, fármacos amplamente empregados no tratamento da osteoporose e de neoplasias ósseas.

A osteonecrose dos maxilares (ONM) associada ao uso de bisfosfonatos representa uma complicação clínica relevante, caracterizada por áreas de osso exposto de difícil cicatrização, dor, infecção e comprometimento funcional. Embora sua prevalência varie conforme a população estudada, relatos indicam taxas entre 1% e 10% em pacientes oncológicos tratados com bisfosfonatos intravenosos, e menores que 1% em pacientes osteoporóticos em uso de formulações orais (Caglayan et al., 2009). A gravidade dessa condição, somada à dificuldade terapêutica, reforça a importância de estratégias preditivas que permitam identificar precocemente pacientes sob maior risco, orientando condutas preventivas mais eficazes. Seguindo essa ideia de identificar pacientes com maior risco de ONM, o estudo de Minari (2022), por meio de métodos estatísticos clássicos, evidenciou que comorbidades metabólicas e o tempo/frequência do medicamento aumentam o risco de ONM, enquanto a avaliação odontológica prévia tem papel crucial na prevenção. Nesse contexto, formula-se a seguinte hipótese: a mineração de dados e a análise preditiva, com suas diferentes metodologias, aplicadas a esse grupo específico de pacientes, podem auxiliar na identificação precoce de fatores de risco que a análise estatística convencional não consegue detectar, favorecendo estratégias preventivas, com impacto direto na prática clínica e na saúde pública.

Diante desta hipótese, este trabalho aplicou técnicas de mineração de dados e modelagem preditiva ao mesmo banco de dados do estudo de MINARI (2022). O objetivo foi

transformar registros clínicos acumulados em *insights* de valor clínico e epidemiológico, que poderiam ter passado despercebidos pela análise estatística padrão.

A organização do restante do presente estudo ocorre da seguinte maneira: o Capítulo 2 apresenta a fundamentação teórica que apoia e sustenta o problema e as escolhas metodológicas realizadas; o Capítulo 3 apresenta os trabalhos relacionados ao tema proposto; o Capítulo 4 descreve o trabalho desenvolvimento em termos de metodologia utilizada para alcançar o objetivo proposto; o Capítulo 5 cita e discute os resultados; e finalmente o Capítulo 6 abrange conclusões e propostas de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica é a base científica deste trabalho, reunindo os conceitos e as evidências da literatura sobre o tema estudado. Esse capítulo apresenta conhecimentos sobre a ONM, lacunas de conhecimento e apoios para a escolha dos métodos utilizados.

2.1 Mineração de Dados e Análise Preditiva

A mineração de dados (*data mining*) consolidou-se como um dos principais ramos da ciência da computação dedicado a extrair conhecimento útil de grandes volumes de dados. Han, Kamber e Pei (2012) destacam que a mineração de dados é a etapa essencial no processo de *Knowledge Discovery in Databases* (KDD), que envolve desde a limpeza e integração dos dados até a seleção de variáveis relevantes, aplicação de algoritmos e interpretação dos resultados. Para os autores, não se trata apenas da execução mecânica de algoritmos, mas de um processo sistemático voltado à descoberta de padrões não triviais, implícitos e potencialmente úteis, com valor científico ou aplicado.

A mineração de dados pode ser dividida em duas abordagens principais: a descritiva, voltada a resumir propriedades dos dados existentes, e a preditiva, direcionada à antecipação de eventos futuros com base em inferências (Patil et al., 2020). Koul (2020) ressalta que o objetivo central desse processo é identificar tendências ocultas que orientem a tomada de decisões estratégicas, independente da área de interesse dos dados. Ogunleye (2022) complementa ao mostrar que tais técnicas já se encontram disseminadas em setores como marketing, segurança cibernética e saúde.

Uma revisão sistemática de NTI et al. (2022) reforça essa percepção, mostrando que os algoritmos mais utilizados em cenários de *Big Data Analytics* (BDA) são as redes neurais profundas (15%), máquinas de vetor de suporte (15%), redes neurais artificiais (14%) e árvores de decisão (12%). Destacam-se as plataformas Apache Spark (34,9%) e Hadoop (30,2%) como ferramentas mais adotadas. O estudo indica que aproximadamente 30% das aplicações de mineração de dados em BDA concentram-se em saúde, o que demonstra a relevância e o potencial do tema para pesquisas clínicas e na odontologia.

Monteiro et al. (2020) aplicaram técnicas de aprendizado de máquina em dados do *PhysioNet Challenge 2012*, com o objetivo de prever mortalidade em UTIs. Os autores utilizaram métodos de pré-processamento como normalização, tratamento de séries temporais e redução de dimensionalidade por PCA e análise fatorial. Os resultados mostraram que o SVM linear alcançou desempenho superior, com AUC de 0,99, usando apenas 12 variáveis críticas em comparação aos 41 atributos originais. Esse estudo exemplifica a importância de selecionar, organizar e preparar dados de forma criteriosa antes de usá-los e da seleção de atributos para garantir modelos preditivos viáveis e clinicamente úteis.

A aplicação de modelos de aprendizado de máquina na predição de riscos em odontologia tem se expandido para diferentes contextos clínicos. Uma revisão sistemática recente realizada por Uppal et al. (2024) analisou 15 estudos sobre o uso de algoritmos de machine learning para estimar a transformação maligna de distúrbios orais potencialmente malignos. Os autores destacam que técnicas como Random Forest, Support Vector Machines, Gradient Boosting, Redes Neurais Artificiais e PCA obtiveram desempenhos relevantes, com índices de acurácia variando entre 74% e 95,4%, e alguns modelos alcançando sensibilidade e especificidade de até 100% em determinados cenários.

2.2 Mineração de Dados e Análise Preditiva na Área da Saúde

A aplicação da inteligência artificial (IA) em saúde tem crescido de forma exponencial, com impactos na detecção precoce de doenças, no prognóstico clínico e na personalização terapêutica. Entretanto, a adoção de modelos de aprendizado de máquina em ambientes clínicos depende de critérios adicionais, como a interpretabilidade e a transparência dos resultados. Em 2016, foi lançado o programa *Explainable Artificial Intelligence (XAI)*, buscando desenvolver algoritmos capazes de fornecer explicações compreensíveis para suas decisões. Adadi e Berrada (2018) contribuíram para o amadurecimento da área ao estruturar uma classificação dos métodos de explicabilidade disponíveis e ao destacar os principais desafios que ainda limitam seu avanço.

Bharati, Mondal e Podder (2024) ampliaram essa perspectiva ao revisar de forma sistemática o uso de XAI em saúde. Para os autores, a acurácia de um modelo é insuficiente se não for

acompanhada da possibilidade de compreender quais variáveis sustentaram a decisão. Técnicas como SHAP, LIME e DeepLIFT permitem identificar os fatores mais relevantes para a predição, enquanto métodos como PCA e ICA (Análise de Componentes Independentes) contribuem para reduzir a dimensionalidade sem comprometer a interpretabilidade.

O princípio conhecido como *Garbage In, Garbage Out* (GIGO) ilustra a importância da qualidade dos dados. Nesse contexto, Costa et al. (2024) destacam o papel crítico da rotulagem precisa e padronizada na criação de bases odontológicas para pesquisa em IA. Os autores desenvolveram um conjunto de 936 radiografias panorâmicas rotuladas por radiologistas, com taxa de discordância inferior a 1%, demonstrando que a consistência das anotações é determinante para o desempenho e a aplicabilidade clínica dos algoritmos. Embora voltado para imagens radiográficas, o princípio é análogo ao que se observa em bases de dados clínicas: sem padronização, curadoria e rastreabilidade, qualquer modelo preditivo se torna vulnerável a vieses e interpretações incorretas. Assim, a etapa de pré-processamento e validação dos dados do presente estudo segue a mesma lógica de rigor metodológico proposta por Costa et al. (2024), buscando assegurar que as predições encontradas reflitam padrões clínicos reais e úteis à prática odontológica.

O campo odontológico tem acompanhado essa evolução. Em análise bibliométrica, Xie et al. (2024) identificaram 651 artigos publicados entre 2000 e 2023 sobre IA em Odontologia, com predominância dos Estados Unidos e da China. As principais aplicações concentram-se em análise de imagens (cáries, fraturas e periodontite), apoio à decisão clínica e implantodontia. Os autores reforçam que a odontologia de precisão já é uma realidade, com forte apoio de algoritmos de *deep learning*.

Avanços metodológicos também são evidentes. Chuang et al. (2023) utilizaram *Large Language Models* (LLMs), como o GPT-4, para extrair informações diagnósticas de notas clínicas odontológicas não estruturadas. Combinando esse recurso a um modelo RoBERTa treinado especificamente em linguagem odontológica, obtiveram acurácia elevada (F1-score $\approx 0,98$), mostrando como técnicas de NLP (*Natural Language Processing* ou Processamento de Linguagem Natural) podem enriquecer os bancos de dados clínicos disponíveis para mineração. Shamout, Zhu e Clifton (2021) reforçam que a qualidade e a organização dos dados são determinantes para a performance dos modelos preditivos em saúde, exigindo padronização e validação externa.

Outros estudos demonstram avanços em diferentes frentes. Como, por exemplo, Mukherjee et al. (2023) desenvolveram o modelo SCOPE, baseado em *Random Forests*, para prever diagnósticos futuros em consultas ambulatoriais, obtendo AUROC de 0,904 em mais de 500 mil pacientes. Pezoulas et al. (2024) destacam a geração de dados sintéticos, via GANs, como estratégia para contornar escassez de amostras em doenças raras. Dash et al. (2019) destacam a importância do Big Data em saúde, que reúne informações de diferentes fontes, como prontuários eletrônicos, exames de imagem, dados biológicos em larga escala (ômicos) e medições feitas por sensores de IoT. Lekadir et al. (2025), por sua vez, estabeleceram diretrizes internacionais (FUTURE-AI) que reforçam princípios de confiabilidade, rastreabilidade e explicabilidade, fundamentais para a consolidação da IA na prática clínica.

2.3 Bisfosfonatos e Osteonecrose dos Maxilares (ONM)

No campo clínico, os bisfosfonatos são amplamente prescritos para osteoporose e doenças metastáticas, mas seu uso prolongado pode ocasionar a osteonecrose dos maxilares. Wei et al. (2024) utilizaram bases de farmacovigilância dos EUA (FAERS) e do Canadá (CVAR) para desenvolver um modelo preditivo de ONM em pacientes tratados com ácido zoledrônico. Por meio de regressão logística com seleção de variáveis via LASSO, identificaram como fatores críticos a presença de metástase óssea, o diagnóstico de osteoporose, a polifarmácia e a dosagem administrada. O modelo apresentou AUC de até 0,899 em validação externa, reforçando sua robustez clínica.

De forma complementar, Kwack e Park (2023) aplicaram a plataforma H2O-AutoML para prever ONM em mulheres com osteoporose submetidas a procedimentos odontológicos. O algoritmo Gradient Boosting Machine (GBM) apresentou os melhores resultados, com AUC de 0,828 em treinamento e 0,753 em validação, tendo como preditores principais o tempo de uso da medicação, a idade e a região cirúrgica. Esses achados reforçam que, mesmo com dados clínicos relativamente simples, é possível desenvolver modelos de apoio à decisão que auxiliem na identificação precoce de pacientes em risco.

Esses estudos dialogam diretamente com a proposta deste trabalho, que visa explorar registros clínicos do sistema ROMEU da FORP-USP para construir modelos preditivos de risco de ONM em pacientes expostos a bisfosfonatos.

2.4 Dados Desbalanceados

Bases clínicas apresentam com frequência problemas de desbalanceamento entre classes, já que os eventos de interesse (como a ONM) são relativamente raros. Essa condição compromete os modelos, que tendem a privilegiar a classe majoritária.

Na tentativa de resolver esse problema, em 2002 Chawla et al. propôs a técnica de SMOTE (Synthetic Minority Oversampling Technique), onde ao invés de simplesmente fazer um oversampling (duplicar exemplos raros), o modelo gera novos exemplos sintéticos. Assim, para cada instância de classe minoritária o SMOTE seleciona os vizinhos mais próximos(geralmente usando $k=5$), cria um novo exemplo interpolando as características encontradas entre o original e os vizinhos, e vai repetindo o processo até equilibrar melhor as classes.

Obviamente que a criação de exemplos sintéticos apesar de útil, tem problemas como o *overfitting* (quando um modelo de aprendizado de máquina aprende tão bem os detalhes e ruídos do conjunto de treino que perde a capacidade de generalizar para novos dados.), ruído amplificado (quando a classe minoritária contém exemplos errados ou ruídos, o SMOTE pode criar novos ruídos sintéticos) e dados irreais (criação de dados fora da realidade ou contexto), Assim, é natural que ao longo do tempo tenham sido propostas variações da técnica como Wibowo, Masruriyah e Rahmawati (2025) que propuseram uma variação do SMOTE chamada ANMWS, orientada por julgamento clínico de especialistas, demonstrando aumento da acurácia e da AUC em modelos para diagnóstico de diabetes. Alsabry et al. (2023) avaliaram 21 algoritmos em um dataset desbalanceado de câncer de mama e mostraram que o uso de SMOTE elevou a acurácia do modelo *Optimized LogitBoost* para 88%. Cicak e Avci (2023), em cenário de manutenção preditiva, mostraram que métodos de *oversampling* e *undersampling* podem aumentar substancialmente o *recall* em desbalanceamentos moderados, embora falhem quando há raridade extrema.

Neste estudo, o uso do SMOTE foi essencial para compensar o forte desbalanceamento dos dados, já que os poucos casos de osteonecrose poderiam prejudicar o desempenho dos modelos. No entanto, o SMOTE não foi aplicado de forma mecânica. Os dados sintéticos foram avaliados com cuidado e rigor, pois eles impactam a interpretação clínica. Gholampour (2024), por exemplo, demonstrou que a geração de pacientes sintéticos pode criar distorções relevantes; em sua análise, a ampliação de 20 casos originais para 62 artificiais resultou em padrões pouco consistentes nas análises de PCA e até em contradições nos parâmetros apontados por métodos

interpretativos como o SHAP. Ou seja, ganhos estatísticos não se traduzem necessariamente em confiabilidade clínica, o que exige interpretação cuidadosa.

O problema do desbalanceamento em dados médicos, entretanto, não é novo. Em uma revisão sistemática recente, Salmi et al. (2024) mostraram que a predominância de classes majoritárias tende a ocultar justamente os desfechos mais relevantes, ou seja, os de baixa frequência. Esses autores organizaram as estratégias de enfrentamento em três níveis: no pré-processamento (resampling, geração de dados sintéticos, seleção de atributos), no próprio algoritmo (modelos sensíveis ao custo e variantes adaptadas) e em soluções que combinam as duas anteriores. Assim, a opção por PCA, Random Forest e SMOTE neste trabalho acompanha práticas aceitas em mineração de dados aplicada à saúde.

