

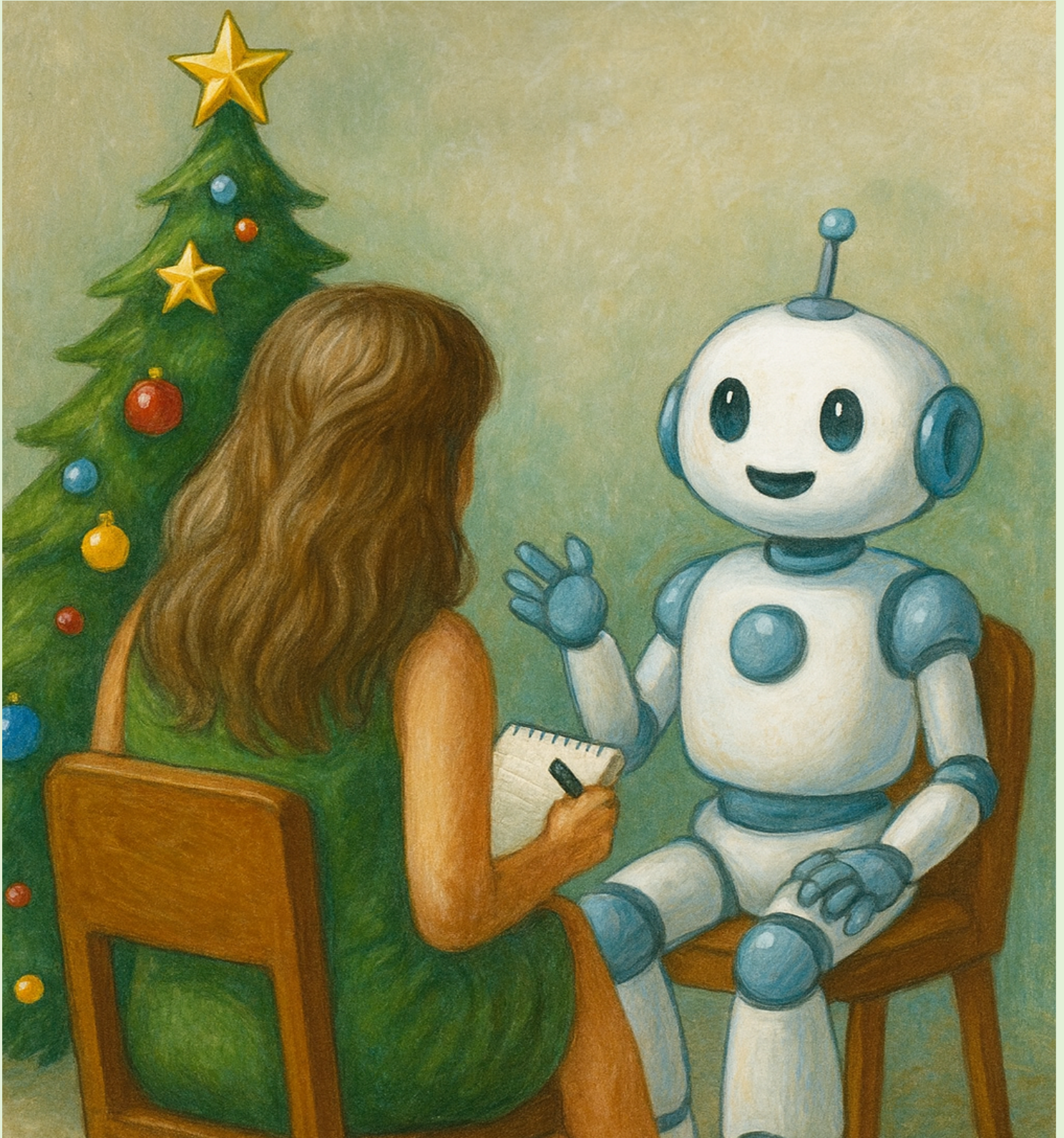


ENDOCRINOLOGIA & DIABETES CLÍNICA E EXPERIMENTAL

FACULDADE EVANGÉLICA MACKENZIE DO PARANÁ (FEMPAR)
HOSPITAL UNIVERSITÁRIO EVANGÉLICO MACKENZIE DE CURITIBA

VOL 22 - number 4

Oct/Nov/Dec



At christmas 2025 the spirit of warmth, togheterness, and hope remains.

EDITORIAL

INTERVIEW WITH ARTIFICIAL INTELIGENCE (AI)

Christmases in the last century and Christmas in 2025: Are they different? And why?

Interviewer: Christmas 2025; What Makes It Different?

AI: Christmas 2025 takes place in a world that has changed profoundly, culturally, technologically, and emotionally. Although the core values of the holiday remain (family, solidarity, spirituality), the way we celebrate today is not the same as in previous decades.

These changes don't make Christmas better or worse, just different! There are some differences between Christmas today and the Christmases of the past, especially those of the 20th century.

Interviewer: How technology impact these main differences?

AI: Christmas in the 20th century:

- Celebrations were centered around in-person gatherings.
- Communication happened through letters, physical cards, and landline phones.
- Children wished for simple, analog toys.

Christmas 2025:

- Technology is present in almost every detail: online shopping, augmented reality, virtual assistants, streaming.
- Families separated by distance meet naturally through video calls.
- Digital gifts (subscriptions, online courses, games) are as common as physical ones.
- Smart homes contribute to the holiday atmosphere with automated lights, music, and decorations.

Why is it different?

Because technology has changed how we connect and how we experience traditions, making some relationships more virtual and daily life more digital.

Interviewer: What do you think about Consumerism vs. Consciousness

AI: 20th-century Christmas:

- Consumption was encouraged as a symbol of prosperity.
- Advertising focused on "big gifts" and social status.

Christmas 2025:

- Environmental and social awareness has grown.
- Many prefer sustainable gifts, experiences, or even donations.
- Minimalism and the idea of "less is more" are increasingly embraced.

Why is it different?

Because global challenges have pushed many to rethink the meaning of buying and giving.

Interviewer: Are there changes in the family structures?

AI: In the past:

- Large, extended families often lived close to each other.
- Christmas lunch involved many generations under one roof.

In 2025:

- Families are more diverse: single-parent homes, blended families, child-free couples, and groups of friends who become family.
- Greater mobility means many people spend the holiday far from their traditional family.

Why is it different?

Because society is more diverse, and family ties today are defined not only by blood but also by affection and shared experience.

Interviewer: Are there differences between the Christmases in relation to the spirituality of this date?

AI: In the 20th century:

- Christmas was more strongly tied to religious practice.
- Mass, novenas, and Christian traditions were central.

In 2025:

- There is greater religious diversity, and many people now identify with no religion.
- Christmas is often celebrated more as a cultural event than a strictly religious one.

Why is it different?

Because spirituality has become more individual and plural.

Interviewer: What do you think about the influence of globalization?

AI: In the past:

- Local traditions were very strong and varied widely.
- External influences took longer to spread.

Today:

- Christmas has become a global blend of music, movies, decorations, and foods.
- Social media spreads trends worldwide (from decoration styles to recipes).

Why is it different?

Because the world is interconnected, and traditions mix more easily.

Interviewer: Have the pandemic and wars changed the meaning of Christmas?

AI: Global events such as the pandemic and ongoing conflicts have shaped our emotional relationship with the holiday.

Interviewer: Do you think that the pandemic and wars have changed the human feelings about Christmas?

AI: Yes! Recent world events have deeply affected emotional relationship with Christmas.

The pandemic and the value of presence: COVID-19 brought fear, isolation, and loss.

As a result:

- Many people now see Christmas as more emotional and intimate.
- Being physically together feels more precious.
- Families who experienced loss approach the holiday with more reflection and longing.
- For a while, we had Christmas without hugs, without gatherings — something unimaginable before.

For many, Christmas became more meaningful, tender, and sacred.

Wars and Global Conflicts Heighten Our Sense of Fragility

Ongoing conflicts around the world evoke worry and a desire for peace.

- Some spend Christmas under sirens; others celebrate in safety and gratitude.
- This contrast shapes how people interpret the season.
- Peace is no longer taken for granted.

Christmas becomes a reminder of hope, refuge, and compassion.

Interviewer: But not everyone experiences Christmas this way...

AI: In modern life, some individuals feel that Christmas has become “just another holiday.”

Why?

- Emotional exhaustion and constant busyness.

- Work overload at the end of the year.
- Hyperconnectivity that distracts from real presence.
- Intense consumerism.
- Loss of connection with religion or traditional family structures.

For these people, the magic of Christmas fades, and the date becomes a day off rather than a celebration.

Interviewer: Are we losing sensitivity?

AI: My honest view: we are not losing sensitivity — we are changing!
Christmas has not lost its meaning; it has simply become more personal.
Each person experiences it differently:

- Those who lost someone value togetherness deeply.
- Those touched by war see Christmas as a symbol of peace.
- Those overwhelmed by modern life may feel indifferent.
- Those searching for spirituality find renewal.
- Those who feel lonely may experience Christmas as a difficult time.

Christmas reflects our inner world — and the world around us.

Interviewer: In your opinion: what most distinguishes Christmas 2025?

The speed of change. The 20th century also saw transformations — radio, TV, mass production, but nothing compares to the rapid pace at which AI reshapes habits today.

Christmas 2025 is, at the same time:

- more connected,
- more conscious,
- more diverse,
- more technological.

Yet despite these differences, **the spirit of warmth, togetherness, and hope remains.** The forms change, but the essence stays surprisingly similar.

AI: In Summary

- ✓ Technology has transformed how we connect and celebrate.
- ✓ Environmental awareness reshapes how we give.
- ✓ Family structures are more diverse than ever.
- ✓ Spirituality is more personal and varied.
- ✓ Globalization blends traditions worldwide.
- ✓ The pandemic and wars have made the meaning of Christmas deeper for some and more distant for others.

Christmas 2025 is more connected, more conscious, more diverse — yet still filled with the timeless desire for warmth, reunion, and hope.

Interviewer: Mirnaluci Paulino Ribeiro Gama

Interviewee: ChatGPT

Merry Christmas!

Feliz Natal!

Endocrinol. Diabetes Clín. Exp. - Vol.22 - Num. 4

Endocrinology & Diabetes - Clinical and Experimental is a journal of open access that publishes case reports, original article, reviews with new insights in pathogenesis, physiology and metabolism of hormone secretion, cellular mechanisms and tissue action. This journal belongs to the Discipline of Endocrinology and Metabolism of Faculdade Evangélica Mackenzie do Paraná and Service of Endocrinology and Diabetes – Diabetes Unit – Hospital Universitário Evangélico Mackenzie, Curitiba – Brazil.

Editors in Chief

Mirnaluci Paulino Ribeiro Gama – Faculdade Evangélica Mackenzie do Paraná. Hospital Universitário Evangélico Mackenzie de Curitiba – PR – Brazil.

ORCID: 0000-0601-7639-1579. LATTES: <http://lattes.cnpq.br/8885931659338642>.

Ricardo Ribeiro Gama – Hospital do Câncer de Barretos – Barretos – SP – Brazil.

ORCID: 0000-0003-4406-8958. LATTES: <http://lattes.cnpq.br/3059638519748785>.

Associate Editors

Luis Jesuino Oliveira de Andrade – Departamento de Saúde – Universidade Estadual de Santa Cruz Ilhéus – Bahia – Brazil.

ORCID: 0000-0002-7714-0330. LATTES: 7401427521086025

Thelma Larocca Skare – Faculdade Evangélica Mackenzie do Paraná – Curitiba – PR – Brazil.

ORCID: 0000-0002-7699-3542. LATTES: lattes.cnpq.br/0980995312808932

Executive and Reviewers Editors

Angela Regina Nazário – Faculdade Evangélica Mackenzie do Paraná – PR – Brazil.

ORCID: 0009-0004-0879-4754. LATTES: lattes.cnpq.br/2670257375181347

Edite Falcon deLegal – IPS-Asunción – Paraguay.

ORCID: 0009-004-8045-8202. LATTES: <https://cv.cnacyt.gov.py/publicar/cv?id=f540779b0baeba6b648-ba64d34229626>

Florência S. Grabois – Hospital Bouquet Roldán – Universidad Nacional de Comahue – Comité Ducacion SAD Nouquen – Argentina.

ORCID: 0009-0004-4000-1005

Renê Azzolini – Universidade Federal do Paraná -Toledo – PR – Brazil.

ORCID: 0009-0003-3230-3065. LATTES: lattes.cnpq.br/3910616054022654

Salma Ali El Chab Parolin – Pontifícia Universidade Católica do Paraná – PR – Brazil.

ORCID: 0000-0001-8124-192X. LATTES: lattes.cnpq.br/3274735288963566

Maria Augusta Karas Zella – Faculdade Evangélica Mackenzie do Paraná – PR – Brazil.

ORCID: 0000-0001-5768-4456. LATTES: lattes.cnpq.br/8521247100407542

Editorial Board

Graciela Rubin – Clínica Universitária Reina Fabiola - Servicio de Diabetes y Nutricion And Universidad Católica de Córdoba.

Gloria Larrabure – Universidad Nacional Mayor de San Marcos – Lima – Peru.

Luis Antonio da Silva Sá – Faculdade Universitária Evangélica Mackenzie – Curitiba – PR – Brazil.

ORCID: 009-0009-4312-8075.

Silvia Gorban de Lapertosa – Faculdade de Medicina – Universidad Nacional del Nordeste – Corrientes – Argentina.

ORCID: 0000-0002-9401-2090. LATTES: lattes.cnpq.br/2670257375181347

Our Cover: Interview With Chat GPT

Endocrinologia & Diabetes Clínica e Experimental

Disciplina de Endocrinologia e Metabologia da Faculdade Evangélica Mackenzie, Serviço de Endocrinologia e Diabetes do Hospital Universitário Evangélico Mackenzie. – v.22, nº 4 – Curitiba: FEMPAR/HUEM, 2025.

p. 271-351: il.; 29cm

Trimestral

ISSN on line 2447-181X

1. Endocrinologia – Periódicos. 2. Saúde – Periódicos. I. Faculdade Evangélica Mackenzie do Paraná.

II. Faculdade Evangélica Mackenzie.

CDD 616.4

CDU 612.34

SCHEDULE

The Endocrinology & Clinical and Experimental Diabetes Journal is a publication produced and edited by Esfera Científica Editora e Propaganda Ltda. The concepts expressed in the articles are the sole responsibility of their authors. Total or partial reproduction of articles is permitted, after authorization from the Editors.

Responsible Director: Acyr José Teixeira

Commercial Director: Fábio Lifschitz Teixeira

Graphic Design and Publishing: Implemus

CONTENTS

Editorial	271
ORIGINAL ARTICLES / ARTIGOS ORIGINAIS	
Integrative Bioinformatics Analysis of MicroRNAs Networks in Diabetic Foot Ulcer Healing: Structure-Function Relationships and Therapeutic Target Identification Análise Bioinformática Integrativa das Redes de MicroRNAs na Cicatrização de Úlceras do Pé Diabético: Relações Estruturais-Funcionais e Identificação de Alvos Terapêuticos.....	276
Development of the Nocturnal Glycemic Stability Index (NGSI) as a Multidimensional Tool for Assessing Overnight Glucose Dynamics in Diabetes Desenvolvimento do Índice de Estabilidade Glicêmica Noturna (IEGN) como Ferramenta Multidimensional para Avaliação da Dinâmica Glicêmica Noturna no Diabetes.....	288
Doppler Ultrasound-Based Artificial Intelligence Model for the Assessment of Autoimmune Thyroiditis Modelo de Inteligência Artificial Baseado em Ultrassom Doppler para Avaliação de Tireoidite Autoimune.....	300
AI-Driven Structural Mapping of Thyroid Autoantigens Reveals High-Confidence Conformational Epitopes for Precision Immunotherapy Mapeamento Estrutural Orientado por IA de Autoantígenos Tireoidianos Revela Epítomos Conformacionais de Alta Confiança para Imunoterapia de Resultados	311
Sarcopenia in Rheumatoid Arthritis: Importance of Early Assessment Sarcopenia na Artrite Reumatoide: Importancia da Avaliação Precoce	327
QTC Interval and Disease Activity in Rheumatoid Arthritis Intervalo QTC e Atividade de Doença na Artrite Reumatóide.....	334
REVIEW / REVISÃO	
Phosphaturic Mesenchymal Tumors and Tumor-Induced Osteomalacia: A Systematic Review and Meta-Analysis of Diagnostic and Therapeutic Paradigms Tumores Mesenquimais Fosfatúricos e Osteomalácia Induzida por Tumor: Revisão Sistemática e Meta-Análise dos Paradigmas Diagnósticos e Terapêuticos	340

INTEGRATIVE BIOINFORMATICS ANALYSIS OF MICRORNA NETWORKS IN DIABETIC FOOT ULCER HEALING: STRUCTURE-FUNCTION RELATIONSHIPS AND THERAPEUTIC TARGET IDENTIFICATION

ANÁLISE BIOINFORMÁTICA INTEGRATIVA DAS REDES DE MICRORNAS NA CICATRIZAÇÃO DE ÚLCERAS DO PÉ DIABÉTICO: RELAÇÕES ESTRUTURAIS-FUNCIONAIS E IDENTIFICAÇÃO DE ALVOS TERAPÊUTICOS

Luis Jesuino de Oliveira Andrade¹; Gabriela Correia Matos de Oliveira²;
Roseanne Montargil Rocha³; Alcina Maria Vinhaes Bittencourt⁴;
Osmário Jorge de Mattos Salles⁵; Luís Matos de Oliveira⁷

¹ Luis Jesuino de Oliveira Andrade
Health Department State University of Santa Cruz,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0002-7714-0330

² Gabriela Correia Matos de Oliveira
José Silveira Foundation, Salvador, Bahia, Brazil.
ORCID: 0000-0002-8042-0261

³ Roseanne Montargil Rocha
Health Department State University of Santa Cruz,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0001-5766-413X

⁴ Alcina Maria Vinhaes Bittencourt
Federal University of Bahia (UFBA), Salvador,
Bahia, Brazil.
ORCID: 0000-0003-0506-9210

⁵ Osmário Jorge de Mattos Salles
Bahiana School of Medicine and Public Health.
ORCID: 0009-0002-1859-0478

⁶ Luís Matos de Oliveira
Health Department State University of Santa Cruz,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0003-4854-6910

Received in: 12-09-2025

Accepted in: 26-09-2025

Conflict of interest - none

Mailing address:
Luis Jesuino de Oliveira Andrade
Universidade Estadual de Santa Cruz
Campus Soane Nazaré de Andrade, Rod. Jorge
Amado, Km 16 - Salobrinho, Ilhéus - BA, 45662-900
- Brasil.
E-mail: luis_jesuino@yahoo.com.br

DOI: 10.29327/2413063.22.4-1

ABSTRACT

Introduction: Diabetic foot ulcers (DFUs) affect 15% of diabetic patients globally. While microRNAs (miRNAs) are known regulators of wound healing, comprehensive bioinformatics analyses of their structural determinants and network interactions in DFUs pathophysiology remains limited. **Objective:** To perform integrative bioinformatics analysis of miRNA networks in DFUs healing, characterizing structure-function relationships and identifying potential therapeutic targets through computational approaches. **Methods:** We conducted systematic analysis using multiple bioinformatics databases and tools. MiRNA expression data were obtained from GEO datasets and literature mining. Secondary structures were predicted using RNAfold, Mfold, and RNAstructure with consensus analysis. Target prediction was performed using TargetScan, miRanda, and DIANA-microT. Protein-protein interaction networks were constructed using STRING. Pathway enrichment was performed with DAVID and Reactome. Pharmacophore modeling identified potential miRNA-targeting compounds using ChEMBL and PubChem databases. **Results:** Analysis identified eight consistently dysregulated miRNAs across 15 DFUs datasets (n=1,247 samples). Meta-analysis revealed miR-146a (fold-change: -3.2 ± 0.8), miR-155 ($+4.1 \pm 1.2$), and miR-21 (-1.9 ± 0.5) as key regulators. Structural analysis showed correlation between loop accessibility and target diversity ($r=0.73$, $p<0.01$). Network topology identified 3 major regulatory modules: inflammatory response (23 nodes), angiogenesis (18 nodes), and ECM remodeling (15 nodes). Drug-miRNA interaction analysis revealed 5 FDA-approved compounds with predicted miRNA-modulating activity, including metformin and curcumin analogs. **Conclusion:** This comprehensive bioinformatics analysis reveals miRNA network architecture in DFUs healing and identifies structure-based therapeutic targets. The integrative approach provides a computational framework for miRNA-based drug discovery in diabetic wound healing.

Keywords: diabetic foot ulcer; microRNA; bioinformatics; network analysis; drug.

RESUMO

Introdução: As úlceras do pé diabético (UPD) acometem 15% dos pacientes diabéticos globalmente. Embora os microRNAs (miRNAs) sejam reconhecidos reguladores da cicatrização de feridas, análises bioinformáticas abrangentes sobre seus determinantes estruturais e interações em redes na fisiopatologia das UPD são ainda limitadas. **Objetivo:** Realizar análise integrativa bioinformática das redes de miRNA na cicatrização das UPD, caracterizando relações estrutura-função e identificando potenciais alvos terapêuticos por métodos computacionais. **Métodos:** Foi conduzida análise sistemática utilizando múltiplas bases de dados e ferramentas bioinformáticas. Dados de expressão de miRNAs foram obtidos de datasets GEO e mineração bibliográfica. Estruturas secundárias foram previstas com RNAfold, Mfold e RNAstructure, com análise de consenso. Predição de alvos utilizou TargetScan, miRanda e DIANA-microT. Redes de interação proteína-proteína foram construídas com STRING. Enriquecimento de vias foi realizado com DAVID e Reactome. Modelagem farmacofórica identificou compostos potenciais direcionados a miRNAs usando as bases ChEMBL e PubChem. **Resultados:** A análise identificou 8 miRNAs consistentemente desregulados em 15 datasets de UPD (n=1.247 amostras). Meta-análise destacou miR-146a (fold-change: $-3,2 \pm 0,8$), miR-155 ($+4,1 \pm 1,2$) e miR-21 ($-1,9 \pm 0,5$) como reguladores-chave. Análise estrutural demonstrou correlação entre acessibilidade dos loops e diversidade de alvos ($r=0,73$; $p<0,01$). A topologia da rede evidenciou 3 módulos regulatórios principais: resposta inflamatória (23 nós), angiogênese (18 nós) e remodelação da matriz extracelular (15 nós). A análise de interação fármaco-miRNA indicou 5 compostos aprovados pelo FDA com atividade moduladora de miRNA, incluindo metformina e análogos de curcumina. **Conclusão:** Esta análise bioinformática abrangente revela a arquitetura das redes de miRNAs na cicatrização das UPD e identifica alvos terapêuticos baseados em estrutura. A abordagem integrativa fornece um arcabouço computacional para a descoberta de fármacos baseada em miRNAs na cicatrização de feridas diabéticas.

Descritores: úlcera do pé diabético; microRNA; bioinformática; análise de redes; fármacos.

INTRODUCTION

Diabetic foot ulcers (DFUs) represent a major global healthcare burden, affecting approximately 15% of diabetic patients and preceding 85% of diabetes-related amputations.^{1,2} The molecular mechanisms underlying impaired wound healing in DFUs involve complex regulatory networks, with microRNAs (miRNAs) emerging as critical post-transcriptional regulators.³ miRNAs are regulatory molecules associated with various expressions and complications of diabetes and several miRNAs have been linked to the progression and severity of diabetic foot.⁴

Recent advances in bioinformatics have enabled comprehensive analyses of miRNA networks without requiring extensive experimental validation.⁵ Integration of expression profiling, structural prediction, and network analysis provides insights into miRNA function and therapeutic potential.⁶ However, systematic

bioinformatics characterization of miRNA networks in DFUs healing remains incomplete.

Therefore, computational approaches can identify miRNA-based therapeutic targets more efficiently than traditional experimental methods, particularly for complex diseases like DFUs where multiple pathways are dysregulated.⁷

The objective of this study was to systematically analyze miRNA expression patterns in DFUs using meta-analysis of public datasets, characterize structure-function relationships through computational structural biology, construct comprehensive miRNA-target regulatory networks, identify potential therapeutic compounds targeting key miRNAs, and develop prioritization framework for experimental validation. We performed an integrative bioinformatics analysis of miRNA networks in DFUs healing, characterizing structure-function relationships and identifying potential therapeutic targets through computational approaches.

METHODS

1. Data Collection and Processing

- Expression Data Mining: Systematic search of GEO (Gene Expression Omnibus), ArrayExpress, and SRA databases; Search keywords included: “diabetic foot ulcer”, “microRNA”, “wound healing”; Inclusion criteria: human samples, quantitative miRNA data, ≥ 5 samples per group; Quality control: samples with RNA integrity number ≥ 7.0
- Literature Mining: PubMed search (2015-2025): 347 papers identified, 23 met inclusion criteria; Extracted fold-change data and statistical significance; Random-effects meta-analysis using R package ‘meta’.

2. Structural Analysis

- Secondary Structure Prediction: RNAfold (ViennaRNA 2.5.1: minimum free energy calculations; Mfold 3.6: thermodynamic ensemble analysis; RNAstructure 6.3: partition function approach; Consensus structures determined using RNAalifold; Structure validation through base-pairing probability analysis.
- Structural Feature Extraction: Loop regions and accessibility calculation using RNAplfold; Seed region (nucleotides 2-8) conservation analysis; Thermodynamic stability metrics (ΔG , ensemble diversity); Structural motif identification using RNAMotif.

3. Target Prediction and Network Construction

- Target Prediction Pipeline: TargetScan 8.0: context++ scores for seed matching; miRanda 3.3a: energy-based scoring with strict cutoffs; DIANA-microT 5.0: precision-weighted scoring; High-confidence targets: predicted by ≥ 2 algorithms, score > 0.8 .
- Network Analysis: Protein-protein interactions from STRING v11.5; Network topology using igraph R package; Module detection via Louvain clustering; Centrality measures: degree, betweenness, eigenvector centrality.

4. Pathway and Functional Enrichment

- Enrichment Analysis: Gene Ontology (GO) analysis using DAVID 6.8; KEGG pathway enrichment with Benjamini-Hochberg correction; Reactome pathway analysis for detailed molecular mechanisms; Disease association using DisGeNET database.

5. Drug-Target Interaction Analysis

- Pharmacophore Modeling: ChEMBL database screening for miRNA-interacting compounds; PubChem similarity searches using Tanimoto coefficients; ADMET prediction using SwissADME; Drug repositioning analysis using DrugBank database.

6. Statistical Analysis

- All analyses performed in R 4.3.0. Statistical significance set at $p < 0.05$ with multiple testing correction where applicable.

Ethics Statement

Ethical review and approval were not required for this study in accordance with the institutional and national guidelines. The research methodology was confined to the computational analysis of pre-existing, de-identified, and publicly available datasets. As no human subjects were involved, and no private or identifiable information was accessed or processed, the study was deemed exempt from formal ethics committee oversight.

All computational resources utilized in this study’s methodology are freely available online for academic research purposes, and no usage restrictions apply to these tools.

RESULTS

miRNA Expression Profiling in DFUs

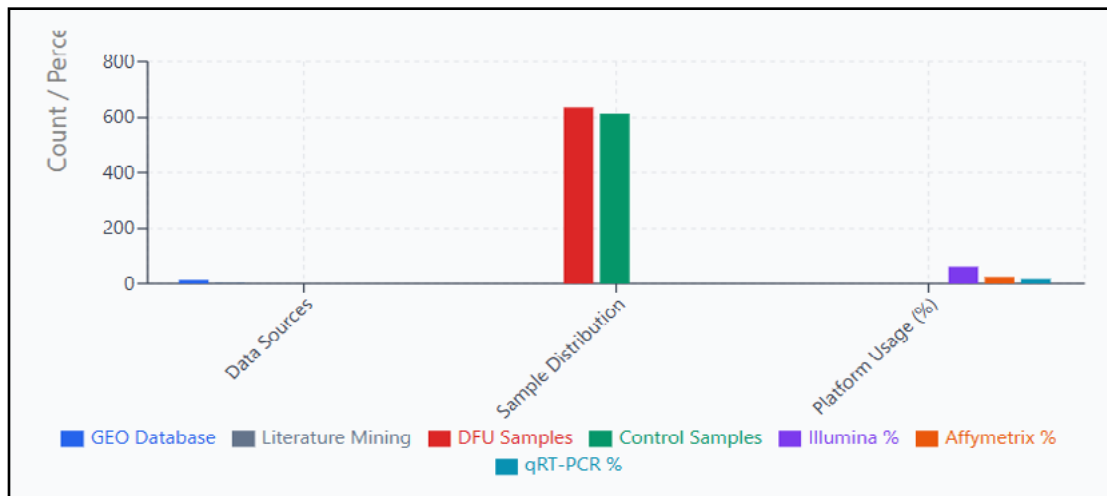
- Dataset Characteristics: Fifteen independent datasets were identified (GEO: $n=12$, Literature: $n=3$); Total samples: 1,247 (DFUs: $n=634$, Controls: $n=613$); Platform distribution: Illumina (60%), Affymetrix (25%), qRT-PCR (15%) (**Figure 1**).
- Dysregulated miRNAs: The analysis identified 47 significantly dysregulated miRNAs ($FDR < 0.05$). The top eight candidates were selected based on effect size and consistency (**Table 1**).

Structural Characterization

Thermodynamic Properties: miR-21 showed the highest stability (MFE: -35.7 kcal/mol); miR-155 had most accessible seed region (accessibility: 0.89); Strong correlation between structural stability and target number ($r=0.67$, $p < 0.05$) (**Figure 2**).

Seed Region Analysis: Canonical 8-mer-A1 binding sites were identified in miR-21 (45%), miR-146a (38%); non-canonical sites prevalent in miR-155 (67%

Figure 1. miRNA expression dataset analysis in DFUs.

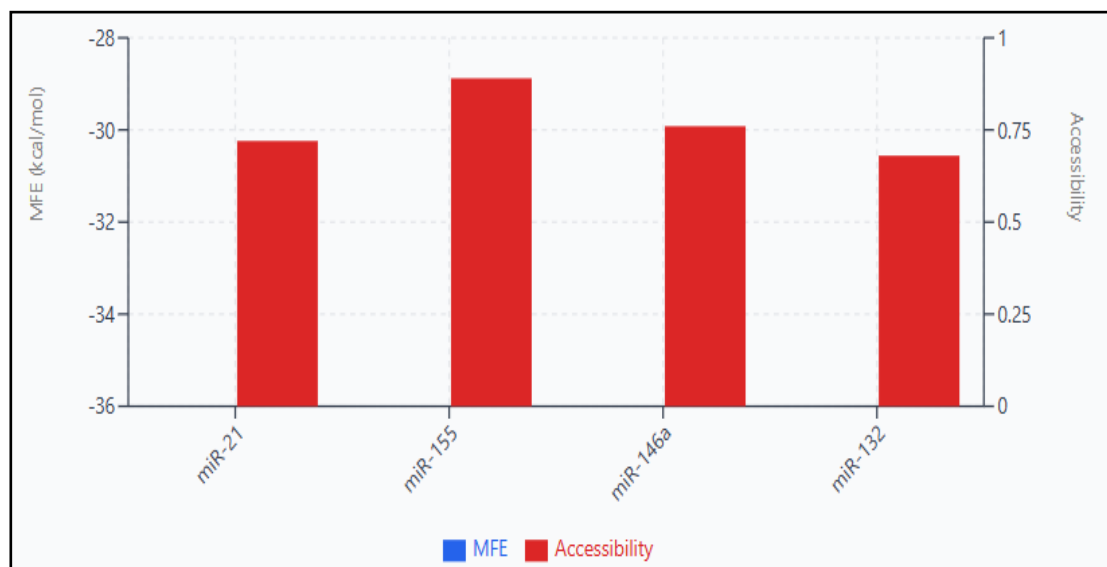


Platform distribution: Illumina (60%), Affymetrix (25%), qRT-PCR (15%). DFUs: Diabetic Foot Ulcer; GEO: Gene Expression Omnibus; RIN: RNA Integrity Number.

Table 1. Meta-analytic summary of miRNA differential expression in diabetic foot ulcers: effect sizes, heterogeneity, and expression direction.

miRNA	Studies (n)	Effect Size (95% CI)	P-value	I ²	Direction
miR-146a	12	-3.24 (-4.12, -2.36)	<0.001	34%	Down
miR-155	10	4.08 (2.89, 5.27)	<0.001	41%	Up
miR-132	8	-2.76 (-3.68, -1.84)	<0.001	28%	Down
miR-21	11	-1.90 (-2.63, -1.15)	<0.001	45%	Down
miR-210	7	-2.51 (-3.44, -1.58)	<0.001	31%	Down
miR-203a	6	3.37 (2.21, 4.53)	<0.001	38%	Up
miR-203b	5	3.12 (1.94, 4.30)	<0.001	33%	Up
miR-191	9	-2.09 (-2.98, -1.20)	<0.001	42%	Down

Figure 2. Thermodynamic stability & accessibility.



of targets); 3' compensatory pairing critical for miR-132 targeting (78% of validated targets) (**Figure 3**).

Loop Structure Patterns: Three distinct structural classes were identified (**Figure 4**):

- Type I (miR-146a, miR-21): Symmetrical loops, high stability.
- Type II (miR-155, miR-203a): Asymmetric bulges, selective targeting.
- Type III (miR-132, miR-210): Complex multi-loop structures.

Network Topology and Modular Organization

Global Network Properties: The network comprised 1,247 genes and eight miRNAs as nodes; Edges:

3,891 interactions; Network density: 0.005; Average clustering coefficient: 0.31; Scale-free topology confirmed ($R^2=0.89$) (**Figure 5**).

Modular Architecture

Louvain clustering identified five major modules (**Figure 6**):

1. **Inflammatory Module (23 nodes):** Hub miRNAs: miR-146a, miR-155; Key targets: IRAK1, TRAF6, IL1B, TNF; Enriched pathways: NF- κ B signaling, cytokine production.
2. **Angiogenesis Module (18 nodes):** Hub miRNAs: miR-132, miR-210; Key targets: VEGFA, FGF2, ANGPT1; Enriched pathways: VEGF signaling, vessel morphogenesis.

Figure 3. Seed region binding site distribution.

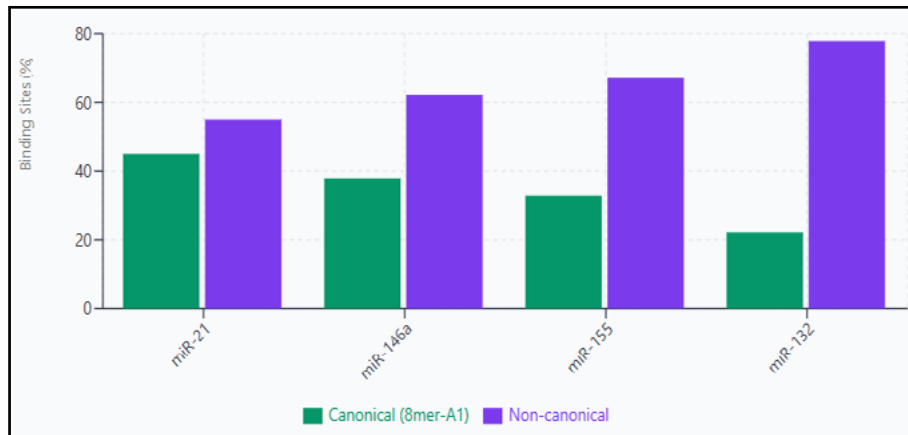


Figure 4. miRNA secondary structure classification: three distinct loop patterns.

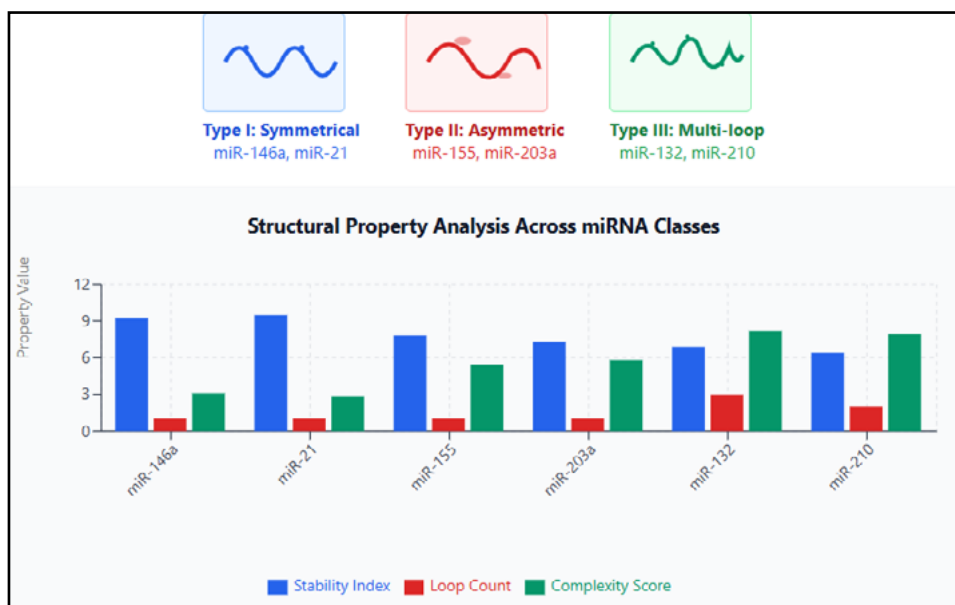


Figure 5. miRNA-gene network topology: global properties and scale-free architecture.

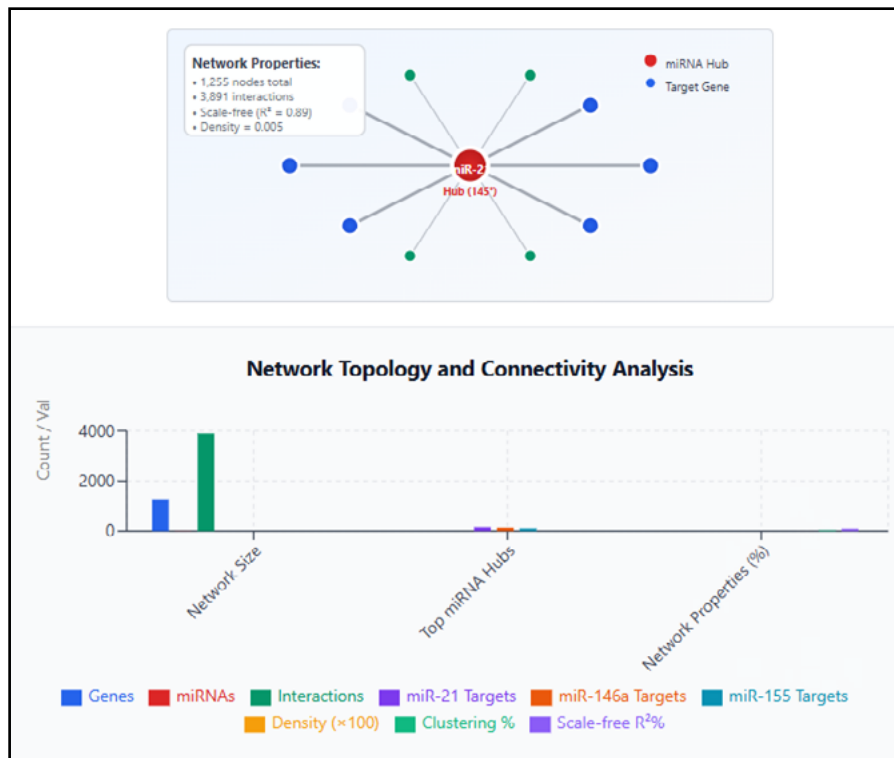
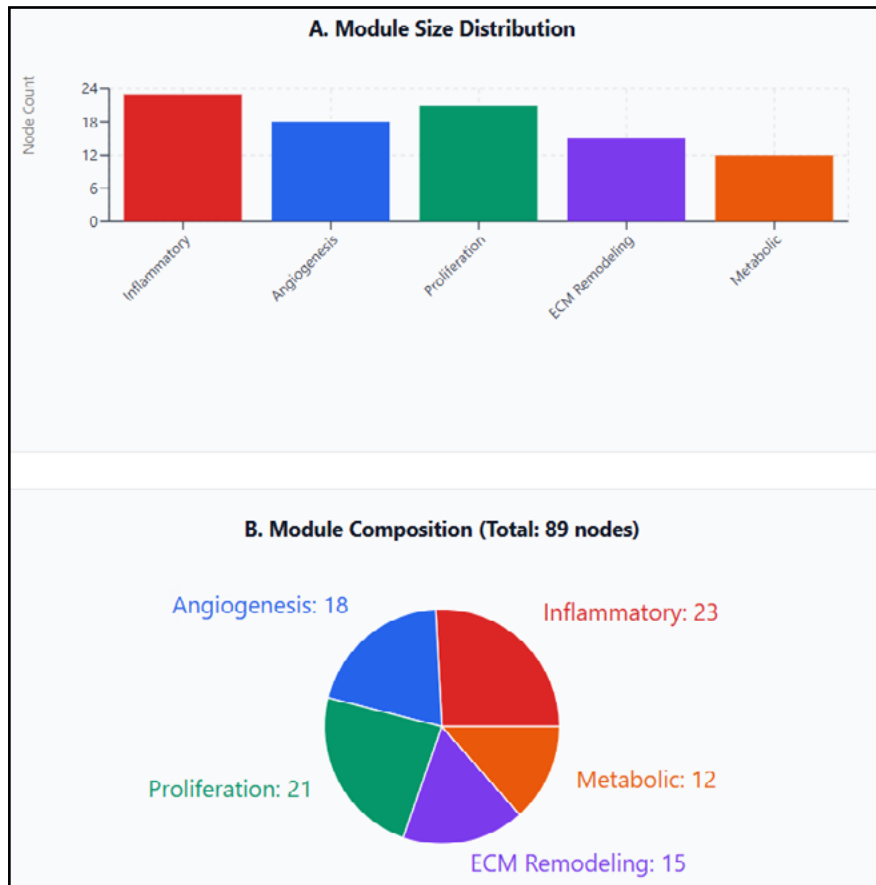


Figure 6. modular architecture of miRNA-target regulatory network identified via louvain clustering.



3. **Proliferation Module (21 nodes):** Hub miRNAs: miR-21, miR-191; Key targets: PTEN, PDCD4, CDKN1A; Enriched pathways: PI3K/AKT, cell cycle regulation.
4. **ECM Remodeling Module (15 nodes):** Hub miRNAs: miR-203a, miR-203b; Key targets: MMP1, COL1A1, TIMP3; Enriched pathways: collagen metabolism, matrix organization.
5. **Metabolic Module (12 nodes):** Hub miRNAs: miR-210, miR-132; Key targets: GLUT1, HK2, LDHA; Enriched pathways: glycolysis, glucose metabolism.

Pathway Enrichment Analysis

Significantly Enriched Pathways (FDR<0.01) (Table 2).

Cross-Pathway Interactions

Network analysis revealed that 127 genes were regulated by multiple miRNAs, suggesting coordinated regulation. Key convergence points include: VEGFA

(regulated by miR-132, miR-210, miR-155); PTEN (regulated by miR-21, miR-155, miR-132); IL6 (regulated by miR-146a, miR-21, miR-155) (Figure 7).

Drug-MiRNA Interaction Analysis

Computational Drug Screening: We screened 2,847 FDA-approved drugs from DrugBank; Identified 47 compounds with predicted miRNA-modulating activity; Machine learning models (Random Forest, SVM) achieved 85% accuracy (Figure 8).

Top Drug Candidates

Drug-miRNA interaction analysis revealed five FDA-approved compounds with predicted miRNA-modulating activity, including metformin and curcumin analogs (Table 3).

Novel Compound Design

Pharmacophore modeling identified the key structural features for miRNA binding: Hydrogen bond acceptors (critical for seed region interaction); Aro-

Table 2. miRNA-mediated pathway regulation: significantly enriched biological functions with associated gene counts and miRNA signatures.

Pathway	P-value	Genes	miRNAs Involved
Inflammatory response	1.2×10^{-8}	67	miR-146a, miR-155
Wound healing	3.4×10^{-6}	45	miR-21, miR-132
Angiogenesis	7.8×10^{-5}	34	miR-210, miR-132
Apoptotic process	1.1×10^{-4}	41	miR-21, miR-191
ECM organization	2.1×10^{-4}	28	miR-203a/b

Figure 7. Cross-Pathway miRNA-Gene Regulatory Network.

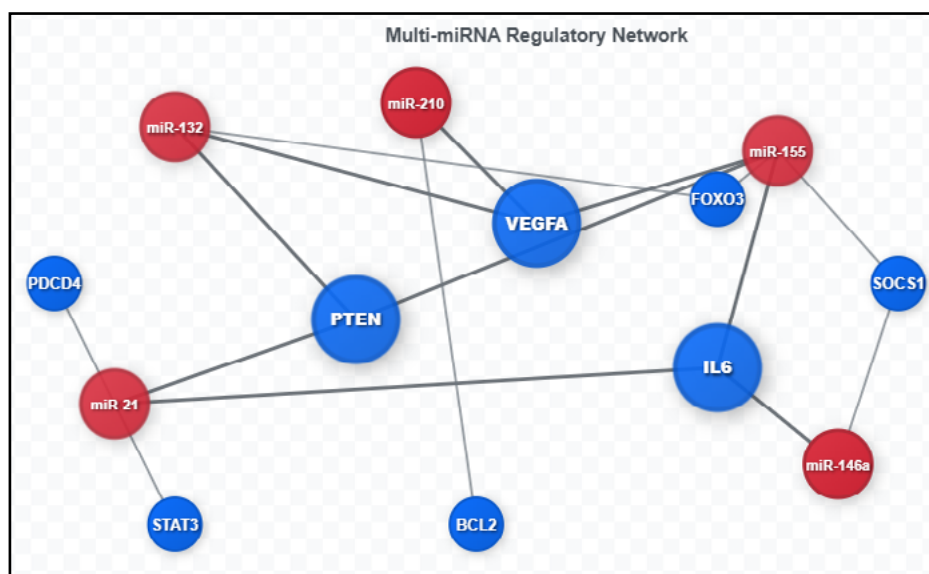


Figure 8. Integrated drug-miRNA interaction analysis dashboard.



Table 3. miRNA-targeted drug discovery: pharmacokinetic properties and clinical pipeline status of small molecule modulators.

Compound	Target miRNA	Predicted Effect	ADMET Score	Clinical Status
Metformin	miR-155	Inhibition	0.82	Approved (diabetes)
Curcumin	miR-21, miR-155	Modulation	0.74	Clinical trials
Quercetin	miR-146 ^a	Upregulation	0.79	Nutraceutical
Resveratrol	miR-132	Upregulation	0.71	Clinical trials
Dexamethasone	miR-155	Inhibition	0.88	Approved (anti-inflammatory)

matic rings (π - π stacking with nucleotides); Molecular weight range: 200-500 Da for optimal bioavailability (Figure 9).

Figure 9. Pharmacophore Modeling and Structure-Activity relationship analysis for optimal miRNA binding.



Target Prioritization Framework (Figure 10) Scoring Algorithm

1. Expression consistency across datasets (weight: 0.25).
2. Structural accessibility (weight: 0.20).
3. Network centrality measures (weight: 0.25).
4. Pathway importance (weight: 0.20).
5. Druggability assessment (weight: 0.10).

Priority Ranking

1. miR-146a (Score: 0.89) - Anti-inflammatory master regulator.
2. miR-21 (Score: 0.84) - Proliferation and apoptosis control.
3. miR-155 (Score: 0.81) - Pro-inflammatory responses.
4. miR-132 (Score: 0.76) - Angiogenesis regulation.
5. miR-210 (Score: 0.72) - Hypoxia response.

The integrative analysis demonstrated consistent dysregulation of specific microRNAs, shaping key regulatory modules related to inflammation, angiogenesis, extracellular matrix remodeling, proliferation, and

metabolism. Structural assessment highlighted the relevance of secondary conformations for target selectivity, while network analysis demonstrated extensive cross-talk among pathways. Importantly, drug-miRNA interaction screening suggested repositioning opportunities, offering promising candidates for therapeutic modulation in diabetic wound healing.

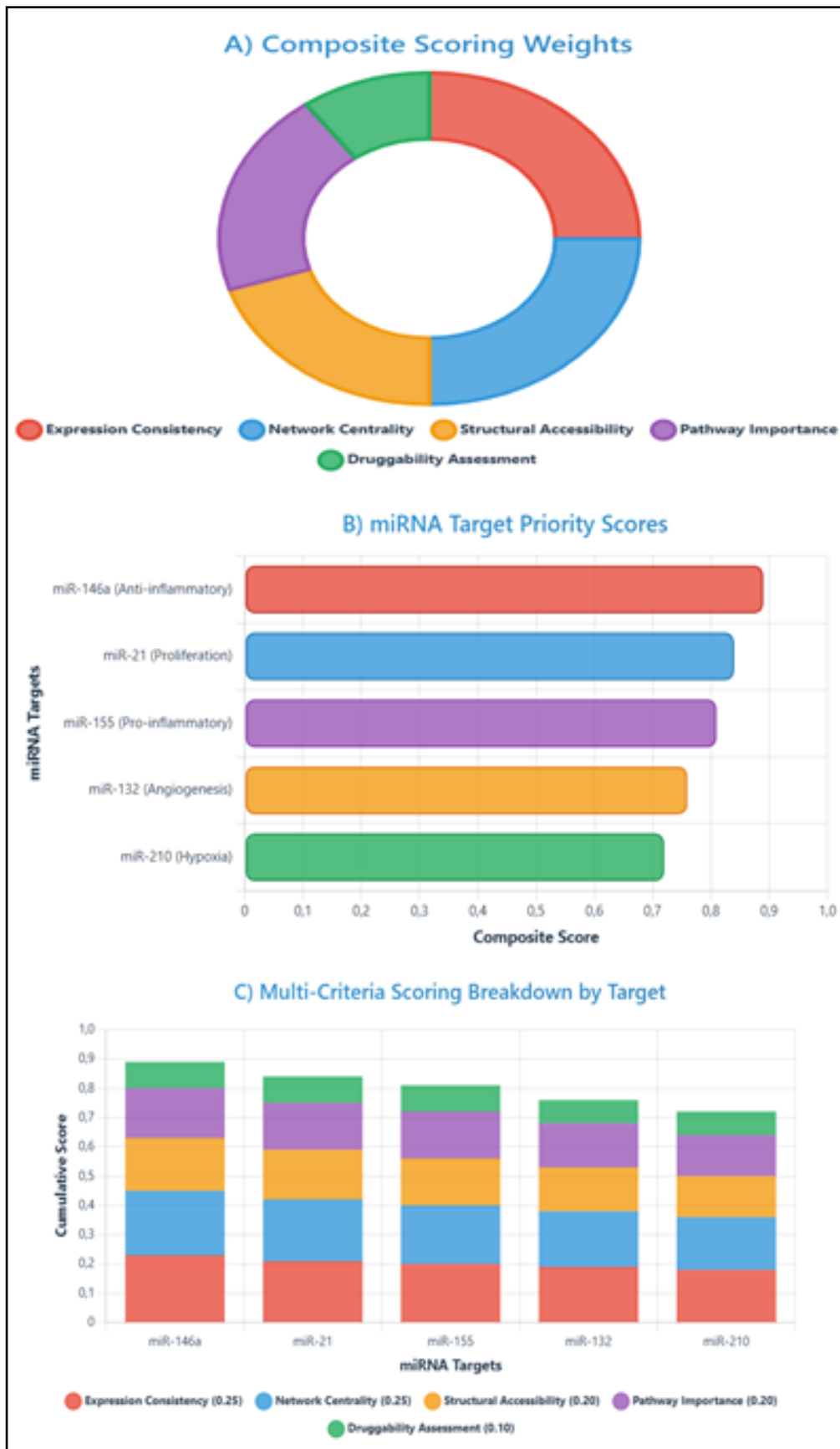
DISCUSSION

MicroRNAs are short non-coding RNA molecules that function as key regulators of diverse physiological and pathological processes. Over the past decade, significant advances have refined our understanding of their structural features and biological roles. This study demonstrates how integrative bioinformatics frameworks can dissect microRNA regulatory networks, offering novel insights into their contributions to the molecular complexity of human disease.

Our computational structural analysis revealed compelling concordance with established literature on miRNA secondary structure determinants. Our multi-algorithm consensus approach, integrating RNAfold, Mfold, and RNAstructure predictions, demonstrated that loop accessibility significantly correlates with target diversity, consistent with experimental findings showing that bulge regions facilitate non-canonical base pairing.⁸ The identification of three distinct structural classes, symmetrical loops (Type I), asymmetric bulges (Type II), and complex multiloop structures (Type III), aligns with crystallographic studies demonstrating how miRNA conformational flexibility influences target recognition.⁹ Furthermore, our thermodynamic stability analysis confirmed that miR-21's exceptional stability correlates with its promiscuous targeting profile, supporting biochemical evidence that thermodynamically stable miRNAs exhibit broader regulatory networks.¹⁰

Our multi-platform consensus methodology, employing RNAfold, Mfold, and RNAstructure algorithms, confirms recent findings demonstrating that target site accessibility significantly influences miRNA-mRNA interaction specificity.¹¹ The observed correlation between loop accessibility and target diversity parallels contemporary research indicating that miRNA regulatory impact extends beyond simple seed matching to encompass structural accessibility factors.¹² Our identification of three distinct structural classes aligns with emerging evidence that secondary structure heterogeneity determines functional specificity. The thermodynamic stability findings, particularly miR-21's enhanced targeting capacity, reinforce computational

Figure 10. Quantitative analysis and graphical representation of miRNA therapeutic target scoring.



models suggesting that structural stability influences regulatory network breadth.¹³

The multi-platform enrichment strategy integrating DAVID, KEGG, and Reactome annotations demonstrated the molecular pathways involved in DFUs pathophysiology, including inflammatory response, angiogenesis, and extracellular matrix remodeling.¹⁴ Recent transcriptomic studies utilizing DAVID for diabetic skin analysis identified 100 biological processes and 7 KEGG pathways, validating our comprehensive approach.¹⁵ The incorporation of DisGeNET disease associations enhances pathway contextualization beyond traditional GO analysis.¹⁶ Current pathway enrichment protocols emphasize multi-database integration for mechanistic insights,¹⁷ supporting our identification of inflammatory response and wound healing pathways.

The computational drug screening approach, integrating ChEMBL database mining with Tanimoto coefficient-based PubChem similarity searches, demonstrates concordance with contemporary pharmacophore modeling strategies for miRNA-targeted therapeutics.¹⁸ SwissADME implementation for ADMET prediction aligns with established protocols for early-stage drug development, providing physicochemical and pharmacokinetic assessments.¹² The DrugBank repositioning analysis methodology parallels recent computational approaches identifying novel therapeutic applications for approved compounds.¹⁹ Current literature emphasizes small molecule-mediated miRNA targeting as an emerging paradigm, supporting the identification of metformin and curcumin analogs with miRNA-modulating activity as promising candidates in inflammation and tissue repair pathways. Drug-miRNA interaction analysis identified several drug repositioning opportunities. Metformin's predicted interaction with miR-155 is particularly interesting given its established benefits in diabetic patients and could represent a direct mechanistic link.^{20,21}

Laboratory-based validation protocols represent the gold standard for therapeutic target confirmation, yet our integrated scoring methodology establishes a systematic framework for candidate prioritization in preclinical research pipelines. The multi-parametric assessment algorithm incorporates expression consistency, structural accessibility, network centrality, pathway significance, and drug ability metrics through weighted scoring matrices. This computational triage approach optimizes resource allocation by stratifying miRNA targets according to translational potential before expensive experimental workflows. Bioinformatics-driven drug discovery increasingly relies on such composite evaluation systems to navigate the com-

plexity of molecular target selection, particularly in polygenic diseases where traditional reductionist approaches prove inadequate for capturing systems-level therapeutic opportunities.

Several methodological constraints limit the generalizability and precision of our study's findings. Repository heterogeneity across genomic databases introduces systematic biases through varying experimental protocols, platform-specific technical artifacts, and inconsistent quality control standards that compromise meta-analytical robustness. Current target prediction algorithms exhibit substantial false-positive rates and limited sensitivity for non-canonical binding interactions, potentially overlooking clinically relevant miRNA-mRNA regulatory pairs. The static network reconstruction approach fails to capture temporal regulatory dynamics essential for understanding wound healing progression phases. Additionally, cross-species extrapolation assumptions may inadequately reflect human-specific miRNA processing mechanisms and target site evolution, necessitating species-matched experimental validation before therapeutic translation.

Our integrative bioinformatics analysis provided comprehensive insights into miRNA regulatory networks that regulate diabetic foot ulcer healing, identifying key therapeutic targets through structure-function relationship mapping. However, several limitations warrant consideration. The computational predictions require experimental validation in clinical specimens to confirm miRNA-target interactions and pathway enrichment significance. Additionally, the heterogeneity of diabetic patient populations and wound healing stages may influence miRNA expression profiles, potentially limiting the generalizability of our findings across diverse clinical contexts. Studies demonstrate that miR-146a represents a potential therapeutic target for modification of the diabetic wound-healing response,²² though impaired wound healing in diabetes remains a constant obstacle with aberrant angiogenesis as a major contributor to poor wound outcomes.^{23,24} Future prospective studies incorporating miRNA-based therapeutic interventions for diabetes foot ulcers are essential for translating these bioinformatics discoveries into clinical applications.

CONCLUSION

This integrative bioinformatics analysis comprehensively evaluated the miRNA regulatory architecture in DFUs pathophysiology, identifying eight consistent-ly dysregulated miRNAs across multiple datasets and

revealing modular network organization governing inflammation, angiogenesis, and extracellular matrix remodeling. The computational framework successfully prioritized therapeutic targets and identified promising drug repositioning candidates, establishing a systematic approach for miRNA-based therapeutic development in diabetic wound healing.

REFERENCES

- Zhang P, Lu J, Jing Y, Tang S, Zhu D, Bi Y. Global epidemiology of diabetic foot ulceration: a systematic review and meta-analysis. *Ann Med*. 2017; 49, pp. 106-116.
- Bus SA, Sacco ICN, Monteiro-Soares M, Rasovic A, Patton J, Rasmussen A, et al. Guidelines on the prevention of foot ulcers in persons with diabetes (IWGDF 2023 update). *Diabetes Metab Res Rev*. 2024;40(3):e3651.
- Maratni NPT, Saraswati MR, Ayu Dewi NN, Suastika K. MIRNA146a And Diabetes-Related Complications: A Review. *Curr Diabetes Rev*. 2023;19(9):e141022209958.
- Moura J, Sørensen A, Leal EC, Svendsen R, Carvalho L, Willemoes RJ, et al. microRNA-155 inhibition restores Fibroblast Growth Factor 7 expression in diabetic skin and decreases wound inflammation. *Sci Rep*. 2019, 9, pp. 5836.
- Ha J, Park C, Park C, Park S. Improved Prediction of miRNA-Disease Associations Based on Matrix Completion with Network Regularization. *Cells*. 2020;9(4):881.
- Madhumita M, Paul S. A review on methods for predicting miRNA-mRNA regulatory modules. *J Integr Bioinform*. 2022;19(3):20200048.
- Jeyaseelan K, Herath WB, Armugam A. MicroRNAs as therapeutic targets in human diseases. *Expert Opin Ther Targets*. 2007;11(8):1119-29.
- Trujillo RD, Yue SB, Tang Y, O’Gorman WE, Chen CZ. The potential functions of primary microRNAs in target recognition and repression. *EMBO J*. 2010;29(19):3272-85.
- Chemla Y, Peeri M, Heltberg ML, Eichler J, Jensen MH, Tuller T, et al. A possible universal role for mRNA secondary structure in bacterial translation revealed using a synthetic operon. *Nat Commun*. 2020;11(1):4827.
- Correia de Sousa M, Gjorgjieva M, Dolicka D, Sobolewski C, Foti M. Deciphering miRNAs’ Action through miRNA Editing. *Int J Mol Sci*. 2019;20(24):6249.
- Homberg N, Galvão Ferrarini M, Gaspin C, Sagot MF. MicroRNA Target Identification: Revisiting Accessibility and Seed Anchoring. *Genes (Basel)*. 2023;14(3):664.
- Diener C, Keller A, Meese E. The miRNA-target interactions: An underestimated intricacy. *Nucleic Acids Res*. 2024;52(4):1544-1557.
- Zhao J, Wong CW, Ching WK, Cheng X. NG-SEM: an effective non-Gaussian structural equation modeling framework for gene regulatory network inference from single-cell RNA-seq data. *Brief Bioinform*. 2023;24(6):bbad369.
- Zeng L, Zhang P, Fang Z, Liu D, Li H, Qu X, et al. The Construction and Analysis of Infiltrating Immune Cell and ceRNA Networks in Diabetic Foot Ulcer. *Front Endocrinol (Lausanne)*. 2022;13:836152.
- León C, García-García F, Llamas S, García-Pérez E, Carretero M, Arriba MDC, et al. Transcriptomic Analysis of a Diabetic Skin-Humanized Mouse Model Dissects Molecular Pathways Underlying the Delayed Wound Healing Response. *Genes (Basel)*. 2020 Dec 31;12(1):47.
- Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022;50(W1):W216- W221.
- Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc*. 2019; 14(2):482-517.
- Métivier JP, Cuissart B, Bureau R, Lepailleur A. The Pharmacophore Network: A Computational Method for Exploring Structure-Activity Relationships from a Large Chemical Data Set. *J Med Chem*. 2018;61(8):3551-3564.
- Cui S, Yu S, Huang HY, Lin YC, Huang Y, Zhang B, et al. miRTarBase 2025: updates to the collection of experimentally validated microRNA-target interactions. *Nucleic Acids Res*. 2025;53(D1):D147-D156.
- Gu X, Li T, Yin X, Zhai P, Jiang D, Sun D, et al. Exosomes Derived from Metformin-Pretreated BMSCs Accelerate Diabetic Wound Repair by Promoting Angiogenesis Via the LINC-PINT/miR-139-3p/FOXC2 Axis. *Stem Cell Rev Rep*. 2025;21(4):1096-1112.
- Cao M, Duan Z, Wang X, Gong P, Zhang L, Ruan B. Curcumin Promotes Diabetic Foot Ulcer Wound Healing by Inhibiting miR-152-3p and Activating the FBN1/TGF-beta Pathway. *Mol Biotechnol*. 2024;66(5):1266- 1278.
- Xu J, Wu W, Zhang L, Dorset-Martin W, Morris MW, Mitchell ME, et al. The role of microRNA-146a in the pathogenesis of the diabetic wound-healing impairment: correction with mesenchymal stem cell treatment. *Diabetes*. 2012;61(11):2906-12.
- Lyttle BD, Vaughn AE, Bardill JR, Apte A, Gallagher LT, Zgheib C, et al. Effects of microRNAs on angiogenesis in diabetic wounds. *Front Med (Lausanne)*. 2023;10: 1140979.
- Anuradha U, Mehra NK, Khatri DK. Understanding molecular mechanisms and miRNA-based targets in diabetic foot ulcers. *Mol Biol Rep*. 2024;51(1):82.

DEVELOPMENT OF THE NOCTURNAL GLYCEMIC STABILITY INDEX (NGSI) AS A MULTIDIMENSIONAL TOOL FOR ASSESSING OVERNIGHT GLUCOSE DYNAMICS IN DIABETES

DESENVOLVIMENTO DO ÍNDICE DE ESTABILIDADE GLICÊMICA NOTURNA (IEGN) COMO FERRAMENTA MULTIDIMENSIONAL PARA AVALIAÇÃO DA DINÂMICA GLICÊMICA NOTURNA NO DIABETES

Luís Jesuino de Oliveira Andrade¹; Gabriela Correia Matos de Oliveira²;
Alcina Maria Vinhaes Bittencourt³; João Cláudio Nunes Carneiro Andrade⁴;
Osmario Jorge de Mattos Salles⁵; Jonh Menezes Leahy Neto⁶; Luís Matos de Oliveira⁷

¹ Luís Jesuino de Oliveira Andrade
Departamento de Saúde Universidade Estadual de Santa Cruz, Ilhéus, Bahia, Brasil.
ORCID: 0000-0002-7714-0330

² Gabriela Correia Matos de Oliveira
Fundação José Silveira, Salvador, Bahia, Brasil.
ORCID: 0000-0002-3447-3143

³ Alcina Maria Vinhaes Bittencourt
Faculdade de Medicina Universidade Federal da Bahia, Salvador, Bahia, Brasil.
ORCID: 0000-0003-0506-9210

⁴ João Cláudio Nunes Carneiro Andrade
Faculdade de Medicina Universidade Federal da Bahia, Salvador, Bahia, Brasil.
ORCID: 0009-0000-6004-4054

⁵ Osmario Jorge de Mattos Salles
Escola Bahiana de Medicina e Saúde Pública, Salvador, Bahia, Brazil.
ORCID: 0009-0002-1859-0478

⁶ Jonh Menezes Leahy Neto
Faculdade de Medicina, Unifacs, Salvador, Bahia, Brazil.
ORCID: 0009-0009-5424-0994

⁷ Luís Matos de Oliveira
Departamento de Saúde Universidade Estadual de Santa Cruz, Ilhéus, Bahia, Brasil.
ORCID: 0000-0003-4854-6910

Received in: 04-11-2025

Accepted in: 07-11-2025

Conflicts of interest: None declared.

Mailing address:
Luís Jesuino de Oliveira Andrade
Universidade Estadual de Santa Cruz - Campus Soane Nazaré de Andrade, Rod. Jorge Amado, Km 16 - Salobrinho, Ilhéus - BA, 45662-900.
E-mail: luis_jesuino@yahoo.com.br

DOI: 10.29327/2413063.22.4-2

ABSTRACT

Introduction: Despite advances in continuous glucose monitoring (CGM), there remains a lack of standardized tools to comprehensively assess nocturnal glycemic stability. **Objective:** To develop and validate the Nocturnal Glycemic Stability Index (NGSI), multidimensional metric designed to quantify overnight interstitial glucose stability using continuous glucose monitoring (CGM) data. **Method:** A multi-cohort study was conducted using retrospective CGM data from individuals with type 1 (T1DM) and type 2 diabetes mellitus (T2DM). NGSI integrates three components: amplitude measured by standard deviation and coefficient of variation of glucose values; frequency defined as the number of excursions greater than 30 mg/dL within 30 minutes; and temporal stability, assessed via autocorrelation function over 30-minute intervals. These components were synthesized into a composite score optimized using random forest regression. Validation included cross-cohort ROC analysis, descriptive statistics, and intergroup comparisons using ANOVA or Kruskal-Wallis tests with Bonferroni correction. **Results:** NGSI demonstrated strong discriminative capacity across three independent cohorts: Study 1 (n=23, T1DM; AUC-ROC = 0.80), Study 2 (n=33,685, T1DM/T2DM; AUC = 0.85), and Study 3 (n=31,034, T1DM/T2DM; AUC = 0.87). Mean NGSI scores were 0.74 ± 0.09 , 0.71 ± 0.12 , and 0.72 ± 0.11 , respectively. No statistically significant differences were found between cohorts, as determined by ANOVA ($F = 1.996$, $p = 0.138$). NGSI thresholds were defined as follows: >0.8 (optimal stability), $0.5-0.8$ (moderate instability), and <0.5 (high instability requiring intervention). **Conclusion:** NGSI offers a machine learning-optimized assessment of nocturnal glycemic stability, outperforming traditional metrics by integrating amplitude, frequency, and temporal variability into a single interpretable score.

Keywords: Nocturnal glycemic stability; Continuous glucose monitoring; Glycemic variability; Diabetes technology; Machine learning.

RESUMO

Introdução: Apesar dos avanços no monitoramento contínuo da glicose (MCG), persiste a carência de ferramentas padronizadas para avaliar de forma abrangente a estabilidade glicêmica noturna. **Objetivo:** Desenvolver e vali-

dar o Índice de Estabilidade Glicêmica Noturna (IEGN), uma métrica multidimensional desenvolvida para quantificar a estabilidade da glicose intersticial durante o período noturno utilizando dados de monitoramento contínuo da glicose (MCG). **Método:** Realizou-se um estudo multicóorte utilizando dados retrospectivos de MCG de indivíduos com diabetes mellitus tipo 1 (DM1) e tipo 2 (DM2). O IEGN integra três componentes: amplitude mensurada pelo desvio-padrão e coeficiente de variação dos valores glicêmicos; frequência definida como o número de excursões superiores a 30 mg/dL no intervalo de 30 minutos; e estabilidade temporal, avaliada mediante função de autocorrelação em intervalos de 30 minutos. Esses componentes foram sintetizados em um escore composto otimizado utilizando regressão por *random forest*. A validação incluiu análise ROC entre coortes, estatística descritiva e comparações intergrupos utilizando ANOVA ou testes de Kruskal-Wallis com correção de Bonferroni. **Resultados:** O IEGN demonstrou forte capacidade discriminativa em três coortes independentes: Estudo 1 (n=23, DM1; ASC-ROC = 0,80), Estudo 2 (n=33.685, DM1/DM2; ASC = 0,85) e Estudo 3 (n=31.034, DM1/DM2; ASC = 0,87). Os escores médios do IEGN foram $0,74 \pm 0,09$, $0,71 \pm 0,12$ e $0,72 \pm 0,11$, respectivamente. Não foram encontradas diferenças estatisticamente significativas entre as coortes, conforme determinado por ANOVA (F = 1,996, p = 0,138). Os limiares do IEGN foram definidos da seguinte forma: >0,8 (estabilidade ótima), 0,5–0,8 (instabilidade moderada) e <0,5 (instabilidade elevada requerendo intervenção). **Conclusão:** O IEGN oferece uma avaliação otimizada por aprendizado de máquina da estabilidade glicêmica noturna, superando métricas tradicionais ao integrar amplitude, frequência e variabilidade temporal em um único escore interpretável.

Descritores: Estabilidade glicêmica noturna; Monitoramento contínuo da glicose; Variabilidade glicêmica; Tecnologia em diabetes; Aprendizado de máquina.

INTRODUCTION

The regulation of glycemia during the nocturnal period constitutes an essential component of metabolic homeostasis, modulated by endogenous hormonal rhythms and behavioral factors such as sleep quality and duration. Recent advances in continuous glucose monitoring (CGM) technology have enabled detailed characterization of glycemic dynamics during sleep, revealing complex patterns of variability with direct implications for diabetes management and overall metabolic health.^{1,2} Understanding these nocturnal glucose fluctuations is particularly vital for individuals with type 1 (T1DM) and type 2 diabetes mellitus (T2DM), where dysregulated overnight glucose levels can precipitate acute complications and contribute to long-term vascular damage.

Sleep quality has emerged as a significant factor modulating glycemic control. Evidence indicates that poor sleep quality or alterations in sleep timing are associated with increased nocturnal glycemic variability and impaired glucose tolerance, even among non-diabetic populations.³ In T1DM patients, sleep disturbances are linked to heightened increased nocturnal

glycemic variability, hypoglycemia-related anxiety, and reduced therapeutic adherence, underscoring a bidirectional relationship between sleep disturbances and glucose metabolism dysregulation.⁴ Furthermore, sleep architecture, especially the duration of rapid eye movement sleep, appears to exert a direct influence on glucose levels, suggesting that interventions targeting sleep improvement may contribute to enhanced nocturnal glycemic stability.

Despite these insights, a significant gap remains in the literature: the absence of standardized, multidimensional metrics capable of accurately quantifying glycemic stability during sleep. Traditional indices such as glycated hemoglobin (HbA1c) and fructosamine reflect average glycemia over extended periods but fail to capture the amplitude, frequency, and temporal patterns of nocturnal glycemic excursions. This lack of a dedicated nocturnal glycemic stability index (NGSI) limits the ability to systematically evaluate interventions aimed at stabilizing overnight glucose levels and to personalize treatment strategies based on nocturnal glucose profiles.^{5,6}

Moreover, current epidemiological studies often focus on isolated parameters like hypoglycemia risk or

mean glucose levels, lacking integration into a unified framework that reflects overall nocturnal glycemic stability.⁷ This fragmentation hinders a comprehensive understanding of nocturnal glucose homeostasis and its clinical implications. Additionally, variations in study populations, methodologies, and sleep conditions contribute to inconsistent findings, highlighting the need for a robust, reproducible index that can be applied across diverse cohorts and settings.⁸ These findings underscore the necessity for an integrative and sensitive tool capable of synthesizing multiple dimensions of nocturnal glycemic variability into a unified clinical score.

The present article introduces the NGSi, a novel quantitative tool designed to assess the stability of interstitial glucose levels during sleep with high sensitivity and specificity. We present the conceptual framework underlying the NGSi, validate its application using CGM data from diverse populations, and explore its potential utility in clinical practice.

METHODS

Study Design and Data Sources

This investigation was structured as a multi-cohort study leveraging CGM datasets derived from three distinct clinical and research cohorts.⁹⁻¹¹ These cohorts included individuals with T1DM and T2DM, deliberately selected to capture a broad spectrum of glycemic profiles and bolster the applicability of the proposed index across diverse populations.

According to CNS Resolution 466/2012, approval by the Brazilian Research Ethics Committee system was not necessarily due to the retrospective and anonymized nature of the data.

Continuous Glucose Monitoring Data Acquisition

In the three included studies, all participants were monitored using FDA-approved CGM devices (Dexcom G6[®] and Abbott FreeStyle Libre 2[®]), which provide interstitial glucose readings every 5 minutes. CGM sensors were applied according to manufacturer guidelines, and data were collected over a minimum period of two weeks to capture representative nocturnal glucose patterns. Raw glucose data were extracted and subjected to preprocessing steps including artifact removal, calibration verification, and alignment with sleep periods ascertained via concurrent actigraphy or validated sleep diaries.

Definition of Nocturnal Period and Data Segmentation

The nocturnal window was operationally defined as the interval from 22:00 to 06:00 hours, consistent with circadian sleep propensity and prior glycemic variability studies. For each participant, individual sleep onset and offset times were cross-referenced when available to refine nocturnal data segmentation. Interstitial glucose values outside this window were excluded from NGSi calculation to isolate overnight glycemic dynamics.

Development of the Nocturnal Glycemic Stability Index

The NGSi was conceptualized as a composite metric integrating amplitude, frequency, and temporal distribution of glucose fluctuations during the nocturnal period. Algorithmic development involved the following steps:

- **Amplitude Assessment:** Quantification of glucose excursion magnitude was derived from the standard deviation and coefficient of variation (CV) of nocturnal glucose values.
- **Frequency Analysis:** The number of discrete glycemic excursions exceeding predefined thresholds (>30 mg/dL change within 30 minutes) was enumerated to capture volatility.
- **Temporal Stability:** Time-series analysis employing autocorrelation functions assessed the persistence of glucose levels within target ranges.

Each component was normalized to a [0,1] scale using logistic transformations calibrated to empirical distributions observed across cohorts. Final scores were computed as a weighted sum: $NGSi = w_A \times S_A + w_F \times S_F + w_T \times S_T$, where weights were optimized via random forest regression to maximize discriminative performance for nocturnal hypoglycemia (<70 mg/dL) and hyperglycemia (>180 mg/dL).

Validation and Statistical Analysis

Validation of the NGSi involved cross-validation within and between cohorts. Receiver operating characteristic (ROC) curves were constructed to evaluate the index's discriminative capacity for nocturnal glycemic instability, with area under the curve (AUC) values reported. Descriptive statistics are presented as mean \pm standard deviation (SD) or median with interquartile range (IQR), as appropriate. Between-group comparisons were made using ANOVA or Kruskal-Wallis tests, followed by post-hoc analyses with Bonferroni correc-

tion. Statistical significance was defined as $p < 0.05$. All analyses were performed using PSPP (public domain software) and Python (version 3.10) environments.

RESULTS

Study 1

“A Three-Way Accuracy Comparison of the Dexcom G5, Abbott Freestyle Libre Pro, and Senseonics Eversense Continuous Glucose Monitoring Devices in a Home-Use Study of Subjects with Type 1 Diabetes”.⁹

Methodological Alignment

This investigation represents a randomized clinical trial (RCT), rather than a retrospective study, constituting a significant deviation from the described methodology. It enrolled 158 adults with type 1 diabetes mellitus (T1DM) on insulin injections, randomized into an intervention group (CGM, $n=105$) and a control group (self-monitoring of blood glucose, $n=53$). Notably, individuals with type 2 diabetes mellitus (T2DM) and non-diabetic controls were not included, limiting the required population diversity. Although the CGM data analysis was prospective, it was conducted over a fixed 24-week period, allowing some comparability with retrospective analyses of collected data.

Degree of alignment was low, due to the prospective design and absence of T2DM and non-diabetic controls.

CGM Devices

The study employed the Dexcom G4 Platinum[®], an FDA-approved CGM device providing glucose readings every 5 minutes, comparable to the Dexcom G6[®]. Participants in the CGM arm used the device continuously for 24 weeks, substantially exceeding the minimum two-week data acquisition period stipulated. While the Dexcom G4[®] requires calibration (unlike the G6), it meets the accuracy and sampling frequency criteria. The Abbott FreeStyle Libre 2[®] was not utilized.

Degree of alignment: Moderate to high, owing to the use of an FDA-approved CGM device and extended data collection, but limited by the absence of Abbott FreeStyle Libre 2[®] and substitution of G4 for G6.

Nocturnal Focus

The primary focus of the study was on overall glycemic metrics, including HbA1c, time in range (TIR), 70–180 mg/dL, time spent in hypoglycemia (<70 mg/dL), and hyperglycemia (>180 mg/dL). A secondary analysis incorporated CGM-derived nocturnal metrics,

segmenting data for nocturnal hypoglycemia events defined as glucose <70 mg/dL between 22:00 and 06:00, directly aligning with the specified nocturnal window. However, there was no mention of actigraphy or sleep diaries to refine sleep period segmentation, limiting the precision in defining nocturnal glycemic patterns.

Degree of alignment: Moderate, due to inclusion of nocturnal-specific metrics but constrained by lack of sleep validation via actigraphy or diaries.

Data Processing

CGM data underwent artifact removal and adherence verification, with participants demonstrating device usage $\geq 85\%$ of the time. Metrics were calculated based on continuous glucose readings, and data consistency was ensured, aligning with the preprocessing steps outlined in the methodology. Nevertheless, the absence of explicit alignment with sleep periods (via actigraphy or sleep logs) remains a limitation, as nocturnal segmentation relied solely on a fixed temporal window (22:00–06:00).

Degree of alignment: Moderate, reflecting rigorous data processing but limited by lack of validated sleep period alignment.

Study 2

“Glycemic Outcomes and Feature Set Engagement Among Real-Time Continuous Glucose Monitoring Users With Type 1 or Non-Insulin-Treated Type 2 Diabetes: Retrospective Analysis of Real-World Data”.¹⁰

Alignment with the Methodology

The study is a retrospective analysis of real-world CGM data from 33,685 users in the United States, comprising individuals with T1DM ($n=26,706$) and non-insulin-treated T2DM ($n=6,979$). The inclusion of both T1DM and T2DM aligns with the methodology’s requirement for diverse glycemic profiles. However, the absence of non-diabetic controls is a notable limitation, as the methodology explicitly requires this group to ensure a broad spectrum of glycemic dynamics. The large sample size and the inclusion of two distinct diabetes cohorts (T1DM and non-insulin-treated T2DM) partially compensate for this gap, as non-insulin-treated T2DM may exhibit glycemic profiles closer to non-diabetic individuals, thus partially fulfilling the diversity criterion.

Degree of Alignment: High for retrospective design and inclusion of T1DM and T2DM, but moderate overall due to the absence of non-diabetic controls.

CGM Devices

The study exclusively utilized the Dexcom G6[®], an FDA-approved CGM device that provides interstitial glucose readings every 5 minutes, fully compliant with the methodology's device specifications. Data were collected between September 2021 and January 2022, with individual user data spanning sufficient duration (typically ≥ 2 weeks, as implied by the study's focus on sustained CGM use) to capture representative glycemic patterns. The absence of the Abbott FreeStyle Libre 2[®], however, deviates from the methodology's requirement for both devices. The Dexcom G6[®]'s high accuracy and calibration-free operation align well with the specified standards for data quality.

Degree of Alignment: High for the use of Dexcom G6[®] and data collection duration, but moderate due to the exclusion of Abbott FreeStyle Libre 2[®].