A literatura recente também mostra exemplos de uso dessas abordagens em cenários distintos. Yang et al. (2025) utilizaram Random Forest, Gradient Boosting e LightGBM para prever demanda de serviços médicos em idosos, lançando mão do SMOTE-TomekLink para balanceamento, e obtiveram AUC acima de 0,90. Já Kivrak et al. (2024), ao trabalhar com a predição de deficiência de testosterona em diabéticos tipo 2, observaram melhora da sensibilidade com SMOTE, mas ao custo de queda na especificidade. Nesses casos, a combinação com algoritmos ensemble (XGBoost, AdaBoost) se mostrou particularmente útil.

O presente estudo seguiu padrão semelhante: o SMOTE elevou a sensibilidade, mas reduziu a especificidade em alguns modelos. Esse resultado reforça a necessidade de leitura crítica: embora o ganho estatístico seja real, ele não garante impacto direto na prática clínica. Portanto, os achados devem ser vistos como um passo dentro de um processo exploratório maior, onde técnicas de balanceamento e *ensemble learning* se mostram promissoras, desde que aplicadas com plena consciência de suas limitações.

3 TRABALHOS RELACIONADOS

Nos últimos anos, diferentes estudos têm avaliado o uso de *machine learning* (ML) na odontologia, em especial para a detecção e previsão de doenças de alta prevalência, como cárie e doença periodontal. Um exemplo é o trabalho de Patel et al. (2022), os pesquisadores utilizaram mais de 18 mil prontuários odontológicos eletrônicos e se propuseram a desenvolver um modelo de predição da doença periodontal. O estudo identificou não apenas fatores clínicos, mas também fatores sociais e hábitos de vida como fatores relevantes, destacando o potencial do ML em revelar associações pouco consideradas na prática clínica. O desempenho moderado do modelo (AUC em torno de 0,72) demonstrou tanto as possibilidades quanto os desafios de aplicar esses métodos em larga escala.

Çiftçi e Aşantoğrol (2024) em um estudo para identificar fatores de risco ligados à saúde bucal, testaram diferentes algoritmos de ML na predição do risco de cárie com base no índice DMFT (dentes cariados, perdidos e obturados) em uma amostra de 2.000 adultos, usando dados coletados por questionários e exames clínicos. Os autores mostraram que redes neurais multicamadas alcançaram desempenho de destaque, com acurácia superior a 95%, superando métodos tradicionais como regressão logística. Além de confirmar variáveis esperadas, como frequência de escovação e consumo de açúcar, o estudo também evidenciou a influência de fatores socioeconômicos e comorbidades médicas, reforçando a multidimensionalidade do risco de cárie.

Além disso, Kang et al. (2023) demonstraram que a qualidade dos atributos utilizados no treinamento é determinante para a performance dos modelos de ML e análise preditiva. Com base em mais de 22 mil crianças coreanas, o estudo propôs um modelo preditivo de cárie que combinou seleção ótima de atributos (mRMR + GINI) com o classificador GBDT. A estratégia permitiu reduzir o conjunto de variáveis de 43 para apenas 18, alcançando acurácia superior a 95% e precisão próxima de 100%. Essa abordagem mostrou que não basta treinar algoritmos complexos: é igualmente essencial selecionar as variáveis mais informativas, o que impacta diretamente na eficiência diagnóstica e no custo operacional.

Uehara, Tirapelli e Macedo (2024) apresentaram a ferramenta Dental Second Look AI, desenvolvida no âmbito do grupo INREDD (Interdisciplinary Research Group in Digital Dentistry) da Universidade de São Paulo, voltada à análise automatizada de radiografias

panorâmicas odontológicas. A solução integra modelos hierárquicos de redes neurais convolucionais, capazes de detectar estruturas bucais, segmentar dentes e classificar lesões cariosas, oferecendo suporte para o diagnóstico clínico e para a obtenção de dados epidemiológicos. Assim como o presente trabalho, os autores usaram a mesma lógica de mineração de dados e análise preditiva de dados odontológicos como apoio à decisão clínica e geração de conhecimento científico, reforçando o potencial da IA na Odontologia, seja na gestão ou na prática clínica

Uma revisão conduzida por Lakhotia et al. (2025) sintetizou mais de 280 estudos que aplicaram ML em diversas áreas da odontologia. Os autores apontaram avanços expressivos, principalmente em radiologia oral, mas também destacaram falhas metodológicas recorrentes, como a baixa atenção a métricas de calibração, à discussão sobre vieses e à reprodutibilidade. A recomendação central do trabalho foi a necessidade de maior padronização e transparência, para que os modelos possam ser realmente incorporados em cenários clínicos.

Avanços recentes em aprendizado profundo têm ampliado o alcance da inteligência artificial também na análise de imagens odontológicas. Em estudo conduzido na Universidade de São Paulo, Carneiro et al. (2025) propuseram um sistema modular baseado em *Deep Learning* capaz de identificar automaticamente as regiões dentomaxilares e segmentar cada dente individualmente em radiografias panorâmicas. O modelo apresentou precisão superior a 90%, superando métodos anteriores de segmentação. Essa abordagem representa um passo importante para a integração entre dados clínicos e variáveis morfológicas obtidas por imagem, possibilitando, no futuro, modelos preditivos mais completos e personalizados dentro da odontologia digital.

Estudos recentes (Wei et al., 2024; Kwack; Park, 2023) demonstram que modelos preditivos baseados em dados clínicos podem identificar fatores de risco importantes e auxiliar na tomada de decisão odontológica. No entanto, a literatura aponta limitações recorrentes, como:

- a escassez de bases de dados estruturadas em Odontologia (Chuang et al., 2023),
- a necessidade de abordagens mais robustas de pré-processamento e seleção de atributos (Monteiro et al., 2020),
- e o problema do desbalanceamento das classes, comum em eventos clínicos raros (Wibowo et al., 2025; Alsabry et al., 2023).

A análise dos estudos até aqui apresentados demonstra a importância e principalmente o potencial das aplicações de IA na área da saúde e em especial na odontologia. As limitações que naturalmente existem vão sendo aos poucos solucionadas, mas sempre com potencial de mais ajustes e melhorias. O próximo capítulo apresenta, portanto, a metodologia utilizada para alcançar os objetivos propostos. Nele são descritas as etapas de preparação e padronização dos dados, as estratégias para minimizar o desafio do desbalanceamento das amostras, os critérios de seleção das variáveis e os modelos preditivos empregados. Cada uma dessas etapas foi planejada para responder às fragilidades identificadas na literatura e chegar a um resultado mais confiável entre o dado clínico e o insight preditivo.

4 DESENVOLVIMENTO

O presente trabalho teve como objetivo desenvolver e validar modelos preditivos para estimar o risco de osteonecrose dos maxilares em pacientes que utilizam bisfosfonatos. Os dados de saúde dos pacientes, incluindo as indicações de dosagem e de tempo de uso dos bisfosfonatos, são dados clínicos estruturados, epidemiológicos e fatores clínicos de risco para o desenvolvimento de ONM, extraídos a partir do trabalho de Minari (2022) com CEP local (CAAE:96418618.4.000.5440, Parecer:3.248.605).

O trabalho a ser descrito neste capítulo diferencia-se por:

- Valorizar um banco de dados institucional de odontologia, ainda pouco explorado para análises avançadas por ferramentas computacionais;
- Aplicar técnicas de tratamento de dados que permitem lidar com a baixa incidência da doença (desbalanceamento de dados) e aumentar a confiabilidade dos resultados;
- Utilizar ferramentas de inteligência artificial explicável (XAI), que oferecem maior clareza sobre como as previsões são geradas, facilitando a adoção dos modelos na prática clínica;
- Gerar evidências aplicáveis à odontologia de precisão, apoiando a detecção precoce de fatores de risco e a personalização da conduta clínica, com benefícios tanto para profissionais quanto para gestores da saúde.

A seguir é detalhada a metodologia seguida no desenvolvimento do trabalho, descrevendo os materiais utilizados e os métodos empregados.

4.1 Materiais

4.1.1 O Conjunto de Dados

Esse trabalho manipulou uma amostra de 354 pacientes oncológicos em uso de ácido zoledrônico intravenoso, acompanhados pelo ambulatório de odontologia hospitalar. Os dados foram obtidos, a partir de prontuários eletrônicos do Hospital das Clínicas da Faculdade de

Medicina de Ribeirão Preto da USP (HCFMRP-USP) de Ribeirão Preto. Trata-se de um estudo baseado na revisão de registros anteriores de pacientes, desenvolvido no próprio hospital após aprovação do CEP local (CAAE:96418618.4.000.5440, Parecer:3.248.605). Em função do perfil retrospectivo e de observação, o estudo obteve dispensa do TCLE.

É a mesma base utilizada nos estudos de Minari (2022) e Minari et al. (2025, no prelo), que avaliaram fatores clínicos de risco por métodos estatísticos convencionais. Os dados foram atualizados por meio de entrevistas com os pacientes (como no número de doses, por exemplo), por Alamino (2025) que associou os dados com achados radiográficos e hematológicos identificando fatores de risco e biomarcadores associados à osteonecrose dos maxilares.

O presente trabalho, contudo, aplicou técnicas de mineração de dados e modelos preditivos de inteligência artificial que não foram utilizados nas análises dos estudos anteriores de Minari (2022) e Alamino (2025).

Todas as informações manipuladas foram previamente anonimizadas pelos estudos anteriores, em conformidade com a Lei Geral de Proteção de Dados (LGPD), garantindo sigilo e confidencialidade.

Os dados originais estavam dispostos em uma tabela com os dados já anonimizados de 354 pacientes, distribuídos em 29 colunas. Contudo, muitas colunas estavam vazias ou com menos de 30% dos dados preenchidos e foram cortadas, uma vez que variáveis com grande quantidade de dados ausentes comprometem a confiabilidade das análises estatísticas, podendo ocasionar desvios e erros nos resultados, e consequentemente, comprometendo o modelo preditivo.

Após a eliminação dessas colunas, os dados encontrados foram:

- idade e sexo;
- a doença base: aquela que originou a indicação do uso do bisfosfonato,
- as comorbidades ou doenças pré existentes;
- tabagismo (sim/não);
- avaliação odontológica prévia (sim/não);
- uso crônico de corticosteróides, e imunossupressores e de antiangiogênicos: o uso de imunossupressores é comum em transplantados e o uso de antiangiogênicos diminui a

abastecimento de sangue em tecidos ou tumores que dependem da criação de novos vasos;

- indicação (tratamento ou prevenção);
- tempo de uso (0 a 5 anos);
- frequência de uso (anual, semestral, trimestral, mensal e uma dose única);
- números de doses acumuladas desde o início do tratamento;
- associação com a medicação via oral, uma vez que no grupo avaliado o medicamento foi ministrado de forma injetável); e
- o desfecho (ONM, sim ou não).

4.2 Métodos

4.2.1 O Ambiente Computacional

As análises foram realizadas em ambiente Python (versão 3.11.0), utilizando bibliotecas amplamente empregadas em ciência de dados, como *pandas*, *numpy*, *scikit-learn* e *imbalanced-learn*. Para visualização, foram utilizadas as seguintes bibliotecas do Python: *matplotlib* e *seaborn*. Todos os experimentos foram executados em ambiente controlado, com semente aleatória fixada e versionamento de código, garantindo a reprodutibilidade dos resultados.

4.2.2 O Fluxo de Trabalho

O fluxo metodológico seguiu as seguintes cinco etapas principais, conforme ilustrado na Figura 1:

1. Preparação e pré-processamento dos dados – inspeção de valores ausentes ou inconsistentes, padronização de variáveis numéricas e codificação das variáveis categóricas. A variável-alvo (“ocorrência de ONM”) foi convertida em formato binário.
2. Análise estatística descritiva – cálculo de frequências, medidas de tendência central e dispersão, além de testes bivariados para explorar associações preliminares.
3. Mineração de dados – aplicação de testes estatísticos (Qui-quadrado, ANOVA), cálculo de *odds ratio*, redução de dimensionalidade por PCA e identificação de padrões por agrupamento (K-Means).
4. Modelagem preditiva – construção de dois modelos de classificação:
 - Random Forest (RF): escolhido por sua capacidade de lidar com variáveis mistas e oferecer medidas de importância das variáveis.
 - Regressão Logística (RL): incluída como modelo de referência (*baseline*), pela ampla utilização em pesquisas clínicas, simplicidade de implementação e interpretabilidade dos coeficientes.
5. Avaliação comparativa dos modelos – baseada em métricas de desempenho (AUC, acurácia, precisão, sensibilidade/recall e F1-score), análise de sobreajuste (*overfitting*) e grau de interpretabilidade.

4.2.3 O Tratamento do Desbalanceamento

A variável-alvo (ONM) apresentou forte desbalanceamento (19 casos positivos em 354 registros, $\approx 5,4\%$). Para reduzir esse efeito, foram aplicadas duas estratégias complementares:

- *SMOTE (Synthetic Minority Over-sampling Technique)*: geração de amostras sintéticas da classe minoritária.
- *Class weights*: atribuição de maior peso às instâncias positivas durante o treinamento de modelos.

O *SMOTE* padrão foi utilizado por meio da classe *SMOTE* da biblioteca *imbalanced-learn*, sem especificação de variantes como *SMOTENC*, *SMOTEN*, *BorderlineSMOTE*, *SVMSMOTE* ou *ADASYN*.

Foram mantidos os parâmetros padrão do algoritmo (*k_neighbors=5* e *sampling_strategy='auto'*), com exceção do *random_state=42*, incluído para garantir reprodutibilidade.

A técnica foi aplicada apenas ao conjunto de treinamento, após a normalização com *StandardScaler*, em conformidade com as boas práticas para evitar um “vazamento de dados” ou *data leakage*.

Com isso, a proporção inicial foi ajustada para um balanceamento 1:1 no conjunto de treino.

A escolha do *SMOTE* padrão mostrou-se adequada neste contexto, uma vez que o conjunto de dados continha apenas variáveis numéricas resultantes do *one-hot encoding*, não havendo necessidade de variantes específicas para dados mistos.

Além disso, trata-se de um método amplamente validado na literatura e reconhecido como eficaz em problemas de classificação binária com desbalanceamento moderado a severo.

Os dados sintéticos gerados foram avaliados clinicamente, a fim de garantir que as combinações criadas fossem plausíveis e coerentes com a realidade biológica e terapêutica dos pacientes, evitando distorções artificiais. Essa etapa reforça a confiabilidade interpretativa do modelo.