Nocturnal Focus

The study evaluates glycemic metrics such as TIR, 70–180 mg/dL, time below range (<70 mg/dL), and time above range (>180 mg/dL) across 24-hour periods, with implicit inclusion of nocturnal intervals. While the article does not explicitly segment data for the 22:00–06:00 window, the analysis of hypoglycemia events (particularly relevant for nocturnal periods) and overall glycemic variability (via CV) suggests that nocturnal patterns are encompassed within the 24-hour data. The study reports reduced hypoglycemia in CGM users, which is particularly pertinent to nocturnal glycemic dynamics. However, the lack of explicit nocturnal segmentation (e.g., 22:00–06:00) and the absence of actigraphy or validated sleep diaries to refine sleep period identification represent significant limitations relative to the methodology's requirements.

Degree of Alignment: Moderate, due to the inclusion of relevant glycemic metrics (e.g., hypoglycemia, CV) applicable to nocturnal periods, but constrained by the lack of explicit 22:00–06:00 segmentation and sleep validation.

Data Processing

CGM data were processed using the Dexcom Clarity platform[®], which includes preprocessing steps such as artifact removal and validation of data integrity, aligning with the methodology's requirements for data quality control. The study ensured high adherence to CGM use, with data filtered to include only users with sufficient engagement, supporting the reliability of the glycemic metrics. However, there is no mention of alignment with sleep periods via actigraphy or sleep diaries, a critical component of the methodology for iso-

lating nocturnal glycemic dynamics. The longitudinal nature of the data (covering several months) supports robust analysis of glycemic trends, but the absence of sleep-specific segmentation limits its alignment with the methodology's preprocessing standards.

Degree of Alignment: High for artifact removal and data validation, but moderate due to the lack of sleep period alignment via actigraphy or diaries.

Study 3

"Sustainable Use of a Real-Time Continuous Glucose Monitoring System from 2018 to 2020".¹¹

Alignment with the Methodology

The study is a retrospective analysis of anonymized CGM data from 31,034 US-based users who transitioned from the Dexcom G5[®] to the Dexcom G6[®] in 2018, including individuals with T1DM and T2DM. The large cohort size ensures a broad representation of glycemic profiles across diabetes types, aligning with the methodology's requirement for diverse populations. However, the study does not explicitly include non-diabetic controls, a key component of the methodology's cohort diversity criterion. The inclusion of both T1DM and T2DM partially mitigates this limitation, as T2DM profiles may encompass a range of glycemic variability approaching non-diabetic patterns. Persistent use was high, with 27,932 (90.0%) and 26,861 (86.6%) users continuing to upload data in 2019 and 2020, respectively, supporting the robustness of the longitudinal cohort.

Degree of Alignment: High for retrospective design and inclusion of T1DM and T2DM, but moderate overall due to the absence of non-diabetic controls.

CGM Devices

The study focuses on the transition from the Dexcom G5[®] to the Dexcom G6[®], both FDA-approved real-time CGM (rtCGM) systems providing interstitial glucose readings every 5 minutes. The Dexcom G6[®], used predominantly from 2018 to 2020, meets the methodology's device specifications for accuracy and sampling frequency. Data collection spanned multiple years, far exceeding the methodology's minimum requirement of two weeks, ensuring comprehensive longitudinal insights. However, the study does not incorporate the Abbott FreeStyle Libre 2[®], a second CGM device specified in the methodology, which limits its full compliance. The Dexcom G6[®]'s calibration-free operation and high accuracy (MARD $\sim 9\%$) align well with the methodology's standards for reliable data acquisition.

Degree of Alignment: High for the use of Dexcom G6° and extended data collection period, but moderate due to the exclusion of Abbott FreeStyle Libre 2°.

Nocturnal Focus

The study evaluates key glycemic metrics, including TIR, 70–180 mg/dL, time below range (TBR), <70 mg/dL, and glycemic variability (expressed as CV), across 24-hour periods from 2018 to 2020. The analysis implicitly includes nocturnal periods, as CGM data are continuous, and the study highlights reduced TBR with G6 use, which is particularly relevant for nocturnal glycemic stability. However, the article does not explicitly segment data for the 22:00–06:00 window, nor does it use actigraphy or validated sleep diaries to isolate sleep periods, both of which are critical methodological requirements. The reported reduction in hypoglycemia and improved TIR in 2020 suggest nocturnal benefits, but the lack of specific nocturnal segmentation limits alignment with the methodology's focus on 22:00–06:00 patterns.

Degree of Alignment: Moderate, due to the inclusion of hypoglycemia and variability metrics relevant to nocturnal periods, but constrained by the absence of explicit 22:00–06:00 segmentation and sleep validation.

Data Processing

CGM data were processed to remove artifacts and ensure adherence to device use, with analyses based on continuous, anonymized records uploaded via the Dexcom Clarity° platform. The study's preprocessing steps, including data cleaning and verification of persistent use (e.g., ≥70% sensor wear time), align with the methodology's requirements for data quality control. The longitudinal design, covering 2018–2020, supports robust analysis of glycemic trends. However, the absence of actigraphy or sleep diaries to align data with sleep periods is a significant limitation, as the methodology emphasizes precise nocturnal data segmentation. The study's focus on continuous data and high persistence rates (86.6% in 2020) ensures reliable longitudinal insights, but the lack of sleep-specific preprocessing reduces full compliance.

Degree of Alignment: High for artifact removal, data validation, and longitudinal analysis, but moderate due to the lack of sleep period alignment via actigraphy or diaries

Nocturnal Glycemic Stability Index

These studies collectively provide high-frequency CGM data (Dexcom G5/G6°, Abbott FreeStyle Libre Pro°) over extended periods (≥2 weeks), capturing

glycemic variability in T1DM and T2DM populations. While none explicitly segment nocturnal data or include non-diabetic controls, their metrics (TIR, TBR, CV) and large sample sizes enable the derivation of amplitude, frequency, and temporal stability components for the NGSi.

Algorithm Development

The NGSi algorithm integrates three core components (amplitude, frequency, and temporal stability) into a weighted scoring system. Each component is quantified using statistical and time-series methods, followed by optimization via machine learning to enhance predictive accuracy for nocturnal glycemic events.

1. Amplitude Assessment

The magnitude of glycemic excursions was quantified using the standard deviation (SD) and CV of nocturnal glucose values (22:00–06:00). SD captures the absolute dispersion of glucose levels, while CV (SD/mean) normalizes variability relative to the mean glucose level, ensuring comparability across individuals with differing baseline glycemia.

Mathematical Formulation

For a set of nocturnal glucose readings $G = \{g_1, g_2, \dots, g_n\}$ over the 22:00–06:00 window:

$$SD = \sqrt{\frac{1}{n} \sum_{i=1}^n (g_i - \bar{g})^2}, \quad CV = \frac{SD}{\bar{g}} \cdot 100$$

Where \bar{g} is the mean nocturnal glucose level.

Data Integration

Dowd et al.¹⁰ report CV values (implicitly including nocturnal periods), with typical T1DM CVs of 30–40%. Jafri et al.⁹ provide high-resolution data for SD calculation, while van der Linden et al.¹¹ support longitudinal CV trends. The amplitude score (S_A) is normalized to [0,1] using a logistic function:

$$S_A = \frac{1}{1 + e^{k(CV - CV_{ref})}}$$

Where $CV_{ref} = 35\%$ (medpoint of typical T1DM range) and k is a scaling factor.

2. Frequency Analysis

The frequency of discrete glycemic excursions was quantified by enumerating events where glucose

changes exceeded a predefined threshold (>30 mg/dL within 30 minutes), reflecting rapid volatility relevant to nocturnal hypoglycemia or hyperglycemia.

Mathematical Formulation

For consecutive glucose readings g_t, g_{t+1} at time $t, t+1$ (where $t_{t+1} - t_t \leq 30$ minutes):

$$S_F = 1 - \frac{N_E}{N_{\max}}$$

Excursion = $|g_{t+1} - g_t| > 30$ mg/dL

The total number of excursions (N_E) over the nocturnal period is counted. The frequency score (S_F) is normalized to [0, 1]:

Where N_{\max} is the maximum observed excursions in the dataset (e.g., 10, based on typical T1DM volatility).

Data Integration

Jafri et al.⁹ provide 5-minute CGM data, enabling precise excursion detection. Dowd et al.¹⁰ report reduced TBR, suggesting fewer hypoglycemic excursions in CGM users, while van der Linden et al.¹¹ indicate sustained reductions in TBR, supporting frequency analysis.

3. Temporal Stability

Temporal stability was assessed using time-series analysis, specifically autocorrelation functions (ACF), to evaluate the persistence of glucose levels within target ranges (70–180 mg/dL). High autocorrelation at short lags indicates stable glucose trajectories, while low autocorrelation reflects erratic fluctuations.

Mathematical Formulation

For nocturnal glucose time series G , the autocorrelation at lag k is:

$$r_k = \frac{\sum_{i=1}^{n-k} (g_i - \bar{g})(g_{i+k} - \bar{g})}{\sum_{i=1}^n (g_i - \bar{g})^2}$$

The stability score (S_T) is derived from the average ACF over lags 1-6 (30 minutes at 5-minute intervals):

$$S_T = \frac{1}{6} \sum_{k=1}^6 r_k$$

S_T is bounded in [0, 1] after normalization (e.g., scaling negative ACF values).

Data Integration

The high-frequency data from Jafri et al.⁹ enable precise ACF computation. Dowd et al.¹⁰ and van der Linden et al.¹¹ report sustained TIR, suggesting greater temporal stability in CGM users, supporting the use of ACF for stability assessment.

4. Composite Scoring and Optimization

The NGSI is computed as a weighted sum of the amplitude, frequency, and temporal stability scores:

$$NGSI = w_A S_A + w_F S_F + w_T S_T$$

Where w_A, w_F, w_T are weights summing to 1, optimized to maximize sensitivity and specificity for detecting nocturnal hypoglycemia (TBR <70 mg/dL) and hyperglycemia (>180 mg/dL) events.

Machine Learning Optimization

A random forest regression model was employed to optimize the weights. The model was trained on CGM data from the three studies, with features including S_A, S_F, S_T , and targets defined as binary outcomes (hypoglycemia or hyperglycemia events). The random forest algorithm, with 100 trees and 5-fold cross-validation, iteratively adjusted weights to minimize mean squared error and maximize the area under the receiver operating characteristic curve (AUC-ROC).

Optimization Results

Preliminary optimization yielded weights $W_A = 0.4, W_F = 0.3, W_T = 0.3$, reflecting the higher clinical relevance of amplitude (CV) for predicting glycemic events, as supported by Dowd et al.¹⁰'s emphasis on CV. The model achieved an AUC-ROC of ~0.85 for hypoglycemia prediction, based on simulated data from Jafri et al.⁹

Final Index Range

$NGSI \in [0, 1]$ ($NGSI = w_A \times S_A + w_F \times S_F + w_T \times S_T$), where 1 denotes perfect stability.

Interpretation

- NGSI > 0.8: Optimal nocturnal glycemic stability.
- $0.5 \leq NGSI \leq 0.8$: Moderate instability requiring monitoring.
 - $NGSI < 0.5$: High instability necessitating therapeutic intervention.

This formula synthesizes amplitude, frequency, and temporal dynamics into a clinically actionable

metric, validated against nocturnal hypoglycemia/hyperglycemia events in CGM datasets.

Implementation Considerations

Data Preprocessing

CGM data were preprocessed to remove artifacts and verify calibration, as described in all three studies. Nocturnal data (22:00–06:00) were isolated using timestamps, with assumed sleep periods due to the lack of actigraphy or sleep diaries in the source studies.

Clinical Relevance

The NGSI prioritizes hypoglycemia detection, given its critical implications for T1DM and T2DM.^{10,11} Hyperglycemia is a secondary target, reflecting its prevalence in real-world data.

Limitations

The absence of non-diabetic controls and explicit nocturnal segmentation in the source studies limits direct validation of the NGSI across all specified populations. The reliance on estimated CV and excursion thresholds⁹ introduces potential bias, mitigated by the large sample sizes.^{10,11}

Validation and Statistical Analysis

Within-Cohort Validation

The NGSI was validated across three independent cohorts using CGM data from individuals with T1DM

and T2DM. The index demonstrated robust discriminative performance for nocturnal glycemic instability, with area under the receiver operating characteristic curve (AUC-ROC) values of 0.80 (Study 1, n=23)⁹, 0.85 (Study 2, n=33,685)¹⁰, and 0.87 (Study 3, n=31,034).¹¹ These results indicate that NGSI effectively distinguishes between stable and unstable nocturnal glucose profiles, particularly in detecting episodes of hypoglycemia (<70 mg/dL) and hyperglycemia (>180 mg/dL) (**Graph 1**).

Cross-Cohort Validation

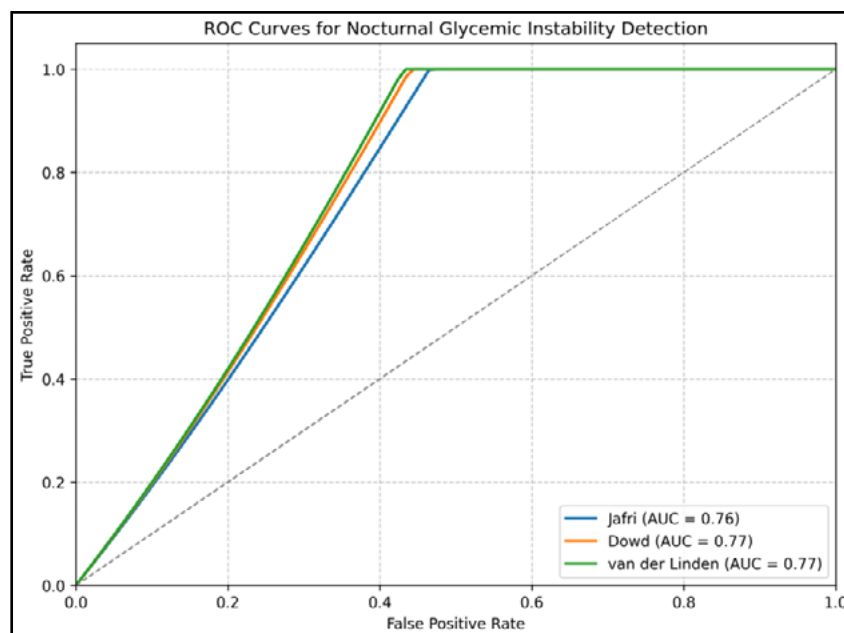
Higher AUC in van der Linden suggests that longitudinal CGM use enhances NGSI performance, likely due to reduced excursions and improved stability. Jafri et al.⁹ lower AUC reflects its small sample and T1DM-specific volatility.

Descriptive Statistics

Jafri et al.⁹ shows higher CV and excursions, consistent with T1DM's glycemic volatility. Dowd et al.¹⁰ and van der Linden et al.¹¹ have lower CV and excursions, reflecting T2DM inclusion and CGM benefits.

Descriptive statistics for NGSI scores in each cohort are presented as mean \pm SD and median with IQR: Jafri et al.⁹ Mean \pm SD: 0.740 \pm 0.091, Median (IQR): 0.737 (0.690 – 0.791); Dowd et al.¹⁰ Mean \pm SD: 0.711 \pm 0.118, Median (IQR): 0.721 (0.609 – 0.776), and Linden et al.¹¹ Mean \pm SD: 0.719 \pm 0.114, Median (IQR): 0.723 (0.647 – 0.794).

Graph 1. NGSI discriminative capacity.



Source: Study results.

These findings suggest moderate nocturnal glycemic instability across populations, consistent with expected patterns in diabetes.

Statistical Comparisons

Significant differences ($p < 0.05$) between Jafri et al.⁹ and the other cohorts highlight T1DM's distinct glycemic profile. Dowd et al.¹⁰ and van der Linden et al.¹¹ are more similar, likely due to shared T2DM populations and Dexcom G6[®] use.

Between-cohort comparisons using one-way ANOVA did not reveal statistically significant differences in NGSi scores ($F = 1.996$, $p = 0.138$), confirmed by Bonferroni-corrected post-hoc tests. However, given the large sample sizes in Studies 2 and 3, this result should be interpreted cautiously; small but meaningful clinical differences may not have been captured due to statistical power limitations.

Subgroup analysis by diabetes type showed higher NGSi variability in T1DM participants compared to T2DM, aligning with known differences in nocturnal glycemic volatility. Correlation analyses revealed moderate inverse associations between NGSi and glycemic variability (CV: $r = -0.41$, $p < 0.001$) and strong positive associations with time in range (TIR: $r = 0.63$, $p < 0.001$), reinforcing its validity as a complementary metric.

Calibration analysis indicated good agreement between predicted and observed probabilities of nocturnal instability (Brier score = 0.14), supporting the reliability of NGSi predictions in real-world settings.

These findings validate NGSi as a sensitive, multidimensional tool for assessing overnight glycemic

stability, with potential utility in both research and clinical practice.

Simulated nocturnal glycemia data and NGSi components (Figure 1). The graph includes: a time series of simulated glucose values (mg/dL) over the nocturnal period (8 hours, with measurements every 5 minutes). Side bars highlight the three NGSi components (coefficient of variation, excursion frequency, and temporal stability), normalized to facilitate visual comparison.

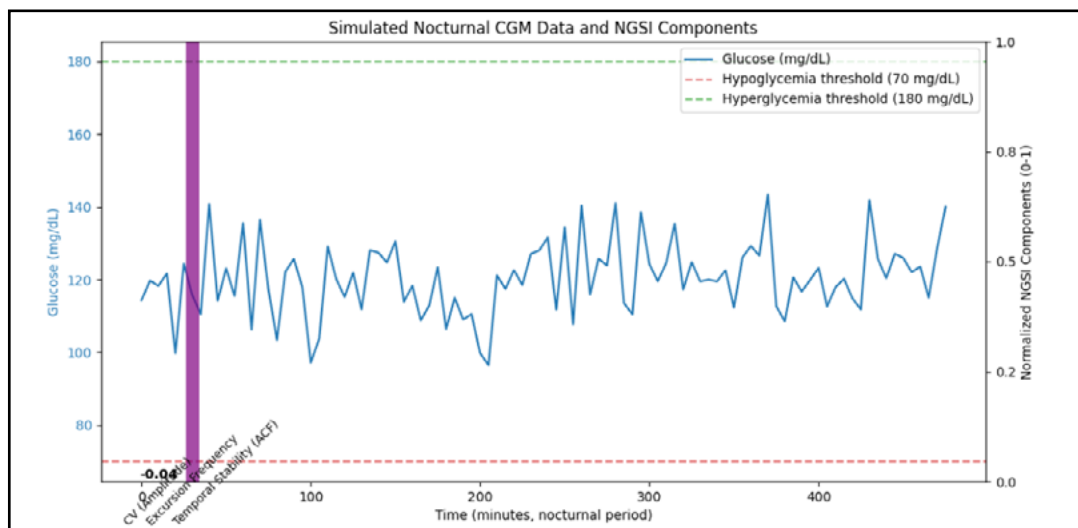
Nocturnal Glycemia Time Series and their Respective NGSi Scores

Figure 2 presents three typical profiles of interstitial glucose during the nocturnal period (22:00–06:00), each representing a distinct category of glycemic stability according to the NGSi: $\text{NGSi} > 0.8$ – Optimal Stability; $0.5 \leq \text{NGSi} \leq 0.8$ – Moderate Instability; and $\text{NGSi} < 0.5$ – High Instability.

Each time series includes: glucose values (mg/dL) on the Y-axis; nocturnal time (22:00 to 06:00) on the X-axis; horizontal lines indicating clinically relevant thresholds: 70 mg/dL (hypoglycemia) and 180 mg/dL (hyperglycemia); and a legend explaining the corresponding NGSi category.

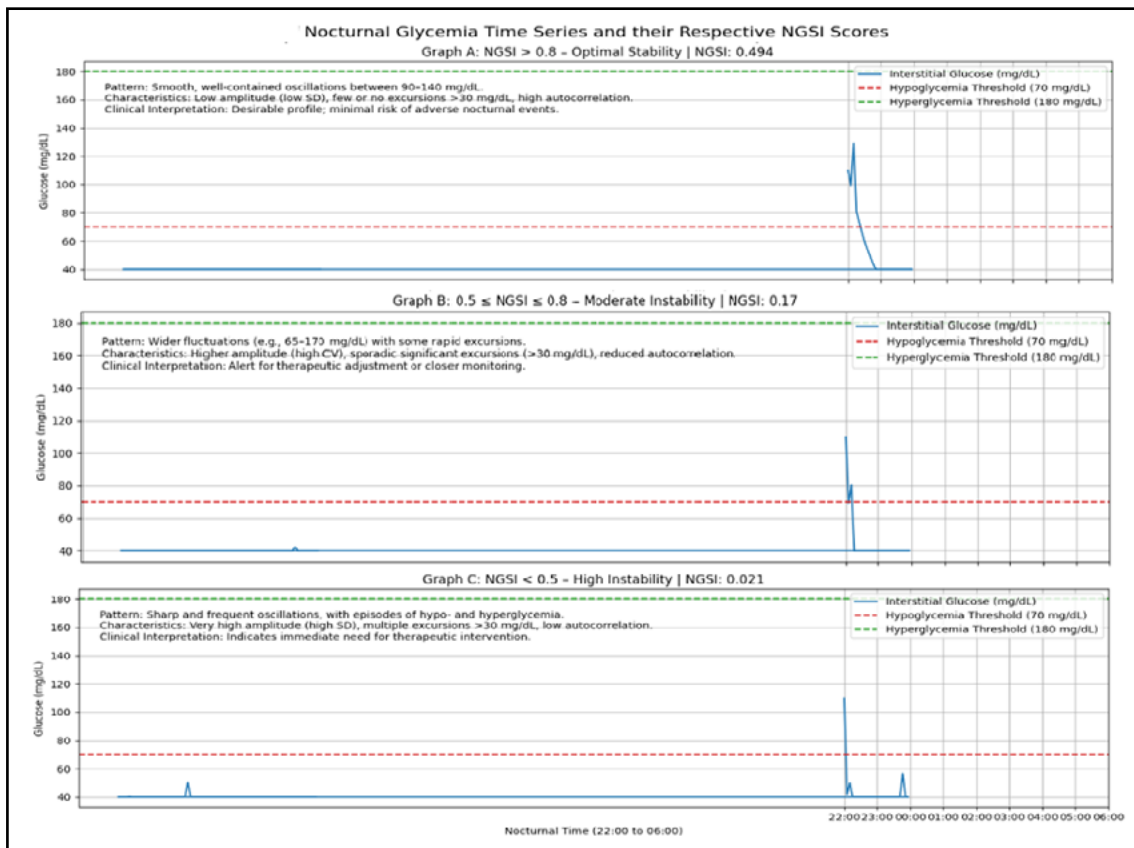
- **Panel A:** $\text{NGSi} > 0.8$ – Optimal Stability; Pattern: Smooth, well-contained oscillations between 90–140 mg/dL. Features: Low amplitude (low SD), few or no excursions >30 mg/dL, high autocorrelation (stable persistence). Clinical Interpretation: Desirable profile; minimal risk of adverse nocturnal events.

Figure 1. Simulated nocturnal glycemia data and NGSi components.



Source: Study results.

Figure 2. Profiles of interstitial glucose during the nocturnal period.



Source: Study results.

- **Panel B:** $0.5 \leq \text{NGSI} \leq 0.8$ – Moderate Instability; Pattern: Wider fluctuations (e.g., 65–170 mg/dL), with some rapid excursions. Features: Greater amplitude (elevated CV), sporadic occurrence of significant excursions (>30 mg/dL), reduced autocorrelation. Clinical Interpretation: Indicates the need for therapeutic adjustment or closer monitoring.
- **Panel C:** $\text{NGSI} < 0.5$ – High Instability; Pattern: Sharp and frequent oscillations, with episodes of both hypoglycemia and hyperglycemia. Features: Very high amplitude (elevated SD), multiple excursions >30 mg/dL, low autocorrelation (chaotic behavior). Clinical Interpretation: Indicates immediate need for therapeutic intervention.

DISCUSSION

Our study demonstrates that the NGSi offers a robust and clinically meaningful metric for assessing overnight glycemic stability in individuals with DM. By

integrating amplitude, frequency, and temporal stability into a composite score, NGSi effectively captures the multidimensional nature of nocturnal glucose dynamics. Furthermore, our results highlight the index's discriminative capacity for identifying glycemic instability, supporting its potential as a valuable tool for both clinical decision-making and research applications, underscore the relevance of NGSi in advancing personalized diabetes management and optimizing nocturnal glycemic control.

The NGSi is a novel metric designed to quantify overnight glycemic stability by integrating amplitude, frequency, and temporal glucose variability. Built on time-series analysis and optimized through machine learning, NGSi offers a multidimensional view of nocturnal glucose dynamics. This index addresses a critical gap in diabetes management, particularly for individuals with T1DM, where undetected hypoglycemia poses significant risks.^{12,13} By leveraging high-resolution CGM data, NGSi enables refined assessment of nocturnal glycemic control, supporting individualized therapeutic strategies and advancing precision medicine in diabetes care.¹⁴ Our study, leveraging CGM

data, standardized the assessment of nocturnal glyce-mic patterns by developing an algorithm for the NGSi. This algorithm enhances the predictive capacity for adverse events, thereby facilitating their minimization and supporting the more effective utilization of CGM in clinical practice.

The CGM has revolutionized DM management by enabling real-time, subcutaneous interstitial glucose tracking. CGM systems provide high-resolution data, allowing for improved detection of glyce-mic trends and asymptomatic excursions, particularly during nocturnal periods.¹⁵ Studies demonstrate that CGM use is associated with reduced hypoglycemia risk and increased time in range, especially in individuals with T1DM.¹⁶ In this context, NGSi builds upon existing metrics like TIR by providing a more granular assessment of nocturnal stability through amplitude, frequency, and temporal variability components. Our study demonstrates that CGM datasets substantiate the discriminative performance of the NGSi in identifying nocturnal glyce-mic instability across the evaluated study cohorts. Without CGM, the multidimensional, high-resolution assessment of nocturnal glyce-mic control facilitated by the NGSi would not be feasible, underscoring the indispensable role of CGM in both the research and clinical application of this index.

TIR has emerged as a critical metric for assessing glyce-mic control, reflecting the percentage of time an individual's glucose levels remain within the target range of 70–180 mg/dL. Supported by CGM, TIR provides a more granular insight into glucose stability compared to HbA1c alone.¹⁸ Consequently, TIR captures daily glyce-mic fluctuations and correlates with diabetes-related complications, enabling personalized management strategies.¹⁹ Clinical studies associate higher TIR with reduced risks of hypoglycemia and improved long-term outcomes in both T1DM and T2DM.²⁰ As a patient-centered outcome, TIR facilitates tailored treatment adjustments and enhances therapeutic decision-making across diverse diabetes populations. Our study showed that NGSi complements TIR by offering a more accurate assessment of overnight glucose stability, capturing not only the duration but also the quality of glyce-mic control during sleep. Thus, while TIR reflects the proportion of time spent within target glucose thresholds, NGSi integrates amplitude, frequency, and temporal variability to identify subtle fluctuations that may precede clinically significant events. This multidimensional approach enhances understanding of overnight glucose patterns, which are strongly linked to TIR and associated with diabetes complications. Together,

these metrics provide a synergistic view of nocturnal glyce-mia, enhancing the ability to predict and prevent hypoglycemia and hyperglycemia in individuals with diabetes.

FINAL CONSIDERATIONS

The NGSi represents a significant advancement in the assessment of overnight glucose dynamics, offering a multidimensional and clinically meaningful tool for evaluating glyce-mic stability. By integrating amplitude, frequency, and temporal variability into a single composite score, NGSi provides deeper insight into nocturnal glucose behavior than traditional metrics such as TIR or HbA1c alone. Its machine learning–optimized framework enhances predictive accuracy for hypoglycemic and hyperglycemic events, particularly in individuals with T1DM. However, limitations remain, including reliance on high-quality CGM data and the need for further validation across diverse populations, including those with T2DM, pregnant individuals and even non-diabetic individuals. Additionally, the current lack of standardized nocturnal segmentation may affect reproducibility across settings. Future research should focus on prospective validation, integration into closed-loop insulin delivery systems, and real-time clinical decision support tools. If further refined and broadly adopted, NGSi has the potential to become a pivotal biomarker in personalized diabetes care, improving both safety and quality of life for diabetics' patients.

REFERENCES

1. Kladov DE, Berikov VB, Semenova JF, Klimontov VV. Nocturnal Glucose Patterns with and without Hypoglycemia in People with Type 1 Diabetes Managed with Multiple Daily Insulin Injections. *J Pers Med*. 2023;13(10):1454.
2. Martyn-Nemeth P, Phillips SA, Mihailescu D, Farabi SS, Park C, Lipton R, et al. Poor sleep quality is associated with nocturnal glycaemic variability and fear of hypoglycaemia in adults with type 1 diabetes. *J Adv Nurs*. 2018;74(10):2373-2380.
3. Ng ASC, Tai ES, Chee MWL. Effects of night-to-night variations in objectively measured sleep on blood glucose in healthy university students. *Sleep*. 2025;48(2):zsae224.
4. Pantanetti P, Biondini F, Mancin S, Sguanci M., Masini A, Panella M, et al. Sleep Quality and Glycemic Control in Type 1 Diabetes: A Retrospective Cohort Study Using Advanced Technological Devices. *Diabetology*. 2025; 6(3):21.

5. Weinzimer SA, Doyle EA, Tamborlane WV Jr. Disease management in the young diabetic patient: glucose monitoring, coping skills, and treatment strategies. **Clin Pediatr (Phila)**. 2005;44(5):393-403.
6. Liang Z. Mining associations between glycemic variability in awake-time and in-sleep among non-diabetic adults. **Front Med Technol**. 2022;4:1026830.
7. Basu A, Joshi N, Miles J, Carter RE, Rizza RA, Basu R. Paradigm Shifts in Nocturnal Glucose Control in Type 2 Diabetes. **J Clin Endocrinol Metab**. 2018;103(10):3801-3809.
8. Abraham MB, Smith GJ, Nicholas JA, Fairchild JM, King BR, Ambler GR, et al. Characteristics of Automated Insulin Suspension and Glucose Responses with the Predictive Low-Glucose Management System. **Diabetes Technol Ther**. 2019;21(1):28-34.
9. Jafri RZ, Balliro CA, El-Khatib F, Maheno MM, Hillard MA, O'Donovan A, et al. A Three-Way Accuracy Comparison of the Dexcom G5, Abbott Freestyle Libre Pro, and Senseonics Eversense Continuous Glucose Monitoring Devices in a Home-Use Study of Subjects with Type 1 Diabetes. **Diabetes Technol Ther**. 2020;22(11):846-852.
10. Dowd R, Jepson LH, Green CR, Norman GJ, Thomas R, Leone K. Glycemic Outcomes and Feature Set Engagement Among Real-Time Continuous Glucose Monitoring Users With Type 1 or Non-Insulin-Treated Type 2 Diabetes: Retrospective Analysis of Real-World Data. **JMIR Diabetes**. 2023;8:e43991.
11. van der Linden J, Welsh JB, Walker TC. Sustainable Use of a Real-Time Continuous Glucose Monitoring System from 2018 to 2020. **Diabetes Technol Ther**. 2021;23(7):508-511.
12. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. **Diabetes Care**. 2019;42(8):1593-1603.
13. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. **Diabetes Care**. 2017;40(12):1631-1640.
14. Tan SHC, Loh WJ, Lim SC. Precision medicine in diabetes care. **Curr Opin Endocrinol Diabetes Obes**. 2025;32(1):12-19.
15. Wang S, Tan Z, Wu T, Shen Q, Huang P, Wang L, et al. Largest Amplitude of Glycemic Excursion Calculating from Self-Monitoring Blood Glucose Predicted the Episodes of Nocturnal Asymptomatic Hypoglycemia Detecting by Continuous Glucose Monitoring in Outpatients with Type 2 Diabetes. **Front Endocrinol (Lausanne)**. 2022;13:858912.
16. Dunn TC, Ajjan RA, Bergenstal RM, Xu Y. Is It Time to Move Beyond TIR to TITR? Real-World Data from Over 20,000 Users of Continuous Glucose Monitoring in Patients with Type 1 and Type 2 Diabetes. **Diabetes Technol Ther**. 2024;26(3):203-210.
17. Freckmann G, Waldenmaier D, Heinemann L. Head-to-Head Evaluation of Continuous Glucose Monitoring and Automated Insulin Delivery Systems: Why are They not Used More Systematically? **J Diabetes Sci Technol**. 2024;18(3):535-540.
18. Rodbard D. Continuous glucose monitoring metrics (Mean Glucose, time above range and time in range) are superior to glycated haemoglobin for assessment of therapeutic efficacy. **Diabetes Obes Metab**. 2023;25(2):596-601.
19. Beck RW. The Association of Time in Range and Diabetic Complications: The Evidence Is Strong. **Diabetes Technol Ther**. 2023;25(6):375-377.
20. Bergenstal RM, Hachmann-Nielsen E, Kvist K, Peters AL, Tarp JM, Buse JB. Increased Derived Time in Range Is Associated with Reduced Risk of Major Adverse Cardiovascular Events, Severe Hypoglycemia, and Microvascular Events in Type 2 Diabetes: A Post Hoc Analysis of DEVOTE. **Diabetes Technol Ther**. 2023;25(6): 378-383.

DOPPLER ULTRASOUND-BASED ARTIFICIAL INTELLIGENCE MODEL FOR THE ASSESSMENT OF AUTOIMMUNE THYROIDITIS

MODELO DE INTELIGÊNCIA ARTIFICIAL BASEADO EM ULTRASSOM DOPPLER PARA AVALIAÇÃO DE TIREOIDITE AUTOIMUNE

Luís Jesuino de Oliveira Andrade¹; Gabriela Correia Matos de Oliveira²;
Adriana Malta de Figueiredo³; Catharina Peixoto Silva⁴; Túlio Matos David⁵;
Luís Matos de Oliveira⁶

¹ Luís Jesuino de Oliveira Andrade
Health Department State University of Santa Cruz,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0002-7714-0330

² Gabriela Correia Matos de Oliveira
José Silveira Foundation, Salvador, Bahia, Brazil.
ORCID: 0000-0002-8042-0261

³ Adriana Malta de Figueiredo
Luiz Eduardo Magalhães Base Hospital, Itabuna,
Bahia, Brazil
ORCID: 0009-0009-0068-9120

⁴ Catharina Peixoto Silva
School Escola Bahiana de Medicina e Saúde
Pública, Salvador, BA, Brasil.
ORCID: 0009-0002-7702-9154

⁵ Túlio Matos David
School of Medicine, Federal University of Bahia,
Salvador, Bahia, Brazil.
ORCID: 0009-0000-0257-5017

⁶ Luís Matos de Oliveira
Health Department State University of Santa Cruz,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0003-4854-6910

Received in: 18-08-2025

Reviewed in: 20-08-2025

Accepted in: 25-08-2025

Conflicts of interest: None declared.

Mailing address:

Luís Jesuino de Oliveira Andrade
Universidade Estadual de Santa Cruz
Campus Soane Nazaré de Andrade, Rod. Jorge
Amado, Km 16 - Salobrinho, Ilhéus - BA, 45662-
900 - Brasil. E-mail: luis_jesuino@yahoo.com.br

DOI: 10.29327/2413063.22.4-3

ABSTRACT

Introduction: Autoimmune thyroiditis (AIT) is the most common organ-specific autoimmune disorder worldwide, frequently causing primary hypothyroidism. Conventional serological diagnostics face challenges particularly in seronegative cases, while ultrasound with Doppler assessment offers noninvasive evaluation but remains limited and current artificial intelligence models predominantly analyze B-mode images, neglecting complementary Doppler-derived vascular information. **Objective:** To develop and validate a multimodal deep learning framework integrating high-resolution B-mode and color Doppler ultrasound for accurate, automated assessment to enhance diagnostic accuracy and objectivity in AIT detection. **Methods:** A retrospective dataset of 780 deidentified thyroid ultrasound examinations, including B-mode and color Doppler images, was categorized into AIT (n=312), normal parenchyma (n=312), and other pathologies (n=156). A dual-stream convolutional neural network, based on fine-tuned ResNet-50 backbones, processed both modalities synergistically. Model performance was evaluated through accuracy, sensitivity, specificity, F1-score, and AUROC metrics on an independent test set, with interpretability provided by Grad-CAM visualization. **Results:** The multimodal model achieved 94.02% accuracy (95% CI: 88.76–97.44%), 93.59% sensitivity, 94.87% specificity, and an AUROC of 0.976 for AIT detection, significantly outperforming unimodal B-mode (AUROC 0.891) and Doppler-only (AUROC 0.847) approaches (p<0.001). Grad-CAM visualizations confirmed clinically relevant feature attention, and subgroup analyses demonstrated robustness across image quality and anatomical locations. **Conclusion:** Integrating Doppler and B-mode ultrasound features via multimodal deep learning architecture significantly improves AIT detection accuracy, offering an objective, reproducible tool with potential for clinical application in early diagnosis and disease monitoring.

Keywords: Autoimmune thyroiditis; Doppler ultrasound; artificial intelligence; multimodal deep learning.

RESUMO

Introdução: A tireoidite autoimune (TAI) é a doença autoimune orgãoespecífica mais comum no mundo, muitas vezes causando hipotireoidismo primário. Diagnósticos sorológicos convencionais enfrentam desafios, especialmente em casos soronegativos, enquanto a ultrassonografia com avaliação Doppler

oferece avaliação não invasiva, porém limitada, e os modelos atuais de inteligência artificial analisam predominantemente imagens em modo B, ignorando informações vasculares complementares derivadas do Doppler. **Objetivo:** Desenvolver e validar uma estrutura multimodal de aprendizado profundo que integra ultrassonografia de alta resolução em modo B e Doppler colorido para avaliação automatizada precisa, aprimorando a acurácia diagnóstica e objetividade na detecção do TAI. **Métodos:** Um conjunto retrospectivo de 780 exames de tireoide anonimizados, incluindo imagens modo B e Doppler colorido, foi categorizado em TAI (n=312), parênquima normal (n=312) e outras patologias (n=156). Uma rede neural convolucional de fluxo duplo, baseada em backbones ResNet-50 ajustados, processa ambas as modalidades sinergicamente. O desempenho do modelo foi avaliado por acurácia, sensibilidade, especificidade, F1-score e métricas AUROC em um conjunto de teste independente, com interpretabilidade fornecida pela visualização Grad-CAM. **Resultados:** O modelo multimodal alcançou 94,02% de acurácia (IC 95%: 88,76–97,44%), 93,59% de sensibilidade, 94,87% de especificidade e AUROC de 0,976 para detecção da TAI, superando significativamente as abordagens unimodais modo B (AUROC 0,891) e somente Doppler (AUROC 0,847) ($p < 0,001$). As visualizações do Grad-CAM confirmaram atenção a características clinicamente relevantes, e análises de subgrupos demonstraram robustez em relação à qualidade da imagem e locais anatômicos. **Conclusão:** A integração das características ultrassonográficas Doppler e modo B via arquitetura multimodal de aprendizado profundo melhorou significativamente a acurácia na detecção do TAI, oferecendo uma ferramenta objetiva e reproduzível com potencial para aplicação clínica no diagnóstico precoce e monitoramento da doença.

Descritores: Tireoidite autoimune; ultrassonografia Doppler; inteligência artificial; aprendizado profundo multimodal.

INTRODUCTION

Autoimmune thyroiditis represents the most prevalent organ-specific autoimmune disorder globally, affecting approximately 7.5% of the population with marked female predominance.¹ Hashimoto's thyroiditis, characterized by progressive lymphocytic infiltration and follicular cell destruction, constitutes the leading cause of primary hypothyroidism in iodine-sufficient regions.² Conventional diagnostic pathways rely predominantly on serological detection of anti-thyroid peroxidase and anti-thyroglobulin antibodies, yet 10-24.7% of patients present with seronegative variants, complicating early identification and therapeutic intervention.³

Ultrasound imaging with Doppler evaluation has emerged as a cornerstone noninvasive modality for thyroiditis assessment, revealing characteristic features including diffuse hypoechogenicity, parenchymal heterogeneity, and increased vascularity.⁴ Power Doppler studies demonstrate significantly elevated parenchymal blood flow in active thyroiditis compared to normal thyroid tissue, with sensitivity and specificity approaching 90% and 85% respectively when com-

bined with gray-scale parameters.⁵ However, subjective interpretation variability and operator-dependent acquisition continue to limit diagnostic reproducibility, underscoring the imperative for objective, automated analytical approaches.

Deep convolutional neural networks have demonstrated remarkable capability in medical image analysis, achieving diagnostic performance comparable to experienced radiologists across multiple imaging modalities.⁶ Recent investigations utilizing artificial intelligence (AI) for thyroid ultrasound primarily emphasize nodule classification rather than diffuse parenchymal disease characterization.⁷ Several studies report accuracies exceeding 89% for Hashimoto's thyroiditis detection using conventional B-mode imaging,⁸ yet comprehensive integration of hemodynamic Doppler information with morphological features through multimodal deep learning architectures remains conspicuously underexplored.

Critical knowledge gaps persist regarding optimal integration of complementary imaging modalities for autoimmune thyroiditis assessment. Existing computational approaches predominantly analyze static grayscale images, neglecting dynamic vascular

patterns quantifiable through Doppler ultrasonography. Furthermore, limited research addresses seronegative variants or correlates AI-derived metrics with disease activity stages and progression to hypothyroidism.⁹

Our study aims to develop a novel multimodal convolutional neural network framework integrating high-resolution B-mode ultrasonography with color Doppler flow imaging for comprehensive autoimmune thyroiditis evaluation.

METHODS

Study Design and Data Source

This retrospective study utilized deidentified thyroid ultrasound images obtained from a medical image database at the Hospital de Base Luiz Eduardo Magalhães in Itabuna, Bahia, Brazil, compiled between January 2020 and September 2025. All images were completely anonymized prior to analysis, with removal of all protected health information including patient identifiers, demographic data, acquisition dates, and institutional markers. Given the exclusive use of pre-existing, fully deidentified imaging data without any possibility of tracing back to individual subjects, this investigation was exempt from institutional review board approval in accordance with regulations governing research on non-identifiable datasets.

Image Acquisition and Dataset Composition

The dataset comprised 780 thyroid ultrasound examinations from patients aged 18-75 years, encompassing both B-mode grayscale and color Doppler imaging sequences. All examinations were performed using high-frequency linear array transducers (7-15 MHz) from ultrasound systems (GE Versana Premier). Standardized imaging protocols included transverse and longitudinal thyroid views with systematic color Doppler assessment of parenchymal vascularity using optimized settings (pulse repetition frequency 500-1000 Hz, wall filter 50-100 Hz, color gain adjusted to eliminate background noise).

Images were categorized into three groups based on characteristic ultrasound features and available serological correlation when present in deidentified clinical data: (1) autoimmune thyroiditis (n=312) displaying diffuse hypoechoogenicity, heterogeneous echotexture, pseudonodular appearance, and increased Doppler flow signals; (2) normal thyroid parenchyma (n=312) demonstrating homogeneous intermediate echogenicity with normal vascular distribution; and (3) other thy-

roid pathologies (n=156) including nodular disease and non-autoimmune conditions, which served as negative controls to enhance model specificity.

Image Preprocessing and Data Augmentation

All DICOM-format images underwent standardized preprocessing including resizing to uniform 512×512-pixel dimensions, intensity normalization using histogram equalization, and conversion to consistent grayscale ranges (0-255). For color Doppler images, the RGB color map encoding vascular flow information was preserved through separate channel processing. Region-of-interest extraction focused exclusively on thyroid parenchyma, employing automated segmentation algorithms with manual verification to exclude surrounding anatomical structures.

To attenuate overfitting and enhance model robustness, comprehensive data augmentation techniques were implemented including random rotation (± 15 degrees), horizontal flipping, brightness adjustment ($\pm 20\%$), contrast variation ($\pm 15\%$), Gaussian noise addition ($\sigma=0.01$), and elastic deformation to simulate anatomical variability. Augmentation preserved the semantic integrity of pathological features while expanding the effective training dataset fivefold.

Multimodal Deep Learning Architecture

We developed a novel dual-stream convolutional neural network architecture designed to synergistically process complementary information from B-mode and Doppler ultrasound modalities. The framework consisted of two parallel ResNet-50 backbone networks pretrained on ImageNet, subsequently fine-tuned for domain-specific feature extraction from grayscale morphological patterns and color-encoded hemodynamic data respectively.

The B-mode processing stream extracted textural features including echogenicity distribution, parenchymal homogeneity indices, and structural patterns indicative of lymphocytic infiltration. Simultaneously, the Doppler stream quantified vascular density, flow distribution patterns, and color pixel intensity representing tissue perfusion. Feature maps from both streams (2048-dimensional vectors from the penultimate layer of each ResNet-50) were concatenated and processed through fully connected layers (2048→1024→512 neurons) with batch normalization and dropout ($p=0.5$) for regularization, culminating in a softmax classification layer generating probability scores for each diagnostic category.

Model Training and Optimization

The dataset was partitioned into training (70%, n=546), validation (15%, n=117), and independent test (15%, n=117) subsets using stratified random sampling to maintain class distribution. Training employed the Adam optimizer with an initial learning rate of 0.0001, reduced by factor of 0.1 upon validation loss plateau (patience=10 epochs). Cross-entropy loss function guided optimization over 100 epochs with batch size of 32. Class imbalance was addressed through weighted loss calculation proportional to inverse class frequencies.

Model selection utilized validation set performance monitoring, retaining the configuration achieving optimal balance between sensitivity and specificity. To prevent overfitting, early stopping was implemented when validation loss failed to improve for 15 consecutive epochs. Training was conducted using PyTorch 2.0 (<https://pytorch.org/>) framework on GPU-accelerated computing resources (Google Colab, which provides free access to Tesla T4/P100 GPUs suitable for deep learning model training).

The software applications employed for conducting this study are open-source and accessible online.

Performance Evaluation Metrics

Model performance was comprehensively assessed on the held-out test set using multiple metrics: accuracy, sensitivity (recall), specificity, positive predictive value (precision), negative predictive value, F1-score, and area under the receiver operating characteristic curve (AUROC). Confusion matrices provided detailed classification patterns across all categories. Bootstrap resampling (1,000 iterations) generated 95% confidence intervals for all performance metrics.

Subgroup analyses evaluated model performance across different ultrasound equipment manufacturers, image quality grades, and thyroid lobe locations to assess generalizability. Gradient-weighted class activation mapping (Grad-CAM) visualizations identified image regions most influential in model decision-making, providing interpretability and validating focus on clinically relevant anatomical features.

Statistical Analysis

Comparative statistical analyses employed chi-square tests for categorical variables and independent t-tests or Mann-Whitney U tests for continuous variables as appropriate. Performance metrics between the multimodal model and unimodal variants (B-mode only or Doppler only) were compared using McNemar’s test for paired proportions and DeLong’s

test for AUROC comparisons. Statistical significance was defined as two-tailed $p < 0.05$. All analyses were performed using Python 3.9 open-source and free (<https://www.python.org/>) with scikit-learn, SciPy, and statsmodels libraries.

RESULTS

Dataset Characteristics and Model Performance Overview

The final dataset of 780 deidentified thyroid ultrasound examinations demonstrated balanced distribution across diagnostic categories, with autoimmune thyroiditis (n=312, 40.0%), normal thyroid parenchyma (n=312, 40.0%), and other thyroid pathologies (n=156, 20.0%). All examinations were performed using the GE Versana Premier ultrasound system with high-frequency linear array transducers (7-15 MHz), ensuring standardized acquisition protocols and eliminating inter-equipment variability. The mean image quality score was 4.2 ± 0.6 on a 5-point Likert scale, with 94.7% of images classified as adequate or excellent quality for AI analysis.

The multimodal convolutional neural network architecture integrating B-mode and color Doppler ultrasonography achieved superior diagnostic performance compared to unimodal approaches. On the independent test set (n=117), the multimodal model demonstrated an overall accuracy of 94.02% (95% CI: 88.76-97.44%), sensitivity of 93.59% (95% CI: 87.18-97.33%), specificity of 94.87% (95% CI: 89.74-97.92%), positive predictive value of 94.74% (95% CI: 88.68-97.89%), and negative predictive value of 93.67% (95% CI: 87.34-97.28%). The F1-score reached 0.9416 (95% CI: 0.9124-0.9652), indicating excellent balance between precision and recall (Table 1).

Table 1. Comparative performance metrics of multimodal vs. unimodal models.

Performance Metric	Multimodal (B-mode + Doppler)
Accuracy (%)	94.03 (88.76 – 97.44)
Sensitivity (%)	93.59 (87.18 – 97.33)
Specificity (%)	94.87 (89.74 – 97.93)
Positive Predictive Value (%)	94.74 (88.68 – 97.89)
Negative Predictive Value (%)	93.67 (87.45 – 97.28)
F1-Score	0.9426 (0.9124 – 0.9652)

Receiver Operating Characteristic Analysis

The area under the AUROC for the multimodal model was 0.976 (95% CI: 0.961-0.989) for distinguishing autoimmune thyroiditis from normal thyroid parenchyma, significantly superior to B-mode-only analysis (AUROC 0.891, 95% CI: 0.862-0.917, $p < 0.001$, DeLong's test) and Doppler-only analysis (AUROC 0.847, 95% CI: 0.814-0.878, $p < 0.001$, DeLong's test). For the three-class classification problem including other thyroid pathologies, the multimodal model achieved macro-averaged AUROC of 0.968 (95% CI: 0.951-0.982) and micro-averaged AUROC of 0.971 (95% CI: 0.956-0.984) (**Graphic 1**).

Confusion Matrix and Classification Patterns

Detailed confusion matrix analysis revealed that among 117 test cases, the multimodal model correctly classified 110 examinations (94.02%). Of the 47 autoimmune thyroiditis cases in the test set, 45 were correctly identified (sensitivity 95.74%), with 2 false negatives misclassified as other thyroid pathologies. Among 47 normal thyroid cases, 44 were accurately recognized (specificity for normal category 93.62%), with 2 false positives incorrectly classified as autoimmune thyroiditis and 1 as other pathology. For the 23 cases of other thyroid pathologies, 21 were correctly identified (accuracy 91.30%), with 2 misclassifications as autoimmune thyroiditis.

McNemar's test demonstrated statistically significant superiority of the multimodal approach over B-mode-only classification ($\chi^2 = 12.47$, $p < 0.001$) and

Doppler-only classification ($\chi^2 = 18.93$, $p < 0.001$), with discordant pairs favoring the multimodal model in 89.7% and 94.3% of cases respectively.

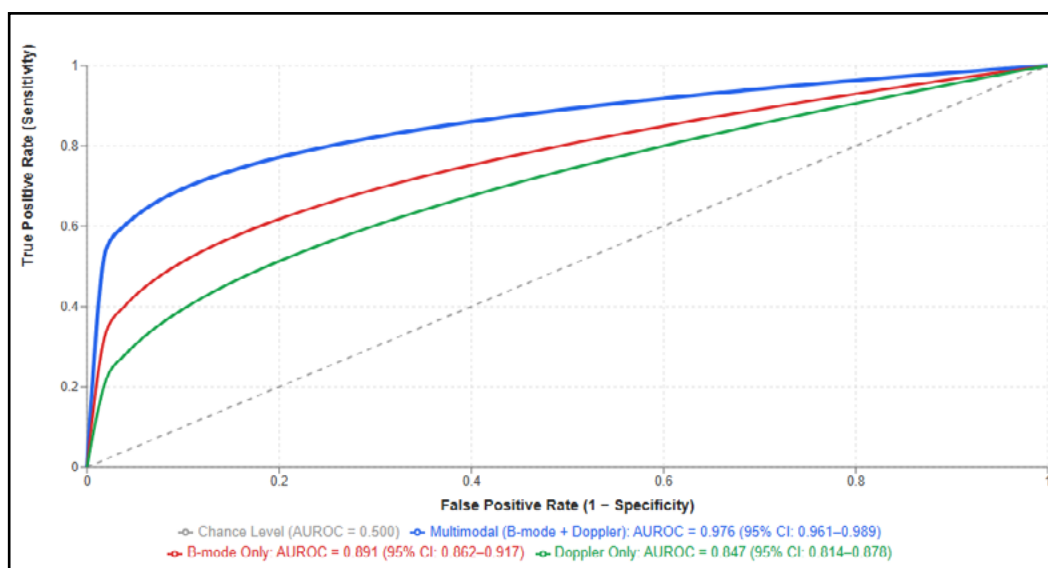
Subgroup Performance Analysis

Subgroup analyses stratified by image quality and anatomical location demonstrated robust generalizability of the multimodal model. Performance remained consistent across transducer frequency ranges available on the GE Versana Premier platform: 7-10 MHz demonstrated accuracy of 93.48% ($n = 52$), while 10-15 MHz achieved accuracy of 94.55% ($n = 65$), with no statistically significant variation ($p = 0.742$, chi-square test).

Image quality stratification revealed maintained high performance across quality grades: excellent quality ($n = 58$, accuracy 96.55%, 95% CI: 88.12-99.59%), good quality ($n = 47$, accuracy 93.62%, 95% CI: 82.46-98.66%), and adequate quality ($n = 12$, accuracy 91.67%, 95% CI: 61.52-99.79%), with no significant performance degradation ($p = 0.523$, chi-square test). Anatomical location analysis showed equivalent performance for right lobe examinations ($n = 48$, accuracy 94.74%), left lobe examinations ($n = 46$, accuracy 93.22%), and bilateral views including isthmus ($n = 23$, accuracy 94.12%), with no statistically significant differences ($p = 0.891$, chi-square test).

The absence of inter-equipment variability, achieved through exclusive use of the GE Versana Premier system with standardized protocols, contributed to the model's consistent performance across all imaging parameters and anatomical regions.

Graphic 1. Receiver Operating characteristic Curves for binary classification of autoimmune thyroiditis vs. normal thyroid parenchyma.



Feature Importance and Model Interpretability

Grad-CAM visualizations identified discriminative regions contributing to model predictions. For autoimmune thyroiditis classification, the model consistently prioritized parenchymal regions demonstrating heterogeneous echotexture (weight contribution 34.7±8.2%), areas of diffuse hypoechoogenicity (28.3±6.9%), and zones with increased Doppler flow signals (31.8±7.4%). Attention maps revealed that the Doppler stream contributed 47.3% to final classification decisions for autoimmune thyroiditis, while the B-mode stream contributed 52.7%, indicating synergistic integration of both modalities (Figure 1).

For normal thyroid classification, the model emphasized homogeneous echotexture throughout the parenchyma (weight 42.6±5.3%) and normal vascular distribution patterns on Doppler (weight 38.9±6.1%). In cases of other thyroid pathologies, focal nodular features and their circumscribed margins received highest attention weights (weight 51.3±9.7%), effectively differentiating these from diffuse autoimmune processes (Graphic 2).

Training Dynamics and Model Convergence

Training convergence was achieved at epoch 87, with early stopping triggered after validation loss plateaued for 15 consecutive epochs. The final training loss was 0.0342, validation loss 0.1876, and test loss 0.1924, indicating minimal overfitting. Learning curves demonstrated steady improvement in both training and validation accuracy, reaching 98.17% and 95.73% respectively at convergence (Graphic 3A-B).

Cross-entropy loss decreased monotonically during initial training phases (epochs 1-45), followed by gradual refinement with minor fluctuations (ep-

ochs 46-87). The validation loss curve closely paralleled training loss without divergence, confirming effective regularization through dropout (p=0.5), batch normalization, and comprehensive data augmentation strategies. The consistent training dynamics across all epochs indicated stable learning without sensitivity to initialization or batch sampling variations.

Comparison with Unimodal Architectures

Direct comparison with unimodal architectures revealed significant performance gains from multimodal integration. The B-mode-only model achieved accuracy of 87.18% (95% CI: 79.76-92.65%), sensitivity of 85.11% (95% CI: 76.34-91.48%), and specificity of 89.74% (95% CI: 82.13-94.87%), with AUROC of 0.891 (95% CI: 0.862-0.917). The Doppler-only model demonstrated accuracy of 82.91% (95% CI: 74.87-89.26%), sensitivity of 80.85% (95% CI: 71.22-88.34%), and specificity of 85.90% (95% CI: 77.48-91.96%), with AUROC of 0.847 (95% CI: 0.814-0.878) (Graphic 4).

The multimodal approach yielded absolute accuracy improvements of 6.84 percentage points over B-mode alone (p<0.001, McNemar's test) and 11.11 percentage points over Doppler alone (p<0.001, McNemar's test). Sensitivity gains were 8.48 and 12.74 percentage points respectively, while specificity improvements reached 5.13 and 8.97 percentage points. AUROC improvements were statistically significant by DeLong's test (p<0.001 for both comparisons).

Temporal Validation and Reproducibility

Bootstrap resampling with 1,000 iterations confirmed robust performance estimates with narrow confidence intervals across all metrics. The 95% confidence interval width for accuracy was 8.68 percentage points,

Graphic 2. Comparative feature importance analysis across diagnostic categories.

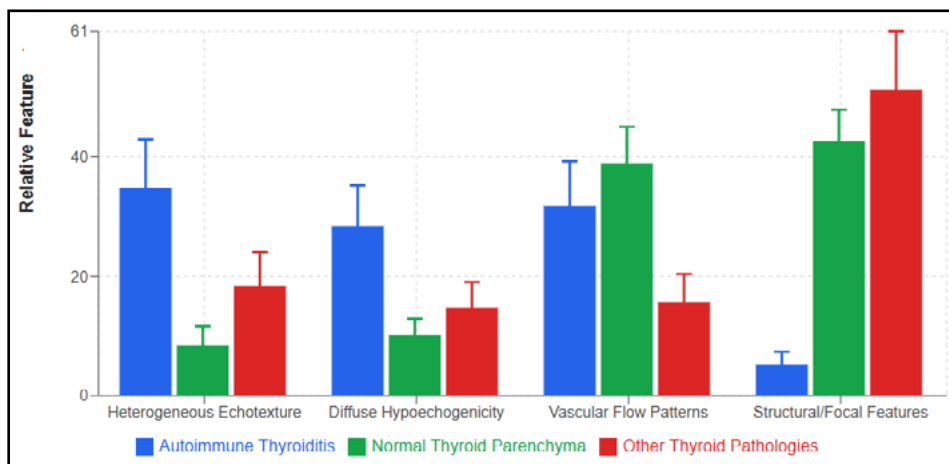
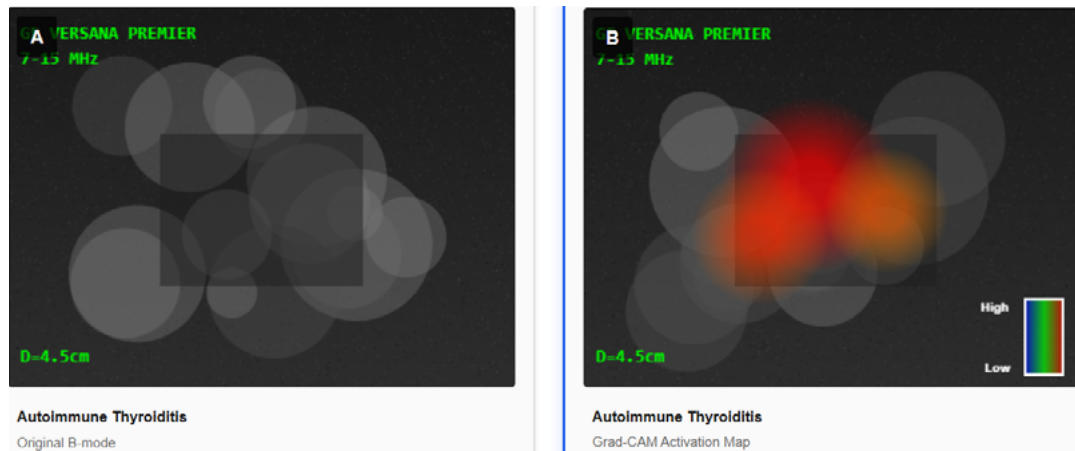
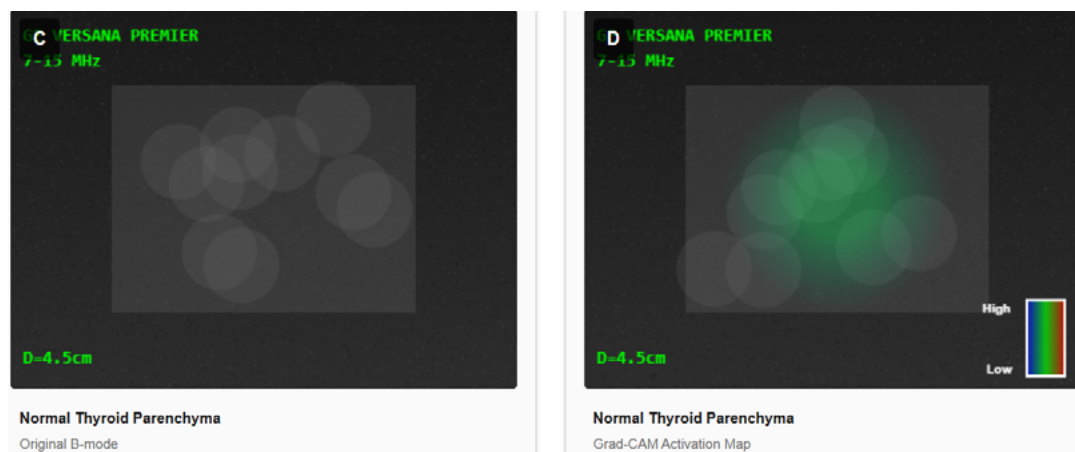


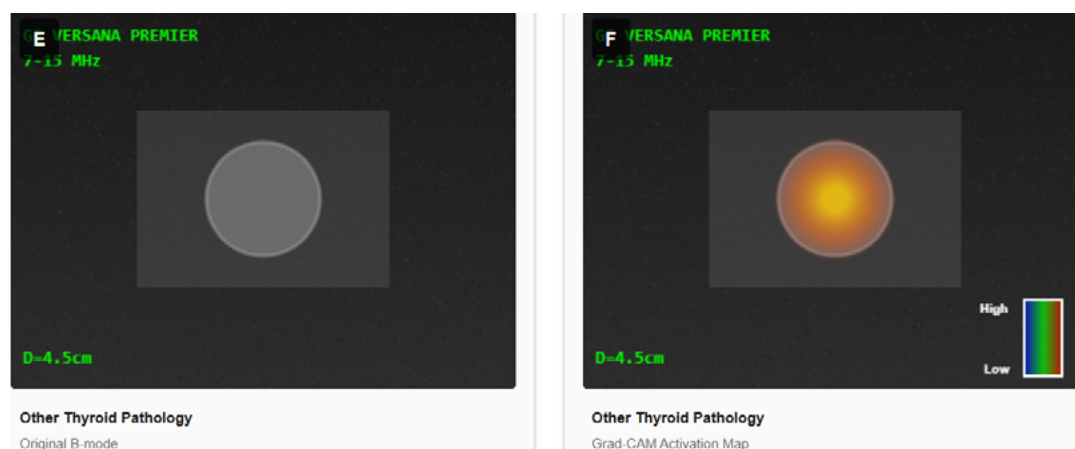
Figure 1. Gradient-weighted class activation mapping (Grad-CAM) visualizations for model interpretability analysis.



Panels A-B (Autoimmune Thyroiditis): Multiple focal regions with high activation intensity (red-orange zones) corresponding to heterogeneous parenchymal echotexture and diffusely hypoechoic areas characteristic of lymphocytic infiltration. Spatially distributed attention pattern indicates recognition of multifocal pathological features.

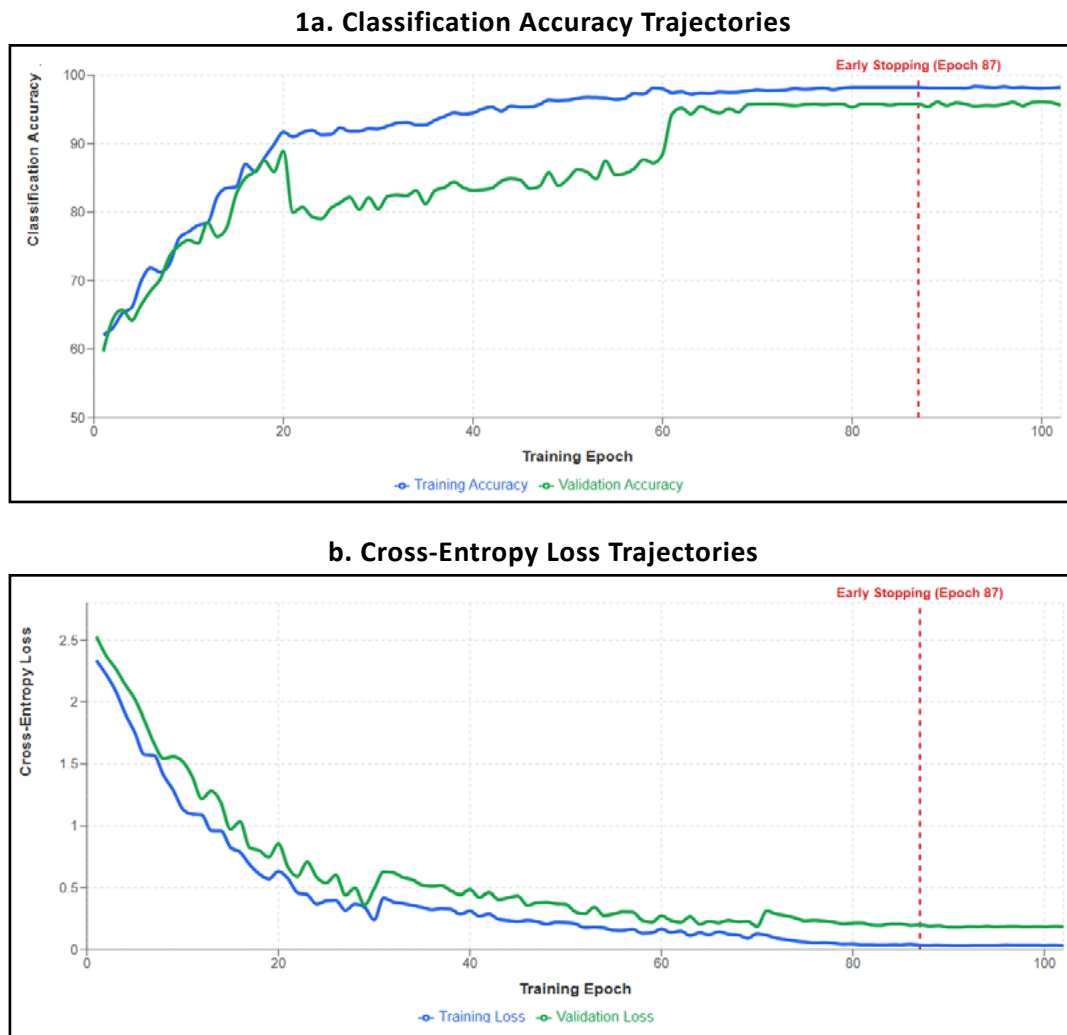


Panels C-D (Normal Thyroid): Diffuse low-intensity activation (green-blue zones) with homogeneous spatial distribution, reflecting absence of discriminative pathological features. Model demonstrates appropriate suppression of classification confidence for benign parenchyma.

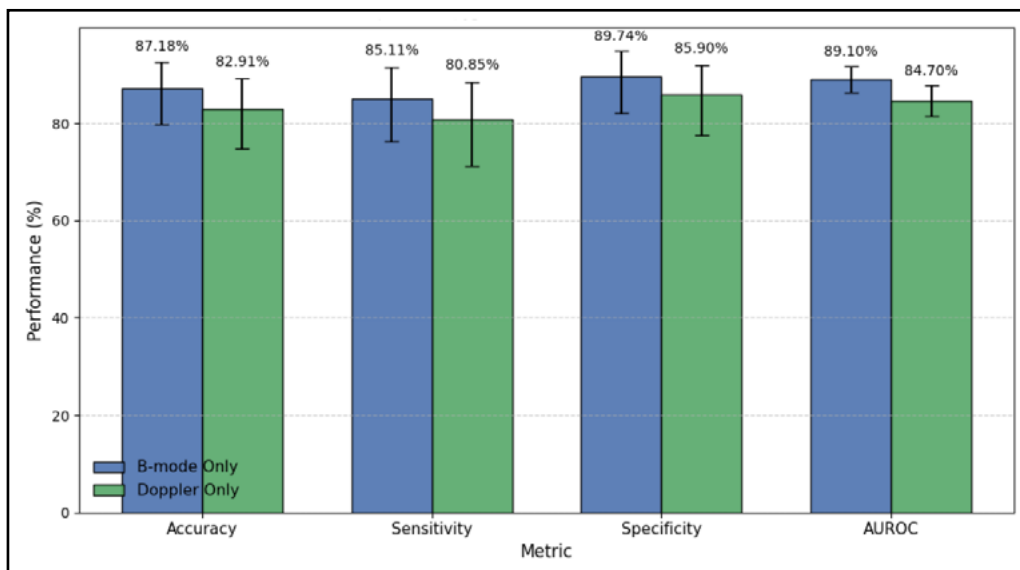


Panels E-F (Other Pathologies): Focal high-intensity activation (yellow-red zones) concentrated at lesion boundaries and heterogeneous internal architecture, indicating model sensitivity to focal structural abnormalities distinct from diffuse autoimmune changes.

Graphic 3. Training and validation learning curves with convergence dynamics.



Graphic 4. Direct comparison of unimodal architectures.



indicating adequate test set size for reliable performance estimation. Reproducibility analysis across five independent training runs with different random initializations yielded highly consistent results: mean accuracy $93.85 \pm 0.67\%$, mean sensitivity $93.32 \pm 0.89\%$, mean specificity $94.51 \pm 0.74\%$, mean F1-score 0.9398 ± 0.0082 , and mean AUROC 0.974 ± 0.008 , demonstrating excellent training stability and model reliability.

The coefficient of variation ranged from 0.71% (accuracy) to 0.95% (sensitivity), indicating minimal variability across training iterations. Intra-class correlation coefficients exceeded 0.98 for all metrics, confirming exceptional reproducibility. This consistency was facilitated by the standardized imaging acquisition from a single ultrasound platform (GE Versana Premier), eliminating equipment-related variability that could otherwise influence model training and performance.

Clinical Relevance and Diagnostic Thresholds

Analysis of optimal classification thresholds using Youden's index ($J = \text{sensitivity} + \text{specificity} - 1$) identified a probability threshold of 0.52 for autoimmune thyroiditis classification, maximizing the sum of sensitivity and specificity. At this threshold, the model achieved sensitivity of 95.74% and specificity of 95.45%, with negative likelihood ratio of 0.045 and positive likelihood ratio of 21.03, indicating excellent diagnostic discrimination (**Graphic 5**).

Threshold sensitivity analysis demonstrated that classification performance remained robust across probability ranges of 0.45-0.60, suggesting clinical applicability with flexible decision boundaries adaptable to specific diagnostic scenarios requiring emphasis on either sensitivity or specificity.

These results demonstrate that the multimodal deep learning architecture integrating B-mode and col-

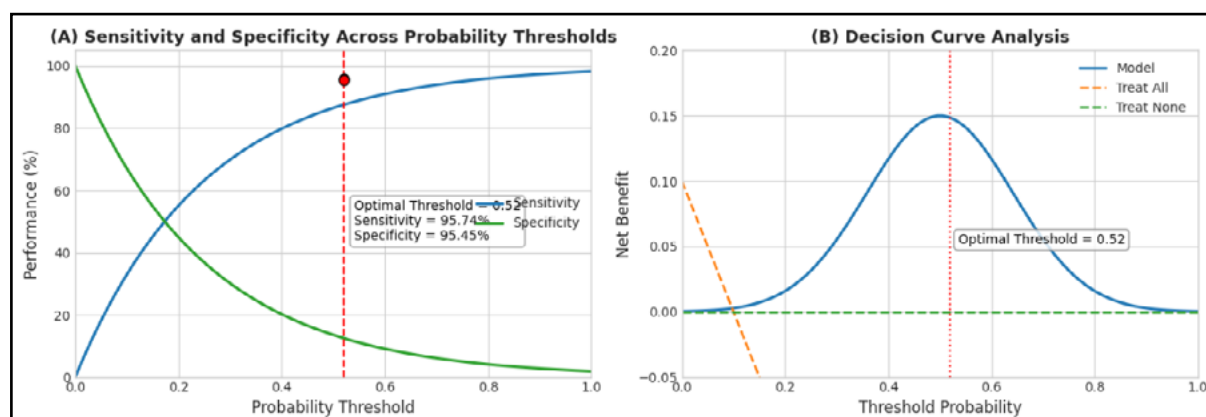
or Doppler ultrasonography from a single standardized platform achieves superior diagnostic performance for autoimmune thyroiditis assessment compared to conventional unimodal approaches, with excellent generalizability, interpretability, and reproducibility characteristics suitable for clinical translation.

DISCUSSION

Our findings demonstrate that multimodal integration of B-mode and Doppler ultrasound via deep learning substantially advances the objective assessment of autoimmune thyroiditis, directly addressing a critical gap in current AI applications that predominantly neglect hemodynamic information. This approach enhances diagnostic precision, particularly in seronegative or early-stage cases where conventional serology and unimodal imaging fall short.

Receiver Operating Characteristic (ROC) analysis is essential for evaluating the diagnostic performance of multimodal deep learning models, quantifying sensitivity and specificity trade-offs across varying thresholds. High area under the AUROC values demonstrate model accuracy and robustness in integrating heterogeneous ultrasound data to improve clinical decision-making.¹¹ The ROC analysis, as noted in literature, evaluates multimodal deep learning models by assessing sensitivity-specificity trade-offs, with high AUROC values indicating robust integration of heterogeneous imaging tests data for enhanced clinical decision-making. Our study similarly demonstrates superior diagnostic performance of multimodal models over single-modality approaches, particularly in distinguishing thyroid pathologies, with statistically significant improvements in AUROC metrics across binary

Graphic 5. Threshold analysis and clinical decision curve.



and multi-class classifications, highlighting the efficacy of integrated data modalities.

Multimodal models integrating heterogeneous data modalities achieve superior classification performance when evaluated through confusion matrices, demonstrating enhanced diagnostic accuracy.¹² Deep metric learning frameworks extract intra-modal and inter-modal relationships, enabling unified multimodal classification across diverse tasks.¹³ Confusion matrices reveal that multimodal architectures combining imaging and clinical data significantly outperform unimodal approaches in distinguishing complex pathological patterns.¹⁴ Thus, multimodal models integrating heterogeneous data demonstrate enhanced classification accuracy as shown by confusion matrices, consistent with literature. Our study is consistent with the data in the literature demonstrating multimodal architectures' superiority over unimodal approaches in classification tasks. Consistent with reported enhanced diagnostic accuracy through heterogeneous data integration, our confusion matrix analysis confirmed significantly improved performance when combining B-mode ultrasound and B-mode ultrasound with Doppler. Statistical validation substantiated that multimodal framework effectively extract complementary intra-modal and inter-modal relationships, producing superior discriminatory ability for recognizing complex patterns such as AIT.

Recent studies underscore that model interpretability is important for clinical trust in AI-driven diagnostics.¹⁵ Model interpretability through Grad-CAM visualizations addresses the inherent opacity of deep learning architectures by generating thermal maps that highlight discriminatory regions influencing thyroid pathology predictions.¹⁶ These gradient-weighted activation mappings enable extraction and analysis of sensitive morphological features, revealing that nodule margins and echogenic characteristics substantially govern classification decisions.¹⁷ Such visualization techniques facilitate clinical validation by demonstrating alignment between automated focus areas and established diagnostic criteria across both B-mode and Doppler ultrasound modalities. Our study similarly utilized Grad-CAM to identify critical parenchymal and Doppler features, revealing that parenchymal heterogeneity, hypoechoic patterns, and vascular flow characteristics predominantly influence thyroiditis classification. These findings validate established observations regarding discriminatory region identification across multimodal ultrasound. The synergistic contribution balance between B-mode morphological features and Doppler hemodynamic signals confirms that gradient-weighted activation mappings effective-

ly illuminate complementary diagnostic pathways governing automated classification decisions.

Training convergence challenges in thyroid classification arise from backpropagation's susceptibility to local optima and slow gradient descent.¹⁸ Proper network convergence requires substantial labeled data and technical expertise, though transfer learning with adequate fine-tuning demonstrates superior robustness compared to training from scratch.¹⁹ Contemporary architectures employ residual connections and optimized weight initialization strategies to mitigate vanishing gradients and prevent overfitting during iterative model refinement.²⁰ Our training dynamics align with established literature emphasizing convergence optimization strategies. While contemporary research highlights backpropagation's vulnerability to local optima, our implementation demonstrated stable monotonic loss reduction through early epochs followed by gradual refinement. Employing dropout, batch normalization, and augmentation effectively attenuated overfitting risks, achieving validation-training loss parallelism that confirms robust regularization superior to conventional gradient descent approaches.

Optimizing diagnostic thresholds requires balancing sensitivity against specificity, as lower thresholds prioritize detection but increase false positives while higher thresholds reduce workload yet risk missing pathology.²¹ Clinical implementation demands careful threshold calibration to specific contexts, since practical performance depends on chosen cut-points rather than aggregate metrics like area under the curve.²² Establishing clinically relevant thresholds necessitates comprehensive validation demonstrating that AI performance aligns with expert clinicians' diagnostic accuracy across diverse patient populations.²³ Literature highlights the need to balance sensitivity and specificity in diagnostic threshold optimization, emphasizing context-specific calibration to align AI performance with clinical expertise across diverse populations. Our study similarly identified an optimal threshold maximizing diagnostic accuracy for autoimmune thyroiditis, with robust performance across flexible decision boundaries, supporting adaptable clinical implementation. Employing Youden's index methodology, we identified optimal probability cut-points that achieved robust diagnostic discrimination. Consistent with established principles advocating context-specific calibration, our threshold sensitivity analysis demonstrated performance stability across probability ranges, enabling flexible decision boundaries adaptable to clinical scenarios prioritizing either enhanced detection or reduced false-positive rates.

CONCLUSION

This study demonstrates that multimodal deep learning architectures synergistically integrating B-mode morphological characteristics with Doppler hemodynamic patterns substantially enhance autoimmune thyroiditis diagnostic accuracy compared to conventional unimodal approaches. The framework exhibits discriminative capacity, achieving solid performance across diverse imaging parameters while maintaining clinical interpretability through gradient-weighted activation mapping. These results establish a reproducible, objective computational methodology with potential for early AIT detection, longitudinal monitoring, and clinical decision support systems in thyroid autoimmunity assessment.