4.2.4 Avaliação e Validação

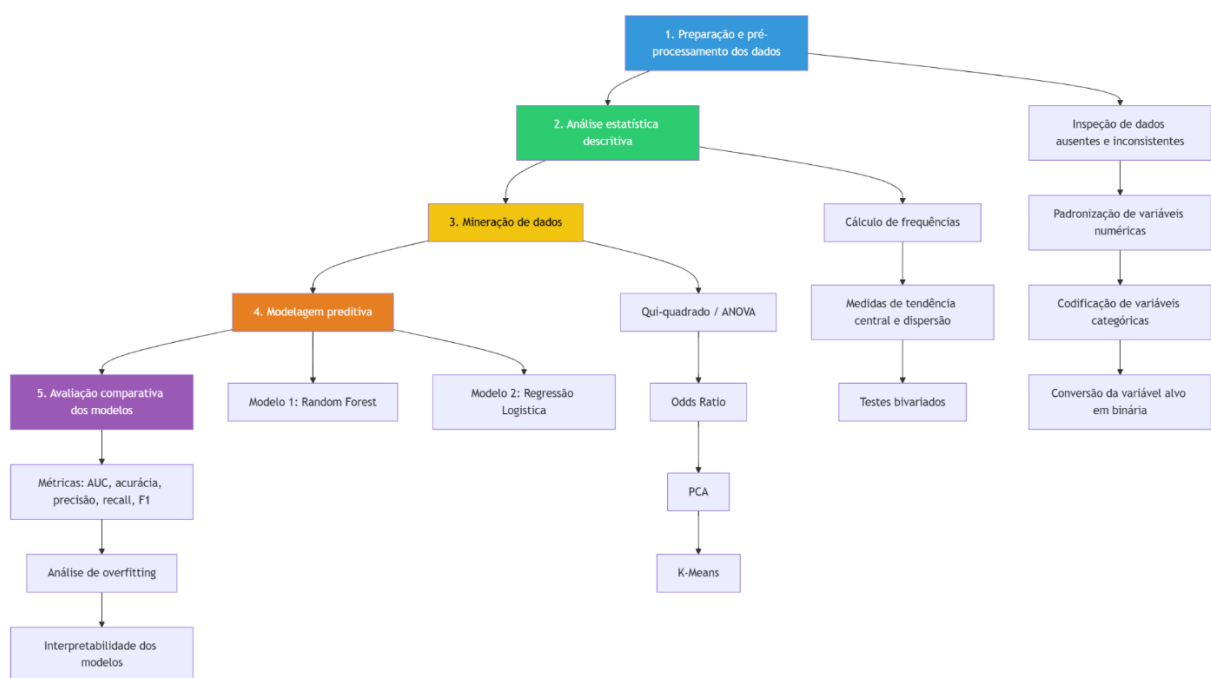
Os modelos foram avaliados em conjunto de teste independente (30% da amostra), com métricas clássicas de classificação. Para o *Random Forest*, foi aplicada validação cruzada estratificada (5 *folds*), reforçando a robustez da estimativa. Os requisitos importantes da avaliação dos resultados foram:

- Interpretabilidade (XAI). A transparência dos modelos foi considerada um ponto central. No *Random Forest*, utilizou-se a análise de importância das variáveis. Na

Regressão Logística, foram examinados os coeficientes dos preditores. Em ambos os casos, empregaram-se valores SHAP (*SHapley Additive exPlanations*), permitindo quantificar a contribuição individual de cada variável para as previsões.

- Reprodutibilidade. Todas as etapas do estudo foram registradas de forma sistemática, com controle de elementos aleatórios e uso de ferramentas de versionamento de código, na tentativa de garantir rastreabilidade e repetibilidade dos experimentos.

Figura 1: Fluxograma de trabalho



4.2.5 O Algoritmo de Análise Preditiva

Para a execução da análise preditiva, foi desenvolvido um algoritmo estruturado em fases sequenciais, que descreve as etapas do pré-processamento, análise estatística, mineração de dados, modelagem preditiva e síntese dos resultados.

Quadro 1: Estrutura do Algoritmo para Análise Preditiva

Entrada: conjunto de dados de pacientes em uso de bisfosfonatos (bisfosfonatos_dados.xlsx)
Saída: modelo preditivo treinado, métricas de desempenho, fatores de risco identificados

// Fase 1 – Preparação e Pré-processamento

1. Carregar o conjunto de dados.
2. Verificar valores ausentes, duplicados e inconsistentes.
3. Converter variável-alvo "ONM?" para formato binário (0 = NÃO; 1 = SIM).
4. Separar variáveis categóricas e numéricas.
5. Dividir o conjunto em treino (70%) e teste (30%) de forma estratificada.
6. No conjunto de treino, aplicar:
 - 6.1. Imputação (média/mediana para variáveis numéricas; moda para categóricas).
 - 6.2. One-Hot Encoding (variáveis categóricas).
 - 6.3. Normalização (StandardScaler).
 - 6.4. Aplicar transformações ao teste sem refazer o ajuste (fit).

// Fase 2 – Análise Estatística Descritiva (dados originais, sem SMOTE)

7. Aplicar para Categóricas: frequências, Qui-quadrado/Fisher, Odds Ratio, IC95%.
8. Aplicar para Numéricas: normalidade (Shapiro), t/ANOVA ou Mann-Whitney/Kruskal, tamanho de efeito.

// Fase 3 – Mineração de Dados (dados originais)

9. Aplicar PCA: variância explicada e biplot.
10. Aplicar K-Means: número ótimo (cotovelo/silhouette), treino, clusters, incidência de ONM.

Observação: PCA e K-Means são exploratórios e não usados nos modelos supervisionados.

// Fase 4 – Modelagem Preditiva

11. Fazer SMOTE no conjunto de treino.
12. rodar Random Forest:
 - 12.1. Cross-validation e ajuste de hiperparâmetros.
 - 12.2. Calibração das probabilidades (Platt/Isotônica).
 - 12.3. Importância das variáveis e gráficos de dependência parcial.
13. Regressão Logística:
 - 13.1. Regularização (L1/L2) e ajuste do parâmetro C.
 - 13.2. Treino com dados balanceados.
 - 13.3. Análise de coeficientes e multicolinearidade (VIF).

// Fase 5 – Avaliação dos Modelos (teste original)

14. Gerar previsões probabilísticas.
15. Avaliar desempenho: AUC-ROC, PR-AUC, Acurácia, Balanced Accuracy, Sensibilidade, Especificidade, Precisão, F1-Score, Kappa de Cohen, Brier Score.
16. Plotar curvas ROC, PR e calibração.
17. Matriz de confusão e otimização do limiar (Youden).

// Fase 6 – Explicabilidade e Síntese dos Resultados

18. Fazer SHAP: contribuições individuais, summary plot e dependence plot.
19. Integrar achados: Odds Ratio, Importância RF, Coeficientes RL, Valores SHAP.
20. Comparar desempenho dos modelos (RF vs RL).
21. Compilar conclusões e recomendações clínicas.

Retornar: modelos calibrados, métricas de desempenho, fatores de risco e recomendações clínicas.

5 RESULTADOS e DISCUSSÃO

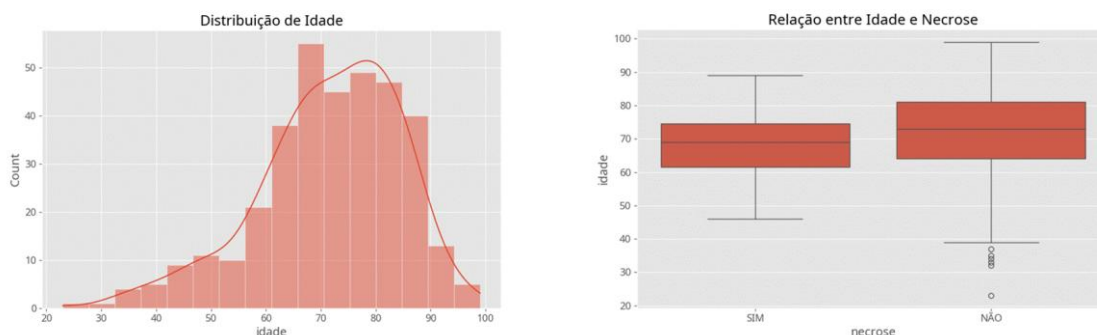
Este capítulo apresenta os resultados obtidos e sua discussão sempre confrontando com a literatura. A organização segue uma sequência lógica onde inicialmente, descrevem-se as características da amostra e os principais fatores clínicos associados à ONM (Seções 5.1 e 5.2). Em seguida, a Seção 5.3 introduz a etapa de mineração de dados, procurando padrões e relações ocultas entre variáveis, por meio de técnicas como correlação, PCA e agrupamento (*clusters*). A seção 5.4 mostra o que foi encontrado na análise preditiva, seguido pela avaliação da performance do modelo (Seção 5.5) e finaliza com a interpretação clínica dos resultados e a discussão das contribuições da IA na saúde (seção 5.6 e 5.7). Essa estrutura tem o objetivo de dar uma visão clínica descritiva dos achados e fechando com a descoberta de insights e suas aplicações.

5.1. Características da População Estudada

A análise foi realizada com base em dados de 354 pacientes que utilizaram bisfosfonatos. A distribuição da variável desfecho (osteonecrose dos maxilares) revelou uma prevalência de 5,4% (19 casos) na população estudada. Esta baixa prevalência está fortemente embasada na literatura (Brunner, 2024) e caracteriza um desbalanceamento significativo entre as classes, o que justificou a adoção de técnicas específicas para lidar com dados desbalanceados nas etapas de modelagem preditiva.

A idade dos pacientes variou entre 23 e 99 anos. Foi observada uma concentração maior de pacientes na faixa de 60 a 80 anos (56,5%), o que é consistente com o perfil epidemiológico das doenças que comumente requerem o uso de bisfosfonatos, como osteoporose e neoplasias (vários tipos de câncer). Já na análise da relação entre idade e ocorrência de ONM foi realizada utilizando o teste t de Student para amostras independentes. Este teste é indicado para a comparação de médias entre dois grupos quando a variável apresenta distribuição aproximadamente normal. Não houve diferença estatisticamente significativa na média de idade entre pacientes com e sem ONM ($p = 0,312$), sugerindo que a idade, isoladamente, pode não ser um fator determinante para o desenvolvimento da complicação. Esses achados estão representados na Figura 2.

Figura 2. Distribuição da idade e relação entre idade e ocorrência dos pacientes. O histograma revela uma concentração maior de pacientes na faixa etária de 60 a 80 anos. O boxplot mostra distribuições semelhantes de idade entre os grupos com e sem ONM, sem diferença estatisticamente significativa ($p = 0,312$).



Quanto ao sexo, a população estudada apresentou predominância feminina (63,8%). Esta distribuição reflete o perfil epidemiológico das principais indicações para uso de bisfosfonatos, especialmente a osteoporose pós-menopausa.

A análise da relação entre sexo e ocorrência de ONM foi realizada utilizando o teste qui-quadrado de Pearson, escolhido por ser o método mais adequado para avaliar associação entre variáveis categóricas em amostras independentes. Não foi encontrada associação estatisticamente significativa entre sexo e desenvolvimento de ONM ($p = 0,427$), indicando que homens e mulheres apresentam riscos semelhantes para esta complicação.

Todas as características demográficas encontradas estão de acordo com a literatura, inclusive em diferentes países como nos trabalhos de Brunner 2024 (Áustria), Şahin 2019 (Turquia) e Teoh 2021 (Australiia).

5.2. Fatores Clínicos Associados à Osteonecrose dos Maxilares

5.2.1. Avaliação Odontológica Prévia

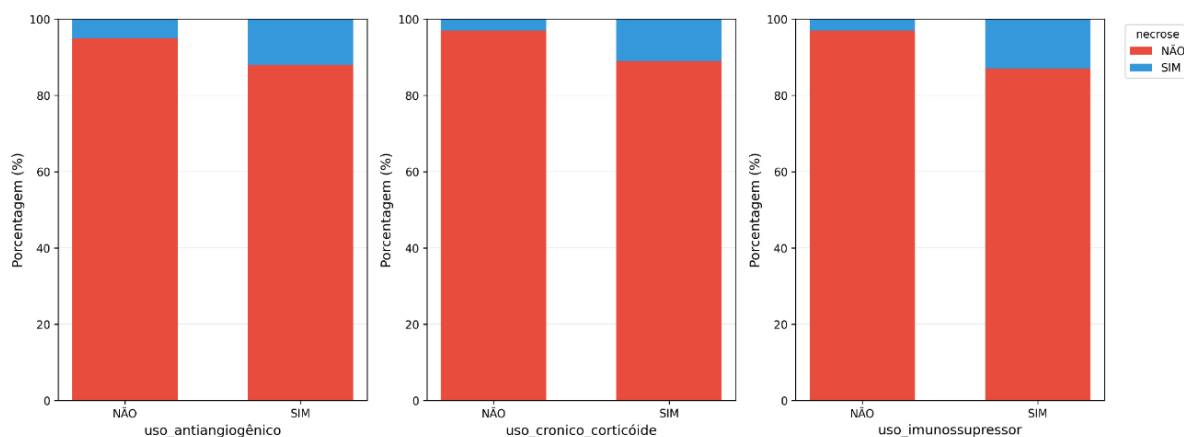
A avaliação odontológica prévia ao início do tratamento com bisfosfonatos mostrou-se significativamente associada à redução do risco de ONM ($p < 0,001$). Este resultado foi obtido por meio do cálculo do teste qui-quadrado de Pearson, que revelou uma forte associação, com apenas 1,2% dos pacientes que realizaram avaliação odontológica prévia desenvolvendo ONM, em comparação com 13,5% daqueles que não realizaram. A avaliação odontológica prévia,

segundo a literatura, é uma das medidas mais eficazes na prevenção da osteonecrose dos maxilares dos maxilares (Brunner,2024). Thomas e Ouanounou (2023) destacam que a prevenção deve ocorrer em todas as fases da terapia (antes, durante e após a administração do medicamento), por meio de exame clínico e radiográfico detalhado, eliminação de focos infecciosos e manutenção de adequada higiene bucal. Kawahara et al. (2021) ressaltam que pacientes oncológicos em uso de altas doses de bisfosfonatos apresentam maior risco para ONM, sendo recomendada a realização de protocolo odontológico prévio que inclua extração de dentes com prognóstico ruim, ajuste de próteses e controle de fatores locais de risco.

5.2.2. Uso de Medicamentos Concomitantes

O uso concomitante de medicamentos que podem afetar o metabolismo ósseo ou a cicatrização foi analisado em relação ao desenvolvimento de osteonecrose dos maxilares. Os resultados para uso crônico de corticoides, imunossupressores e antiangiogênicos são apresentados na Figura 3.

Figura 3. Os gráficos mostram maior proporção de ONM entre pacientes em relação ao uso antiangiogênicos ($p = 0,014$), imunossupressores ($p=0,002$) e a de corticoide ($p=0,003$).



Para todas estas análises, foi utilizado o teste qui-quadrado de Pearson, escolhido por sua adequação para avaliar associação entre variáveis categóricas binárias. Os resultados revelaram associações estatisticamente significativas entre o uso destes medicamentos e o desenvolvimento de osteonecrose dos maxilares.

5.2.3. Quimioterapia Prévia e Indicação de Uso

A realização de quimioterapia prévia mostrou associação significativa com o desenvolvimento de ONM ($p = 0,007$). Esta análise foi realizada utilizando o teste qui-quadrado de Pearson.