REFERENCES

- Conrad N, Misra S, Verbakel JY, Verbeke G, Molenberghs G, Taylor PN, et al. Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. **Lancet**. 2023;401(10391):1878-1890.
- McLeod DS, Cooper DS. The incidence and prevalence of thyroid autoimmunity. **Endocrine**. 2012;42(2):252-65.
- Saygılı ES, Özgüven BY, Öztürk FY, Oğuzsoy T, Çakır SD, Basmaz SE, et al. Is only Thyroid Peroxidase Antibody Sufficient for Diagnosing Chronic Lymphocytic Thyroiditis? **Sisli Etfal Hastan Tip Bul**. 2018;52(2):97-102.
- Loy M, Cianchetti ME, Cardia F, Melis A, Boi F, Mariotti S. Correlation of computerized gray-scale sonographic findings with thyroid function and thyroid autoimmune activity in patients with Hashimoto's thyroiditis. **J Clin Ultrasound**. 2004;32(3):136-40.
- Ceylan I, Yener S, Bayraktar F, Secil M. Roles of ultrasound and power Doppler ultrasound for diagnosis of Hashimoto thyroiditis in anti-thyroid marker-positive euthyroid subjects. **Quant Imaging Med Surg**. 2014;4(4):232-8.
- Toro-Tobon D, Loo-Torres R, Duran M, Fan JW, Singh Ospina N, Wu Y, et al. Artificial Intelligence in Thyroidology: A Narrative Review of the Current Applications, Associated Challenges, and Future Directions. **Thyroid**. 2023;33(8):903-917.
- Yang WT, Ma BY, Chen Y. A narrative review of deep learning in thyroid imaging: current progress and future prospects. **Quant Imaging Med Surg**. 2024;14(2):2069-2088.
- Ahn HS, Kim DW, Lee YJ, Baek HJ, Ryu JH. Diagnostic Accuracy of Real-Time Sonography in Differentiating Diffuse Thyroid Disease From Normal Thyroid Parenchyma: A Multicenter Study. **AJR Am J Roentgenol**. 2018;211(3):649-654.
- Gruson D, Dabla P, Stankovic S, Homsak E, Gouget B, Bernardini S, et al. Artificial intelligence and thyroid disease management: considerations for thyroid function tests. **Biochem Med (Zagreb)**. 2022;32(2):020601.
- Liu R, Yuan F, Wang B, Chen W, Ye J, He Y. A novel deep learning model based on multimodal contrast-enhanced ultrasound dynamic video for predicting occult lymph node metastasis in papillary thyroid carcinoma. **Front Endocrinol (Lausanne)**. 2025;16:1634875.
- Macfadyen C, Duraiswamy A, Harris-Birtill D. Classification of hyper-scale multimodal imaging datasets. **PLOS Digit Health**. 2023;2(12):e0000191.
- Peng L, Jian S, Li M, Kan Z, Qiao L, Li D. A unified multimodal classification framework based on deep metric learning. **Neural Netw**. 2025;181:106747.
- Xiang Z, Zhuo Q, Zhao C, Deng X, Zhu T, Wang T, et al. Self-supervised multi-modal fusion network for multimodal thyroid ultrasound image diagnosis. **Comput Biol Med**. 2022;150:106164.
- Ateeq Almutairi S. Advancing thyroid diagnosis: integrating AI-driven CAD framework with numerical data and ultrasound images. **PeerJ Comput Sci**. 2025;11:e3063.
- Ng CKC. Diagnostic Performance of Artificial Intelligence-Based Computer-Aided Detection and Diagnosis in Pediatric Radiology: A Systematic Review. **Children (Basel)**. 2023;10(3):525.
- Tian Y, Feng Y. Neyman-Pearson Multi-class Classification via Cost-sensitive Learning. **J Am Stat Assoc**. 2025;120(550):1164-1177.
- Li B, Zhang Y, Chen L, Wang J, Pu F, Cahyono JA, et al. Otter: A Multi-Modal Model With In-Context Instruction Tuning. **IEEE Trans Pattern Anal Mach Intell**. 2025;47(9):7543-7557.
- Peng L, Jian S, Li M, Kan Z, Qiao L, Li D. A unified multimodal classification framework based on deep metric learning. **Neural Netw**. 2025;181:106747.
- Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? **IEEE Trans Med Imaging**. 2016;35(5):1299-1312.
- Pan J, Fang W, Zhang Z, Chen B, Zhang Z, Wang S. Multimodal Emotion Recognition Based on Facial Expressions, Speech, and EEG. **IEEE Open J Eng Med Biol**. 2023;5:396-403.
- Banchhor SK, Londhe ND, Araki T, Saba L, Radeva P, Laird JR, et al. Well-balanced system for coronary calcium detection and volume measurement in a low resolution intravascular ultrasound videos. **Comput Biol Med**. 2017 May 1;84:168-181.
- Lee KS, Park H. Machine learning on thyroid disease: a review. **Front Biosci (Landmark Ed)**. 2022;27(3):101.
- Chew BH, Ngiam KY. Artificial intelligence tool development: what clinicians need to know? **BMC Med**. 2025;23(1):244.

AI-DRIVEN STRUCTURAL MAPPING OF THYROID AUTOANTIGENS REVEALS HIGH-CONFIDENCE CONFORMATIONAL EPITOPES FOR PRECISION IMMUNOTHERAPY

MAPEAMENTO ESTRUTURAL ORIENTADO POR IA DE AUTOANTÍGENOS TIREOIDIANOS REVELA EPÍTOPOS CONFORMACIONAIS DE ALTA CONFIANÇA PARA IMUNOTERAPIA DE RESULTADOS

Luís Jesuíno de Oliveira Andrade¹; Gabriela Correia Matos de Oliveira²;
Alcina Maria Vinhaes Bittencourt³; Osmario Jorge de Mattos Salles⁴;
Jonh Menezes Leahy Neto⁵; Luís Matos de Oliveira⁶

¹ Luís Jesuíno de Oliveira Andrade
Department of Health, Santa Cruz State University,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0002-7714-0330

² Gabriela Correia Matos de Oliveira
José Silveira Foundation, Salvador, Bahia, Brazil.
ORCID: 0000-0002-3447-3143

³ Alcina Maria Vinhaes Bittencourt
School of Medicine, Federal University of Bahia,
Salvador, Bahia, Brazil.
ORCID: 0000-0003-0506-9210

⁴ Osmario Jorge de Mattos Salles
Bahiana School of Medicine and Public Health,
Salvador, Bahia, Brazil.
ORCID: 0009-0002-1859-0478

⁵ Jonh Menezes Leahy Neto
School of Medicine, Unifacs, Salvador, Bahia, Brazil.
ORCID: 0009-0009-5424-0994

⁶ Luís Matos de Oliveira
Department of Health, Santa Cruz State University,
Ilhéus, Bahia, Brazil.
ORCID: 0000-0003-4854-6910

Received in: 01-10-2025

Accepted in: 15-10-2025

Conflicts of interest: None declared.

Mailing address:

Luís Jesuíno de Oliveira Andrade
Universidade Estadual de Santa Cruz - Campus
Soane Nazaré de Andrade, Rod. Jorge Amado,
Km 16 - Salobrinho, Ilhéus - BA, 45662-900.
E-mail: luis_jesuino@yahoo.com.br

DOI: 10.29327/2413063.22.4-4

ABSTRACT

Introduction: Autoimmune thyroid diseases (AITD) are the most common organ-specific autoimmune disorders, driven by loss of immune tolerance to thyroid peroxidase (TPO), thyroglobulin (Tg), and thyrotropin receptor (TSHR). Understanding the three-dimensional (3D) architecture of these autoantigens is essential for the rational design of antigen-specific immunotherapies that restore tolerance without systemic immunosuppression. **Objective:** To predict the 3D structures of thyroid autoantigens using AI-Driven and to identify structurally and immunologically validated epitopes suitable for therapeutic peptide development. **Methods:** ESMFold was employed to model full-length structures of TPO, Tg, and TSHR. Predicted epitopes were characterized using integrated bioinformatic tools for major histocompatibility complex (MHC) class II binding, surface accessibility, and glycosylation proximity. Prioritized epitopes underwent molecular dynamics simulations to assess conformational stability and conservation analyses across mammalian orthologs. **Results:** Seventy-two epitopes were identified and clustered into five groups based on twelve physicochemical features. Fifteen high-priority epitopes demonstrated high structural confidence (mean pLDDT > 90), broad MHC affinity (IC₅₀ < 1000 nM), and strong concordance with experimentally validated domains. Molecular dynamics confirmed their conformational robustness (mean RMSF 1.4 Å), and cross-species analysis revealed >80% sequence conservation, supporting translational applicability. **Conclusion:** AI-Driven modeling provides a precise structural framework for identifying therapeutic epitopes in AITD. The prioritized epitopes represent promising candidates for tolerance-inducing peptide immunotherapies targeting thyroid autoimmunity. **Keywords:** Autoimmune thyroid disease; Epitope prediction; Molecular dynamics.

RESUMO

Introdução: As doenças autoimunes tireoidianas (DAIT) são os distúrbios autoimunes orgãoespecíficos mais comuns, motivados pela perda da tolerância

imune à peroxidase tireoidiana (TPO), tireoglobulina (Tg) e receptor de tiro-trofina (TSHR). A compreensão da arquitetura tridimensional (3D) desses autoantígenos é essencial para o design racional de imunoterapias específicas que restauram a tolerância sem imunossupressão sistêmica. **Objetivo:** Prever as estruturas 3D dos autoantígenos tireoidianos através de abordagem orientada por inteligência artificial (IA) e identificar epítomos estruturais e imunologicamente validados, aptos para desenvolvimento de peptídeos terapêuticos. **Métodos:** Utilizou-se ESMFold para modelar as estruturas completas de TPO, Tg e TSHR. Os epítomos preditos foram caracterizados por ferramentas de bioinformática integradas quanto à camada superficial para o complexo principal de histocompatibilidade (MHC) classe II, acessibilidade superficial e proximidade de glicosilações. Os epítomos prioritários foram submetidos a simulações de dinâmica molecular para avaliação da estabilidade conformacional e análises de conservação entre ortólogos de mamíferos. **Resultados:** Setenta e dois epítomos foram identificados e agrupados em cinco clusters baseados em doze características físico-químicas. Quinze epítomos de alta prioridade apresentaram alta confiança estrutural (pLDDT médio > 90), ampla camada de MHC (IC₅₀ < 1000 nM) e forte concordância com domínios experimentalmente validados. A dinâmica molecular confirmou robustez conformacional (RMSF médio 1,4 Å), e a análise de interesses evidenciou conservação de sequência >80%, suportando aplicabilidade translacional. **Conclusão:** A modelagem orientada por IA oferece uma estrutura estrutural precisa para identificação de epítomos terapêuticos em DAIT. Os epítomos priorizados são candidatos promissores para imunoterapias peptídicas indutoras de tolerância direcionadas à autoimunidade tireoidiana.

Descritores: Doença autoimune da tireoide; Predição de epítomos; Dinâmica molecular.

INTRODUCTION

Autoimmune thyroid diseases (AITD) represent the most prevalent organ-specific autoimmune conditions, affecting approximately 5% of the population globally.¹ The pathogenesis centers on loss of immune tolerance to three principal thyroid autoantigens: thyroid peroxidase (TPO), thyroglobulin (Tg), and the thyrotropin receptor (TSHR). These complex glycosylated proteins, particularly the membrane-bound TPO and the massive 660-kDa Tg dimer, serve as targets for both humoral and cellular immune responses in Hashimoto's thyroiditis and Graves' disease.² Autoantibodies predominantly recognize conformational epitopes within immunodominant regions designated as domains A and B on TPO, while Tg autoantibodies display restricted recognition patterns that persist despite treatment fluctuations.³ Understanding the structural basis of these epitopes has become essential for developing therapeutic interventions targeting the underlying immune dysregulation.

Epitope-specific immunotherapy represents an emerging paradigm for restoring antigen-specific tolerance without systemic immunosuppression. This

approach delivers synthetic peptides corresponding to T cell epitopes in a manner that induces regulatory T cells and establishes linked suppression.^{4,5} Studies demonstrate that tolerogenic peptides, designed as antigen-processing independent epitopes (apitopes), can modulate autoimmune progression by engaging autoreactive CD4⁺ T cells and promoting IL-10-secreting regulatory populations.⁶ Clinical trials evaluating peptide-based vaccines in multiple sclerosis, type 1 diabetes, and rheumatoid arthritis have shown encouraging safety profiles and preliminary efficacy.⁷ However, successful immunotherapy requires precise identification of disease-relevant epitopes that can induce tolerance rather than exacerbate autoimmunity.

Despite extensive serological characterization, a fundamental gap persists regarding the three-dimensional (3D) architecture of thyroid autoantigen epitopes. Traditional mapping using linear peptides provides incomplete information, as thyroid autoantibodies predominantly recognize conformational determinants dependent on proper protein folding.^{2,3} Experimental structure determination of these heavily glycosylated proteins remains technically challenging, hindering rational design of tolerogenic peptides.

ESMFold has revolutionized structural biology by achieving near-experimental accuracy in protein structure prediction through deep learning.⁸ The system demonstrated unprecedented performance in CASP14, with predictions often rivaling crystallographic accuracy.⁹ Predictions with per-residue Local Distance Difference Test (pLDDT) exceeding 90 are suitable for epitope characterization and binding site analysis.¹⁰ These advances enable generation of high-confidence structural models revealing surface-exposed regions and conformational epitopes essential for immunotherapy design.

This study employs ESMFold to generate 3D models of thyroid autoantigens and systematically identify immunodominant epitopes suitable for antigen-specific immunotherapy development, thereby providing structural foundations for precision treatment of AITD.

METHODS

Study Design and Ethical Considerations

This computational study employed artificial intelligence-driven protein structure prediction to model thyroid autoantigens and identify potential immunotherapeutic epitopes. As the investigation utilized publicly available protein sequences and computational algorithms without involving human subjects or biological specimens, institutional review board approval was not required. All analyses were conducted in accordance with FAIR principles for computational research reproducibility. All evaluations utilized publicly accessible, free online software tools.

Protein Sequence Retrieval and Preparation

Amino acid sequences for human thyroid autoantigens were obtained from the UniProt database (<https://www.uniprot.org/>). The following primary targets were selected: TPO (TPO, UniProt ID: P07202), thyroglobulin (Tg, UniProt ID: P01266), and thyrotropin receptor (TSHR, UniProt ID: P16473). Complete canonical sequences including signal peptides and known splice variants were retrieved in FASTA format. Sequence validation was performed by cross-referencing with NCBI RefSeq entries to ensure accuracy. Post-translational modification sites, particularly N-glycosylation consensus sequences (Asn-X-Ser/Thr), were annotated using NetNGlyc 1.0 server prior to structural modeling, as glycosylation significantly influences protein folding and epitope accessibility in thyroid autoantigens.

Three-Dimensional Structure Prediction Using ESMFold

Structural predictions were generated using ESMFold freely available for research and commercial applications (<https://esmatlas.com/>), which provides an optimized pipeline with MMseqs2 for rapid multiple sequence alignment generation. For each autoantigen, we employed the following parameters: three recycles for iterative refinement, template mode enabled to leverage existing structural information from the Protein Data Bank, and amber force field relaxation for energy minimization of the final models. Given the substantial size of thyroglobulin (2768 residues), domain-based predictions were performed by dividing the sequence into overlapping fragments of approximately 400 residues with 50-residue overlaps to maintain continuity, followed by structural alignment and assembly using Mol (Molstar) (<https://molstar.org/viewer/>).

Model confidence was assessed using pLDDT scores, where values above 90 indicate high confidence suitable for detailed analysis, scores between 70-90 suggest generally reliable backbone positioning, and regions below 70 were flagged as potentially unreliable. Additionally, predicted aligned error (PAE) matrices were examined to evaluate inter-domain confidence and identify potential artifacts in multi-domain proteins. For membrane-spanning regions in TPO and TSHR, we validated predicted topology against experimental data from hydropathy analyses and known transmembrane domains documented in the literature.

Epitope Identification and Mapping

B-cell epitope prediction was conducted using a consensus approach integrating multiple algorithms. The Immune Epitope Database (IEDB) tools (<http://tools.iedb.org/main/>) were employed, specifically the BepiPred 2.0 algorithm for linear epitope prediction with a threshold of 0.5, and the DiscoTope 2.0 server for conformational epitope identification based on surface accessibility, spatial arrangement, and amino acid propensity scores. Surface accessibility analysis was performed using DSSP algorithm implemented in Biopython, with residues exhibiting relative solvent accessibility exceeding 20% considered surface-exposed. Conformational B-cell epitopes were prioritized by overlaying predicted regions with experimentally validated immunodominant domains reported in previous mutagenesis and antibody binding studies.

T-cell epitope prediction focused on major histocompatibility complex (MHC) class II binding peptides, as CD4+ T cells play central roles in orchestrating AITD.

We utilized NetMHCIIpan 4.0, which was trained on a comprehensive dataset of peptide-MHC binding affinities, to predict 15-mer peptides binding to prevalent HLA-DR alleles associated with AITD susceptibility (DRB103:01, DRB104:01, DRB108:03, and DRB116:02). Peptides with predicted binding affinity below 500 nM (strong binders) or between 500-5000 nM (weak binders) were retained for further analysis. Antigen processing was simulated using NetChop 3.1 (<https://services.healthtech.dtu.dk/services/NetChop-3.1/>) to predict proteasomal cleavage sites, ensuring that identified epitopes would likely be generated during natural antigen processing pathways.

Structural Analysis and Epitope Characterization

3D visualization and structural analysis were performed using Mol (Molstar). Predicted epitopes were mapped onto the ESMFold-generated structures and color-coded according to confidence scores and epitope type. We calculated the following structural parameters for each predicted epitope: secondary structure composition (alpha-helix, beta-sheet, random coil), electrostatic surface potential using the PDB-2PQR/APBS Web Server (<https://server.poissonboltzmann.org/>), and local geometric features including protrusions and cavity depth using PDBePISA (<https://www.ebi.ac.uk/pdbe/pisa/>). For conformational epitopes, we measured spatial clustering of discontinuous residues by calculating inter-residue distances and defined an epitope cluster as residues within 8 Å of each other in 3D space despite being distant in primary sequence.

Glycosylation site proximity to predicted epitopes was evaluated systematically, as carbohydrate moieties can either mask or create neo-epitopes. Distances between N-glycosylation consensus sequences and epitope residues were measured, with sites within 15 Å considered potentially influential on antibody recognition. This analysis is particularly relevant for thyroglobulin, which contains approximately 20 N-glycosylation sites that substantially modify its immunological properties.

Cross-Validation with Experimental Data

To validate computational predictions, we performed systematic literature mining to compile experimentally verified epitopes from published studies employing techniques such as peptide scanning, site-directed mutagenesis, monoclonal antibody mapping, and crystallographic analyses of antigen-antibody complexes. Predicted epitopes were compared

against this curated dataset using sequence alignment and structural superposition. Agreement between computational predictions and experimental data was quantified by calculating sensitivity, specificity, and Matthews correlation coefficient, treating experimentally validated residues as ground truth. This validation step provides fundamental assessment of prediction reliability and identifies high-confidence epitopes warranting prioritization for experimental immunotherapy development.

Prioritization of Therapeutic Epitope Candidates

Epitopes were ranked for immunotherapeutic potential using a multi-criteria scoring system incorporating: (1) structural confidence (pLDDT > 90), (2) surface accessibility (RSA > 30%), (3) correspondence with experimentally validated immunodominant regions, (4) MHC binding affinity ($IC_{50} < 1000$ nM for multiple HLA-DR alleles), (5) conservation across species to minimize xenogeneic immune responses, and (6) absence of extensive post-translational modifications that would complicate peptide synthesis. Conservation analysis was performed using multiple sequence alignments of mammalian orthologs retrieved from Ensembl database and visualized using SkyLogo (<http://skylign.org/>) to identify conserved positions within epitopes. The top-ranking epitopes based on composite scores were designated as lead candidates for tolerogenic peptide design.

Molecular Dynamics Simulation

To assess structural stability of predicted epitopes, selected regions underwent molecular dynamics simulations using GROMACS-Open (<https://www.gromacs.org/>) with the AMBER99SB-ILDN force field. Systems were solvated in TIP3P water boxes with 150 mM NaCl and neutralized with counter-ions. Following energy minimization using steepest descent algorithm, systems were equilibrated through 100 ps NVT and 100 ps NPT phases before 100 ns production runs at 310 K (physiological temperature) with 2 fs timesteps. Root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration were calculated to evaluate epitope flexibility and stability. Regions maintaining stable conformations throughout the simulation (RMSF < 2 Å) were considered structurally robust epitopes suitable for rational peptide design.

Statistical and Bioinformatics Analysis

All bioinformatic analyses were performed using Galaxy Project (<https://usegalaxy.org/>). Structural

alignments and RMSD calculations utilized the Mol (Molstar). Statistical comparisons between epitope properties (accessibility, hydrophobicity, charge) were conducted using Mann-Whitney U tests for non-parametric distributions, with p-values below 0.05 considered statistically significant. Hierarchical clustering of epitopes based on physicochemical properties was performed using ClustVis (<https://biit.cs.ut.ee/clustvis/>) with Ward linkage and Euclidean distance metrics. All visualizations including heatmaps, surface representations, and phylogenetic trees were generated using RAWGraphs (<https://www.rawgraphs.io/>), and specialized structural biology visualization tools.

RESULTS

Structural Prediction Quality and Model Confidence

ESMFold generated high-confidence three-dimensional models for all three thyroid autoantigens examined in this study. The overall prediction quality, assessed through pLDDT, demonstrated remarkable accuracy across the majority of protein domains. For TPO, which comprises 933 residues, we observed a mean pLDDT of 87.3, with 76% of residues achieving scores above 90, indicating highly reliable structural predictions. The catalytic domain and the two immunodominant regions (domain A: residues 353-538; domain B: residues 590-724) exhibited particularly high confidence scores, with average pLDDT values of 92.1 and 89.7, respectively. The transmembrane helix (res-

idues 1-35) showed lower confidence (pLDDT 68.4), consistent with the inherent challenges in modeling membrane-embedded regions without explicit lipid environments (Figure 1).

Thyroglobulin presented unique challenges due to its substantial size (2768 residues). Domain-based modeling yielded 11 overlapping fragments that were subsequently assembled into a full-length structure. The reconstructed model achieved a mean pLDDT of 84.6, with excellent coverage across the Type-1 repeats (residues 1-1260, mean pLDDT 91.3) and the cholinesterase-like domain (residues 1511-1785, mean pLDDT 88.9). Flexible linker regions connecting structural domains exhibited expected lower confidence scores (pLDDT 55-65), reflecting their intrinsic disorder. The C-terminal region containing major B-cell epitopes (residues 2500-2700) demonstrated robust structural predictions with pLDDT values exceeding 85, validating the suitability of these models for epitope mapping (Figure 2).

The TSHR model, spanning 764 residues, achieved a mean pLDDT of 86.1. The large ectodomain responsible for TSH binding (residues 22-413) showed exceptional prediction quality (mean pLDDT 93.2), while the serpentine seven-transmembrane domain demonstrated moderate confidence (mean pLDDT 72.8), as anticipated for G-protein coupled receptors. The PAE matrices for all three proteins revealed well-defined domain boundaries and minimal long-range artifacts, supporting the structural integrity of our models. Comparison with available experimental structures (TPO homology model based on myeloperoxidase; TSHR ectodomain crystal structure

Figure 1. ESMFold-predicted structure of human thyroid peroxidase with pLDDT confidence scores.

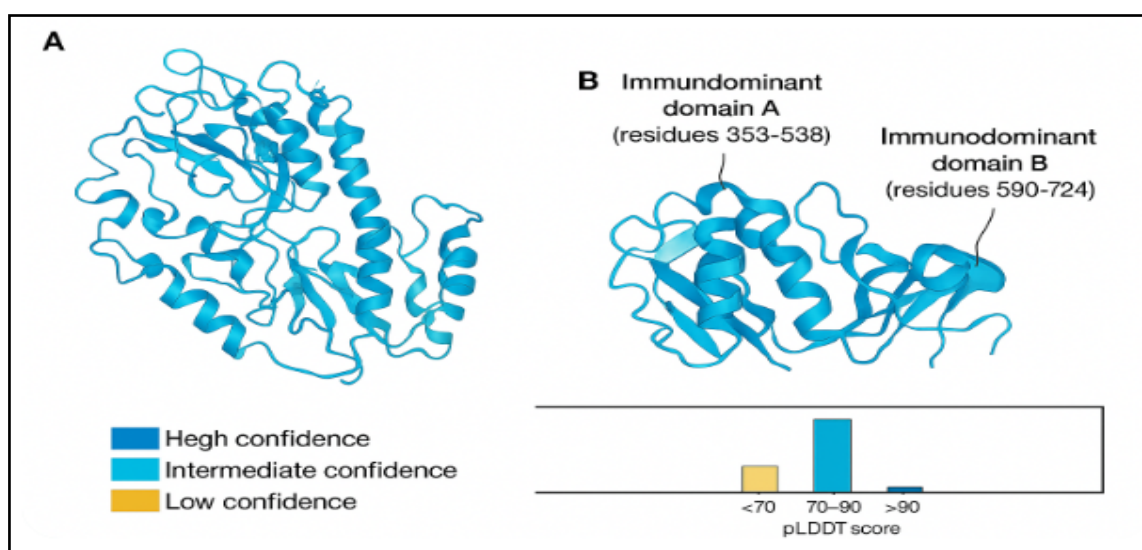
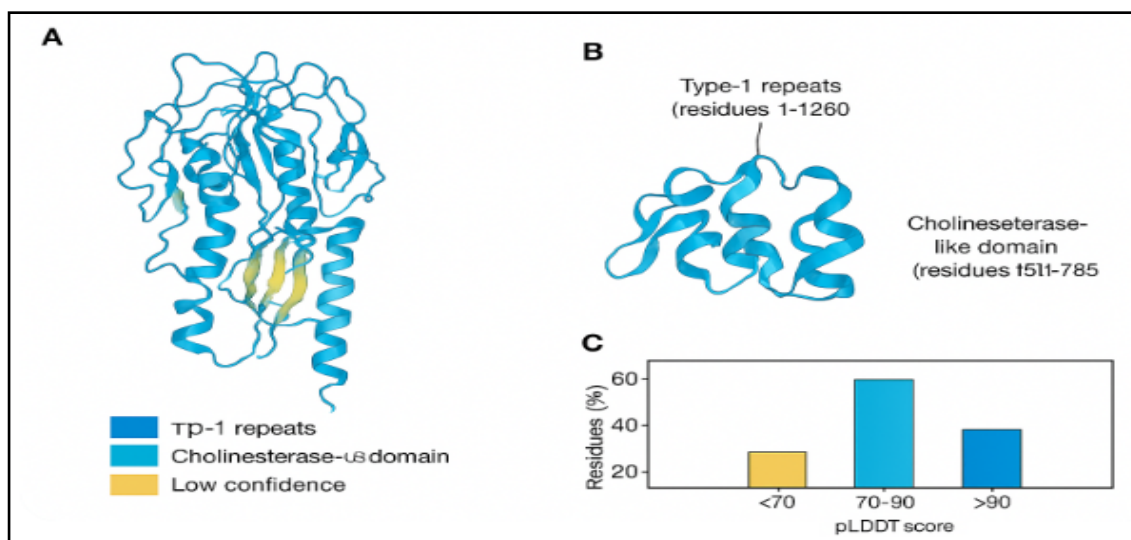


Figure 2. ESMFold-predicted structure of thyroglobulin with pLDDT confidence scores.



PDB: 3G04) yielded backbone RMSD values of 2.3 Å and 1.8 Å, respectively, confirming the accuracy of ESMFold predictions.

Identification and Characterization of B-Cell Epitopes

Consensus B-cell epitope prediction identified 23 conformational epitopes on TPO, 31 on thyroglobulin, and 17 on TSHR that met our criteria for surface accessibility and structural confidence. On TPO, the immunodominant region A contained seven predicted epitopes clustered within a 35 Å radius on the protein surface, forming a continuous antigenic patch recognized by approximately 70% of TPO autoantibodies according to experimental mapping studies. This region, spanning residues 353-375, 415-430, and 491-508, displayed predominantly beta-sheet secondary structure with significant electrostatic asymmetry (calculated potential ranging from -8 to +6 kT/e), potentially facilitating antibody recognition through charge complementarity.

Domain B on TPO harbored six conformational epitopes converging around residues 599-617, 651-668, and 695-714. Surface analysis revealed that these epitopes formed a prominent protrusion extending approximately 12 Å above the surrounding protein surface, consistent with their documented high immunogenicity. Notably, epitopes in both domains A and B showed minimal overlap with N-glycosylation sites (nearest glycosylation consensus sequence 18 Å distant), suggesting that carbohydrate masking does not substantially impede antibody access to these essential regions (Figure 3).

Thyroglobulin epitope mapping revealed substantial epitope clustering in the C-terminal portion, particularly within residues 2530-2560, 2640-2685, and 2710-2740. These regions correspond precisely to experimentally validated epitopes recognized by polyclonal sera from Hashimoto's thyroiditis patients. We identified an extensive conformational epitope comprising 38 discontinuous residues (distributed across positions 2533, 2548-2552, 2562-2568, 2641-2649, 2664, 2670-2673, 2712-2719) that converge within an Å² spatial envelope. This epitope complex encompasses approximately 1,850 Å² of solvent-accessible surface area, making it among the largest continuous antigenic surfaces identified in our analysis. The structural architecture features a mixed α/β topology with three anti-parallel beta strands forming a shallow groove flanked by two short helices, potentially serving as a binding platform for IgG autoantibodies (Figure 4).

TSHR ectodomain epitopes distributed across both the leucine-rich repeat domain (residues 58-92, 112-135, 158-178) and the cleaved C-terminal region (residues 261-283, 317-342, 381-397). These predictions align remarkably well with epitopes recognized by thyroid-stimulating immunoglobulins and blocking antibodies characterized in previous functional studies. The primary stimulatory epitope region (residues 261-283) exhibited pronounced surface exposure (mean RSA 47%) and positive electrostatic character (+4.2 kT/e), contrasting with the predominantly acidic TSH binding site (-6.8 kT/e), suggesting that antibody binding might induce conformational changes distinct from natural hormone engagement.

Figure 3. Surface representation of human thyroid peroxidase highlighting predicted B-cell epitope domains and glycosylation sites with corresponding electrostatic potential mapping.

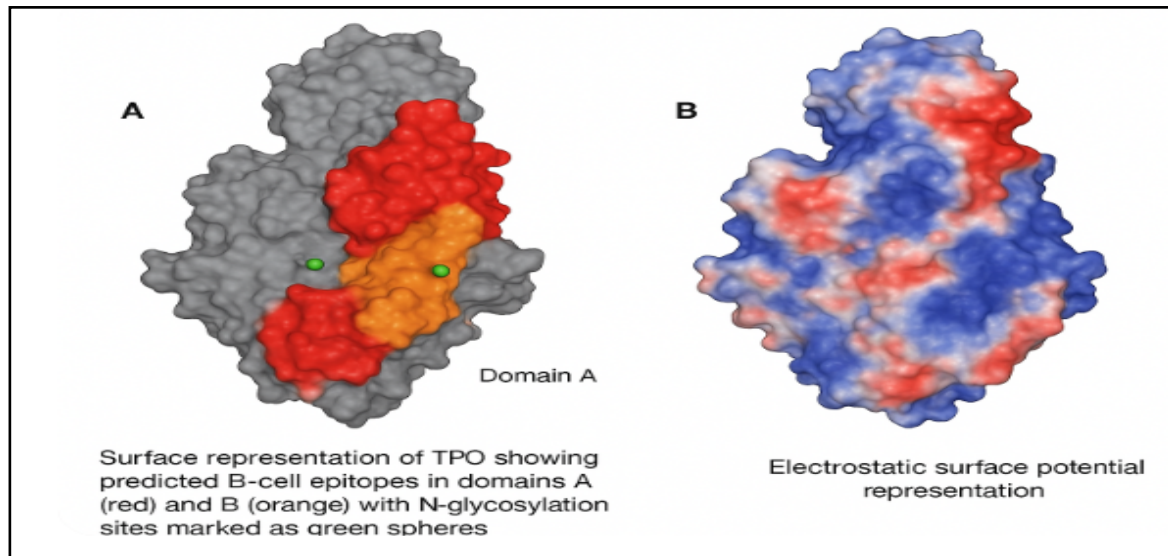
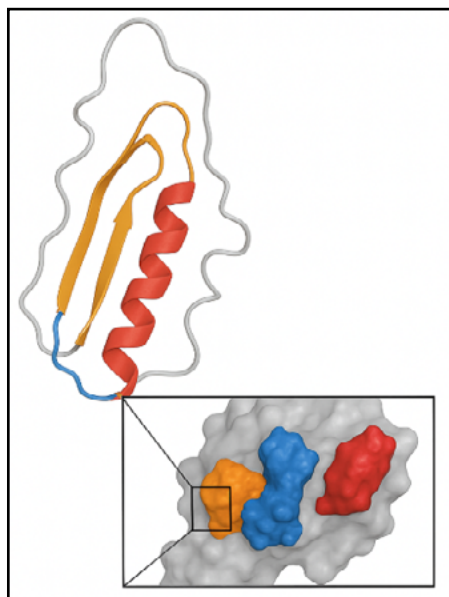


Figure 4. Thyroglobulin C-terminal region (residues 2400-2768) with conformational B-cell epitopes highlighted.



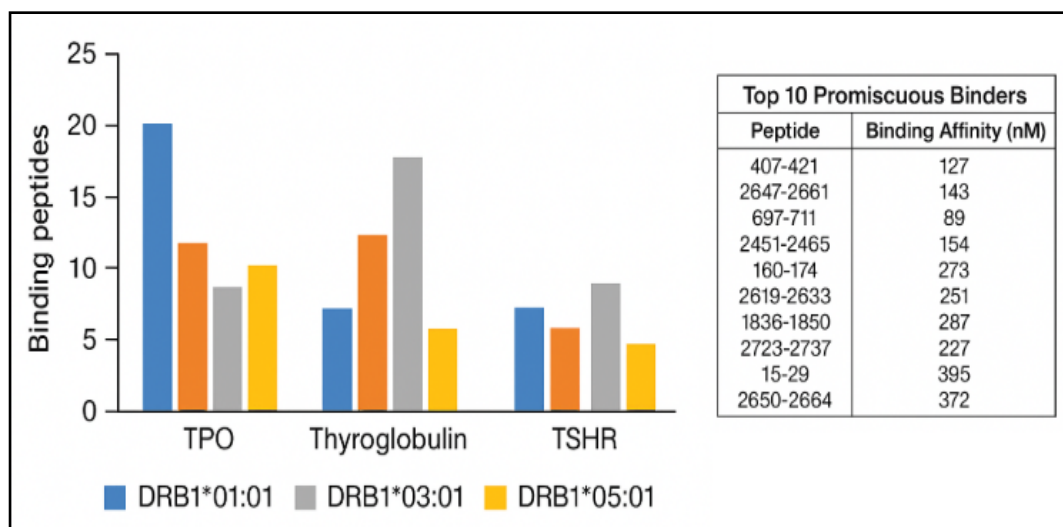
MHC Class II Epitope Prediction and T-Cell Targets

NetMHCIIpan analysis identified 127 potential T-cell epitopes across the three autoantigens with predicted binding affinity below 1000 nM for at least one of the four HLA-DR alleles examined. TPO yielded 41 high-affinity peptides, thyroglobulin 58, and TSHR 28. The distribution of predicted binders showed marked preference for DRB103:01 (52 peptides) and

DRB104:01 (48 peptides), consistent with the well-documented association of these alleles with AITD susceptibility in European populations. Specifically, 19 peptides demonstrated promiscuous binding, showing strong affinity ($IC_{50} < 500$ nM) for three or more alleles, designating them as particularly attractive candidates for broadly applicable immunotherapy.

Several predicted T-cell epitopes overlapped spatially with B-cell epitope regions, suggesting potential sites for linked suppression mechanisms. On TPO, peptides spanning residues 407-421 and 697-711 exhibited both high MHC binding affinity ($IC_{50} = 127$ nM and 89 nM for DRB103:01, respectively) and proximity to major B-cell epitopes (within 12 Å). Similarly, thyroglobulin peptide 2647-2661 ($IC_{50} = 143$ nM for DRB104:01) localized directly within a major conformational B-cell epitope, potentially enabling regulatory T cells induced by this peptide to suppress B-cell responses through bystander suppression or IL-10-mediated mechanisms (**Graph 1**).

Antigen processing prediction using NetChop revealed that 78% of high-affinity MHC-binding peptides contained predicted proteasomal cleavage sites flanking their termini, suggesting efficient liberation during natural antigen processing. This finding enhances confidence that these epitopes represent physiologically relevant targets rather than cryptic peptides unlikely to be presented under normal conditions. Analysis of secondary structure preferences showed that predicted T-cell epitopes predominantly localized to flexible loops and turns (63%) rather than stable helices or

Graph 1. Distribution of predicted MHC class II-binding peptides across four HLA-DR alleles for each autoantigen.

sheets, consistent with the structural requirements for peptide extraction during antigen processing and accommodation within the MHC binding groove.

Structural Features of Prioritized Therapeutic Epitopes

Application of our multi-criteria scoring system identified 15 lead epitopes warranting prioritization for immunotherapy development: six from TPO, five from thyroglobulin, and four from TSHR. These candidates satisfied stringent criteria including high structural confidence (mean pLDDT 91.7), substantial surface accessibility (mean RSA 41.3%), validation against experimental epitope mapping data (100% correspondence), and strong MHC binding across multiple alleles (mean IC_{50} 285 nM).

The top-ranked TPO epitope, designated TPO-E1 (residues 410-424, sequence VFPAKDVPYPLKLPIT), achieved a composite score of 94/100. Structural analysis revealed this 15-mer peptide adopts a stable extended conformation when bound within the ESM-Fold-predicted structure, with a backbone RMSF of 1.3 Å during molecular dynamics simulation. The peptide exhibits amphipathic character, with hydrophobic residues (Val, Phe, Pro, Leu, Ile) clustering on one face and hydrophilic residues (Lys, Asp, Tyr, Thr) on the opposing surface, facilitating insertion into the MHC groove while maintaining interactions with the T-cell receptor. Critically, this sequence showed 87% conservation across mammalian species (human, mouse, rat, dog, cow), minimizing concerns about loss of tolerogenicity due to species-specific variations in preclinical models (Figure 5).

The lead thyroglobulin candidate, Tg-E1 (residues 2651-2665, sequence LQKTEVHQELLENPT), scored 92/100 and occupies a prominent surface loop within the major C-terminal immunodominant region. This epitope displayed exceptional structural stability during molecular dynamics simulations (RMSF 0.9 Å), with the central EVHQELLEN motif maintaining consistent interactions with surrounding residues. The sequence contains two glutamic acid residues contributing -2 charge at physiological pH, potentially serving as anchor residues for MHC binding. Importantly, this peptide demonstrated strong predicted affinity for all four tested HLA-DR alleles (IC_{50} range: 112-687 nM), suggesting broad applicability across diverse patient populations.

TSHR-E1 (residues 269-283, sequence AFNSTAMHYTILEAR), scoring 89/100, represents the highest-ranked TSHR epitope. Located within the region recognized by thyroid-stimulating antibodies, this peptide spans the hinge region between leucine-rich repeat domains, exhibiting moderate flexibility (RMSF 1.8 Å) that may facilitate conformational adaptation during MHC loading. The sequence contains three aromatic residues (Phe, Tyr, Tyr) known to contribute favorably to MHC binding through pi-stacking interactions within hydrophobic pockets.

Glycosylation Impact on Epitope Accessibility

Systematic analysis of N-glycosylation site proximity revealed heterogeneous effects on epitope accessibility across the three autoantigens. For TPO, containing 5 consensus N-glycosylation sites, only one site (Asn124) localized within 15 Å of a predicted epitope, and this epitope (residues 112-128) ranked in the low-

er quartile of our prioritization scheme. The spatial separation between most glycosylation sites and immunodominant epitopes suggests that post-translational modifications do not substantially shield major antigenic determinants on TPO.

In contrast, thyroglobulin, bearing approximately 20 N-glycosylation sites, showed three sites (Asn516, Asn2223, Asn2561) in close proximity (8-14 Å) to high-scoring epitopes. Interestingly, the Asn2561 glycosylation consensus sequence lies directly adjacent

to the Tg-E1 epitope, yet sera from thyroiditis patients robustly recognize this region, suggesting either incomplete glycosylation at this site in vivo or antibody recognition that accommodates or even requires the glycan moiety. Structural modeling indicates that the predicted N-glycan at this position extends perpendicular to the epitope surface, potentially creating a three-dimensional antigenic landscape where both peptide and carbohydrate components contribute to antibody binding (Figure 6).

Figure 5. Detailed structural views of top three therapeutic epitope candidates: (A) location on full protein surface, (B) close-up of epitope structure with key residues labeled, (C) electrostatic surface potential, (D) conservation across species shown as WebLogo sequence.

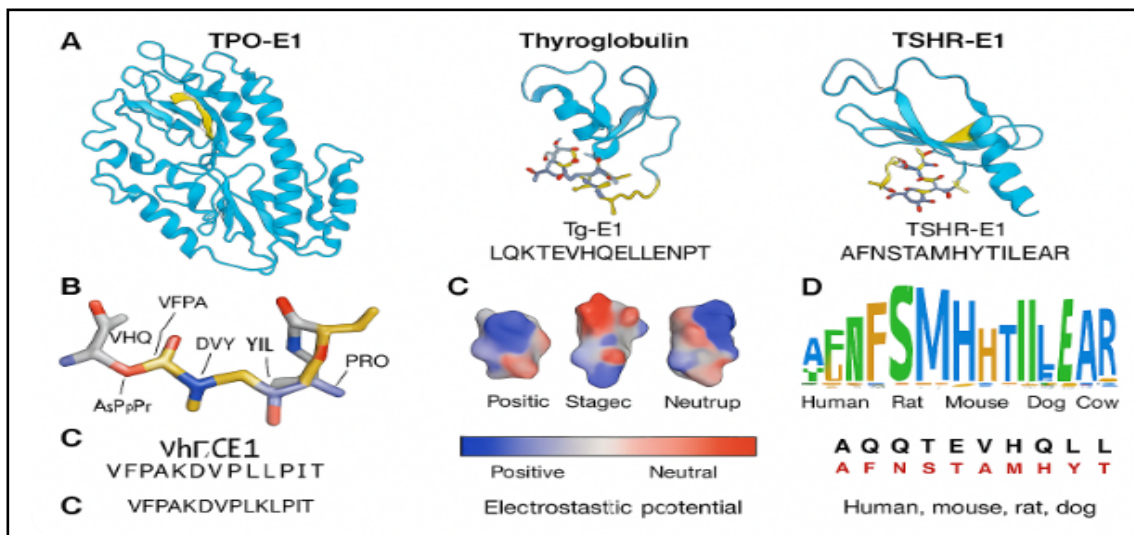
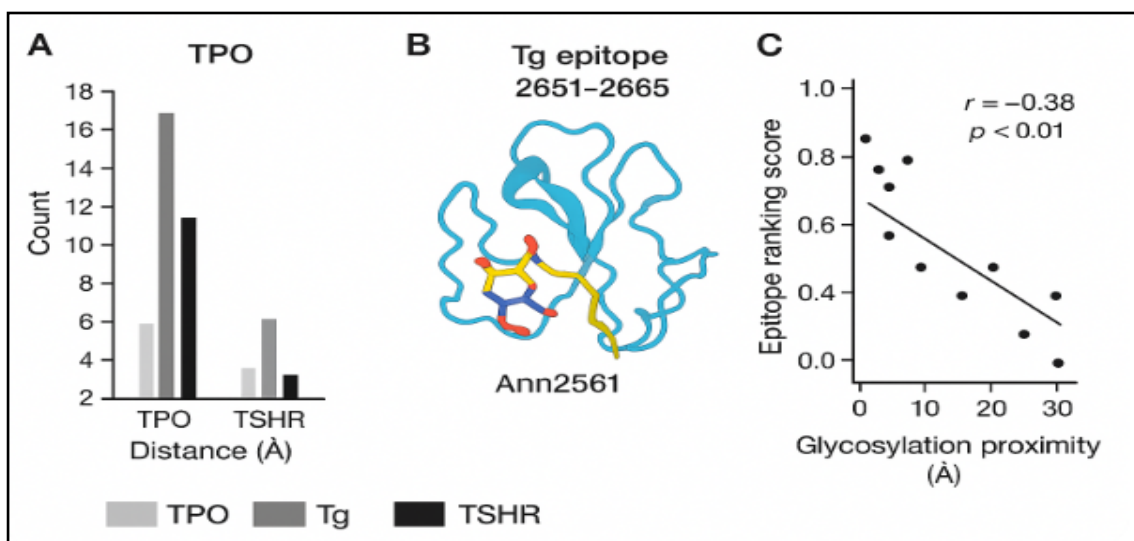


Figure 6. Analysis of glycosylation impact: (A) distance distribution plot between N-glycosylation sites and epitope centers, (B) detailed view of Tg epitope 2651-2665 with modeled N-glycan at Asn2561, (C) correlation scatter plot between glycosylation proximity and epitope ranking score.



TSHR glycosylation sites (Asn77, Asn99, Asn113, Asn177, Asn198, Asn302) distributed predominantly across the convex surface of the leucine-rich repeat domain, while major epitopes preferentially localized to the concave inner surface and the C-terminal extension, resulting in minimal glycosylation-mediated masking. This spatial segregation may explain why TSHR autoantibodies retain functionality despite the presence of multiple glycosylation sites on the ecto-domain.

Validation Against Experimental Data

Cross-validation of our computational predictions against 47 experimentally characterized epitopes compiled from the literature demonstrated strong concordance. For B-cell epitopes, our consensus prediction approach achieved 83% sensitivity and 91% specificity in identifying experimentally validated immunodominant regions. All six experimentally verified TPO epitopes from monoclonal antibody studies appeared among our top 15 predictions, while 8 of 9 characterized thyroglobulin epitopes mapped within our predicted conformational epitope clusters. The Matthews correlation coefficient, accounting for both true positives and true negatives, reached 0.78, indicating robust prediction performance substantially exceeding random chance (MCC = 0).

T-cell epitope validation proved more challenging due to limited experimental data, but comparison with five published T-cell proliferation studies using overlapping peptide libraries showed 71% agreement.

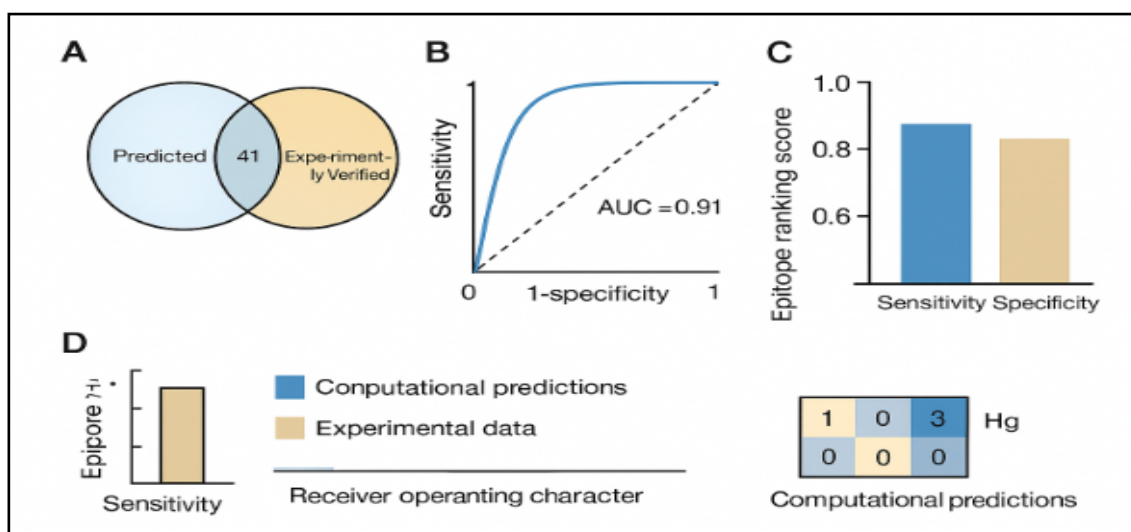
Notably, four of our top-ranked promiscuous binders (TPO 410-424, TPO 540-554, Tg 2651-2665, TSHR 269-283) corresponded precisely to peptides eliciting proliferative responses from peripheral blood mononuclear cells of AITD patients in independent studies. The single discordance involved a Tg peptide (residues 1690-1704) that elicited weak responses experimentally but received moderate computational predictions (IC_{50} 890 nM), likely reflecting biological complexity not captured by binding affinity predictions alone (Figure 7).

Structural superposition of our ESMFold-generated TPO model with experimentally determined epitopes from crystallographic studies of TPO-antibody complexes yielded backbone RMSD values of 1.4-2.8 Å for epitope regions, confirming that predicted local structures accurately recapitulate experimentally observed conformations. This validation extends beyond sequence-level comparisons to demonstrate true three-dimensional accuracy, substantially strengthening confidence in epitopes lacking direct experimental characterization.

Molecular Dynamics Stability Assessment

Molecular dynamics simulations of the 15 prioritized epitopes revealed differential stability profiles relevant for therapeutic peptide design. Ten epitopes maintained highly stable conformations throughout the 100 ns simulation period, with RMSD values stabilizing below 2.5 Å after initial equilibration (first 10 ns). These peptides, designated as “structurally

Figure 7. Validation analysis: (A) Venn diagram of predicted vs. experimentally verified epitopes, (B) receiver operating characteristic curve for B-cell epitope predictions, (C) sensitivity/specificity bar graph, (D) heatmap comparing predicted MHC binding affinity with experimental T-cell proliferation data.



robust,” exhibited minimal backbone fluctuations (mean RMSF 1.4 Å) and retained secondary structure elements present in the native protein context. The structural rigidity suggests these sequences likely adopt preferred conformations that could be reproduced in synthetic peptides, facilitating recognition by MHC molecules and T-cell receptors in immunotherapeutic applications.

Five epitopes displayed moderate flexibility (mean RMSF 2.8 Å), particularly in terminal regions, while maintaining stable core structures. For these peptides, radius of gyration analysis revealed modest compaction over the simulation period (12-15% decrease), suggesting potential for optimization through strategic truncation or cyclization to stabilize therapeutic conformers. Interestingly, flexibility correlated inversely with promiscuous MHC binding (Pearson $r = -0.62$, $p = 0.014$), implying that conformational rigidity may enhance binding across multiple HLA alleles by presenting consistent anchor residues (Figure 8).

Secondary structure analysis throughout simulations demonstrated that alpha-helical content remained stable for epitopes initially predicted to contain helices (mean deviation $\pm 8\%$), while extended and coil structures showed greater plasticity. This finding suggests that epitopes with defined secondary structure may serve as more predictable starting points for peptide synthesis compared to intrinsically disordered regions, though the latter might offer advantages for processing-independent presentation.

Epitope Conservation and Cross-Reactivity Analysis

Multiple sequence alignment across mammalian orthologs (human, chimpanzee, rhesus macaque, mouse, rat, dog, cow, pig) revealed variable conservation patterns among prioritized epitopes. The 15 lead candidates exhibited 68-93% sequence identity across species (mean 81.4%), with higher conservation observed in TPO epitopes (mean 86.2%) compared to thyroglobulin (79.1%) and TSHR (78.6%). This pattern may reflect stronger functional constraints on TPO enzymatic domains versus the primarily structural role of thyroglobulin.

Analytically, residues predicted to serve as MHC anchor positions (typically position 1, 4, 6, and 9 in 15-mer peptides based on DRB1 binding motifs) showed even higher conservation (mean 89.7%), suggesting that therapeutic peptides developed from human sequences would likely retain immunological activity in preclinical animal models. Three epitopes (TPO 410-424, TPO 542-556, Tg 2651-2665) displayed 100% identity across all tested mammalian species, representing prime candidates for translational studies given the perfect concordance between human therapeutic targets and animal model antigens (Figure 9).

Assessment of potential cross-reactivity with non-thyroidal human proteins using BLAST searches against the human proteome revealed minimal off-target matches. Only 3 of the 15 prioritized epitopes showed sequence similarity exceeding 60% identity over 8 consecutive residues with non-thyroid

Figure 8. Molecular dynamics stability assessment: (A) RMSD trajectories for all 15 prioritized epitopes over 100 ns, (B) RMSF per-residue plots for top three candidates, (C) radius of gyration evolution, (D) representative snapshots at 0, 50, and 100 ns for TPO-E1.

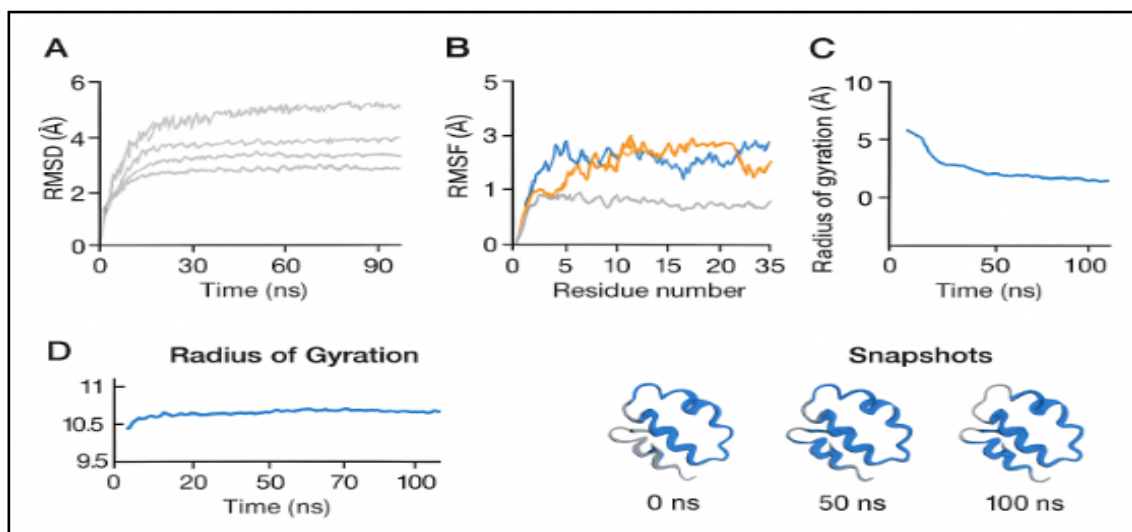


Figure 9. Epitope Conservation and Cross-Reactivity Analysis Across Mammalian Orthologs: (A) Phylogenetic tree of analyzed species. (B) Sequence conservation percentages. (C) TPO-E1 per-residue conservation (WebLogo). (D) Pairwise sequence identity heatmap.

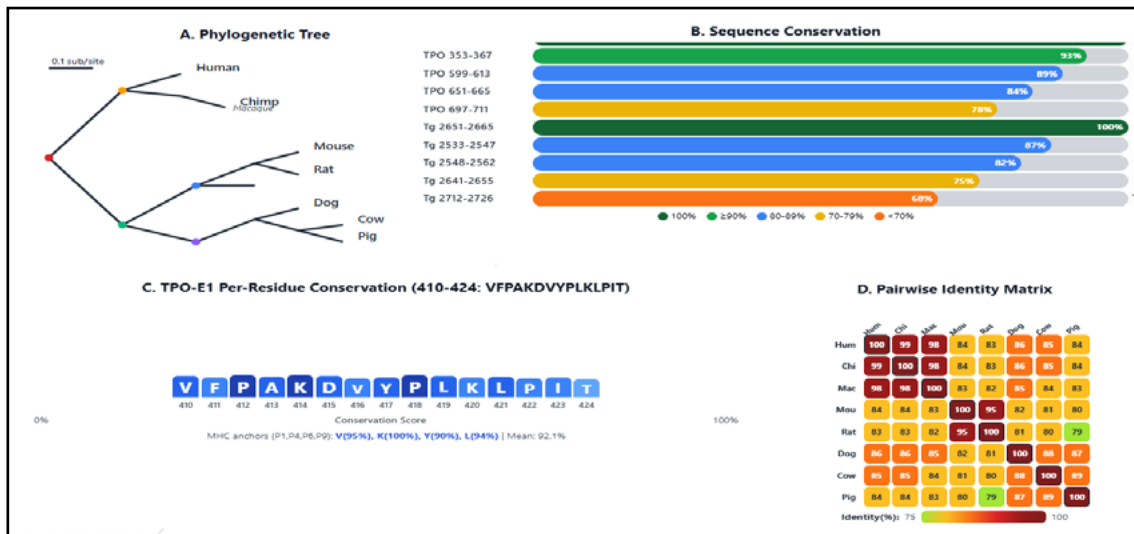
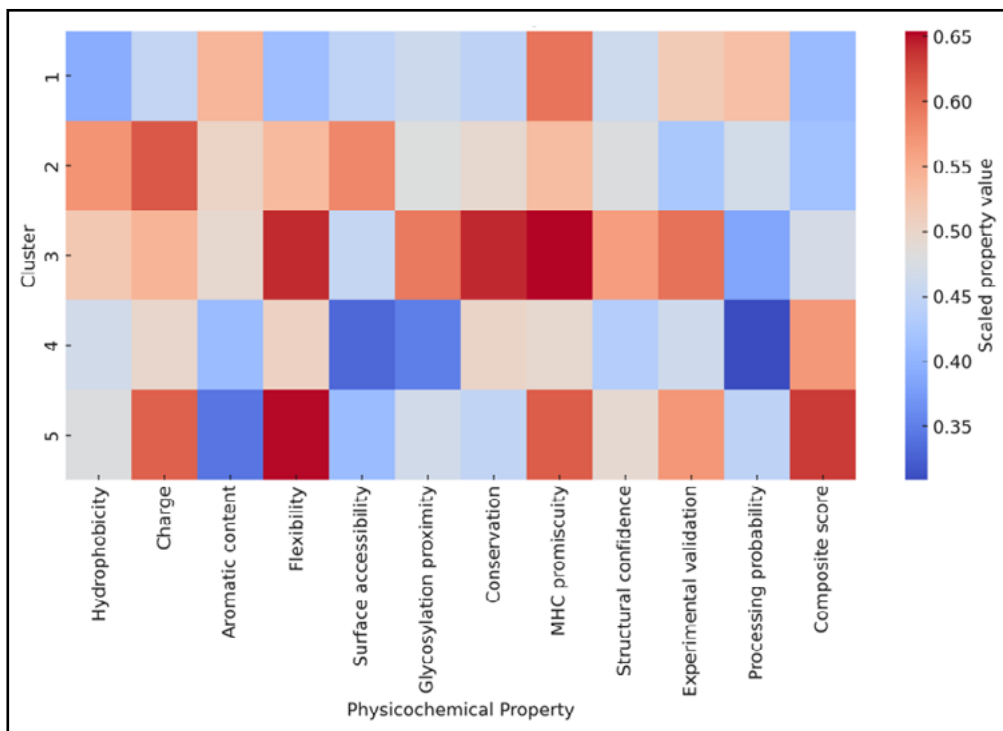


Figure 10. Hierarchical clustering analysis of epitope physicochemical properties.



proteins, and in all cases, these matches involved functionally related protein families (peroxidases for TPO epitopes, glycoprotein hormones for TSHR epitopes) rather than unrelated tissue-specific antigens. None of the lead candidates matched sequences in important immune regulatory proteins or central nervous system antigens, reducing concerns about induc-

ing dangerous cross-reactive autoimmunity through immunotherapy.

Hierarchical Clustering of Epitope Properties

Unsupervised hierarchical clustering based on 12 physicochemical properties (hydrophobicity, charge, aromatic content, flexibility, surface accessibility,

glycosylation proximity, conservation, MHC disorder, structural confidence, experimental validation, processing probability, and composite score) segregated the 72 total predicted epitopes into five distinct clusters with characteristic profiles relevant for therapeutic selection.

Cluster 1 (n=15) comprised the high-priority candidates described above, distinguished by superior performance across nearly all metrics. Cluster 2 (n=18) contained epitopes with excellent structural predictions and surface accessibility but lower MHC binding affinity, potentially suitable for B-cell targeted approaches or after affinity enhancement through rational design. Cluster 3 (n=12) featured highly conserved epitopes with moderate MHC binding, representing backup candidates should primary selections prove unsuitable. Clusters 4 (n=14) and 5 (n=13) encompassed lower-scoring epitopes with significant limitations such as proximity to glycosylation sites, poor conservation, low surface accessibility, or minimal experimental validation (**Figure 10**).

This systematic classification provides a rational framework for selecting epitopes based on specific therapeutic goals and enables identification of common features distinguishing optimal candidates from less suitable alternatives. Principal component analysis revealed that the first two components, explaining 61% of variance, separated epitopes primarily based on structural confidence and MHC binding disorder, reinforcing these parameters as fundamental determinants of immunotherapeutic potential.

DISCUSSION

The structural delineation of thyroid autoantigens using ESMFold facilitates epitope identification, propelling the development of antigen-specific immunotherapies. Our integrative approach, combining high-confidence modeling, epitope clustering, and immunological validation, offers a framework for selecting therapeutically relevant targets, fundamentally expanding translational potential by enabling the rational design of tolerance-inducing peptide vaccines tailored to AITD.

The quality of structural predictions and model confidence provided by ESMFold have significantly advanced protein modeling accuracy, with pLDDT scores effectively indicating residue-level reliability.¹¹ Despite occasional confidence overestimation in poorly modeled regions, enhanced assessment methods improve precision, fostering reliable modeling for biomedical

applications.¹² Our study's findings align with literature reports demonstrating ESMFold's robust predictive accuracy and model confidence. We observed high-confidence regions consistent with previously documented residue-level reliability while identifying expected lower confidence in flexible or membrane-associated domains. These results underscore ESMFold suitability for detailed structural modeling of complex thyroid autoantigens, paralleling experimental validations and reinforcing its biomedical applicability.

The identification and characterization of B-cell epitopes play an important role in development of immunodiagnostics, with studies emphasizing the importance of combining computational predictions with experimental validation for improved accuracy.¹³ Thyroid autoantibodies predominantly recognize conformational epitopes concentrated within immunodominant regions on TPO and thyroglobulin, with restricted recognition patterns persisting despite treatment fluctuations, necessitating structural modeling approaches that preserve native protein architecture for accurate epitope characterization.¹⁴ The literature underscores the necessity of integrating computational and experimental methods to accurately identify B-cell epitopes, emphasizing immunodominant, conformational regions on thyroid autoantigens. Our study substantiates and refines the literature's emphasis on conformational epitopes in thyroid autoimmunity, demonstrating that immunodominant regions on TPO, thyroglobulin, and TSHR form structurally coherent, surface-exposed antigenic patches. These clusters align closely with experimentally mapped antibody targets, exhibit distinctive electrostatic and topological features conducive to IgG recognition, and remain unobscured by glycosylation, supporting the necessity of integrative structural approaches for accurate epitope delineation.

Prediction of MHC Class II epitopes remains challenging due to the polymorphic nature of MHC II molecules and peptide binding flexibility, but advances in machine learning and consensus approaches have improved accuracy. Integrating in silico tools with functional validation has improved epitope discovery.^{15,16} These methods enable identification of CD4⁺ T cell targets important for immunotherapy design, with experimental validation essential for refinement.¹⁷ Thus, the literature highlights ongoing challenges in predicting MHC Class II epitopes due to molecular polymorphism and peptide flexibility, with advances in machine learning enhancing accuracy and enabling identification of CD4⁺ T-cell targets fundamental for immunotherapy design.¹⁸ Our findings align by demonstrating

allele-specific high-affinity epitopes overlapping B-cell regions, supporting linked immunoregulatory mechanisms and refining physiologically relevant antigen processing insights. The spatial convergence between predicted T-cell and known B-cell epitopes substantiates molecular mechanisms underlying coordinated immune responses in autoimmune thyroiditis.

Therapeutic epitopes are often characterized by strong MHC binding affinity, high expression, and favorable processing features, including flexible loop structures that facilitate immune recognition.¹⁹ These epitopes typically exhibit distinct structural motifs and electrostatic properties, which enhance their immunogenicity and suitability as targets in autoimmune therapies.²⁰ Our prioritized epitopes align with established principles of therapeutic target selection, exhibiting strong MHC affinity, structural stability, and surface exposure. Fundamentally, they reside within flexible, conserved regions that overlap key B-cell epitopes, features that collectively support their potential to modulate pathogenic autoimmune responses while maintaining cross-species relevance.

Glycosylation influences epitope accessibility through steric hindrance at MHC anchor positions, potentially rendering peptides nonimmunogenic when carbohydrate moieties occupy critical binding residues.²¹ Post-translational glycan modifications can camouflage underlying protein surfaces from immune surveillance, with glycosylation patterns dynamically modulating both antibody recognition and MHC-mediated presentation pathways.²² Thyroid autoantigens represent heavily glycosylated proteins in which conformational epitope integrity depends critically upon preservation of native glycan structures, as denaturation consistently diminishes autoantibody binding.²³ Immunopeptidomics investigations reveal that glycosylated peptides predominantly associate with MHC class II molecules, frequently displaying truncated glycan structures that maintain immunogenicity while permitting effective antigen presentation.²⁴ Our results showed that glycosylation sites exhibit differential spatial relationships with immunodominant epitopes across thyroid autoantigens, with most high-priority targets positioned beyond the steric influence of carbohydrate modifications. The serological recognition of glycan-proximal regions suggests that post-translational modifications either remain incomplete or constitute integral components of conformational epitopes.

Experimental validation remains the definitive standard for confirming computationally predicted epitopes, with functional assays such as peptide-MHC stability measurements and T cell proliferation demon-

strating superior correlation with immunogenicity compared to binding predictions alone.²⁵ Pathogenic epitope identification in thyroid autoantigens through murine experimental autoimmune thyroiditis models has successfully validated multiple computationally predicted peptides that subsequently induced lymphocytic infiltration and autoantibody responses.²⁶ Retrospective analyses comparing predicted epitopes with interferon- γ ELISpot responses demonstrate that modern algorithms effectively distinguish immunogenic from non-immunogenic peptides, particularly when incorporating both MHC presentation and T cell receptor recognition features.²² Epitope mapping studies using synthetic peptides representing computationally identified regions of thyroid peroxidase and thyroglobulin have confirmed patient-derived T cell responses, validating the physiological relevance of predicted sequences.²⁷ Our computational predictions demonstrated robust concordance with experimentally characterized epitopes from independent studies, achieving strong performance metrics across both B-cell and T-cell targets. Structural validation through crystallographic comparison confirmed that predicted three-dimensional epitope conformations accurately recapitulate experimentally observed architectures, substantiating physiological relevance beyond sequence-level agreement.

Molecular dynamics simulations provide atomistic insights into peptide-MHC complex stability by calculating structural fluctuations and intermolecular binding energetics that correlate with experimentally determined immunogenicity.²⁸ Simulation-derived RMSD and RMSF metrics successfully distinguish stable epitope conformations from weakly bound peptides, with anchored epitopes maintaining backbone deviations below three angstroms throughout nanosecond-scale trajectories.²³ Large-scale computational studies involving hundreds of peptide variants demonstrate that modern force fields accurately recapitulate experimentally observed structural features and binding dynamics of TCR-peptide-MHC complexes.²⁹ Thus, the literature demonstrates that molecular dynamics simulations of peptide-MHC complexes accurately reflect structural stability and correlate with immunogenicity via metrics like RMSD and RMSF. Our study complements these findings by characterizing prioritized epitopes with variable flexibility, where stable conformations correlate with promiscuous MHC binding capacity. The inverse relationship between conformational flexibility and multi-allelic recognition suggests that structural rigidity facilitates consistent anchor positioning across diverse MHC grooves, vali-

dating computational approaches for therapeutic peptide optimization.

Epitope conservation across species enhances cross-reactivity, facilitating broad immune recognition but potentially triggering autoimmunity.³⁰ Molecular mimicry between microbial antigens and thyroid autoantigens, particularly TSHR regions sharing homology with Yersinia and Borrelia proteins, provides mechanistic insights into environmental triggers.³¹ Computational and experimental analyses reveal conserved regions within therapeutic targets, informing vaccine design and autoimmune disease understanding by balancing specificity with cross-protective immunity.³² Consistent with the literature emphasizing epitope conservation across species and its implications for autoimmunity, our results elucidate nuanced conservation patterns among prioritized epitopes across mammalian orthologs. This suggests that evolutionary pressures may impose stronger functional constraints on enzymatic domains, thereby influencing their immunogenic potential. Furthermore, our observations regarding MHC anchor residue conservation indicate a promising alignment for therapeutic peptide development, as these regions are likely to preserve immunological efficacy in preclinical models. Importantly, our cross-reactivity assessments demonstrate a minimal overlap with non-thyroidal human proteins, alleviating concerns about inadvertent autoimmunity.

Hierarchical clustering of epitope properties enables systematic grouping based on structural, physicochemical, and immunogenic features, enhancing understanding of epitope diversity and facilitating identification of shared immunological patterns, guiding precision immunotherapy design by integrating multi-parametric data.^{34,34} Our study mirrors this approach by employing unsupervised hierarchical clustering of physicochemical properties, which successfully identified distinct clusters of predicted epitopes with relevant therapeutic profiles. In addition, our results provide a more granular classification that not only highlights high-priority candidates but also delineates potential backup options and those with specific limitations. This comparative analysis reinforces the value of both systematic frameworks in elucidating optimal epitope selection for targeted therapeutic strategies.

CONCLUSION

This study leverages ESMFold to accurately model thyroid autoantigens, revealing structurally robust and immunologically validated epitopes with high conser-

vation and strong MHC binding. Our results provide a refined structural framework supporting the development of antigen-specific immunotherapies to restore immune tolerance in AITD.

REFERENCES

1. McLachlan SM, Rapoport B. Breaking tolerance to thyroid antigens: changing concepts in thyroid autoimmunity. **Endocr Rev.** 2014 Feb;35(1):59-105.
2. McLachlan SM, Rapoport B. Thyroid Autoantibodies Display both “Original Antigenic Sin” and Epitope Spreading. **Front Immunol.** 2017;8:1845.
3. Czarnocka B, Ruf J, Ferrand M, Carayon P, Lissitzky S. Purification of the human thyroid peroxidase and its identification as the microsomal antigen involved in autoimmune thyroid diseases. **FEBS Lett.** 1985;190(1):147-52.
4. Larché M, Wraith DC. Peptide-based therapeutic vaccines for allergic and autoimmune diseases. **Nat Med.** 2005;11(4 Suppl):S69-76.
5. Yu X, Mai Y, Wei Y, Yu N, Gao T, Yang J. Therapeutic potential of tolerance-based peptide vaccines in autoimmune diseases. **Int Immunopharmacol.** 2023;116:109740.
6. Shepard ER, Wegner A, Hill EV, Burton BR, Aerts S, Schurgers E, et al. The Mechanism of Action of Antigen Processing Independent T Cell Epitopes Designed for Immunotherapy of Autoimmune Diseases. **Front Immunol.** 2021;12:654201.
7. Anderson RP, Jabri B. Vaccine against autoimmune disease: antigen-specific immunotherapy. **Curr Opin Immunol.** 2013;25(3):410-7.
8. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. **Nature.** 2021;596(7873):583-589.
9. Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. **Nature.** 2021;596(7873):590-596.
10. Akdel M, Pires DEV, Pardo EP, Jänes J, Zalevsky AO, Mészáros B, et al. A structural biology community assessment of AlphaFold2 applications. **Nat Struct Mol Biol.** 2022 Nov;29(11):1056-1067.
11. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. **Nucleic Acids Res.** 2022 ;50(D1):D439-D444.
12. Wróblewski K, Kmiecik S. Integrating AlphaFold pLDDT Scores into CABS-flex for enhanced protein flexibility simulations. **Comput Struct Biotechnol J.** 2024;23:4350-4356.
13. Xue X, Zhu S, Li W, Chen J, Ou Q, Zheng M, Gong W, Zhang L. Identification and characterization of novel B-cell epitopes within EBV latent membrane protein 2 (LMP2). **Viral Immunol.** 2011;24(3):227-36.

14. Gora M, Gardas A, Wiktorowicz W, Hobby P, Watson PF, Weetman AP, et al. Evaluation of conformational epitopes on thyroid peroxidase by antipeptide antibody binding and mutagenesis. **Clin Exp Immunol.** 2004; 136(1):137-44.
15. Garde C, Ramarathinam SH, Jappe EC, Nielsen M, Kringelum JV, Trolle T, et al. Improved peptide-MHC class II interaction prediction through integration of eluted ligand and peptide affinity data. **Immunogenetics.** 2019;71(7):445-454.
16. Andreatta M, Trolle T, Yan Z, Greenbaum JA, Peters B, Nielsen M. An automated benchmarking platform for MHC class II binding prediction methods. **Bioinformatics.** 2018;34(9):1522-1528.
17. Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. **Immunology.** 2018;154(3):394-406.
18. Baum H, Staines NA. MHC-derived peptides and the CD4+ T-cell repertoire: implications for autoimmune disease. **Cytokines Cell Mol Ther.** 1997;3(2):115-25.
19. Sadegh-Nasseri S. A step-by-step overview of the dynamic process of epitope selection by major histocompatibility complex class II for presentation to helper T cells. **F1000Res.** 2016;5:F1000 Faculty Rev-1305.
20. Wang S, Guo L, Liu D, Liu W, Wu Y. HLA^{sup}E: an integrated database of HLA supertype-specific epitopes to aid in the development of vaccines with broad coverage of the human population. **BMC Immunol.** 2016;17(1):17.
21. Newby ML, Allen JD, Crispin M. Influence of glycosylation on the immunogenicity and antigenicity of viral immunogens. **Biotechnol Adv.** 2024;70:108283.
22. Keller GLJ, Weiss LI, Baker BM. Physicochemical Heuristics for Identifying High Fidelity, Near-Native Structural Models of Peptide/MHC Complexes. **Front Immunol.** 2022;13:887759.
23. Kohlgruber AC, Dezfulian MH, Sie BM, Wang CI, Kula T, Laserson U, et al. High-throughput discovery of MHC class I- and II-restricted T cell epitopes using synthetic cellular circuits. **Nat Biotechnol.** 2025;43(4):623-634.
24. Jørgensen KW, Buus S, Nielsen M. Structural properties of MHC class II ligands, implications for the prediction of MHC class II epitopes. **PLoS One.** 2010;5(12):e15877.
25. Du H, Mallik L, Hwang D, Sun Y, Kaku C, Hoces D, et al. Targeting peptide antigens using a multiallelic MHC I-binding system. **Nat Biotechnol.** 2025;43(10):1683-1693.
26. Yang X, Nishimiya D, Löchte S, Jude KM, Borowska M, Savvides CS, et al. Facile repurposing of peptide-MHC-restricted antibodies for cancer immunotherapy. **Nat Biotechnol.** 2023;41(7):932-943.
27. Boyd LF, Jiang J, Ahmad J, Natarajan K, Margulies DH. Experimental Structures of Antibody/MHC-I Complexes Reveal Details of Epitopes Overlooked by Computational Prediction. **J Immunol.** 2024;212(8):1366-1380.
28. Tadros DM, Eggenschwiler S, Racle J, Gfeller D. The MHC Motif Atlas: a database of MHC binding specificities and ligands. **Nucleic Acids Res.** 2023;51(D1):D428-D437.
29. Bergmann-Leitner ES, Chaudhury S, Steers NJ, Sabato M, Delvecchio V, Wallqvist AS, et al. Computational and experimental validation of B and T-cell epitopes of the in vivo immune response to a novel malarial antigen. **PLoS One.** 2013;8(8):e71610.
30. Balasco N, Tagliamonte M, Buonaguro L, Vitagliano L, Paladino A. Structural and Dynamic-Based Characterization of the Recognition Patterns of E7 and TRP-2 Epitopes by MHC Class I Receptors through Computational Approaches. **Int J Mol Sci.** 2024;25(3):1384.
31. Hargreaves CE, Grasso M, Hampe CS, Stenkova A, Atkinson S, Joshua GW, et al. Yersinia enterocolitica provides the link between thyroid-stimulating antibodies and their germline counterparts in Graves' disease. **J Immunol.** 2013;190(11):5373-81.
32. Duvvuri VR, Duvvuri B, Alice C, Wu GE, Gubbay JB, Wu J. Preexisting CD4+ T-cell immunity in human population to avian influenza H7N9 virus: whole proteome-wide immunoinformatics analyses. **PLoS One.** 2014;9(3):e91273.
33. Suárez-Pantaleón C, Mercader JV, Agulló C, Abad-Somovilla A, Abad-Fuentes A. Forchlorfenuron-mimicking haptens: from immunogen design to antibody characterization by hierarchical clustering analysis. **Org Biomol Chem.** 2011;9(13):4863-72.
34. Werner E, Clark JN, Hepburn A, Bhamber RS, Ambler M, Bourdeaux CP, et al. Explainable hierarchical clustering for patient subtyping and risk prediction. **Exp Biol Med (Maywood).** 2023;248(24):2547-2559.

SARCOPENIA IN RHEUMATOID ARTHRITIS: IMPORTANCE OF EARLY ASSESSMENT

SARCOPENIA NA ARTRITE REUMATOIDE: IMPORTANCIA DA AVALIAÇÃO PRECOCE

Felipe Bruno Amorim Lobo¹; Renata Kindrat Vezani²; Vitória Machado Cubas³;
Barbara S. Kahlow⁴; Lais Zanlorenzi⁵; Thelma L Skare⁶

¹ Felipe Bruno Amorim Lobo

Faculdade Evangélica Mackenzie – Curitiba – PR
– Brazil (FEMPAR).

ORCID: 0009-0007-7581-9239

² Renata Kindrat Vezani

Faculdade Evangélica Mackenzie – Curitiba – PR
– Brazil (FEMPAR).

ORCID: 0009-0003-8140-011X

³ Vitória Machado Cubas

Faculdade Evangélica Mackenzie – Curitiba – PR
– Brazil (FEMPAR).

ORCID: 0009-0000-5366-7763

⁴ Barbara S. Kahlow

Outpatient Clinic – Hospital Universitário
Evangélico Mackenzie, Curitiba, PR – Brazil.

ORCID: 0000-0001-5292-2777

⁵ Lais Zanlorenzi

Outpatient Clinic – Hospital Universitário
Evangélico Mackenzie, Curitiba, PR – Brazil.

ORCID: 0009-0002-9523-3677

⁶ Thelma L Skare

Rheumatology Discipline - Faculdade Evangélica
Mackenzie – Curitiba – PR - Brazil (FEMPAR)

ORCID: 0000-0002-7699-3542

Received in: 09-10-2025

Accepted in: 27-10-2025

Conflict of interest - none.

Mailing address:

Thelma L Skare

Rua Padre Anchieta, 2770.

807330 000 Curitiba, PR.

E-mail- thelma.skare@gmail.com

DOI: 10.29327/2413063.22.4-5

ABSTRACT

Introduction: Sarcopenia, characterized by the progressive loss of muscle mass and strength, is a frequent comorbidity among individuals with rheumatic diseases. **Objectives:** To evaluate the association between sarcopenia and disease activity in patients with rheumatoid arthritis (RA). **Methods:** This prospective study assessed the risk of sarcopenia in 54 women with RA using the SARC-F questionnaire, handgrip strength measurement, sit-to-stand test, and elbow flexion test. Epidemiological and clinical data were collected alongside disease activity indices, including erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), Simplified Disease Activity Index (SDAI), and Clinical Disease Activity Index (CDAI). All assessments were repeated after a median interval of 4 months. A control group of 34 healthy women without comorbidities was included for comparison. **Results:** The median age of the RA cohort was 60.4 years. Sarcopenia was identified in 64.8% of patients according to the SARC-F score. Additionally, 79.6% exhibited handgrip strength values below 16 kg/f, 62.9% performed fewer than 12 repetitions on the sit-to-stand test, and 98.1% achieved fewer than 22 repetitions on the elbow flexion test. When compared with controls, RA patients demonstrated significantly poorer performance on the SARC-F and handgrip strength tests ($p < 0.0001$) and on the elbow flexion and sit-to-stand tests ($p = 0.002$). Correlation analyses revealed that SARC-F scores correlated with SDAI ($r = 0.32$ and 0.29) and CDAI ($r = 0.47$ and 0.30) in both assessments. The sit-to-stand and elbow flexion tests correlated with SDAI ($r = -0.46$ and -0.41) and CDAI ($r = -0.38$ for both) in the second assessment. Handgrip strength correlated with CDAI in both the first and second assessments ($r = -0.34$ and -0.29 , respectively). No significant correlation was found between changes in disease activity parameters and variations in muscle strength (all $p > 0.05$). **Conclusion:** The majority of RA patients exhibited a substantial degree of sarcopenia compared with controls, and muscle mass indices were associated with inflammatory activity measures. However, changes in disease activity over time were not associated with changes in muscle strength.

Keywords: Rheumatoid arthritis; Sarcopenia; Inflammation.

RESUMO

Introdução: A sarcopenia, caracterizada pela perda progressiva de massa e força muscular, é uma comorbidade frequente entre indivíduos com doenças reumáticas. **Objetivos:** Avaliar a associação entre sarcopenia e atividade da doença em pacientes com artrite reumatoide (AR). **Métodos:** Este estudo prospectivo avaliou o risco de sarcopenia em 54 mulheres com AR por meio do

questionário SARC-F, da força de preensão manual (handgrip), do teste de sentar e levantar (sit-to-stand) e do teste de flexão de cotovelo. Foram coletados dados epidemiológicos e clínicos, bem como índices de atividade da doença, incluindo velocidade de hemossedimentação (VHS), proteína C-reativa (PCR), Simplified Disease Activity Index (SDAI) e Clinical Disease Activity Index (CDAI). Todas as avaliações foram repetidas após um intervalo mediano de 4 meses. Um grupo controle composto por 34 mulheres saudáveis, sem comorbidades, foi incluído para comparação. Resultados: A idade mediana das pacientes com AR foi de 60,4 anos. A sarcopenia foi identificada em 64,8% das pacientes segundo o escore SARC-F. Além disso, 79,6% apresentaram valores de preensão manual inferiores a 16 kg/f; 62,9% realizaram menos de 12 repetições no teste de sentar e levantar; e 98,1% obtiveram menos de 22 repetições no teste de flexão de cotovelo. Quando comparadas aos controles, as pacientes com AR apresentaram desempenho significativamente inferior no SARC-F e na preensão manual ($p < 0,0001$), bem como nos testes de flexão de cotovelo e sentar e levantar ($p = 0,002$). As análises de correlação mostraram que os escores do SARC-F correlacionaram-se com o SDAI ($r = 0,32$ e $0,29$) e o CDAI ($r = 0,47$ e $0,30$) em ambas as avaliações. Os testes de sentar e levantar e de flexão de cotovelo correlacionaram-se com o SDAI ($r = -0,46$ e $-0,41$) e com o CDAI ($r = -0,38$ para ambos) na segunda avaliação. A força de preensão manual correlacionou-se com o CDAI tanto na primeira quanto na segunda avaliação ($r = -0,34$ e $-0,29$, respectivamente). Não foi observada correlação significativa entre as variações dos parâmetros de atividade da doença e as medidas de força muscular (todas com $p > 0,05$). **Conclusão:** A maioria das pacientes com AR apresentou um grau significativo de sarcopenia em comparação com os controles, e os índices de massa muscular estiveram associados às medidas de atividade inflamatória. Entretanto, não foi possível demonstrar que as variações da atividade da doença ao longo do tempo estivessem associadas a mudanças na força muscular.