Diversos estudos ressaltam que o risco de desenvolvimento de osteonecrose dos maxilares não está associado apenas ao uso isolado de bisfosfonatos, mas se intensifica na presença de terapias concomitantes. Kemp et al. (2022) demonstraram que pacientes em uso de bisfosfonatos apresentaram maior incidência de ONM quando submetidos, ao mesmo tempo, a quimioterapia e corticoides, além de protocolos de infusão mensais, condições que amplificam a imunossupressão e comprometem a reparação óssea. Thomas e Ouanounou (2023) destacam que antiangiogênicos e imunossupressores, quando utilizados em associação a anti-reabsortivos, aumentam de forma significativa o risco de ONM, reforçando a etiologia multifatorial da doença.

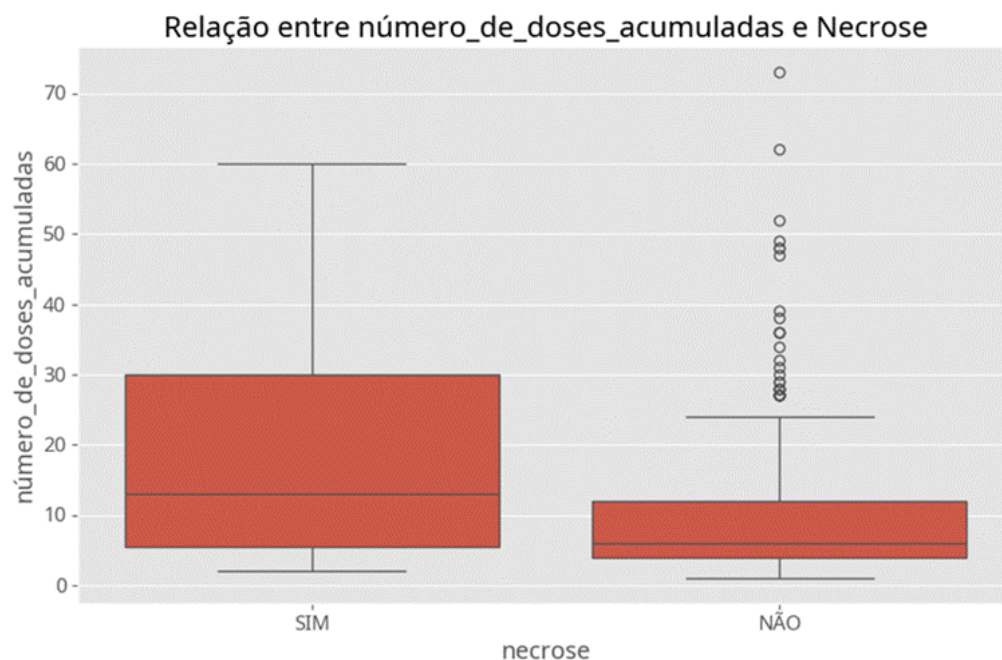
Esses achados confirmam a análise do presente trabalho, em que a investigação dos fatores predisponentes deve considerar não apenas a droga primária em estudo, mas também a utilização de outras drogas típicas dos pacientes oncológicos e osteometabólicos, fator crítico para a uma correta avaliação de risco e para estratégias preventivas individualizadas.

5.2.4. Características do Tratamento com Bisfosfonatos

Diversos aspectos relacionados ao uso de bisfosfonatos foram analisados quanto à sua associação com o desenvolvimento de ONM incluindo o número de doses acumuladas, tempo de uso e frequência de administração.

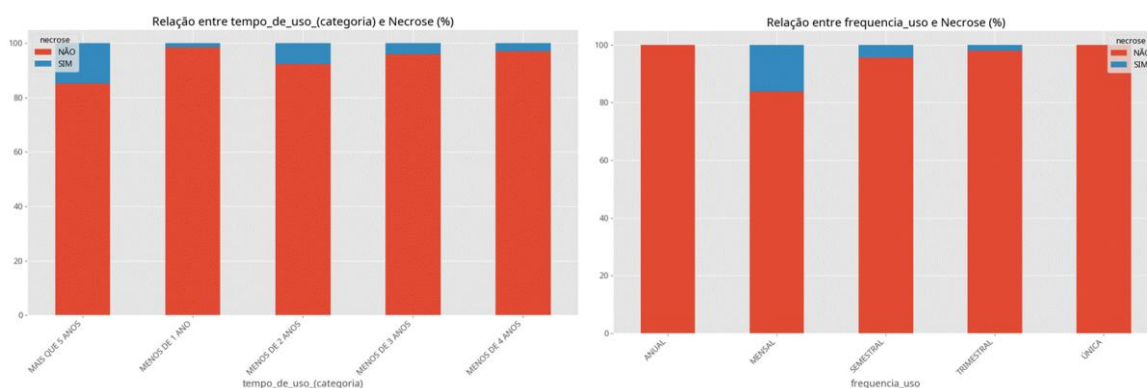
O número de doses acumuladas mostrou associação significativa com o desenvolvimento de ONM ($p < 0,001$), como evidenciado na Figura 4. Esta análise foi realizada utilizando o teste t de Student para amostras independentes, após verificação da normalidade da distribuição por meio do teste de Shapiro-Wilk.

Figura 4. Relação entre número de doses acumuladas de bisfosfonatos e ocorrência de ONM. O boxplot evidencia maior número de doses entre pacientes que desenvolveram ONM ($p < 0,001$).



O tempo de uso categorizado e a frequência de uso também apresentaram associação significativa com o desenvolvimento de ONM ($p = 0,002$ e $p = 0,018$, respectivamente). Para esta análise, foi utilizado o teste qui-quadrado de Pearson, evidenciando maior risco em pacientes com uso prolongado (mais de 5 anos) e com administração mensal do medicamento, conforme a Figura 5.

Figura 5. Relação do tempo e da frequência de uso de bisfosfonatos na ocorrência de ONM. Os gráficos evidenciam maior risco com administração mensal ($p = 0,018$) e uso por mais de 5 anos ($p = 0,002$).



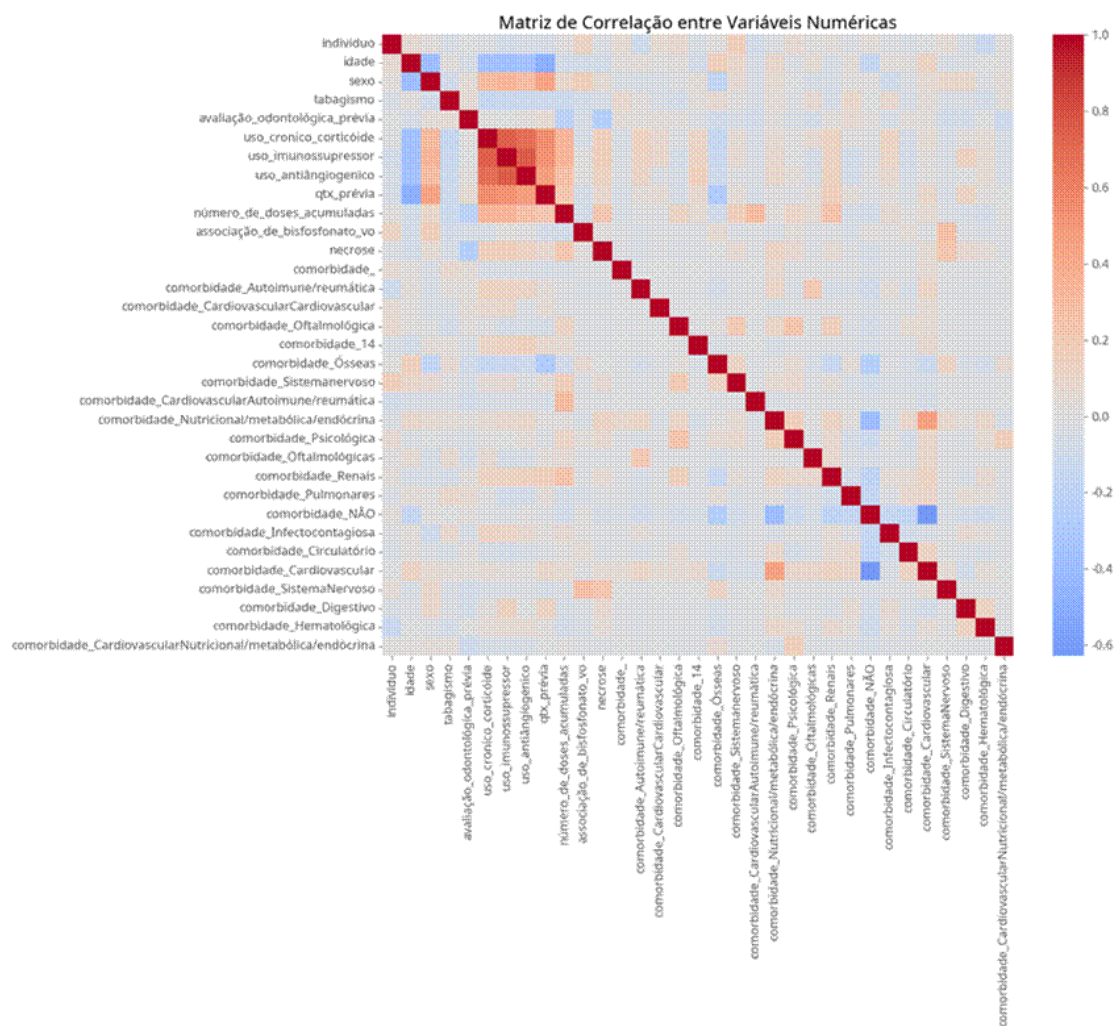
5.3. Mineração de Dados

5.3.1. Análise de Correlação

A análise de correlação foi realizada para identificar relações lineares entre as variáveis do estudo. A matriz de correlação é uma ferramenta estatística que ajuda a identificar quais variáveis se movem juntas (ou em direções opostas) e quais não têm relação (ver Figura 6). A escolha da análise de correlação como etapa inicial da mineração de dados justifica-se por sua capacidade de fornecer uma visão geral das relações entre variáveis, identificando potenciais preditores para análises mais aprofundadas. As correlações mais fortes identificadas foram:

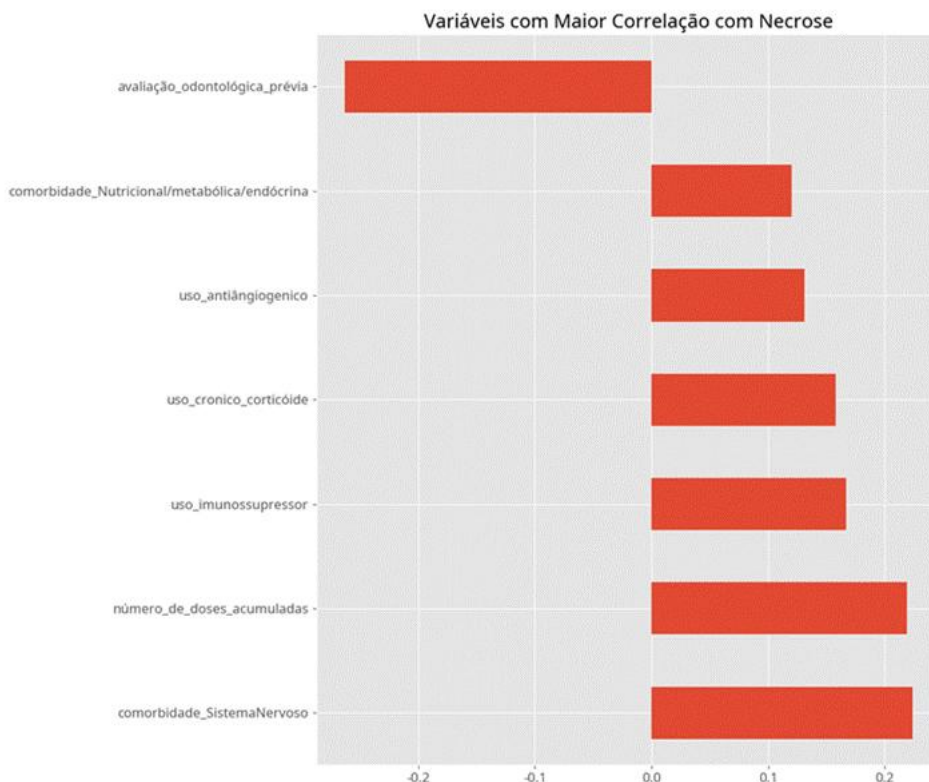
- Avaliação odontológica prévia: correlação negativa (-0,26), indicando efeito protetor ou preventivo
- Comorbidades do sistema nervoso: correlação positiva (0,22), indicando maior risco
- Número de doses acumuladas: correlação positiva (0,22), indicando maior risco com doses maiores
- Uso de imunossupressor: correlação positiva (0,17), indicando maior risco
- Uso crônico de corticoide: correlação positiva (0,16), indicando maior risco

Figura 6. Matriz de correlação entre as variáveis do estudo. As cores mais intensas indicam correlações mais fortes (positivas em vermelho, negativas em azul).



Para facilitar a interpretação, as correlações mais relevantes com a variável desfecho (ONM?) foram extraídas e são apresentadas na Figura 7. Esta análise permitiu reunir as variáveis com maior associação linear com o desenvolvimento de osteonecrose dos maxilares.

Figura 7. Principais variáveis correlacionadas com a ocorrência de ONM. Destaca-se a correlação negativa com avaliação odontológica prévia e positiva com comorbidades do sistema nervoso e número de doses acumuladas.



A análise de correlação realizada neste estudo forneceu não apenas uma visão estatística das relações lineares entre variáveis, mas também uma base para interpretações clínicas relevantes. A variável avaliação odontológica prévia apresentou correlação negativa com a ocorrência de ONM (-0,26), sugerindo um possível efeito protetor. Este achado é consistente com a literatura (Thomas e Ouanouno, 2023), que aponta a avaliação odontológica como medida preventiva capaz de identificar e corrigir fatores de risco locais, reduzindo a necessidade de procedimentos invasivos em pacientes expostos a bisfosfonatos.

Entre os fatores predisponentes à osteonecrose, o número de doses acumuladas destacou-se (0,22), reforçando a relação dose-dependente já descrita. Quanto maior a exposição à medicação, maior o seu depósito no organismo e a consequente inibição da remodelação óssea, aumentando o risco de necrose óssea. O mesmo acontece com o uso de imunossupressores (0,17) e o uso crônico de corticoides (0,16) que mostraram correlações

positivas. A literatura confirma que a ação dessas drogas compromete a resposta imunológica e a reparação óssea, potencializando os efeitos indesejados dos bisfosfonatos (Şahin,2019, Teoh ,2023).

A presença de comorbidades do sistema nervoso também demonstrou correlação positiva (0,22). Esse achado pode estar relacionado tanto ao uso concomitante de outros medicamentos (como antiepilépticos ou imunossupressores) quanto a limitações funcionais e cognitivas que podem dificultar a manutenção da higiene oral, favorecendo quadros infecciosos e a progressão de necrose óssea.

Em síntese, os resultados da correlação destacam variáveis suportadas pela literatura, sugerindo que fatores relacionados tanto ao manejo odontológico preventivo quanto à exposição medicamentosa e às condições sistêmicas desempenham papel importante no risco de desenvolvimento de osteonecrose dos maxilares. Estes achados reforçam a necessidade de uma abordagem integrada, envolvendo monitoramento da dose cumulativa, controle rigoroso de comorbidades e avaliação odontológica sistemática antes e durante o tratamento com bisfosfonatos.

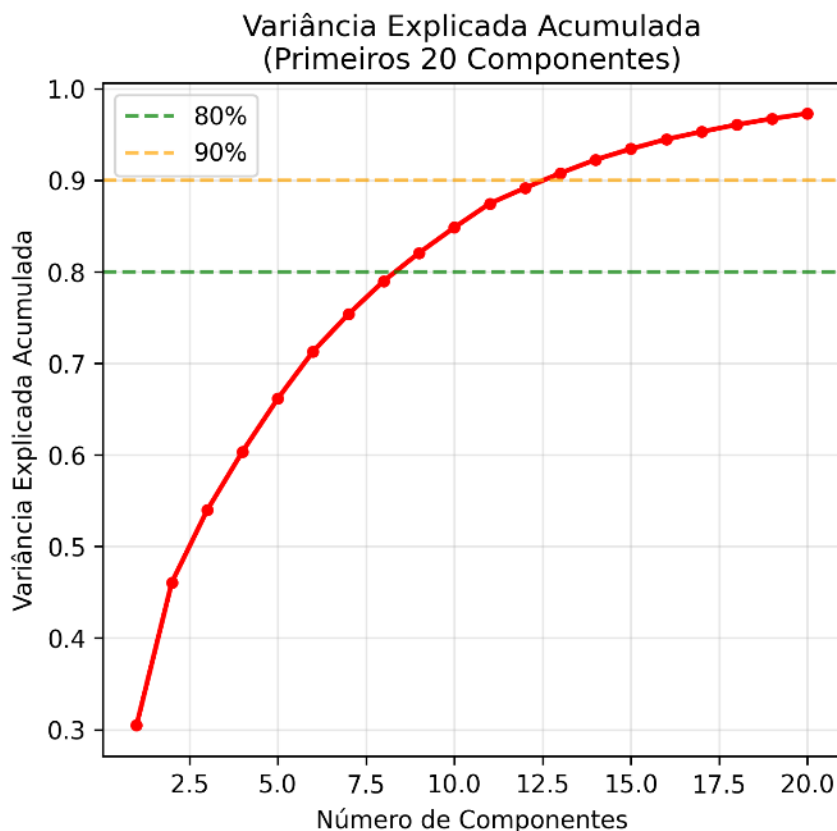
5.3.2. Análise de Componentes Principais (PCA)

Após o pré-processamento dos dados iniciais com as 37 variáveis do *dataset* e a realização das correlações apontadas na seção anterior (Seção 6.3.1), realizou-se o desmembramento de todas as possíveis comorbidades em novas variáveis, chegando-se a 52 variáveis, sendo 2 variáveis numéricas (idade e número de doses acumuladas) e 50 categóricas. A análise de Componentes Principais (PCA) foi aplicada para simplificar o conjunto de dados. Em vez de manipular dezenas de variáveis possivelmente correlacionadas, a Análise de Componentes Principais (PCA) reúne estatisticamente essas variáveis em novos componentes, independentes entre si. Isso facilita a identificação de padrões escondidos e torna a análise mais clara, sem comprometer informações importantes.

A análise de PCA revelou que os primeiros dez componentes principais explicam 84,83% da variância total dos dados. Esta descoberta valida a eficácia da técnica para

identificação de padrões em *datasets* de alta dimensionalidade, sem perda significativa de informações (Figura 8).

Figura 8. Variância explicada por cada componente principal. O gráfico mostra que são necessários 10 componentes para explicar mais de 80% da variância total dos dados.



5.3.3. Análise de *Clusters*

A análise de *clusters* foi realizada para identificar grupos naturais de pacientes com características similares. O método K-means foi escolhido por sua capacidade de lidar com grandes conjuntos de dados.

A determinação do número ideal de *clusters* foi realizada utilizando o método do Cotovelo (Elbow Method) e a análise do Coeficiente de Silhueta (Silhouette Score), como mostrado na Figura 9. Os valores de Inércia (WCSS - Within-Cluster Sum of Squares) e do

Coeficiente de Silhueta médio, calculados para cada valor de K, estão detalhados na tabela 1. Estes dados correspondem aos pontos plotados nos gráficos da Figura 11. Esta abordagem permitiu identificar o número de *clusters* que maximiza a coesão interna e a separação entre grupos.

Figura 9. Método do cotovelo para determinação do número ótimo de clusters. O gráfico sugere 2 clusters como número ótimo, baseado no ponto de inflexão da curva de inércia e no score de silhueta.

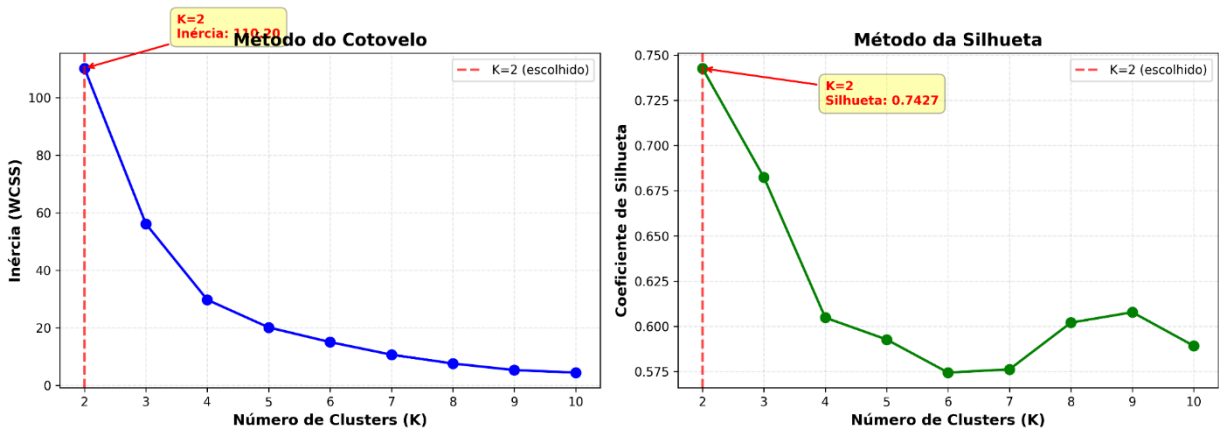


TABELA 1: Os valores de Inércia e do Coeficiente de Silhueta médio, calculados para cada valor de K. Estes dados correspondem aos pontos plotados nos gráficos da Figura 11

K (Número de Clusters)	Inércia (WCSS)	Coeficiente de Silhueta
2	110.20	0.7427
3	56.14	0.6824
4	29.73	0.6048
5	20.12	0.5927
6	15.02	0.5744
7	10.64	0.5762
8	7.58	0.6020
9	5.29	0.6078
10	4.42	0.5892

No método de cotovelo, o gráfico da inércia (à esquerda) exibe uma queda acentuada ao passar de $K=2$ para $K=3$, mas a redução mais drástica ocorre antes, formando um "cotovelo" nítido em $K=2$. A partir deste ponto, a curva começa a achatar, indicando que adicionar mais clusters (de 3 em diante) não resultaria em ganhos significativos. A Inércia cai de 110.20 para 56.14 (uma redução de aproximadamente 49%), mas as reduções subsequentes são progressivamente menores. Isso sugere que a divisão em dois grupos já captura uma porção significativa dos dados sem levar a uma complexidade desnecessária.

Na análise do Coeficiente de Silhueta (gráfico à direita) confirma a conclusão da técnica de cotovelo. O valor máximo do coeficiente (0.7427) é alcançado precisamente em $K=2$. Um coeficiente de silhueta mais alto indica que os clusters são, em média, mais densos e mais bem separados entre si. Para todos os valores de K superiores a 2, o coeficiente médio diminui, o que sugere que os clusters começam a se sobrepor ou que os pontos não estão tão claramente agrupados quanto na configuração com $K=2$.

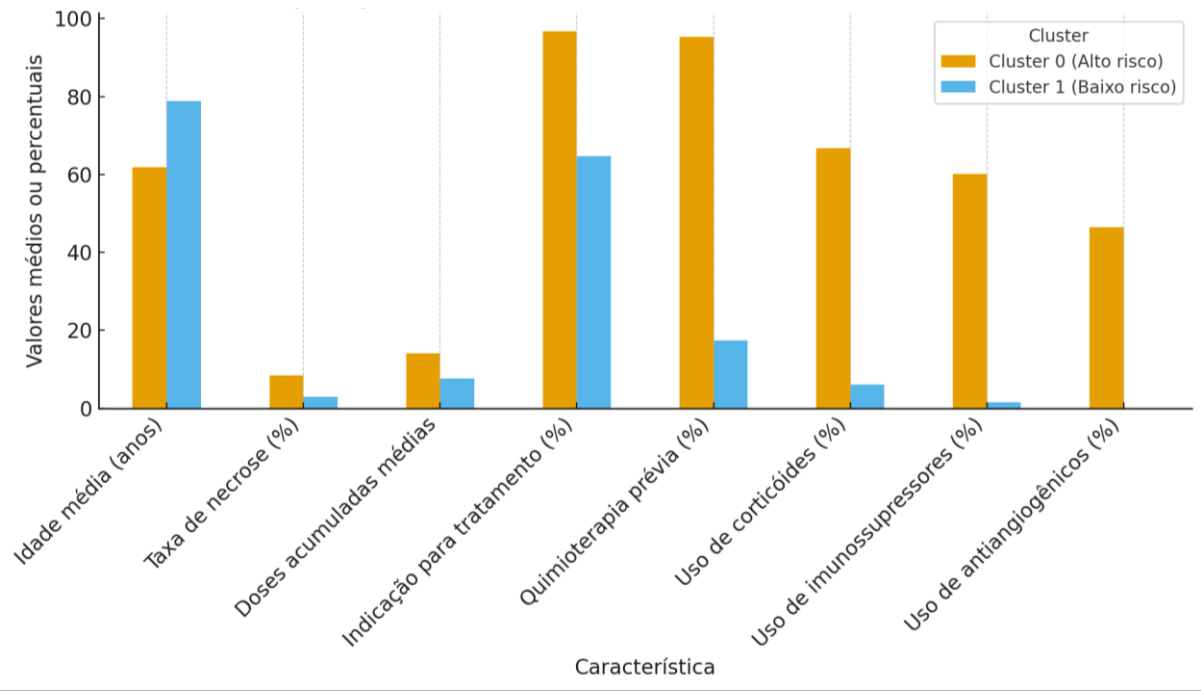
Assim, a convergência dos resultados de ambos os métodos fornece uma forte evidência metodológica para a seleção de $K=2$. O método do cotovelo identifica $K=2$ como o ponto de equilíbrio ideal entre a complexidade do modelo e a variância, enquanto o método da silhueta confirma que $K=2$ produz a configuração de clusters com a melhor definição e separação estrutural.

A análise de clusters identificou dois grupos principais de pacientes e as suas diferenças podem ser visualizadas na TABELA 2 e na FIGURA 10.

TABELA 2: Detalhamento das características dos Clusters 0 (alto risco de ONM) e Cluster 1 (baixo risco)

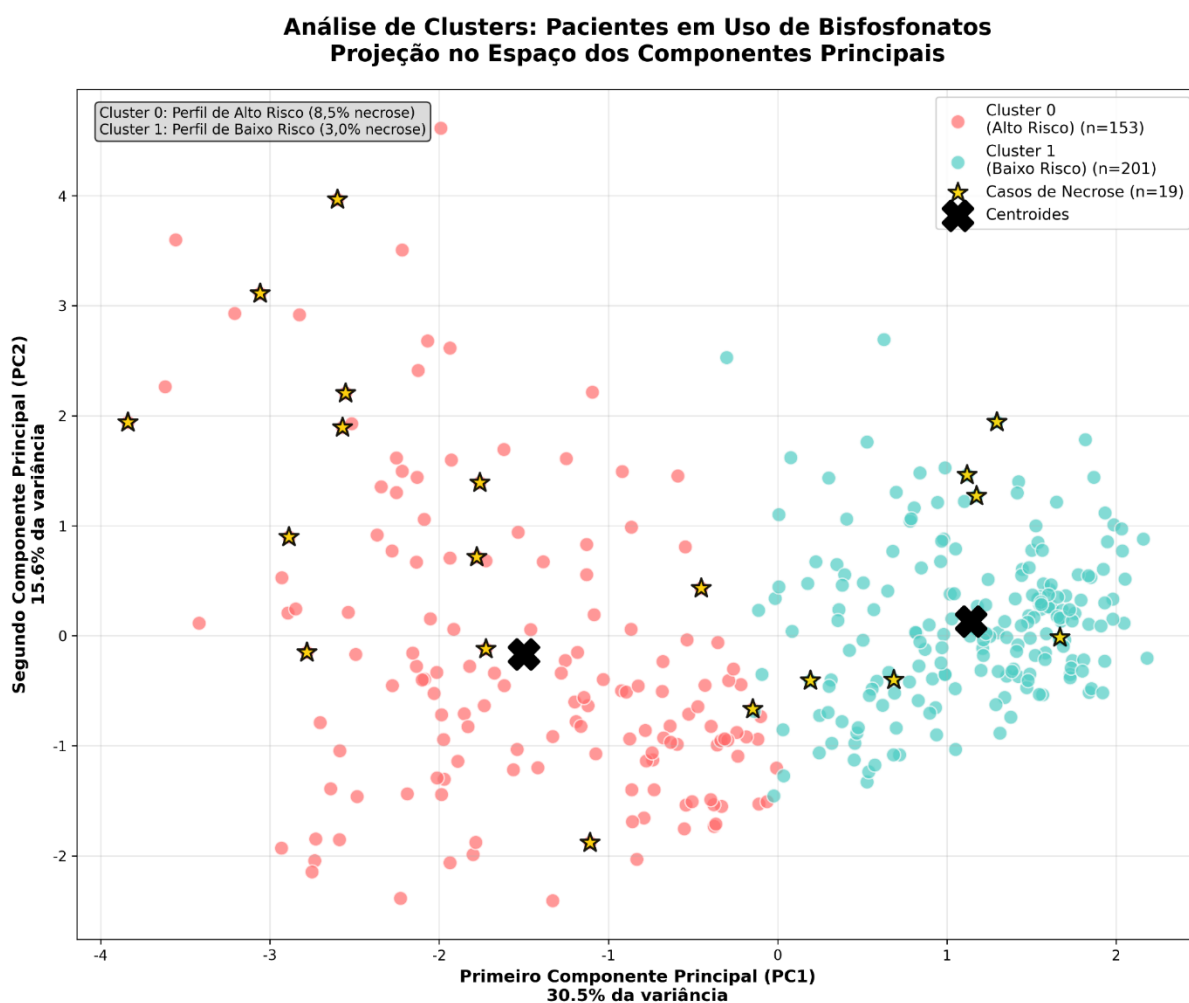
Característica	Cluster 0 (Alto risco)	Cluster 1 (Baixo risco)
Número de Pacientes	153 (43,2%)	201 (56,6%)
Idade média (anos)	61.9	78.8
Sexo predominante	Feminino	Masculino
Taxa de ONM (%)	8.5	2.99
Doses acumuladas médias	14.1	7.7
Frequência principal	Trimestral/Mensal	Semestral
Indicação para tratamento (%)	96.7	64.7
Quimioterapia prévia (%)	95.4	17.4
Uso de corticóides (%)	66.7	6.0
Uso de imunossupressores (%)	60.1	1.5
Uso de antiangiogênicos (%)	46.4	0.0

Figura 10: Comparação das características numéricas dos cluster 0 (alto risco) e 1 (baixo risco)



A visualização dos clusters nos dois primeiros componentes principais é apresentada na Figura 11. Esta representação permite avaliar visualmente a separação entre os grupos identificados.