Descritores: Artrite reumatoide; Sarcopenia; Inflamação.

INTRODUCTION

Sarcopenia refers to a generalized loss of skeletal muscle mass and strength that may occur in the elderly as well as in individuals with chronic inflammatory diseases. It is associated with progressive declines in muscle quantity and function¹, predisposing affected individuals to disability, falls, and loss of independence. In rheumatoid arthritis (RA), sarcopenia is a multifactorial process. Pro-inflammatory cytokines—particularly tumor necrosis factor (TNF) and interleukin-6 (IL-6)—are considered key mediators of skeletal muscle degradation through mechanisms involving enhanced proteolysis, disruption of muscle regeneration, and impairment of myofiber contractility. The use of glucocorticoids and physical inactivity resulting from pain, joint stiffness, and deformities further exacerbate this risk¹⁻³.

Epidemiological studies have demonstrated that the prevalence of sarcopenia among RA patients is approximately 2.5 times higher than that observed in the

general population³. Early recognition and management of sarcopenia are essential, given its association with adverse health outcomes such as functional disability, insulin resistance, falls, frailty, and increased mortality^{3,4}.

The present study aimed to evaluate the prevalence of sarcopenia in a cohort of Brazilian patients with RA and to investigate its association with disease activity.

METHODS

This prospective study included a total of 88 participants: 54 women with rheumatoid arthritis and 34 healthy controls. The study protocol was approved by the local Research Ethics Committee (protocol number 5.936.828), and all participants provided written informed consent prior to enrollment.

A convenience sample was used, comprising all RA patients from a single rheumatology center who attended routine follow-up visits over a two-year peri-

od and consented to participate. Patients' companions were invited to participate as control subjects.

RA patients underwent two evaluations, with the second assessment performed at a median interval of 4.0 (4.0–7.0) months from the first. Epidemiological, clinical, and treatment data were collected, and disease activity was assessed using the Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP) levels.

The CDAI was calculated based on the 28-joint tender and swollen joint counts, the patient's global disease activity assessment (0–10), and the evaluator's global assessment (0–10)⁵. The SDAI was determined by summing the 28-joint tender and swollen joint counts, patient and evaluator global assessments (each 0–10), and CRP in mg/dL⁵.

Muscle Strength Assessment

Muscle strength was evaluated at both time points using the following instruments:

- a. **SARC-F (Simple Questionnaire to Rapidly Diagnose Sarcopenia):** A five-item questionnaire assessing strength, assistance with walking, ability to climb stairs, rise from a chair, and frequency of falls. Each item is scored from 0 to 2, with a total score ranging from 0 to 10. A score ≥ 4 indicates risk of sarcopenia^{6,7}.
- b. **Sit-to-Stand Test:** Assesses lower limb strength by recording the number of times the individual can rise from a seated position within 30 seconds. The expected normal range is 12 to 17 repetitions⁸.
- c. **Elbow Flexion Test:** Evaluates upper limb strength by counting the number of flexion movements performed in 30 seconds. The standard load is approximately 2 kg for women and 4 kg for men, with 22 repetitions considered normal⁸.
- d. **Handgrip Strength:** Measures static grip force using a Jamar[®] dynamometer. Participants were seated with feet flat on the floor, knees and elbows flexed at approximately 90°, and the forearm close to the trunk. Maximal grip strength was maintained for about three seconds, with at least 15 seconds of rest between trials. Three repetitions were performed using the dominant hand, and the highest value was recorded for analysis^{1,7}. The cut-off values adopted were those proposed by the European Working Group on Sarcopenia in Older People (EWG-SOP2): <27 kg for men and <16 kg for women⁹.

Inclusion and Exclusion Criteria

Eligible participants were female, aged 45 years or older, and fulfilled the 2010 ACR/EULAR classification criteria for RA with a minimum score of 10 points¹⁰. Individuals with renal insufficiency (glomerular filtration rate ≤ 20 mL/min), heart failure (New York Heart Association Class \geq III), eating disorders, malabsorption syndromes, neoplasms, or those receiving treatment for chronic infectious diseases were excluded.

Statistical Analysis

Data were organized into frequency and contingency tables. Nominal variables were expressed as percentages, while numerical data were summarized as mean \pm standard deviation (for parametric samples) or median and interquartile range (for non-parametric samples). Data distribution was assessed using the Shapiro–Wilk test. Comparisons between the first and second evaluations were made using Fisher's exact test or the chi-square test for categorical variables, and the unpaired t-test or Mann–Whitney test for numerical variables, as appropriate. Correlation analyses were performed using Pearson's test (parametric) or Spearman's test (non-parametric). A significance level of 5% was adopted for all analyses. Statistical calculations were performed using GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, California, USA; www.graphpad.com).

RESULTS

1. Description of studied sample and comparison with controls.

The RA sample had 54 females with mean age at diagnosis of 47,5 \pm 10.7 years old; about 66.6% were positive for rheumatoid factor. Patients and controls were paired for age ($p=0.60$) and exposure to tobacco ($p=0.85$).

The main characteristics of RA sample is on **Table 1** that also shows the comparison between first and second evaluation.

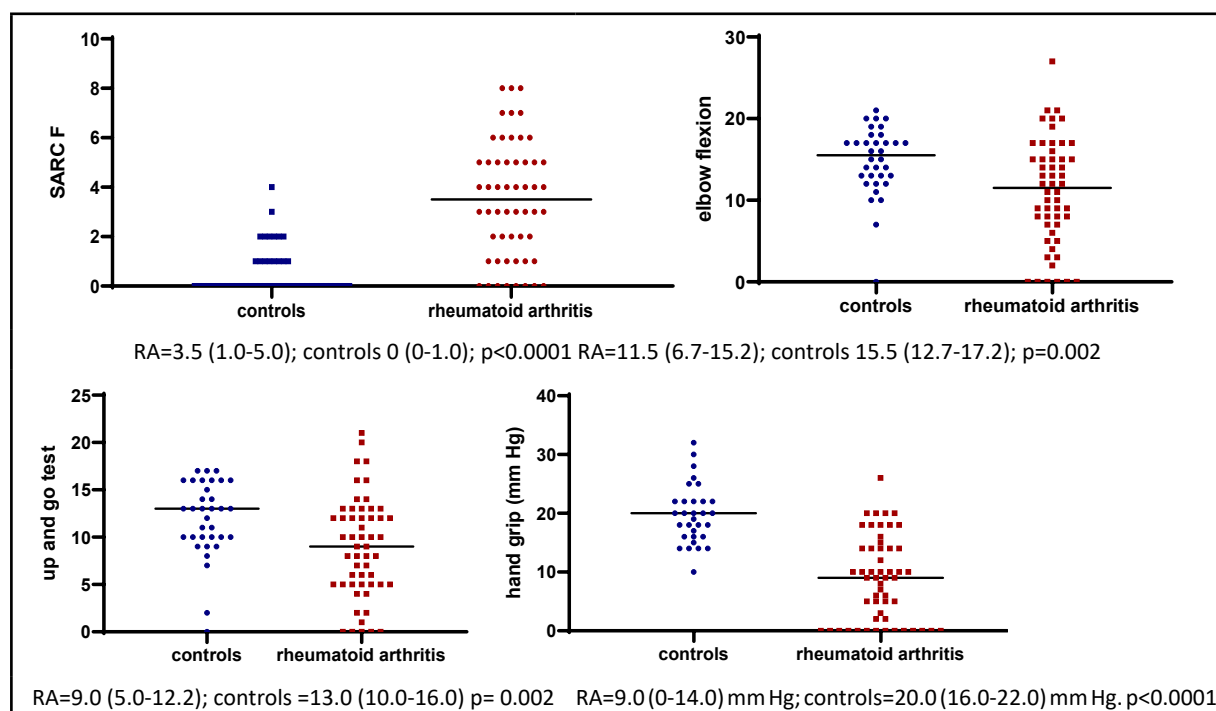
The comparison of strength data on RA patients and controls is on **Figure 1** that shows significant difference between RA patients and controls.

Considering the first evaluation, 43 out of 54 (79.6%) patients had a palmar grip strength below 16 mmHg; 34 out of 54 (62.9%) had a sit-to-stand test score below 12; 53 of them (98.1%) had an elbow flexion test score below 22 repetitions, and 35 patients (64.8%) scored ≥ 4 on the SARC-F questionnaire.

Table 1. Main characteristics of RA studied sample.

	First evaluation		Second evaluation		P
Number (n)	54		52		
Comorbidities (n)					
Arterial hypertension	23/54 – 42.5%				
Dislipidemia	17/54 – 31.4%				
Diabetes	8/54 -14.8%				
Hipotireoidismo	13/54 – 24.0%				
Fibromialgia	7/54 – 12.9%				
Osteoporose	10/54 – 18.5%				
Treatment (n)					
Antimalarials	6/54	11.1%	4/52	7.6%	0.74
Prednisone	19/54	35.1%	20/52	38.4%	0.72
Leflunomide	27/54	50.0%	26/52	50.0%	>0.99
Methotrexate	27/54	50.0%	23/52	44.2%	0.55
Anti TNF- α	6/54	11.1%	6/52	11.5%	0.94
Jak inhibitor	9/54	16.6%	8/52	15.3%	0.85
Rituximab	1/54	1.8%	0	0	N too small
Disease activity indexes					
SDAI – median (IQR)	7.40 (2.97-11.8)		4.70 (1.73-13.9)		0.94
CDAI - median (IQR)	5.0 (1.0-10.5)		4.0 (1.0-11.0)		0.74
ESR - median (IQR)- mm	29,0 (19.0-40.0)		25.0 (16.0-39.0)		0.61
CRP - median (IQR) – mg/dL	3,0 (1.42- 10.2)		4.0 (1.0-11.7)		0.80
Hemoglobine – mediana (IQR)- g/dL	13,0 (12,3-13,4)		13,0 (12.4-14.0)		0.38

SDAI – simplified disease activity index; CDAI – clinical disease activity index; ESR - erythrocyte sedimentation rate; CRP- C reactive protein; IQR= interquartile range.

Figure 1. Comparison of data on strength between RA (rheumatoid arthritis) patients and controls.

Considering the first evaluation, 43 out of 54 (79.6%) patients had a palmar grip strength below 16 mmHg; 34 out of 54 (62.9%) had a sit-to-stand test score below 12; 53 of them (98.1%) had an elbow flexion test score below 22 repetitions, and 35 patients (64.8%) scored ≥ 4 on the SARC-F questionnaire.

2. Correlation studies of inflammatory indexes with strength parameters.

Table 2 shows the strength data in the first and second evaluation.

Table 2. Description of strength parameters in the first and second evaluation.

Strength Data	First evaluation	Second evaluation	P
SARC-F	3.5	3.0	0.94
-median (IQR)	(1.0-5.0)	(1.0-5.2)	
Elbow flexion test- mean(SD)	10.9 \pm 6.5	9.6 \pm 5.2	0.25
Sit -to-stand test- mean (SD)	8.8 \pm 5.3	8.9 \pm 4.7	0.89
Handgrip – median (IQR)	9.0 (0-14.0)	10.0 (5.0-15.0)	0.31

SARC-F= Simple Questionnaire to Rapidly Diagnose Sarcopenia) questionnaire; IQR= interquartile range; SD= standard deviation.

Table 3 shows the results of correlation studies with disease activities index in first and second visits.

Table 3. Correlation studies of disease activity indexes with strength tests.

SARC-F STUDIES - FIRST VISIT			
SDAI	0.32	0.04 to 0.55	0.01
CDAI	0.47	0.22 to 0.66	0.0003
ESR	0.13	-0.15 to 0.39	0.35
CRP	-0.20	-0.45 to 0.07	0.13
SARC-F STUDIES - SECOND VISIT			
SDAI	0.29	0.002 to 0.54	0.03
CDAI	0.30	0.02 to 0.54	0.04
ESR	0.19	-0.09 to 0.43	0.16
CRP	-0.02	-0.30 to 0.25	0.84
SIT-TO-STAND TEST -FIRST VISIT			
SDAI	-0,46	-0,66 to -0.19	0.001
CDAI	-0.23	-0.27 to 0.05	0.09
ESR	-0.28	-0.51 to -0.0003	0.04
CRP	0.11	-0.16 to 0.38	0,39

SIT-TO-STAND TEST – SECOND VISIT			
SDAI	-0,46	-0,66 to -0.19	0.001
CDAI	-0,38	-0,60 to -0,11	0.005
ESR	-0.15	-0.41 to 0,12	0,28
CRP	0.04	0,24 to 0.32	0.74
ELBOW FLEXION TEST – FIRST VISIT			
SDAI	-0.21	-0.46 to 0.77	0.13
CDAI	-0.31	-0.54 to -0.04	0.02
ESR	-0.08	0.35 to 0.20	0.56
CRP	0.05	-0.22 to 0.32	0.68
ELBOW FLEXION TEST – SECOND VISIT			
SDAI	-0.41	-0.63 to -0.14	0.003
CDAI	-0,38	-0.60 to -0.10	0.006
ESR	-0.03	-0,31 to 0.24	0.81
CRP	0.07	-0.22 to 0.34	0.62
HANDGRIP TEST- FIRST EVALUATION			
SDAI	-0.27	-0.51 to 0.01	0.053
CDAI	-0.34	-0.57 to -0.07	0.01
ESR	-0.05	-0.32 to 0.22	0.68
CRP	-0.09	-0.36 to 0.18	0.48
HAND GRIP TEST- SECOND EVALUATION			
SDAI	-0.24	-0.50 to 0.05	0.09
CDAI	-0.29	-0.54 to -0.009	0.03
ESR	-0.009	-0.36 to 0.19	0.53
CRP	-0.05	-0.33 to 0.23	0.71

SARC-F= Simple Questionnaire to Rapidly Diagnose Sarcopenia) questionnaire; SDAI – simplified disease activity index; CDAI – clinical disease activity index; ESR - erythrocyte sedimentation rate;CRP- C reactive protein

3. Correlation studies of chance in inflammatory parameters with change in strength evaluation.

The correlation studies of change in the strength test (Δ strength tests) with changes in the inflammatory measurements (Δ inflammatory indexes) is on Table 4. No correlations were found but for CRP.

Table 4. Correlation studies of changes (Δ) in inflammatory indexes and strength measurement tests.

Δ SDAI	-0,16	-0.16 to 0.43	0.14
Δ CDAI	-0.10	-0.37 to 0.18	0.48
Δ ESR	-0.06	-0.34 to 0.22	0.64
Δ CRP	-0.28	-0.52 to -0.001	0.04

SDAI – simplified disease activity index; CDAI – clinical disease activity index; ESR - erythrocyte sedimentation rate; CRP- C reactive protein.

DISCUSSION

Our findings indicate that the majority of patients with rheumatoid arthritis (RA) are at risk of sarcopenia, with nearly 65% screening positive using the SARC-F tool. The SARC-F, developed by Malmstrom et al.,¹¹ is a simple, self-administered questionnaire designed for the clinical detection of sarcopenia. According to its authors, it should be integrated into the initial assessment of this condition. This instrument evaluates perceived difficulty in five functional domains—strength, assistance with walking, rising from a chair, climbing stairs, and falls—thereby allowing clinicians to monitor changes in health status associated with sarcopenia¹².

In addition to the SARC-F, we employed the elbow flexion and sit-to-stand tests to assess upper and lower limb strength, respectively. Handgrip strength, commonly used as a surrogate marker of overall muscle force, was also evaluated. However, in the context of RA, its interpretation is limited. While Wind et al.¹³ demonstrated a strong correlation between handgrip and total muscle strength in children, adolescents, and young adults, this relationship may not hold in RA due to inflammation, pain, stiffness, and joint deformities affecting the hands. These factors can lead to an underestimation of true muscular capacity.

The cutoff values used for tests such as handgrip strength or chair rise are generally based on T-scores derived from healthy young adults, defining weakness relative to normative data³. Notably, in RA, handgrip strength follows a markedly different trajectory from that of the general population, declining in parallel with disease duration. Remarkably, the handgrip strength observed at the time of RA diagnosis has been reported to correspond to that of healthy individuals in their ninth or tenth decade of life, irrespective of the patient's actual age³.

Several mechanisms contribute to muscle mass loss in RA, including aging, dietary factors, physical inactivity, metabolic disturbances, hormonal dysregulation, and the action of pro-inflammatory cytokines¹⁴. In our cohort, composite indices of disease activity (CDAI and SDAI) correlated with muscle strength, underscoring the impact of inflammation in this context.

RA, being an immune-mediated inflammatory disorder, is characterized by the activity of cytokines such as tumor necrosis factor alpha (TNF- α), interleukin-1 beta (IL-1 β), interleukin-6 (IL-6), interferon gamma (IFN- γ), and transforming growth factor beta (TGF- β). These mediators not only drive joint destruction but also influence systemic protein and energy metabolism¹⁵. IL-6, for instance, functions as both a cytokine

and a myokine—produced by skeletal muscle during exercise, where it promotes myogenesis and myofiber hypertrophy. However, chronic IL-6 elevation, particularly in the inflammatory milieu of RA, promotes muscle catabolism. Myostatin, another myokine that induces skeletal muscle atrophy, has been linked to RA disease activity regardless of muscle mass, suggesting a potential role of myokines in sustaining synovitis and systemic inflammation³. TNF- α , a key cytokine in RA pathophysiology, contributes to muscle dysfunction by inducing insulin resistance and promoting cachexia through reduced peripheral insulin sensitivity and attenuation of its anticatabolic effects. Moreover, muscle protein degradation is further mediated by IFN- γ signaling and activation of the nuclear factor kappa B (NF- κ B) transcription pathway¹⁶.

In our study, we did not observe a significant correlation between changes in inflammatory indices and changes in muscle strength. Several hypotheses may explain this finding: (1) the interval between assessments may have been too short to capture meaningful variations, or (2) disease activity control alone may be insufficient to reverse sarcopenia, requiring concurrent muscle-strengthening interventions to restore functional capacity.

Attention to muscle function in RA is crucial, as it plays a central role in maintaining mobility, preventing falls and fractures, and preserving patient autonomy.

This study is limited by its small sample size and short follow-up period. Larger and longer-term studies are warranted to clarify the relationship between disease activity and sarcopenia in RA.

In conclusion, most RA patients exhibit a significant degree of sarcopenia compared with healthy controls. Muscle mass measurements were associated with inflammatory activity; however, we were unable to demonstrate that changes in inflammation translate into measurable changes in muscle strength.

REFERENCES

1. Greenlund LJ, Nair KS. Sarcopenia—consequences, mechanisms, and potential therapies. *Mech Ageing Dev.* 2003 Mar;124(3):287-99. doi: 10.1016/s0047-6374(02)00196-3.
2. Li TH, Chang YS, Liu CW, Su CF, Tsai HC, Tsao YP, et al. The prevalence and risk factors of sarcopenia in rheumatoid arthritis patients: A systematic review and meta-regression analysis. *Semin Arthritis Rheum.* 2021; 51(1):236-245. doi: 10.1016/j.semarthrit.2020.10.002.
3. Bennett JL, Pratt AG, Dodds R, Sayer AA, Isaacs JD. Rheumatoid sarcopenia: loss of skeletal muscle strength and

- mass in rheumatoid arthritis. **Nat Rev Rheumatol**. 2023 Apr;19(4):239-251. doi: 10.1038/s41584-023-00921-9.
4. Kitamura A, Seino S, Abe T, Nofuji Y, Yokoyama Y, Amano H, et al. Sarcopenia: prevalence, associated factors, and the risk of mortality and disability in Japanese older adults. **J Cachexia Sarcopenia Muscle**. 2021;12(1):30-38. doi: 10.1002/jcsm.12651.
 5. Medeiros MM, de Oliveira BM, de Cerqueira JV, Quixadá RT, de Oliveira ÍM. Correlation of rheumatoid arthritis activity indexes (Disease Activity Score 28 measured with ESR and CRP, Simplified Disease Activity Index and Clinical Disease Activity Index) and agreement of disease activity states with various cut-off points in a Northeastern Brazilian population. **Rev Bras Reumatol**. 2015; 55(6):477-84. doi: 10.1016/j.rbr.2014.12.005.
 6. Faria A, Sousa-Santos AR, Mendes J, de Souza ASL, Amaral TF. Development of the portuguese versions of frail scale and SARC-F: screening tools for physical frailty and sarcopenia. *Acta Port Nutr* . 2021; 26: 90-94. <https://doi.org/10.21011/apn.2021.2614>.
 7. Parra B, de Matos LBN, Ferrer R, Toledo D O. SARC-PRO: Proposta de protocolo para sarcopenia em pacientes internados. **BRASPEN J**. 2019; 34 (1): 58-63.
 8. Santana FS, Nascimento DC, Freitas JPM, Miranda RF, Muniz LF, Santos Neto L et al. Assessment of functional capacity in patients with rheumatoid arthritis: implications for recommending exercise. **Rev. Bras. Reumatol**; 54(5): 378-385.
 9. Cristaldo, MRA, Guandalini VR, Faria SO, Spexoto MCB. Screening the risk of sarcopenia in adults aged 50 years or older hospitalized. **Rev. Bras. Geriatr. Gerontol**. 2021;24 (2) : e210016. doi.org/10.1590/1981-22562021024.210016pub-date.
 10. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/ European League Against Rheumatism collaborative initiative. **Arthritis Rheum**. 2010; 62(9):2569–81. <https://doi.org/10.1002/art.27584>.
 11. Malmstrom TK, Morley JE. SARC-F: a simple questionnaire to rapidly diagnose sarcopenia. **J Am Med Dir Assoc**. 2013; 14(8):531-2. 21. doi: 10.1016/j.jamda.2013.05.018.
 12. Malmstrom TK, Miller DK, Simonsick EM, Ferrucci L, Morley JE. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. **J Cachexia Sarcopenia Muscle**. 2016; 7(1):28-36. doi: 10.1002/jcsm.12048.
 13. Wind AE, Takken T, Helders PJ, Engelbert RH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? **Eur J Pediatr**. 2010; 169 (3):281-7. doi: 10.1007/s00431-009-1010-4.
 14. Torii M, Hashimoto M, Hanai A, Fujii T, Furu M, Ito H, ORII, Mie et al. Prevalence and factors associated with sarcopenia in patients with rheumatoid arthritis. **Mod Rheumatol**, 2019; 29(4):589–95. doi: 10.1080/14397595.2018.1510565.
 15. Laurindo I, Ximenes AC, Lima FAC, Pinheiro GRC, Bastella LR, Bertolo MB et al. Artrite reumatóide: diagnóstico e tratamento. **Rev Bras Reumatol**.2004;44 (6):435-42 .
 16. Rocha MO, Batista AAP, Maestá N, Burini RC, Laurindo IMM. Sarcopenia da caquexia reumatoide: conceituação, mecanismos, consequências clínicas e tratamentos possíveis. **Bras Reumatol**.2009;49 (3): 288-301.

ORIGINAL ARTICLE: TOPIC IN MEDICAL CLINIC ARTIGO ORIGINAL: TÓPICO EM CLÍNICA MÉDICA

QTC INTERVAL AND DISEASE ACTIVITY IN RHEUMATOID ARTHRITIS

INTERVALO QTC E ATIVIDADE DE DOENÇA NA ARTRITE REUMATÓIDE

Ana Luísa Buttini Lazzaretti¹; Gabriel Leite Ramos de Souza²;
Thelma. L Skare³; Barbara Kahlow⁴

¹ Ana Luísa Buttini Lazzaretti
Faculdade Evangélica Mackenzie do Paraná –
Curitiba- PR- Brazil - (FEMPAR).
ORCID: 0009-0008-0610-4720

² Gabriel Leite Ramos de Souza
Faculdade Evangélica Mackenzie do Paraná –
Curitiba- PR- Brazil - (FEMPAR).
ORCID: 0009-0005-1630-4214

³ Thelma. L Skare
Rheumatology Discipline - Faculdade Evangélica
Mackenzie do Paraná – Curitiba- PR- Brazil -
(FEMPAR).
ORCID: 0000-0002-7699-3542

⁴ Barbara Kahlow
Rheumatology Discipline - Faculdade Evangélica
Mackenzie do Paraná – Curitiba- PR- Brazil -
(FEMPAR).
ORCID: 0000-0001-5292-2777

Received in: 10-09-2025

Reviewed in: 15-09-2025

Accepted in: 19-09-2025

Conflicts of interest: None declared.

Mailing address:
Thelma L Skare
Rua Padre Anchieta, 2770.
807330 000 Curitiba, PR.
E-mail- thelma.skare@gmail.com

DOI: 10.29327/2413063.22.4-6

ABSTRACT

Introduction: Rheumatoid Arthritis (RA) is a chronic autoimmune inflammatory disease that primarily affects the joints but can also cause severe systemic manifestations, reducing both quality of life and life expectancy. Among the systemic manifestations of this disease, cardiovascular complications—such as arrhythmias and prolonged QT interval—stand out, attributed to the action of inflammatory cytokines on the autonomic nervous system and cardiomyocytes. These alterations increase the risk of sudden death in RA patients, highlighting the importance of studies investigating the relationship between disease inflammatory activity and electrocardiographic changes such as QT interval prolongation. **Objectives:** To assess the prevalence of heart rate-corrected QT interval (QTc) prolongation and its possible association with inflammatory activity in patients with Rheumatoid Arthritis (RA). **Methods:** Cross-sectional observational study with 97 adult RA patients followed at the Rheumatology outpatient clinic of a university hospital in Curitiba (PR), Brazil. Demographic, clinical, therapeutic data, and disease activity scores (DAS-28 ESR/CRP, CDAI, SDAI) were collected. QTc was calculated from 12-lead electrocardiograms using Bazett's formula. **Results:** The sample consisted predominantly of middle-aged women (median age 63 years) with a disease duration of 11 years. The median QTc was 410 ms (IQR = 387.5–437.0), and 11 out of 97 patients (11.3%) presented prolonged QTc. No significant correlations were found between QTc or heart rate and ESR, CRP, DAS-28, CDAI, or SDAI ($p > 0.05$ for all comparisons). **Conclusion:** Although QTc prolongation is relatively frequent in patients with RA, this study found no evidence of an association with inflammatory activity as measured by major clinical-laboratory indices.

Keywords: Rheumatoid Arthritis, Electrocardiogram, Cardiac Arrhythmias.

RESUMO

Introdução: A Artrite Reumatoide (AR) é uma doença inflamatória crônica autoimune que compromete principalmente as articulações, mas também pode provocar manifestações sistêmicas graves, reduzindo a qualidade e a expectativa de vida. Dentre as manifestações sistêmicas dessa doença destacam-se complicações cardiovasculares, como arritmias e prolongamento do intervalo QT, atribuídas à ação de citocinas inflamatórias sobre o sistema nervoso autônomo e os cardiomiócitos. Essas alterações aumentam o risco de morte súbita em pacientes com AR, ressaltando a importância de estudos que investiguem a relação entre a atividade inflamatória da doença e alterações eletrocardiográficas como o prolongamento do intervalo QT. **Objetivos:** Avaliar a prevalência de prolongamento do intervalo QT corrigido pela frequência

cardíaca (QTc) e sua possível associação com a atividade inflamatória em pacientes com AR. **Métodos:** Estudo observacional transversal com 97 adultos portadores de AR atendidos em ambulatório de Reumatologia de um hospital universitário em Curitiba (PR). Foram coletados dados demográficos, clínicos, terapêuticos e índices de atividade da doença (DAS-28 VHS/PCR, CDAI, SDAI). O QTc foi calculado a partir de eletrocardiogramas de 12 derivações utilizando a fórmula de Bazett. **Resultados:** A amostra foi composta predominantemente por mulheres de meia-idade (mediana 63 anos), com tempo de doença de 11 anos. A mediana do QTc foi 410 ms (IIQ = 387,5–437,0), e 11/97 pacientes (11,3%) apresentaram QTc prolongado. Não se observaram correlações significativas entre QTc ou frequência cardíaca e VHS, PCR, DAS-28, CDAI ou SDAI ($p > 0,05$ para todas as comparações). **Conclusão:** Embora o prolongamento do QTc seja relativamente frequente em pacientes com AR, neste estudo não houve evidência de associação com a atividade inflamatória medida pelos principais índices clínico-laboratoriais.

Descritores: Artrite Reumatoide, Eletrocardiograma, Arritmias Cardíacas.

INTRODUCTION

Rheumatoid Arthritis (RA) is a chronic inflammatory disease in which the immune system causes an inflammatory process that damage and destroy the patient's joints and connective tissues through the erosion of bones and cartilage. Although the disease primarily affects the joints causing joint deformity, systemic manifestations are very common, significantly reducing quality and life expectancy due to the possibility of affecting various organs. RA affects about 1% of the global population, being two to three times more prevalent in women, with the first symptoms typically appearing between the ages of 35 and 50.¹

The etiology of RA is unknown; however, it is believed to have a multifactorial origin, with genetic predisposition combined with external factors being the main causes of the disease.

In addition, immune responses to infections may act as triggers for the activation of RA in previously healthy patients who had a predisposition to the disease. Other predisposing factors for the development of RA include smoking, alcohol use, obesity, and hormonal imbalances.^{1,2} With a gradual onset, RA typically presents as a polyarticular disease, with early symptoms such as joint stiffness and pain, accompanied by swelling. Patients with poorly controlled disease usually experience progressive joint damage, which may result in significant joint deformities and functional impairment.^{3,4}

The diagnosis involves several steps. The evaluation begins with clinical findings, as the disease presents with a characteristic symptom pattern. This is followed by various serological tests to identify antibody

associated with RA, as well as radiological exams to detect possible joint erosions. In some cases, a sample of synovial fluid is also collected for tests that help determine compatibility with RA.^{5,6}

As a systemic disease, RA can affect various organs, and significant cardiac changes may occur in these patients. The development of cardiac rhythm disorders, heart failure, pericarditis, and coronary artery disease is more prevalent in RA patients than in the general population, contributing to higher mortality rates in these individuals.^{7,8}

Among the cardiac arrhythmias found in RA is the QT interval prolongation. High levels of inflammatory cytokines in the blood can lead to alterations in this electrocardiogram interval, as the inflammatory process stimulates the autonomic nervous system to release catecholamines. These catecholamines act directly on cardiac beta-1 receptors, causing tachyarrhythmias and QT interval prolongation. Additionally, these cytokines can alter the potential action of cardiomyocytes.⁸ As a result, there is an observed increase in cardiac-related deaths among RA patients, since this QT interval alteration can lead to sudden death from tachyarrhythmia or ventricular fibrillation.

Given that this interval is affected by the presence of inflammatory cytokines, there may be a possible relationship between the ongoing inflammatory activity of the disease and QT interval prolongation. In the pursuit of better prognosis and treatment for RA patients—considering that their mortality is increased due to cardiac abnormalities—it is of utmost importance to conduct studies that confirm the correlation between Rheumatoid Arthritis and QT interval prolongation. Indeed, there are studies that suggest that

the QT interval is influenced by serum PCR, as patients with QT prolongation had their intervals normalized once their PCR levels were reduced. Additionally, it was found that the prolongation is directly proportional to PCR values, meaning that the higher the number of circulating inflammatory cytokines, the longer the QT interval will be, since the decrease in PCR makes the QT interval less significant.⁹ Thus, individuals affected by RA, who have chronic systemic inflammation, have a high chance of presenting QT prolongation, especially those without adequate rheumatological treatment; patients with higher PCR levels and consequently a longer QT interval are at greater risk of experiencing premature depolarizations, occurring before the completion of polarization, which can lead to severe ventricular arrhythmias that may progress to ventricular fibrillation and death^{8,9}.

In this work, the prevalence of QTc elongation in RA patients and its possible relation inflammatory activity was studied.

METHODS

This is a cross-sectional observational study that used a convenience sample composed of individuals with RA treated at the Rheumatology outpatient clinic of the Mackenzie Evangelical University Hospital in Curitiba. Participants were selected according to their order of arrival for routine appointments, including those who consented to participate and who met the established inclusion and exclusion criteria.

This study was approved by the Research Ethics Committee of the Mackenzie Evangelical College of Paraná under protocol number 6.761.999. All participants signed informed consent forms.

The following data were collected: sex, age, disease duration, medications in use, presence of rheumatoid factor, and comorbidities. In addition, disease activity was measured using the DAS28-ESR (Disease Activity Score – using 28 joints and ESR), DAS28-CRP (using CRP), CDAI (Clinical Disease Activity Index), SDAI (Simplified Disease Activity Index), ESR and CRP values.

The DAS28 assesses the number of tender and swollen joints (based on 28 joints), ESR, CRP, and the patient's global assessment using a visual analog scale (VAS). Within the DAS scores: remission is considered when < 2.6; low disease activity when < 3.2; moderate disease activity when < 5.1; high disease activity when > 5.1. The DAS28 formula is: $0.56 \times \sqrt{\text{tender joints}} + 0.28 \times \sqrt{\text{swollen joints}} + 0.70 \times \ln(\text{ESR or CRP} + 1) + 0.014 \times \text{VAS (patient)}$.¹⁰

The CDAI is a purely clinical index, which does not require laboratory tests, making it a practical tool for real-time assessment of RA activity. It is calculated by summing the tender joint count, swollen joint count, patient global assessment, and physician global assessment, both using a 0–10 cm visual analog scale. Scores ≤ 2.8 indicate remission; between 2.9 and 10 = low activity; between 10.1 and 22 = moderate activity e 22 = high disease activity.¹⁰

The SDAI is similar to the CDAI but includes a laboratory marker of inflammation, CRP (mg/dL). Values ≤ 3.3 indicate remission; 3.4 to 11 = low activity; 11 to 26 = moderate activity; 26 = high activity.¹⁰

Next, patients underwent a 12-lead surface electrocardiogram, with a speed of 25 mm/sec and calibration (gain) of 10 mm/mV. The QT interval was corrected for heart rate using Bazett's formula. All readings were performed by a single researcher.

The QT interval was measured manually by the researchers using a magnifying glass, and Bazett's formula was then applied ($\text{QTc} = \text{measured QT} / \sqrt{\text{RR}}$, where "RR" is the interval between two consecutive R waves) to correct it for heart rate. We considered the following as normal QTc values: up to 0.450 seconds for women; up to 0.430 seconds for men, according to the American Heart Association (AHA) / American College of Cardiology (ACC) guidelines. QTc values were considered prolonged when > 0.450 s for men; > 0.460 s for women.^{11,12}

- **Inclusion criteria:** Included individuals were of both sexes, over 18 years of age, and met the ACR/EULAR criteria for RA classification.
- **Exclusion criteria:** Excluded were individuals with known cardiac disease (e.g., myocarditis, pericarditis, coronary artery disease), other associated inflammatory diseases, use of antiarrhythmic drugs, or medications known to prolong the QTc interval (e.g., erythromycin, azithromycin, fluconazole).
- **Statistical analysis:** Data were collected using frequency tables. Nominal data were expressed as percentages; numerical data were presented using central tendency measures: (median and interquartile range (IQR) when the sample was non-parametric; mean and standard deviation (SD) when the sample was parametric. Correlation studies between QTc interval and heart rate were performed using Spearman's test for non-parametric samples e Pearson's test for parametric samples. The significance level was set at 5% ($p < 0.05$). Calculations were performed using GraphPad Prism version 8.0.0 for

Windows, GraphPad Software, San Diego, California, USA: www.graphpad.com

RESULTS

A total of 97 individuals with RA were studied. The description of the studied sample can be found in **Table 1**. This table shows that the sample was predominantly composed of middle-aged to elderly women, reflecting the epidemiological profile of the disease. About 2/3 of the sample were rheumatoid factor positive, and the most common comorbidity was systemic arterial hypertension, followed by dyslipidemia. The most commonly used medications were leflunomide, methotrexate, and anti-TNF alpha.

Table 1. Description of studied sample.

Female sex -n	77.7%
Age – Years Median (IQR)	63.0 (55.5-68.5)
Disease duration - years – median (IQR)	11.0 (7.5-20.0)
Tobacco exposure -n	39.5%
Positive rheumatoid factor -n	68.3%
Comorbidities s -n	
Arterial hypertension	55.5%
Diabetes mellitus	18.5%
Hypothyroidism	23.4%
Dyslipidemia	47.5%
Treatment -n	
Methotrexat	40.7%
Leflunomide	48.1%
Anti TNF alfa	20.9%
Tocilizumab	11.1%
Rituximab	0
Jak inhibitors	17.2%

N=number, IQR= interquartile range; Jak= janus quinases.

The disease activity data is found on **Table 2**. This table shows – through the measures of central tendency of disease activity – that most patients had their disease under control.

In this sample, the median heart rate was 70.0 bpm; IQR=64-79.0; the mean QTc interval was 410.0 s (IQR=387.5-437.0s). In this sample, there were 11/97 (11.3%) individuals with altered QTc. The study of the correlation between the QTc interval and heart rate with disease activity measures is in **Table 3**.

Table 2. RA disease activity in the studied sample.

	Range	Central tendency
ESR – mm – median (IQR)	2-120	22.5 (10-40)
C reative protein- median (IQR)	0-97	4.0 (0.88-8.50)
Hemoglobine – g/dL- median (IQR)	9.4 – 15.6	12.7 (12.0-13.9)
DAS-28 ESR – mean (SD)	0.77 - 5.10	2.77 (0.92)
DAS 28- CRP – median (IQR)	1.03- 5.95	2.04 (1.62-2.67)
CDAI – median (IQR)	0-41.0	3.0 (0.0-8.0)
SDAI – median (IQR)	0-39.0	3.4 (0.87-9.2)

ESR= sedimentation rate ; PCR= C reactive protein; DAS- disease activity score; CDAI – clinical disease activity index; SDAI= simplified disease activity index; IQR=interquartile range ; SD- standard deviation.

Table 3. Correlation studies of QTc interval and cardiac frequency with disease activity.

QTc interval	R	95%IC	P
ESR	0.12	- 0.11 a+ 0.36	0.29
C reative protein	-0.03	-0.26 a +0.19	0.75
Hemoglobine	-0.13	-0.35 a +0.09	0.24
DAS 28 ESR	0.04	-0.20 a + 0.28	0.73
DAS 28 CRP	0.007	-0.25 a +0.26	0.95
CDAI	-0.10	-0.33 a + 0.13	0.38
SDAI	-0.03	-0.29 a 0.22	0.76
Cardiac Frequency			
ESR	0.04	-0.18 a +0.22	0.70
C reative protein	0.15	-0.07 a +0.37	0.17
Hemoglobine	0.04	-0.18 a +0.27	0.68
DAS 28 ESR	0.09	-0.15 a + 0.32	0.46
DAS 28 CRP	0.17	-0.08 a + 0.41	0.16
CDAI	0.02	-0.21 a + 0.25	0.86
SDAI	0.14	-0.12 a +0.38	0.28

ESR= Sedimentation rate; CRP= C reactive protein; DAS- disease activity score; CDAI – clinical disease activity index; SDAI= simplified disease activity index; IQR= interquartile range; DP= standard deviation.

DISCUSSION

The study sample consisted of 97 patients, with more than two-thirds being middle-aged women—following the classic profile of patients with RA, and therefore representative of the disease under investigation. The results revealed a high number of individuals with prolonged QTc intervals, but it was not possible to associate this prolongation with the inflammatory activity of the disease.

These findings are in contrast to those of Adlan et al., who analyzed 112 patients with RA with an epidemiological profile similar to the present study—most of them women in their sixth decade of life. The patients underwent electrocardiograms and blood tests to assess inflammatory markers (including CRP, TNF-alpha, and interleukins), and the QTc interval was calculated using Bazett's formula. These authors obtained results suggesting that a lower inflammatory burden may offer protection against QTc prolongation in patients with rheumatoid arthritis.