Figura 11. Visualização dos clusters nos dois primeiros componentes principais. Observa-se os centroides bem afastados e a distribuição dos casos com osteonecrose.

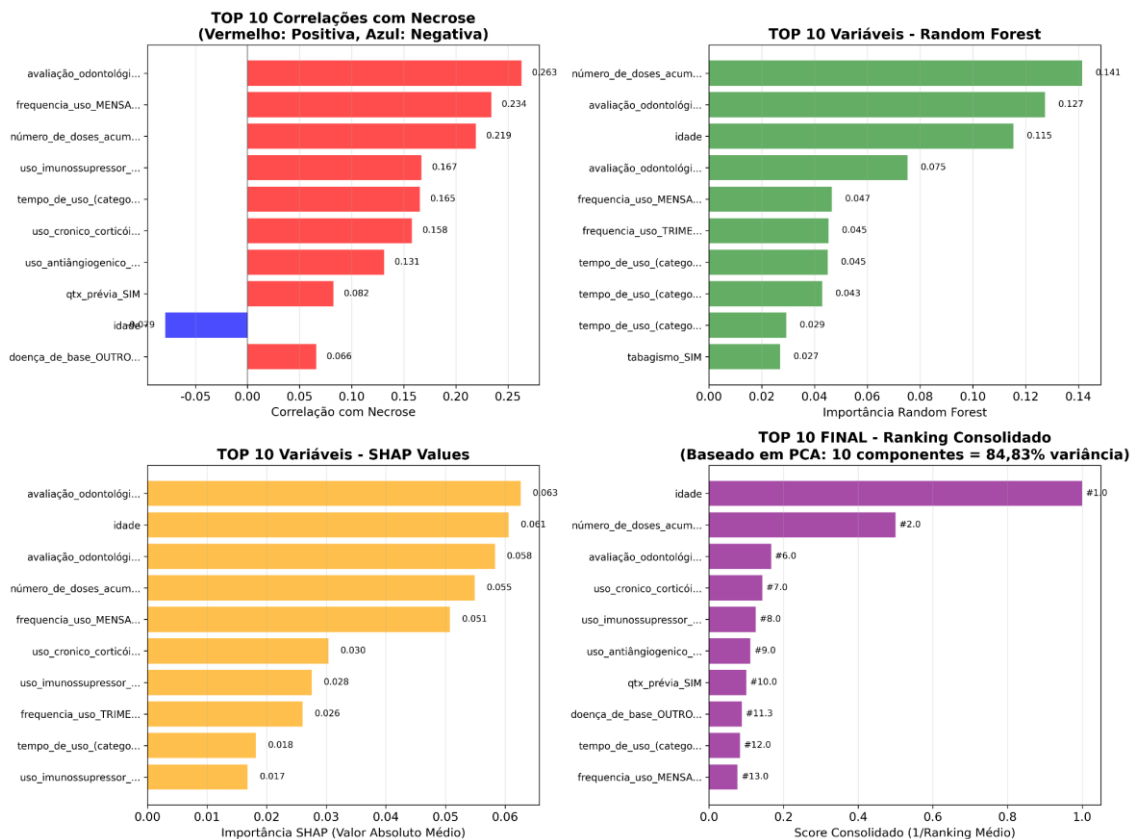


A escolha da análise de *clusters* como complemento ao PCA justifica-se pela capacidade desta técnica de identificar subgrupos na população estudada, potencialmente revelando perfis de pacientes com diferentes riscos para o desenvolvimento de ONM. Os resultados mostraram que o *Cluster 0* apresentou prevalência para ONM 2,85 vezes maior que o *Cluster 1* (8,5% vs. 2,99%), confirmando as correlações já levantadas anteriormente e sugerindo que as

características específicas deste grupo como o uso de outros medicamentos e comorbidades estão associadas ao maior risco.

A análise de importância de variáveis, para selecionar aquelas que foram efetivamente usadas foi uma das fases mais críticas da mineração de dados. O principal desafio foi avaliar aquelas de maior relevância clínica juntamente com o valor preditivo. Por essa razão, foram usados métodos estatísticos tradicionais, como correlação, testes de hipóteses e medidas de efeito, aliados a técnicas de *machine learning*, como o *Random Forest* e os valores SHAP. Essa integração foi importante porque cada método “enxergou” o problema por um ângulo diferente. Quando todos apontaram para os mesmos fatores, demonstrou-se que esses achados eram precisos e não fruto do acaso (Figura 12).

Figura 12: Comparação entre quatro métodos de análise (correlações, Random Forest, SHAP e ranking consolidado) mostrando as 10 variáveis mais importantes para predição de ONM. A convergência entre métodos valida os achados.



Os valores plotados no gráfico da análise inicial de correlação, estão na TABELA 3. Esses resultados reforçam a importância do acompanhamento odontológico antes e durante o tratamento.

Tabela 3 - Top 10 Correlações com ONM, observou-se que a ausência de avaliação odontológica prévia foi o fator mais fortemente ligado ao desenvolvimento de ONM, seguida pela frequência mensal de uso do medicamento e pelo número de doses acumuladas

Variável	Correlação	P-value	Interpretação
Ausência de avaliação odontológica	0,263	<0,001	Moderada positiva
Frequência de uso mensal	0,234	<0,001	Moderada positiva
Número de doses acumuladas	0,219	<0,001	Moderada positiva
Uso de imunossupressor	0,167	0,002	Moderada positiva
Tempo de uso > 5 anos	0,165	0,002	Moderada positiva
Uso crônico de corticóide	0,158	0,003	Moderada positiva
Uso de antiangiogênico	0,131	0,014	Fraca positiva
Quimioterapia prévia	0,082	0,122	Fraca positiva
Idade	-0,079	0,138	Fraca negativa
Doença de base outros	0,066	0,217	Fraca positiva

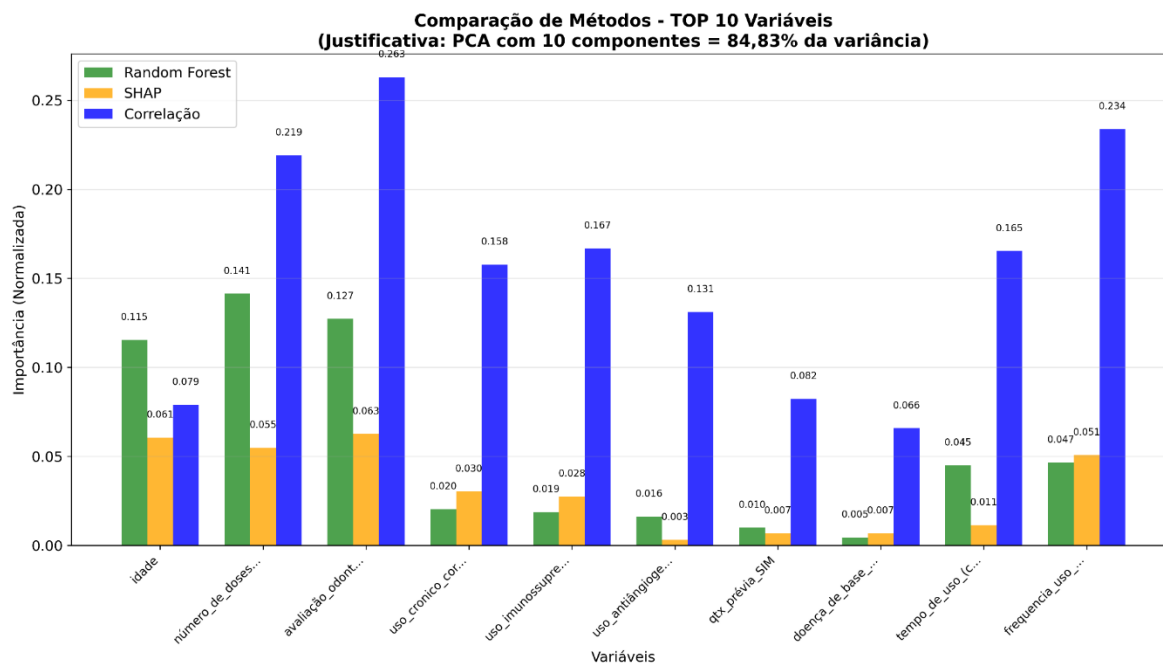
Os testes estatísticos confirmaram esses achados e estão demonstrados na TABELA 4. O número de doses acumuladas, por exemplo, mostrou diferença significativa entre pacientes com e sem ONM, portanto, não se trata apenas de um resultado “estatisticamente significativo”, mas também de relevância clínica.

Tabela 4 - Resultados dos testes estatísticos, confirmando os achados da análise das correlações com ONM

Variável	Teste	Estatística	P-value	Effect Size	Interpretação
Número de doses acumuladas	Teste t	4,213	<0,001	d = 0,994	Efeito grande
Idade	Teste t	-1,487	0,138	d = -0,351	Efeito pequeno
Avaliação odontológica prévia	Qui-quadrado	22,063	<0,001	V = 0,250	Efeito moderado
Frequência de uso	Qui-quadrado	20,525	<0,001	V = 0,241	Efeito moderado
Uso de imunossupressor	Qui-quadrado	8,263	0,004	V = 0,153	Efeito pequeno
Uso crônico de corticóide	Qui-quadrado	7,377	0,007	V = 0,144	Efeito pequeno

Já o *Random Forest* trouxe uma hierarquia das variáveis, destacando novamente o número de doses acumuladas, a ausência de avaliação odontológica e a idade como fatores centrais. Para aprofundar a interpretação, aplicamos os valores SHAP, que funcionam como uma lupa sobre cada variável, mostrando o quanto ela contribuiu individualmente para a previsão de risco (Figura 13). Nessa análise, os cuidados odontológicos preventivos voltaram a aparecer como determinantes críticos.

Figura 13: Gráfico de barras agrupadas comparando a importância das 10 principais variáveis segundo Random Forest, SHAP e correlações. A convergência entre métodos confirma os achados.



Por fim, foi elaborada uma lista ordenada da consolidação das variáveis, unindo os resultados de todas as abordagens, representados na TABELA 5. Esse cruzamento mostrou consistência notável: idade, doses acumuladas e ausência de avaliação odontológica se mantiveram no topo em diferentes análises. Isso reforça que não estamos diante de achados isolados, mas de fatores realmente centrais no processo.

Tabela 5: Ranking consolidado das variáveis TOP 10

Posição	Variável	Ranking Médio	RF Importância	SHAP Importância	Correlação
1	Idade	1,0	0,115	0,061	-0,079
2	Número de doses acumuladas	2,0	0,141	0,055	0,219
3	Ausência de avaliação odontológica	6,0	0,127	0,063	0,263
4	Uso crônico de corticóide	7,0	0,020	0,030	0,158
5	Uso de imunossupressor	8,0	0,019	0,028	0,167
6	Uso de antiangiogênico	9,0	0,016	0,003	0,131
7	Quimioterapia prévia	10,0	0,010	0,007	0,082
8	Doença de base outros	11,3	0,005	0,007	0,066
9	Tempo de uso > 5 anos	12,0	0,045	0,011	0,165
10	Frequência mensal	13,0	0,047	0,051	0,234

Do ponto de vista clínico, os resultados confirmam três grandes grupos de fatores de risco:

1. Cuidados preventivos – ausência de avaliação odontológica prévia;
2. Intensidade do tratamento – doses acumuladas e frequência de uso;
3. Complexidade clínica – uso concomitante de corticoides, imunossupressores e antiangiogênicos.

Esses achados não apenas confirmam recomendações já presentes em diretrizes, como também abrem espaço para a aplicação da medicina personalizada em odontologia. Identificar pacientes de maior risco com base em um conjunto claro de variáveis pode guiar condutas preventivas e personalizar o acompanhamento clínico.

5.4 Análise Preditiva

A análise preditiva representa um dos principais objetivos da aplicação de inteligência artificial em contextos clínicos, transformando conhecimento estatístico em ferramentas práticas de apoio à decisão do profissional da área da saúde. Neste estudo, o desenvolvimento de modelos preditivos para ONM associada a bisfosfonatos criaram ferramentas que permitem avaliar o risco de cada paciente de forma individual, ajudando a aplicar protocolos personalizados tanto na odontologia quanto na medicina.

A estratégia escolhida buscou equilibrar a clínica e o rigor científico, reconhecendo que, para os modelos de machine learning serem aceitos na prática odontológica, é essencial que os processos de decisão sejam transparentes. Essa visão acompanha as recomendações atuais de explicabilidade em inteligência artificial (XAI), que defendem modelos compreensíveis para os profissionais de saúde.

5.4.1 Seleção de Variáveis e Preparação dos Dados

O processo de construção do modelo preditivo envolveu duas etapas complementares: análise de dimensionalidade e seleção de variáveis.

Primeiramente, aplicou-se a Análise de Componentes Principais (PCA) apenas com o propósito de avaliar a estrutura e redundância dos dados, verificando que dez componentes principais explicavam 84,83% da variabilidade total do conjunto. Essa etapa teve caráter exploratório, indicando que as variáveis originais continham informações suficientes para representar adequadamente o fenômeno estudado.

Em seguida, foi realizada a seleção de variáveis propriamente dita, baseada em medidas de importância obtidas nos modelos e em critérios clínicos. Foram mantidas as dez variáveis mais relevantes e consideradas essenciais para a interpretação e aplicação prática do modelo. As variáveis selecionadas abrangeram quatro domínios clínicos fundamentais: características demográficas (idade), tempo e frequência do tratamento (número de doses acumuladas, frequência e duração de uso), cuidados preventivos (avaliação odontológica prévia) e cofatores clínicos (uso concomitante de corticóides, imunossupressores, antiangiogênicos e quimioterapia, além do tipo de doença de base). Essa classificação mostra que a osteonecrose dos maxilares é um problema com muitas causas e permite entender o conjunto de fatores que influenciam seu desenvolvimento.

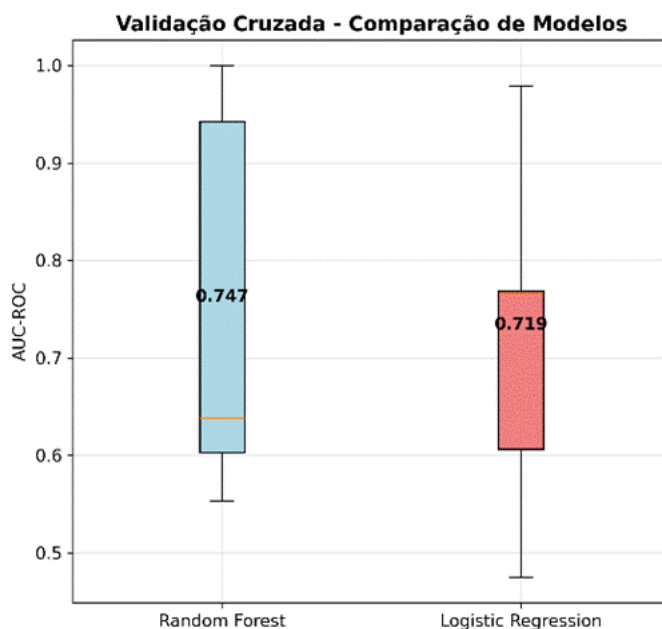
Na etapa de preparação, os dados foram tratados seguindo práticas de *machine learning*. As variáveis numéricas foram padronizadas com o *StandardScaler* e as categóricas convertidas em variáveis binárias pelo método *one-hot*. Ao final, o conjunto de dados passou a ter 43 variáveis (2 numéricas e 41 categóricas). Para garantir equilíbrio entre os grupos, a divisão destinou 247 amostras ao treinamento (69,8%) e 107 ao teste (30,2%), mantendo a proporção original da variável alvo.

A modelagem comparou de forma sistemática dois algoritmos complementares: Random Forest e Regressão Logística. Essa escolha buscou equilibrar capacidade preditiva e interpretabilidade, já que ambos permitem compreender os fatores que influenciam as previsões.

Nos testes de validação cruzada, o Random Forest apresentou desempenho superior, com AUC-ROC médio de $0,747 \pm 0,186$, frente a $0,719 \pm 0,170$ da Regressão Logística. Embora a diferença numérica seja pequena, ela reflete a maior habilidade do Random Forest em capturar interações não lineares entre variáveis, algo comum com dados clínicos complexos.

A Figura 14 ilustra a comparação detalhada entre os modelos por meio de box plots, mostrando não apenas a média mais alta do Random Forest, mas também sua maior estabilidade, evidenciada pela menor variação. Essa consistência é especialmente importante no contexto clínico, onde as previsões precisam ser confiáveis em diferentes perfis de pacientes.

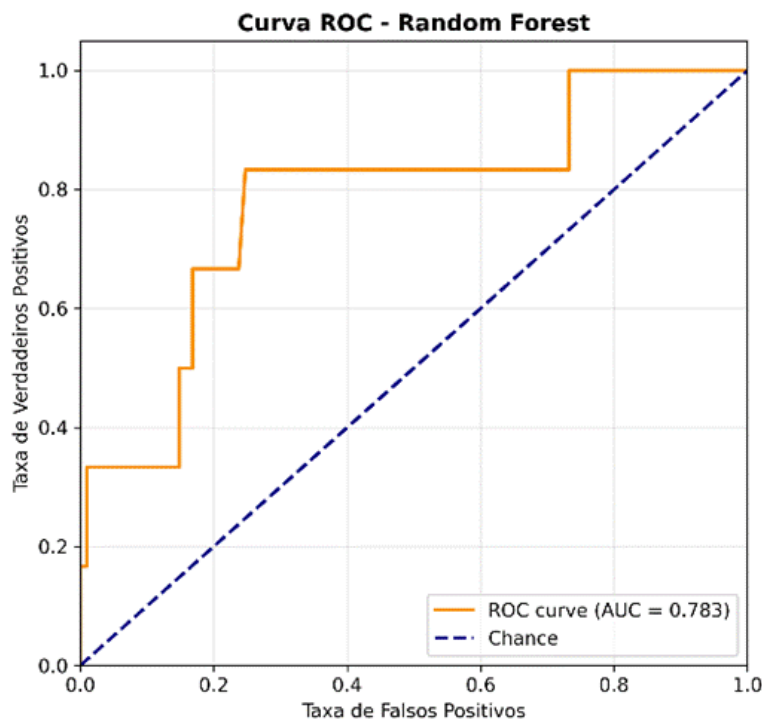
Figura 14: Comparação de modelos usando a validação cruzada. Desempenho dos algoritmos Random Forest e Regressão Logística, representado por boxplots do AUC-ROC. O Random Forest teve melhor média e menor variação entre os testes, ou seja, é mais estável e confiável.



5.4.2. Random Forest

O modelo Random Forest final apresentou desempenho muito bom no conjunto de teste independente, alcançando AUC-ROC de 0,783, valor superior ao limiar de 0,7, tradicionalmente associado a boa capacidade discriminativa. A Figura 15 mostra a curva ROC completa, evidenciando desempenho acima do acaso em toda a faixa de pontos de corte, com maior eficiência na região de alta especificidade.

Figura 15: Curva ROC do modelo Random Forest. Relação entre sensibilidade e especificidade, com destaque para a área sob a curva (AUC) como medida de desempenho global.

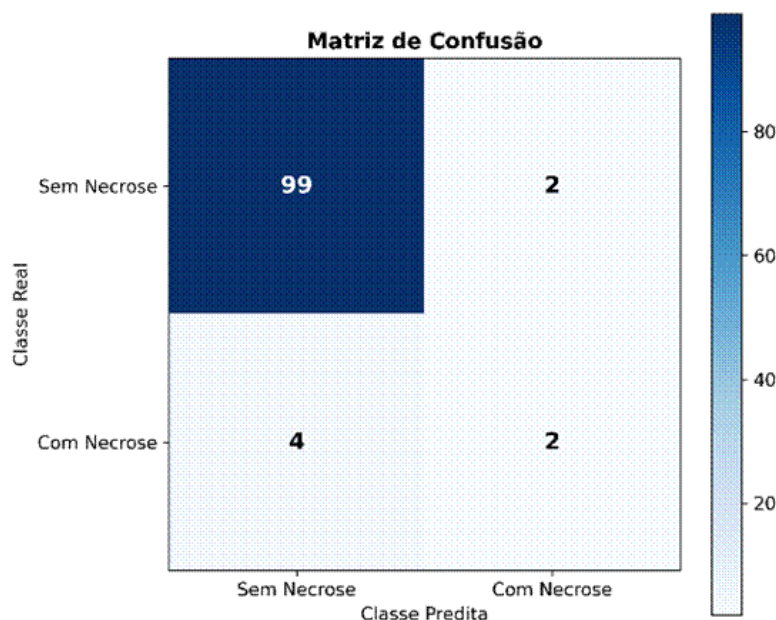


Nas métricas de classificação, a sensibilidade de 33,3% indica que o modelo identifica um terço dos casos de ONM. Embora pareça limitada, essa taxa deve ser analisada à luz da baixa prevalência do evento (5,4%). Já a especificidade de 98,0% demonstra notável capacidade de excluir corretamente pacientes sem risco, minimizando falsos alarmes e ansiedade desnecessária.

O valor preditivo negativo de 96,1% se destaca pela relevância clínica: quando o modelo prediz ausência de risco, está correto em mais de 96% dos casos. Isso permite classificar pacientes de baixo risco com alta confiança, possibilitando protocolos de acompanhamento menos intensivos.

A matriz de confusão (Figura 16) mostra que, entre 107 casos do conjunto de teste, houve apenas 2 falsos positivos (pacientes sem ONM classificados como de risco) e 4 falsos negativos. Essa distribuição de erros é clinicamente aceitável: os falsos positivos implicam apenas em monitoramento adicional, enquanto os falsos negativos, embora mais críticos, permaneceram em número reduzido.

Figura 16: Matriz de confusão detalhada. Distribuição das classificações corretas e incorretas do modelo Random Forest no conjunto de teste.



5.4.3 Interpretabilidade e Importância das Variáveis

A análise de importância confirmou e detalhou achados anteriores. A Figura 17 mostra como os principais fatores associados ao risco foram:

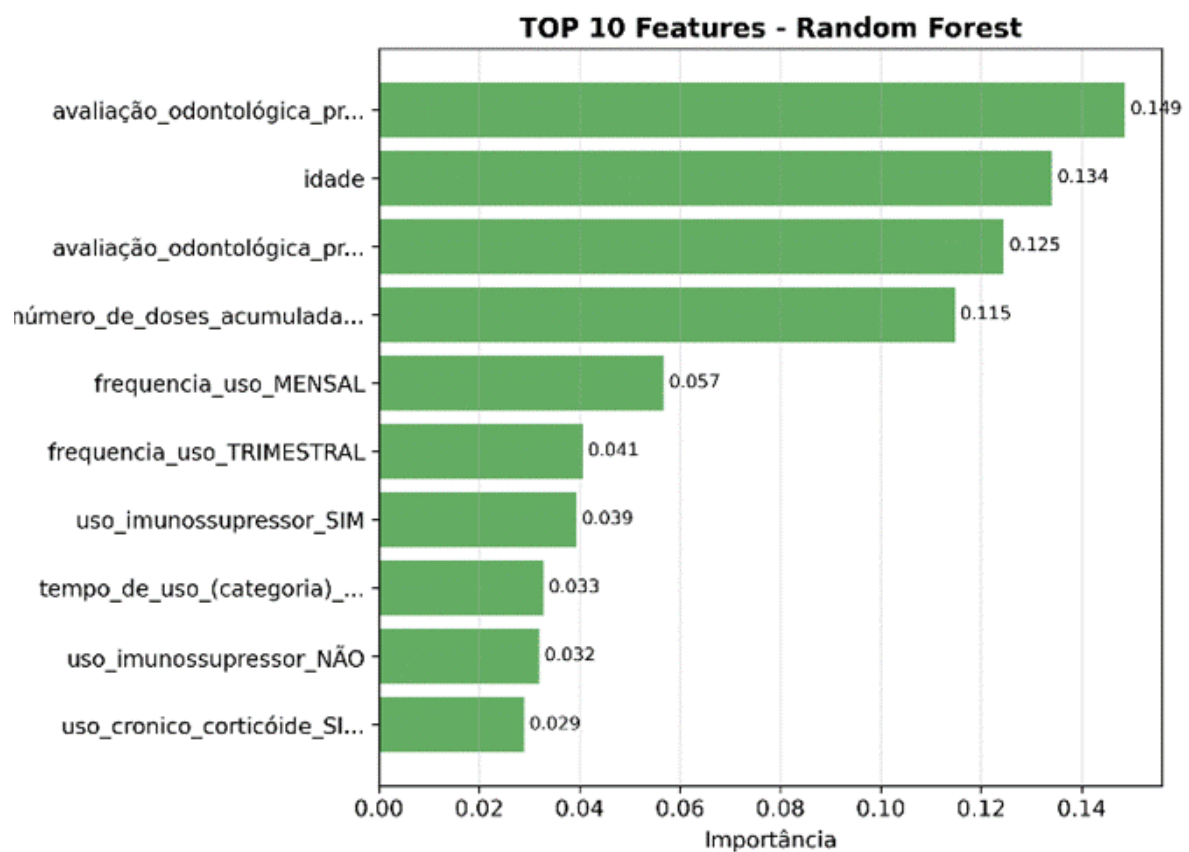
- Avaliação odontológica prévia (presença: 14,9%);
- Idade (13,4%);
- Ausência de avaliação odontológica (12,5%);
- Número de doses acumuladas (11,5%);
- Frequência mensal de administração (5,7%).

Essas cinco variáveis concentram cerca de 58% da importância total, mostrando que um pequeno grupo de fatores explica a maior parte do risco preditivo.

A duplicidade entre presença e ausência de avaliação odontológica decorre da codificação binária, mas reforça o papel central dos cuidados preventivos.

A convergência entre Random Forest, análises de correlação e valores SHAP fortalece a confiabilidade dos resultados, indicando que os padrões encontrados refletem relações reais e não meros artefatos de técnica.

FIGURA 17: Importância relativa das variáveis no modelo Random Forest. Distribuição das principais variáveis preditoras de osteonecrose dos maxilares segundo a análise de importância. Observa-se destaque para a avaliação odontológica prévia, idade, ausência de avaliação odontológica, número de doses acumuladas e frequência mensal de administração. Juntas, essas cinco variáveis concentram cerca de 58% da importância total do modelo, reforçando seu papel central na predição do risco.



5.5 Performance do Modelo Selecionado

Como já citado anteriormente, o modelo Random Forest apresentou desempenho muito bom no conjunto de teste, acima do limiar de 0,7, considerado indicativo de boa discriminação.

Nas métricas de classificação, a sensibilidade foi de 33,3%, o que significa que o modelo identificou um terço dos casos de ONM. Embora esse valor pareça baixo, ele deve ser entendido dentro do contexto da raridade do evento (prevalência de 5,4%). Já a especificidade atingiu 98,0%, demonstrando excelente capacidade de reconhecer pacientes sem risco, reduzindo falsos alarmes e ansiedade desnecessária.

O valor preditivo negativo de 96,1% se mostrou a métrica mais relevante do ponto de vista clínico: quando o modelo prevê ausência de risco, essa previsão está correta em mais de 96% dos casos. Isso permite classificar pacientes como baixo risco com alta confiança, viabilizando protocolos de acompanhamento menos intensivos.

A matriz de confusão ilustra que, entre os 107 casos testados, apenas dois pacientes sem ONM foram incorretamente classificados como de risco (falsos positivos) e quatro casos de Osteonecrose dos maxilares não foram detectados (falsos negativos). Essa distribuição de erros é considerada aceitável: falsos positivos implicam apenas monitoramento adicional, enquanto falsos negativos, embora mais preocupantes, permaneceram em número reduzido.

5.6. Interpretação Clínica dos Principais Resultados

Os resultados encontrados mostram que consultas odontológicas antes do início da terapia com bisfosfonatos funcionam como medida preventiva fundamental. O exame clínico pode identificar focos infecciosos, dentes com mau prognóstico e necessidade de adequação da higiene, reduzindo drasticamente a chance de complicações futuras. Assim, a aplicação de protocolos institucionais deveria tornar obrigatória a avaliação odontológica prévia antes da prescrição de bisfosfonatos.

Associação significativa do uso prévio de corticoides com ONM ($p < 0,005$) confirma que esses medicamentos retardam a cicatrização e o metabolismo ósseo. O risco maior em usuários crônicos sugere necessidade de vigilância odontológica contínua nesses pacientes. Dessa maneira, pacientes poli medicados (com terapias oncológicas, reumatológicas ou imunológicas) devem ser classificados como alto risco e acompanhados com maior rigor.

O número de doses acumuladas e tempo de uso dos bisfosfonatos também foi um fator com forte associação com a ONM. Esse achado confirma que a exposição cumulativa é um dos principais fatores de risco. Quanto maior a carga total de medicamento, maior o dano ósseo potencial. Justificando um monitoramento odontológico mais frequente em pacientes com tratamentos prolongados, e sempre que possível deve-se discutir interrupções ou ajustes na posologia com a equipe médica.

Outro fator importante encontrado, foi a frequência de administração do bisfosfonato. A administração mensal aumenta o risco de ONM. Esses protocolos de maior intensidade parecem acelerar o acúmulo do efeito deletério no osso. Isso sugere que regimes menos frequentes podem ter menor toxicidade óssea, devendo essa informação ser considerada em decisões terapêuticas personalizadas.

Pacientes oncológicos submetidos a múltiplas terapias como a quimioterapia estão mais fragilizados e apresentam tecidos menos responsivos à cicatrização. Já existe protocolo de avaliação odontológica prévia para os tratamentos de quimioterapia e radioterapia. Mesmo assim, nesses casos, deve haver planejamento odontológico ainda mais rigoroso antes da introdução de bisfosfonatos.

5.7. Contribuições da Inteligência Artificial na Saúde

Este estudo mostra que é viável aplicar técnicas de *machine learning* em cenários clínicos específicos, mantendo o equilíbrio entre sofisticação técnica e interpretabilidade. A metodologia proposta pode servir como referência para outras áreas médicas que lidam com eventos raros e exigem estratificação de risco.

A combinação de estatística clássica, modelos preditivos e ferramentas de explicabilidade (XAI) fortalece a confiança nos achados, pois diferentes métodos apontaram para os mesmos resultados.

Ao valorizar transparência e aplicabilidade clínica, sem perder o rigor científico, este trabalho contribui para o avanço da IA explicável em saúde. Esse equilíbrio é essencial para conquistar a confiança dos profissionais e favorecer a adoção dessas tecnologias na prática clínica.

6 CONCLUSÃO

O presente trabalho teve como objetivo desenvolver e validar modelos preditivos para estimar o risco de osteonecrose dos maxilares em pacientes que utilizam bisfosfonatos, utilizando como fonte os dados clínicos estruturados. De acordo com os resultados expostos, o modelo preditivo apresentou alta capacidade de distinguir pacientes em risco, com valor preditivo negativo de 96,1%. Assim é possível concluir que o modelo tem grande utilidade para identificar pacientes de baixo risco, ajudando o clínico a direcionar recursos preventivos e acompanhamento intensivo apenas aos casos realmente necessários.

6.1. Limitações e Considerações

Apesar da performance satisfatória, algumas limitações precisam ser destacadas. A sensibilidade de 33,3% indica que dois terços dos casos de ONM não são identificados. Esse resultado é esperado por ser uma situação de baixa prevalência, mas deve ser comunicado de forma clara: um paciente classificado como baixo risco ainda pode desenvolver ONM.

Outro ponto é o tamanho limitado da amostra, especialmente o baixo número de casos positivos (19), o que restringe a complexidade do modelo e reduz a precisão das estimativas. Estudos com amostras maiores devem aprimorar a sensibilidade sem comprometer a especificidade.

Importante lembrar que os dados utilizados neste estudo vêm do trabalho de Minari (2022) e Minari et al. (2025, no prelo), que reuniu informações de pacientes oncológicos em uso de bisfosfonato intravenoso atendidos no Hospital das Clínicas de Ribeirão Preto-USP. Essa amostra foi escolhida por conveniência e passou por filtros rigorosos, incluindo apenas prontuários completos e pacientes acompanhados pelo serviço de odontologia.

Esses critérios aumentam a confiabilidade dos registros, mas também limitam a generalização dos resultados.

Por isso, os achados devem ser vistos como representativos desse grupo específico e não de todos os usuários de bisfosfonatos. Em estudos futuros, a inclusão de amostras maiores e mais variadas poderá ampliar a precisão e aplicabilidade dos modelos preditivos desenvolvidos,

já que os resultados podem variar conforme características demográficas, protocolos de tratamento ou prevalência da ONM

6.2 Trabalhos futuros: Implicações Clínicas e Perspectivas de Implementação

O modelo desenvolvido contribui para a prática clínica ao oferecer uma ferramenta objetiva de avaliação de risco baseada em evidências. A alta especificidade e o valor preditivo negativo o tornam particularmente útil para identificar pacientes que podem receber acompanhamento menos intensivo, otimizando recursos e reduzindo ansiedade desnecessária.

Embora o presente estudo tenha se concentrado na comparação entre *Random Forest* e Regressão Logística, outros algoritmos poderiam ser explorados em trabalhos futuros, como o XGBoost (*Extreme Gradient Boosting*). Esse método baseia-se em um conjunto sequencial de árvores de decisão, nas quais cada nova árvore corrige os erros da anterior, o que frequentemente resulta em maior capacidade preditiva, especialmente em bases de dados clínicas desbalanceadas. Além disso, o XGBoost mantém compatibilidade com técnicas de interpretabilidade, como os valores SHAP, possibilitando compreender a contribuição individual de cada variável para a previsão. Assim, seu uso poderia aprofundar a análise e servir como comparação adicional à robustez dos resultados obtidos com o Random Forest.

Na prática, o modelo pode ser implementado de diferentes formas:

- calculadoras web para uso em consultório;
- integração em sistemas de prontuário eletrônico;
- aplicativos móveis voltados a profissionais de saúde.

A simplicidade das variáveis de entrada, todas acessíveis na rotina clínica, favorece a adoção ampla. Nesse sentido, a ferramenta se insere no movimento de medicina personalizada em odontologia, permitindo estratégias individualizadas que podem melhorar desfechos clínicos, aumentar a satisfação dos pacientes e tornar o uso de recursos de saúde mais eficiente.

7 REFERÊNCIAS

1. ADADI, A.; BERRADA, M. *Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)*. **IEEE Access**, v. 6, p. 52138-52160, 2018. DOI: 10.1109/ACCESS.2018.2870052.
2. ALAMINO, I. C. P. S. *Análise clínica, radiográfica e de exames hematológicos em pacientes oncológicos sob uso de bisfosfonatos injetáveis*. Dissertação (Mestrado) – USP FORP, 2025.
3. ALSABRY, A. et al. *Enhancing prediction models' performance for breast cancer using SMOTE technique*. In: **IEEE eSmarTA 2023**. p. 1-6.
4. BHARATI, S.; MONDAL, M. R. H.; PODDER, P. *A review on explainable artificial intelligence for healthcare: Why, how, and when?* **IEEE Trans. on Artificial Intelligence**, v. 5, n. 4, p. 1429-1440, 2024. DOI: 10.1109/TAI.2023.3266418.
5. CARNEIRO, R. M. et al. *Transformative deep learning modular segmentation for dental panoramic radiographs*. In: **IEEE CBMS 2025**, São Paulo.
6. CHAWLA, N. V. et al. *SMOTE: Synthetic Minority Over-sampling Technique*. **Journal of Artificial Intelligence Research**, v. 16, p. 321-357, 2002.
7. CHUANG, Y.-S. et al. *Artificial intelligence in extracting diagnostic data from dental records*. **Journal of Dentistry**, 2023. DOI: 10.1016/j.jdent.2023.104747.
8. CİCAK, D.; AVCI, S. *Handling imbalanced data in predictive maintenance: A resampling-based approach*. **Machines**, v. 11, n. 2, p. 239, 2023.
9. ÇİFTÇİ, Y.; AŞANTOĞROL, N. *Machine learning algorithms to predict dental caries risk*. **Journal of Oral Health AI**, 2024.
10. COSTA, Eliana Dantas; GAÊTA-ARAUJO, Hugo; CARNEIRO, José Andery; ZANCAN, Breno Augusto Guerra; BARANAUSKAS, José Augusto; MACEDO, Alessandra Alaniz; TIRAPELLI, Camila. **Development of a dental digital data set for research in artificial intelligence: the importance of labeling performed by radiologists**. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, v. 138, n. 1, p. 205–213, 2024. DOI: 10.1016/j.oooo.2023.12.006
11. DASH, S.; SHAKYAWAR, S. K.; SHARMA, M. *Big data in healthcare: management, analysis and future prospects*. **Journal of Big Data**, v. 6, p. 54, 2019. DOI: 10.1186/s40537-019-0217-0.

12. DUCLOS, V. *Big Data*. **Anthropen**, v. 1, n. 1, 2023. DOI: 10.47854/anthropen.v1i1.52071.
13. EKWONWUNE, E.; UBOCHI, C.; DUROHA, A. *Data Mining as a Technique for Healthcare Approach*. **International Journal of Communications, Network and System Sciences**, v. 15, p. 149-165, 2022. DOI: 10.4236/ijcns.2022.159011.
14. FUTURO DA SAÚDE. *Saúde digital: conheça os avanços e as principais tendências*. Disponível em: <https://futuodasaude.com.br/saude-digital/?form=MG0AV3>. Acesso em: 8 jan. 2025.
15. GHOLAMPOUR, S. *Impact of nature of medical data on machine and deep learning for imbalanced datasets*. **Machine Learning and Knowledge Extraction**, v. 6, n. 2, p. 827-841, 2024.
16. HAN, J.; KAMBER, M.; PEI, J. *Mineração de Dados: Conceitos e Técnicas*. 3. ed. Rio de Janeiro: Elsevier, 2012.
17. KANG, H. et al. *Predictive modeling of dental caries in Korean children using GBDT*. **Healthcare Informatics Research**, 2023.
18. KIVRAK, M. et al. *The impact of the SMOTE method on machine learning and ensemble learning performance...* **Diagnostics**, v. 14, n. 2634, 2024.
19. KOUL, R. *Overview of Data Mining*. **International Journal of Trend in Scientific Research and Development**, v. 4, n. 4, p. 1333–1336, jun. 2020.
20. KUMARI, S.; CHAKRABORTY, D. *A review on Big Data Analytics and Its Tools*. **International Journal of Scientific Research in Science and Technology**, 10 dez. 2023.
21. KWACK, D. W.; PARK, S. M. *Prediction of medication-related osteonecrosis of the jaw (MRONJ) using automated machine learning in patients with osteoporosis*. **Journal of the Korean Association of Oral and Maxillofacial Surgeons**, v. 49, n. 3, p. 135–141, 2023.
22. LAKHOTIA, A. et al. *Artificial intelligence in dentistry: a systematic mapping review*. **Artificial Intelligence in Medicine**, 2025.
23. LEKADIR, K. et al. *FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare*. **Nature Medicine**, v. 31, p. 1-14, 2025.
24. MINARI, V. A. *Avaliação dos fatores de risco para o desenvolvimento de osteonecrose dos maxilares induzida por bisfosfonatos*. Dissertação (Mestrado) – USP FORP, 2022.

25. MINARI, V. A.; DE SANTIS, G. C.; DE MACEDO, L. D.; ORELLANA, M. D.; INNOCENTINI, L.; FERRARI, T. C.; RICZ, H.; CARUSO, S. R.; FERNANDES, T. R.; COVAS, D. T. *Dental evaluation prior to the use of bisphosphonate is a protective factor for jaw osteonecrosis. Supportive Care in Cancer*, Ribeirão Preto, 2025. No prelo
26. MONTEIRO, F.; MELONI, F.; BARANAUSKAS, J. A.; MACEDO, A. A. *Prediction of mortality in Intensive Care Units: a multivariate feature selection. Journal of Biomedical Informatics*, v. 107, p. 103456, 2020. DOI: 10.1016/j.jbi.2020.103456.
27. MUKHERJEE, P. et al. *SCOPE: predicting future diagnoses in office visits using electronic health records. Scientific Reports*, v. 13, n. 11005, 2023.
28. NTI, I. K. et al. *A mini-review of machine learning in big data analytics: applications, challenges, and prospects. Big Data Mining and Analytics*, v. 5, n. 2, p. 81–97, jun. 2022. DOI: 10.26599/BDMA.2021.9020028.
29. OGUNLEYE, J. O. *The Concept of Data Mining. In: Artificial Intelligence*. IntechOpen, 2022. DOI: 10.5772/intechopen.99417.
30. PATEL, P. et al. *Prediction of periodontal disease using ML on dental EHRs. Journal of Public Health Dentistry*, 2022.
31. PATIL, P. B. et al. *Data Mining and Knowledge Discovery in Databases. International Journal of Computer Applications*, v. 175, n. 2, p. 1–6, 2020.
32. PEZOULAS, V. C. et al. *Synthetic data generation methods in healthcare: A review on open-source tools and methods. Computational and Structural Biotechnology Journal*, v. 22, p. 3592–3608, 2024.
33. SADIKU, M. N. O.; EZE, K. G.; MUSA, S. M. *Data Mining in Healthcare. International Journal of Advances in Scientific Research and Engineering (IJASRE)*, v. 4, n. 9, p. 90–92, 2018. DOI: 10.31695/IJASRE.869.
34. SALMI, M. et al. *Handling imbalanced medical datasets: review of a decade of research. Artificial Intelligence Review*, v. 57, p. 273–329, 2024.
35. SHAMOUT, F. E.; ZHU, T.; CLIFTON, D. A. *Machine learning for clinical outcome prediction. IEEE Reviews in Biomedical Engineering*, v. 14, p. 16–30, 2021.
36. UEHARA, V.; TIRAPELLI, C.; MACEDO, A. A. *Dental Second Look AI. INREDD Group, USP*, 2024.
37. UPPAL, S. et al. *Machine learning methods in predicting the risk of malignant transformation of oral potentially malignant disorders: a systematic review.*

- International Journal of Medical Informatics**, v. 186, p. 105421, 2024. DOI: 10.1016/j.ijmedinf.2024.105421.
38. WEI, Z. et al. *Development and validation of a clinical prediction model for osteonecrosis of the jaw in patients receiving zoledronic acid using FAERS and Canadian databases*. **Frontiers in Pharmacology**, v. 15, p. 1456900, 2024.
 39. WIBOWO, S.; MASRURIYAH, S.; RAHMAWATI, H. *Refining diabetes diagnosis models...* **Int. J. of Advances in Intelligent Informatics**, v. 11, n. 1, p. 64-78, 2025.
 40. XIE, B. et al. *Artificial intelligence in dentistry: A bibliometric analysis from 2000 to 2023*. **Journal of Dental Sciences**, v. 19, p. 1722-1733, 2024. DOI: 10.1016/j.jds.2023.10.025.
 41. YANG, G. et al. *Utilizing SMOTE-TomekLink and machine learning to construct a predictive model for elderly medical and daily care services demand*. **Scientific Reports**, v. 15, n. 8446, 2025.