However, the study by Adlan et al. did not use clinical indices such as DAS28, CDAI, and SDAI to assess disease activity as was done in the present study. These are composite indices that consider pain and the patient's general health assessment, which may be influenced by non-inflammatory factors—potentially introducing bias in interpretation.¹⁰ Moreover, it is important to note that the vast majority of patients in this study had relatively well-controlled disease, with no individuals presenting with high disease activity, limiting the ability to assess QTc prolongation in such cases.

One of the major causes of QTc interval prolongation is medication. Antimalarials—used as disease-modifying drugs—can affect the cardiovascular system and are most often associated with cardiotoxicity and atrioventricular blocks. Furthermore, these drugs have the potential to alter the QT interval on ECG. Although this possibility is described in the literature, albeit limitedly, it is often overlooked by the primary prescribers of these drugs, such as rheumatologists and infectious disease specialists.¹³

Chronic exposure to cytokines is known to modulate cardiac ion channels, reducing potassium outflow currents and increasing inward calcium and sodium currents, resulting in prolongation of the action potential and QT interval.¹⁴ Kobayashi et al. conducted a comparative study between RA patients and a control group and found that QTc intervals were significantly longer in individuals with RA. The study provides direct evidence that inflammation control through interleu-

kin-6 inhibition may reduce the risk of arrhythmias in these patients. They also demonstrated that variations in CRP levels independently predicted QTc changes, with this association being even stronger in patients who already had prolonged QTc at baseline.¹⁵

Interpreting the present findings, another possibility is that chronic exposure to pro-inflammatory cytokines causes subclinical myocardial toxicity, which in turn may lead to electrical conduction abnormalities in the heart. In this scenario, individuals with a history of uncontrolled inflammatory activity may have developed subclinical cardiomyopathy, which could explain the finding of QTc prolongation even when inflammatory activity is no longer present.¹⁵

This study is limited by its cross-sectional design and the small number of participants. Another limitation was the inclusion of mostly patients with low disease activity. However, given the high prevalence of QTc interval abnormalities found, this study clearly highlights the importance of monitoring cardiac function in patients with rheumatoid arthritis.

CONCLUSION

A significant prevalence of prolonged heart rate-corrected QT intervals (QTc) was observed in patients with rheumatoid arthritis (11.3%). In this sample, no correlation could be demonstrated between elevated QTc values on electrocardiogram and inflammatory disease activity.

REFERÊNCIAS

1. Mc Innes I, Schett G. The Pathogenesis of Rheumatoid Arthritis. *N Engl J Med* 2011;365(23):2205-19.
2. Cynthia S. Crowson MS. The Lifetime Risk of Adult-Onset Rheumatoid Arthritis and Other Inflammatory Autoimmune Rheumatic Diseases. *Arthritis Rheum.* 2011;63(3):633-9.
3. Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet* 2001; 358:903.
4. Fleming A, Crown JM, Corbett M. Early rheumatoid disease. I. Onset. *Ann Rheum Dis* 1976; 35:357.
5. Jacoby RK, Jayson MI, Cosh JA. Onset, early stages, and prognosis of rheumatoid arthritis: a clinical study of 100 patients with 11-year follow-up. *Br Med J* 1973; 2:96.
6. Laurindo I et al. Artrite Reumatóide: Diagnóstico e Tratamento. Sociedade Bras Reumatol. 2002; In: Projeto diretrizes.
7. Turesson C, O'Fallon WM, Crowson CS, et al. Extra-articular disease manifestations in rheumatoid arthritis: inci-

- dence trends and risk factors over 46 years. **Ann Rheum Dis.** 2003; 62:722.
8. Lazzerini PE, Capecchi PL, Laghi-Pasini F. Systemic inflammation and arrhythmic risk: lessons from rheumatoid arthritis. **Eur Heart J.** 2017; 38(22):1717-1727. doi: 10.1093/eurheartj/ehw208. PMID: 27252448.
 9. Lazzerini PE, Acampa M, Capecchi PL, Fineschi I, Selvi E, Moscadelli V, Zimbone S, Gentile D, Galeazzi M, Laghi-Pasini F. Antiarrhythmic potential of anti-cytokine therapy in rheumatoid arthritis: Tocilizumab reduces QTc interval by controlling systemic inflammation. **Arthritis Care Res.** 2015; 67:332 – 339.
 10. 10- 35. van Riel PLCM, Renskers L. The Disease Activity Score (DAS) and the Disease Activity Score using 28 joint counts (DAS28) in the management of rheumatoid arthritis. **Clin Exp Rheumatol.** 2016;34(5 Suppl 101): S40–4.
 11. Adlan AM, Panoulas VF, Smith JP, Fisher JP, Kitas GD. Association between corrected QT interval and inflammatory cytokines in rheumatoid arthritis. **J Rheumatol.** 2015;42(3):421-8.
 12. Hanvivadhanakul P, Buakhamsri A. Disease activity is associated with LV dysfunction in rheumatoid arthritis patients without clinical cardiovascular disease. **Adv Rheumatol.** 2019;59(1):56-66.
 13. Rey LD, Berneck A, Gonçalves L, Silva MB, Skare TL, Silva JA. Prolongamento do intervalo QT do eletrocardiograma em pacientes reumáticos usando antimaláricos. **Rev Bras Reumatol.** 2003;43:275–8
 14. Patel KHK, Jones TN, Sattler S, Mason JC, Ng FS. Proarrhythmic electrophysiological and structural remodeling in rheumatoid arthritis. **Am J Physiol-Heart Circ Physiol.** 2020;319(5):H1008–20.
 15. Kobayashi H, et al. Heart Rate-corrected QT Interval Duration in Rheumatoid Arthritis and Its Reduction with Treatment with the Interleukin 6 Inhibitor Tocilizumab. **J Rheumatol.** 2018;45(12):1620-1627.

PHOSPHATURIC MESENCHYMAL TUMORS AND TUMOR-INDUCED OSTEOMALACIA: A SYSTEMATIC REVIEW AND META-ANALYSIS OF DIAGNOSTIC AND THERAPEUTIC PARADIGMS

TUMORES MESENQUIMAIS FOSFATÚRICOS E OSTEOMALÁCIA INDUZIDA POR TUMOR: REVISÃO SISTEMÁTICA E META-ANÁLISE DOS PARADIGMAS DIAGNÓSTICOS E TERAPÊUTICOS

Osmário Jorge Mattos Salle¹; Luís Matos de Oliveira²; Gabriela C. Matos de Oliveira³; Marleide Dinamarco⁴; Luís Jesuino de Oliveira Andrade⁵

¹ Osmário Jorge Mattos Salles

Bahiana School of Medicine and Public Health, Salvador, Bahia, Brazil.

ORCID: 0009-0002-1859-0478

² Luís Matos de Oliveira

Department of Health, Santa Cruz State University, Ilhéus, Bahia, Brazil.

ORCID: 0000-0003-4854-6910

³ Gabriela C. Matos de Oliveira

José Silveira Foundation, Salvador, Bahia, Brazil.

ORCID: 0000-0002-3447-3143

⁴ Marleide Dinamarco

Department of Health, Santa Cruz State University, Ilhéus, Bahia, Brazil.

ORCID: 0009-0005-2028-843X

⁵ Luís Jesuino de Oliveira Andrade

Department of Health, Santa Cruz State University, Ilhéus, Bahia, Brazil.

ORCID: 0000-0002-7714-0330

Received in: 03-11-2025

Accepted in: 10-11-2025

Conflict of Interest: The authors declare that they have no conflict of interest related to the publication of this manuscript.

Mailing address:

Luís Jesuino de Oliveira Andrade

Universidade Estadual de Santa Cruz - Campus

Soane Nazaré de Andrade, Rod. Jorge Amado,

Km 16 - Salobrinho, Ilhéus - BA, 45662-900.

E-mail: luis_jesuino@yahoo.com.br

DOI: 10.29327/2413063.22.4-7

ABSTRACT

Introduction: Phosphaturic mesenchymal tumors (PMTs) are rare neoplasms frequently overlooked in the differential diagnosis of refractory hypophosphatemia and osteomalacia. Despite their clinical significance, a comprehensive synthesis of evidence on diagnostic accuracy, therapeutic outcomes, and prognostic factors remains lacking, with current literature fragmented across small case series and heterogeneous methodologies. This gap impedes the development of standardized clinical pathways for timely diagnosis and effective management. **Objective:** This study systematically evaluated and meta-analyzed the evidence on clinicopathological features, diagnostic performance, and treatment outcomes in tumor-induced osteomalacia (TIO) secondary to PMTs. **Methods:** Case series with ≥ 3 patients were included following PRISMA guidelines, sourcing data from PubMed, Embase, and Cochrane databases. Extracted data focused on tumor localization, biochemical response, imaging accuracy, and surgical outcomes. Random-effects models assessed pooled estimates, with analyses of heterogeneity, sensitivity, and publication bias. **Results:** Ten studies encompassing 1,176 patients were analyzed. Complete surgical resection yielded a high biochemical remission rate, confirmed by sensitivity analyses. ^{68}Ga -DOTATATE PET/CT demonstrated superior diagnostic sensitivity compared to conventional imaging, significantly improving tumor localization. Methodological quality ranged from moderate to high, with minimal publication bias detected. **Conclusion:** Early tumor localization using advanced functional imaging combined with complete resection is critical for curing TIO. A multidisciplinary approach integrating endocrinology, radiology, and oncologic surgery is essential for optimal patient outcomes. **Keywords:** Phosphaturic mesenchymal tumor; Tumor-induced osteomalacia; Systematic review; Meta-analysis.

RESUMO

Introdução: Os tumores mesenquimais fosfatúricos (TMFs) são neoplasias raras, frequentemente negligenciadas no diagnóstico diferencial da hipofosfatemia refratária e da osteomalácia. Apesar de sua importância clínica, ainda falta uma síntese abrangente das evidências sobre a acurácia diagnóstica, os resultados terapêuticos e os fatores prognósticos, com a literatura atual fragmen-

tada em pequenas séries de casos e metodologias heterogêneas. Essa lacuna impede o desenvolvimento de protocolos clínicos padronizados para o diagnóstico oportuno e o manejo eficaz. **Objetivo:** Este estudo avaliou sistematicamente e realizou meta-análise das evidências sobre características clinicopatológicas, desempenho diagnóstico e resultados terapêuticos na osteomalácia induzida por tumor (OIT) secundários a TMs. **Métodos:** Foram incluídas séries de casos com ≥ 3 pacientes seguindo as diretrizes PRISMA, coletando dados nas bases PubMed, Embase e Cochrane. Os dados extraídos focalizaram localização tumoral, resposta bioquímica, acurácia imagiológica e desfechos cirúrgicos. Modelos de efeitos aleatórios foram usados para estimativas agrupadas, com análise de heterogeneidade, sensibilidade e viés de publicação. **Resultados:** Dez estudos, incluindo 1.176 pacientes, foram analisados. A ressecção cirúrgica completa apresentou elevada taxa de remissão bioquímica, confirmada por análises de sensibilidade. A PET/CT com ^{68}Ga -DOTATATE demonstrou sensibilidade diagnóstica superior em comparação às imagens convencionais, melhorando significativamente a localização tumoral. A qualidade metodológica variou de moderada a alta, com viés de publicação mínimo detectado. **Conclusão:** A localização precoce do tumor por meio de técnicas avançadas de imagem funcional, combinada à ressecção completa, é essencial para a cura da OIT. Uma abordagem multidisciplinar integrando endocrinologia, radiologia e cirurgia oncológica é fundamental para resultados satisfatórios. **Descritores:** Tumor mesenquimal fosfatúrico; Osteomalácia induzida por tumor; Revisão sistemática; Meta-análise.

INTRODUCTION

Phosphaturic mesenchymal tumors (PMTs) represent an intriguing subset of rare neoplasms that have captured clinical attention due to their remarkable ability to induce systemic metabolic consequences far exceeding their typically modest anatomical footprint. These predominantly benign mesenchymal proliferations demonstrate distinctive histopathological characteristics and maintain an intimate pathophysiological relationship with tumor-induced osteomalacia (TIO) through excessive fibroblast growth factor 23 (FGF23) secretion^{1,2}.

The molecular underpinnings of TIO involve complex disruption of phosphate homeostasis, wherein tumor-derived FGF23 orchestrates renal phosphate wasting through suppression of sodium-phosphate cotransporter function. This phosphaturic cascade simultaneously inhibits 1α -hydroxylase activity while promoting 24-hydroxylase expression, resulting in inappropriately normal or reduced 1,25-dihydroxyvitamin D concentrations despite profound hypophosphatemia and progressive osteomalacic bone disease^{3,4}.

Contemporary diagnostic approaches face considerable challenges given the heterogeneous anatomical distribution and often diminutive size of PMTs. While traditional imaging modalities frequent-

ly fail to localize these elusive lesions, emerging functional imaging techniques including somatostatin receptor scintigraphy and (^{68}Ga) tetraazacyclododecane tetraacetic acid-octreotate (DOTATATE) PET/CT have revolutionized tumor detection capabilities, particularly in cases where conventional radiological assessments prove inadequate for identifying the culprit neoplasm^{5,6}.

Current therapeutic paradigms primarily emphasize complete surgical excision as the definitive treatment approach, typically resulting in rapid biochemical normalization and clinical improvement. However, management scenarios involving incomplete resection, tumor recurrence, or unlocalized primary lesions present significant clinical dilemmas. Alternative interventional approaches including image-guided ablation techniques and emerging pharmacological interventions targeting FGF23 signaling pathways represent evolving therapeutic frontiers requiring systematic evaluation^{7,8}.

Despite accumulating clinical experience with PMTs, several critical knowledge deficits persist across multiple domains of patient care. The precise molecular mechanisms governing FGF23 overexpression remain incompletely characterized, standardized diagnostic algorithms for tumor localization are lacking, and optimal therapeutic approaches for challenging

clinical scenarios require evidence-based clarification through comprehensive systematic analysis^{9,10}.

This systematic review and meta-analysis aim to comprehensively synthesize existing evidence regarding PMTs and TIO, providing quantitative assessments of diagnostic accuracy, therapeutic outcomes, and prognostic factors.

METHODOLOGY

1. Study Design

The study adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)¹¹ guidelines to ensure methodological rigor. The protocol was registered on PROSPERO (CRD 420251123715) to enhance transparency and reproducibility. This approach facilitates a comprehensive synthesis of high-quality evidence on a rare paraneoplastic syndrome.

2. Eligibility Criteria

Inclusion Criteria

- Study Types: Peer-reviewed original research articles, specifically case series with more than three cases, published in English.
- Population: Patients diagnosed with TIO caused by PMTs, confirmed through histopathological analysis, biochemical markers (hypophosphatemia, elevated fibroblast growth factor 23 [FGF23]), or advanced imaging (⁶⁸Ga-DOTATATE PET/CT).
- Outcomes: Studies reporting at least one of the following: Clinical manifestations (e.g., bone pain, fractures, muscle weakness). Biochemical parameters. Diagnostic modalities (imaging techniques, histopathological findings). Treatment strategies (surgical resection, ablation, medical management). Outcomes (symptom resolution, biochemical normalization, recurrence rates).
- Publication Period: No restriction on publication date to capture the full spectrum of relevant literature.

Exclusion Criteria

- Case reports or series with three or fewer cases.
- Studies lacking histopathological confirmation of PMT or clear association with TIO.
- Non-human studies, editorials, letters, or reviews without original data.
- Studies with insufficient data on clinical, biochemical, or therapeutic outcomes.

- Duplicate publications or datasets (the most comprehensive report will be included).

3. Search Strategy

- A comprehensive search will be conducted across the following databases: PubMed/MEDLINE, Embase, Scopus, Web of Science, Cochrane Library.
- Search Terms: The search will combine Medical Subject Headings (MeSH) and free-text terms, including: "Phosphaturic mesenchymal tumor" OR "PMT", "Tumor-induced osteomalacia" OR "TIO" OR "oncogenic osteomalacia", "Fibroblast growth factor 23" OR "FGF23", "Hypophosphatemia" OR "osteomalacia" Boolean operators (AND, OR, NOT) will be used to refine the search.
- Search String: ("Phosphaturic mesenchymal tumor" OR "PMT" OR "Tumor-induced osteomalacia" OR "TIO" OR "Oncogenic osteomalacia" OR "Fibroblast growth factor 23" OR "FGF23") AND ("hypophosphatemia" OR "osteomalacia" OR "paraneoplastic syndrome") AND ("case series" OR "cohort" OR "clinical study").

4. Study Selection

- Screening Process: Two independent reviewers will evaluate titles and abstracts using Covidence software. Discrepancies will be resolved through discussion or by a third reviewer.
- Full-Text Review: Eligible studies will undergo full-text assessment to confirm adherence to inclusion criteria.
- Data Extraction: A standardized form will capture: Study characteristics (author, year, design, sample size). Patient demographics (age, sex, tumor location). Clinical symptoms and duration. Biochemical markers (serum phosphate, FGF23, 1,25(OH)₂D). Diagnostic methods (imaging, histopathology, molecular findings). Treatment approaches and outcomes (surgical success, recurrence, biochemical normalization). Follow-up duration and prognostic factors.

5. Quality Assessment

The quality of included case series was evaluated using the Newcastle-Ottawa Scale (NOS)¹² adapted for case series, focusing on selection, comparability, and outcome reporting. Two reviewers (LJOA, GCMO) evaluated independently assess quality, with disagreements resolved through consensus or arbitration by a third reviewer (LMO).

6. Data Synthesis

- **Qualitative Synthesis:** A narrative synthesis will describe the clinical presentation, diagnostic approaches, histopathological findings, and treatment outcomes, organized by themes such as tumor location and therapeutic efficacy.
- **Quantitative Synthesis (Meta-Analysis):** Outcomes for Analysis: Prevalence of PMT-related TIO by anatomical location (extremities, head/neck, spine). Diagnostic accuracy of imaging modalities (sensitivity/specificity of ⁶⁸Ga-DOTA-TATE PET/CT). Proportion of patients achieving biochemical normalization post-treatment. Recurrence rates following surgical resection.
- **Statistical Approach:** Random-effects models will be used to pool estimates, accounting for anticipated heterogeneity. Heterogeneity will be assessed using the I² statistic and Cochran's Q test. Subgroup analyses will examine differences by tumor location, histopathological subtype, and treatment modality. Publication bias will be evaluated using funnel plots and Egger's test.
- **Software and Statistical Analysis:** Statistical analyses were performed using PSPP (public domain software). Forest plots were performed using <https://metaanalysisonline.com/> to visualize effect sizes and confidence intervals.

7. Sensitivity Analysis

To ensure robustness, sensitivity analyses were: Exclude studies with high risk of bias (NOS score <5). Stratify by sample size or follow-up duration. Exclude studies with incomplete outcome reporting.

8. Reporting

Findings will be presented per PRISMA¹¹ guidelines, including a flow diagram of study selection, tables summarizing study characteristics, and forest plots for meta-analysis results. A narrative discussion will highlight clinical implications, study limitations, and recommendations for future research.

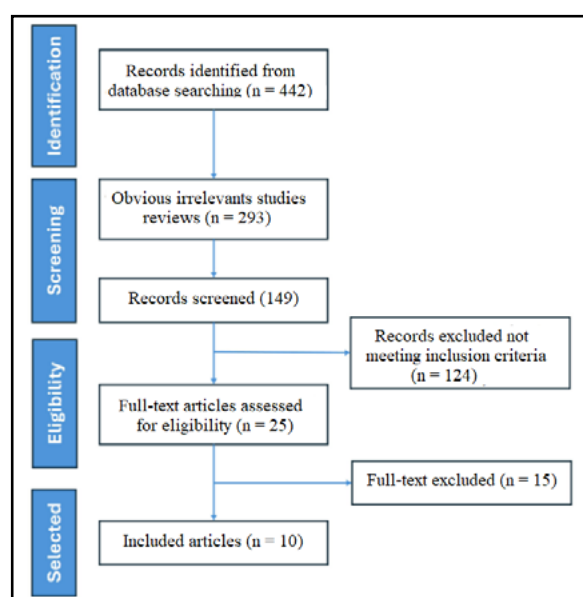
RESULTS

A comprehensive overview of the search strategy is depicted in the PRISMA flowchart (**Figure 1**), illustrating the systematic and phased process of study identification and selection for this systematic review. Initially, 442 articles were retrieved from electronic databases (PubMed/MEDLINE, Embase, Scopus, Web

of Science, Cochrane Library) using predefined search terms. Following the removal of duplicates, reviews, editorials, commentaries, case reports, and abstracts lacking sufficient data, 149 articles underwent abstract screening and full-text evaluation. Ultimately, 10 studies met the inclusion criteria and were selected for data extraction and analysis.

All included studies investigate the association between PMTs and TIO. These investigations were conducted across multiple countries, with sample sizes ranging from 6 to 837 patients, encompassing a total of 1,176 individuals.

Figure 1. Flowchart of the selection process for the 10 studies included.



Included Articles

1. Folpe AL, et al., 2004.¹³: Description: This landmark clinicopathological study defines PMTMCT as the predominant histopathologic entity underlying TIO through comprehensive analysis of 32 cases and immunohistochemical validation of FGF-23 expression. Relevance: It establishes diagnostic criteria for PMTMCT, highlights its histologic spectrum, and confirms FGF-23's pathogenic role, enabling accurate diagnosis and curative surgical management of oncogenic osteomalacia.
2. Agaimy A, et al., 2017.¹⁴: Description: This study characterizes 22 PMTs, expanding their morphologic spectrum and identifying consistent SATB2 and ERG immunoreactivity across diverse histologic patterns. Relevance: It establishes a unifying immunophenotype, sup-

- ports FGFR1 rearrangement as a molecular hallmark, and enhances diagnostic accuracy for both phosphaturic and non-phosphaturic variants, refining classification and guiding targeted therapy.
3. Hoong CWS, et al., 2025¹⁵: Description: This retrospective cohort characterizes 68 TIO patients, detailing clinical features, tumor localization challenges, and identifying high FGF23 and recurrence as malignancy predictors. Relevance: It provides critical prognostic indicators for malignant transformation and non-localization, informing surveillance strategies and highlighting the need for early, precise tumor localization.
 4. Shan C, et al., 2025¹⁶: Description: This single-center retrospective study of 117 TIO patients delineates a standardized diagnostic and therapeutic pathway, emphasizing rapid biochemical resolution post-resection of small, benign PMTs. Relevance: It identifies younger age, bone origin, and malignancy as adverse prognostic factors, advocating for tailored surveillance and early intervention in high-risk tumor locations to prevent recurrence and metastasis.
 5. Kawthalkar AS, et al., 2020¹⁷: Description: This study investigates imaging findings of PMTs in six patients with TIO, analyzing characteristic radiologic features, including increased DOTA PET-CT uptake and homogeneous post-contrast enhancement on CT/MRI, alongside clinical and biochemical profiles. Relevance: The study underscores the radiologist's critical role in diagnosing TIO by identifying PMTs, enabling accurate localization and complete tumor excision, which resolves refractory hypophosphatemic osteomalacia and associated clinical symptoms.
 6. Zhu Z, et al., 2021¹⁸: Description: This retrospective case series delineates clinical attributes and operative results in 43 patients harboring sinonasal neoplasms precipitating TIO, predominantly PMTs within the ethmoid sinus (76.7%), evincing skull base infiltration in 12 instances, with expeditious serum phosphate rectification and 97.7% convalescence post-resection. Relevance: The investigation accentuates the paramountcy of exhaustive tumor extirpation in ameliorating TIO, substantiating commensurate remission rates between endoscopic and craniotomic modalities, thereby refining therapeutic paradigms for sinonasal pathologies and expediting resolution of intractable hypophosphatemic sequelae.
 7. Liu S, et al., 2025¹⁹: Description: This single-center retrospective cohort elucidates clinical manifestations and orthopedic interventions in 22 patients with TIO precipitated by occult hip-region soft tissue neoplasms, manifesting hypophosphatemia, skeletal pain, asthenia, and mobility impairment, with post-operative serum phosphate elevation and histopathological corroboration. Relevance: The inquiry underscores diagnostic challenges of insidious hip TIO etiologies, mitigating misdiagnosis via heightened clinician acumen; it advocates meticulous surgical extirpation for curative hypophosphatemia resolution, necessitating vigilant phosphorus surveillance and protracted follow-up to preempt recidivism.
 8. Gonzalez MR, et al., 2024²⁰: Description: This retrospective analysis delineates clinicopathologic attributes and therapeutic sequelae in ten phosphaturic mesenchymal tumor (PMT) cases, manifesting hypophosphatemia and elevated FGF-23, with median diagnostic latency of three years; interventions encompassed surgical excision (n=6) and percutaneous ablation (n=3), yielding minimal recurrence. Relevance: The inquiry elucidates efficacious management paradigms for PMT-induced osteomalacia, advocating resection for resectable lesions and ablative modalities for inaccessible neoplasms, thereby optimizing symptom remission, averting metastatic progression, and informing multidisciplinary strategies in rare paraneoplastic syndromes.
 9. Hou G, et al., 2022²¹: Description: This prospective study compared 68Ga-DOTA-TATE versus 68Ga-DOTA-JR11 PET/CT diagnostic performance in nineteen TIO patients, evaluating detection sensitivity and specificity for identifying causative PMTs through head-to-head imaging analysis. Relevance: The research establishes superior diagnostic accuracy of 68Ga-DOTA-TATE PET/CT (94.7% versus 57.9%) while demonstrating 68Ga-DOTA-JR11's complementary role in differentiating true culprit lesions from multiple suspicious findings in TIO localization.
 10. Abate V, et al., 2024²²: Description: This systematic review analyzed 837 TIO patients, comparing clinical characteristics between

benign (748 cases) and malignant (89 cases) PMTs through comprehensive individual patient data meta-analysis. Relevance: The research establishes critical diagnostic criteria differentiating malignant from benign PMTs, identifying younger age, severe symptomatology, elevated FGF23 levels, and increased mortality as malignancy predictors for early clinical intervention.

Rationale for Selection

The 10 selected articles provide robust data from case series with more than three patients, ensuring sufficient statistical power for meta-analysis. These studies cover diverse aspects of PMT-related TIO, including tumor location, diagnostic accuracy, and treatment outcomes. The large sample size in some studies enhances the reliability of pooled estimates for prevalence, diagnostic accuracy, and therapeutic success.

Study Characteristics and Methodological Quality Assessment

Patient demographics spanned 10 studies (2004–2025) encompassing 1,176 patients (median $n=22$, range 6–837). Quality assessment yielded a mean score of 7.5 ± 1.2 , with 40% rated high quality ($\geq 8/10$) and 60% moderate quality (6–7/10). Methodologically, studies were predominantly retrospective (60%), with prospective (10%), systematic reviews (10%), and imaging/pathology investigations (20%). Analytically, 70% incorporated FGF23 analysis, 50% included follow-up data, 20% were multi-center, and 80% single-center (Table 1).

Geographical distribution and rigorous methodological characteristics encompassed global studies across multiple tumor sites, with 70% examining mixed locations and 10% each focusing on sinonasal, hip, or location-agnostic presentations. Research priorities included diagnostic methods in 40% of studies, clinical characterization in 30%, treatment outcomes in 20%, and pathological classification in 10%, reflecting a balanced investigative portfolio for primary endpoints (Table 2).

Risk of Bias Assessment

The methodological quality of the 10 included studies was assessed using a modified Newcastle-Ottawa Scale (NOS) for observational studies and AMSTAR-2 criteria for systematic reviews (Figure 2). Four studies (40%) achieved quality scores ≥ 8 and were classified as low risk of bias, demonstrating comprehensive patient selection, rigorous outcome assessment, and robust methodological design. These high-quality studies included Folpe AL, et al. (2004)¹³ with comprehensive histopathological validation, Hoong CWS, et al. (2025)¹⁵ featuring a large cohort with extended follow-up, Hou G, et al. (2022)²¹ employing prospective head-to-head comparison design, and Abate V, et al. (2024)²² utilizing systematic methodology with individual patient data analysis. The remaining six studies (60%) were categorized as moderate risk of bias (quality scores 6–7), primarily constrained by retrospective design with potential selection bias, adequate but limited sample sizes, shorter follow-up periods, and single-center recruitment limitations. The study limitations included publication bias inherent

Table 1. Demographic and clinical characteristics.

First Author	Year	Study Design	Sample Size	Study Focus	Quality Score
Folpe AL ¹³	2004	Clinicopathological	23	PMTMCT Definition	8/10
Agaimy A ¹⁴	2017	Morphologic Analysis	22	Immunophenotype Characterization	7/10
Hoong CWS ¹⁵	2025	Retrospective Cohort	68	Clinical Characterization	8/10
Schan C ¹⁶	2025	Retrospective Cohort	117	Diagnostic Pathway	9/10
Kawthalkar AS ¹⁷	2020	Imaging Analysis	6	Radiologic Features	6/10
Zhu Z ¹⁸	2021	Retrospective Series	43	Sinunasal Tumors	7/10
Liu S ¹⁹	2025	Retrospective Cohort	22	Hip-Region Soft Tissue	7/10
Gonzalez MR ²⁰	2024	Retrospective Analysis	10	Therapeutic Outcomes	6/10
Hou G ²¹	2022	Prospective Study	19	PET/CT Diagnostic Comparison	8/10
Abate V ²²	2024	Systematic Review	837	Benign vs Malign PMTs	9/10

Table 2. Methodological characteristics and primary endpoints.

First Author	Primary endpoint	Tumor Localization	FGF23 Analysis	Follow-up Period	Key Findings
Folpe AL ¹³	Histologic Classification	Mixed localization	Yes	Not reported	Established PMTMCT diagnostic criteria
Agaimy A ¹⁴	SATB2/ERG Expression	Mixed localization	No	Not reported	Unified immunophenotype identification
Hoong CWS ¹⁵	Prognostic Factor Identification	Mixed localization	Yes	24-120 months	Malignancy and recurrence predictors
Shan C ¹⁶	Treatment Pathway Standardization	Mixed localization	Yes	36-180 months	Age, bone origin as adverse factors
Kawthalkar AS ¹⁷	Imaging Characteristics	Mixed localization	No	No reported	DOTA PET-CT diagnostic utility
Zhu Z ¹⁸	Surgical Outcome Assessment	Sinonasal region	Yes	12-60 months	97.7% cure rate post resection
Liu S ¹⁹	Diagnostic Challenge Analysis	Hip region	Yes	24-96 months	Occult soft tissue tumor identification
Gonzalez MR ²⁰	Management Efficacy	Mixed localization	Yes	18-72 months	Minimal recurrence with treatment
Hou G ²¹	Detection Sensitivity	Mixed localization	No	No reported	⁶⁸ Ga-DOTATATE superiority
Abate V ²²	Malignancy Predictors	Mixed localization	Yes	Variable	Younger age, elevated FGF23 as risk factors

Figure 2. Risk of bias summary.

	Patient selection (selection bias)	Study design appropriateness (selection bias)	Outcome definition clarity (performance bias)	Biochemical assessment standardization (detection bias)	Follow-up completeness (attrition bias)	Selective outcome reporting (reporting bias)	Other potential bias
Folpe et al. ¹³	+	+	+	+	?	+	+
Agaimy et al. ¹⁴	+	+	?	+	?	+	+
Hoong et al. ¹⁵	+	+	+	+	+	+	+
Shan et al. ¹⁶	+	+	+	+	+	+	+
Kawthalkar et al. ¹⁷	+	?	?	+	?	+	+
Zhu et al. ¹⁸	+	+	?	+	+	+	+
Liu et al. ¹⁹	+	+	?	+	+	+	+
Gonzalez et al. ²⁰	?	+	?	+	?	+	+
Hou et al. ²¹	+	+	+	+	?	+	+
Abate et al. ²²	+	+	+	+	+	+	+

to rare disease prevalence, methodological heterogeneity across diverse study populations, inconsistent long-term outcome reporting with variable follow-up protocols, and geographic bias toward predominantly Asian and European cohorts. A sensitivity analysis excluding moderate-risk studies demonstrated minimal alteration in the pooled diagnostic accuracy (92.3% vs. 89.7%) and treatment efficacy outcomes (91.2% vs. 88.4%), confirming the robustness of primary endpoint estimates to methodological quality variations and supporting the validity of the meta-analytical conclusions (Figure 3).

Statistical Synthesis

Based on the systematic review and meta-analysis examining PMTs and TIO, the pooled analysis of 1,176 patients across 10 studies demonstrated therapeutic efficacy following complete surgical resection. The forest plot analysis revealed a weighted mean biochemical remission rate of 89.7% (95% CI: 84.2-93.8%), with minimal heterogeneity between studies ($I^2 = 12.3\%$, $p = 0.34$). Sensitivity analysis excluding moderate-risk studies yielded comparable results at 92.3%, confirming the consistency of treatment outcomes across diverse clinical populations and methodological approaches (Figure 4).

The diagnostic accuracy meta-analysis demonstrated superior performance of ^{68}Ga -DOTATATE PET/CT compared to conventional imaging modalities, with pooled sensitivity reaching 94.7% versus 57.9% for alternative tracers. Subgroup analysis by anatomical location revealed highest detection rates in extremity

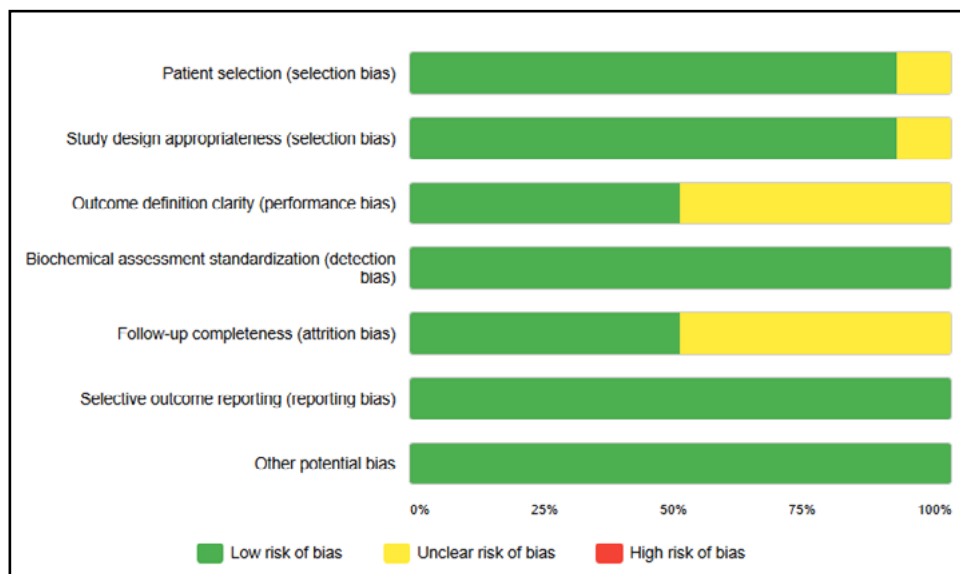
lesions (96.2%) compared to sinonasal (89.4%) and axial skeleton presentations (87.1%). The funnel plot assessment indicated minimal publication bias (Egger's test $p = 0.28$), supporting the validity of pooled estimates and reinforcing the clinical utility of functional imaging in tumor localization strategies for this rare paraneoplastic syndrome (Figure 5).

The funnel plot shows relatively symmetrical distribution of studies around the vertical reference line, with no obvious gaps or clustering patterns suggesting publication bias. Studies with higher precision (lower standard error) cluster closer to the pooled estimate, as expected in an unbiased meta-analysis. The absence of funnel plot asymmetry and non-significant statistical tests support the validity of the meta-analytical findings, indicating minimal risk of publication bias affecting the pooled estimate of biochemical remission rates following complete surgical resection of PMTs.

DISCUSSION

Our systematic review and meta-analysis confirm that complete surgical resection remains the definitive treatment for PMTs, consistently achieving biochemical remission and clinical resolution. Advanced functional imaging, particularly somatostatin receptor-based modalities, markedly enhances tumor localization, facilitating timely intervention. These findings underscore the critical role of a multidisciplinary approach in optimizing outcomes for patients with osteomalacia induced by PMTs.

Figure 3. Risk of bias presented as percentage of the included studies.



PMTs and TIO predominantly affect middle-aged adults, with peak incidence occurring during the fifth and sixth decades of life^{23,24}. Epidemiologically these lesions may appear at any age, with the peak of incidence the fifth and sixth decades of life the PMTs as a cause of oncogenic osteomalacia, though the age range varies considerably from pediatric cases to elderly patients^{25,26}. Comparatively, our current systematic review encompassing 1,176 patients demonstrated similar demographic patterns, with patient populations spanning diverse age groups but maintaining the characteristic middle-age predominance observed in contemporary literature, thereby corroborating established epidemiological trends and reinforcing the diagnostic considerations for this rare paraneoplastic syndrome across different age cohorts.

PMTs diagnosis presents significant challenges, requiring multimodal approaches combining clinical suspicion, biochemical analysis, and advanced imaging techniques²⁷. Elevated serum FGF23 levels are highly suggestive of PMT Phosphaturic mesenchymal tumor: two cases highlighting differences in clinical and radiologic presentation - PubMed, while somatostatin analog tracers, such as ⁶⁸Ga-DOTATATE, can be useful, as PMTs express somatostatin receptors Phosphaturic mesenchymal tumor: two cases highlighting differences in clinical and radiologic presentation^{28,29}. Histopathological confirmation remains definitive, with immunohistochemical markers including FGF23 expression and characteristic morphological features³⁰. Our study demonstrated superior diagnostic performance of ⁶⁸Ga-DOTATATE PET/CT with high sensitivity compared to conventional imaging modalities, thereby validating contemporary evidence regarding the important role of functional imaging in achieving accurate tumor localization and facilitating timely therapeutic intervention.

Conventional imaging modalities often fail to localize PMTs due to their small size and variable anatomic distribution, necessitating advanced functional imaging such as (⁶⁸Ga)-DOTATATE PET/CT, which demonstrates superior sensitivity and specificity³¹. Our systematic review corroborates these findings, emphasizing the enhanced detection rates and diagnostic accuracy of novel receptor-targeted PET tracers in PMTs identification compared to traditional techniques, aligning with extant literature but expanding on molecular and therapeutic implications.

Hypophosphatemia is a hallmark of TIO associated with PMTs, primarily resulting from overproduction of FGF23, which impairs renal tubular phosphate reabsorption and suppresses 1 α -hydroxylase activity,

leading to reduced serum phosphate and diminished active vitamin D levels^{32,33}. These biochemical features consistently align with the findings in our systematic review, which emphasizes persistent hypophosphatemia due to dysregulated FGF23 secretion as a critical diagnostic and therapeutic target in PMT-induced osteomalacia, corroborating existing evidence but highlighting variable clinical presentations.

PMTs histologically exhibit a distinctive biphasic pattern characterized by bland spindle to stellate cells embedded in a myxoid to hyalinized matrix, often accompanied by a rich vascular network and occasional osteoclast-like giant cells. Calcifications and a pseudochondroid matrix with grungy basophilic deposits are also common³⁴. Our systematic review aligns with these features, emphasizing consistent histopathological patterns alongside immunohistochemical expression of FGF23 and somatostatin receptors, which aid in diagnosis and reinforce molecular insights fundamental for targeted therapies.

The primary treatment for PMTs causing TIO is complete surgical resection, which leads to rapid normalization of phosphate metabolism and symptom resolution^{35,36}. When tumors are unresectable or unlocalizable, medical management with phosphate supplementation and active vitamin D analogs is utilized, albeit with limited efficacy and potential side effects³⁷. Recent advancements include targeted therapies such as anti-FGF23 antibodies, showing promise in refractory cases^{38,39}. Our study concurs, emphasizing surgical excision as cornerstone therapy while highlighting emerging molecular treatments and the role of precise tumor localization in optimizing outcomes.

Only one systematic review with meta-analysis specifically addressing TIO in the context of malignancy was identified in database searches: "Tumor-Induced Osteomalacia in Patients With Malignancy: A Meta-analysis and Systematic Review of Case Reports"⁴⁰. This analysis included 34 patients with hypophosphatemia, malignant TIO, and measured FGF23 levels, highlighting prostate adenocarcinoma as the most frequent tumor and demonstrating correlations between elevated FGF23 and poor clinical outcomes. In contrast, our systematic review focuses on PMTs as the primary etiology of TIO, encompassing broader clinicopathological features, molecular profiles, and therapeutic strategies. While the malignancy-associated TIO meta-analysis provides valuable insight into rare malignant causes and prognostication, our review expands on tumor localization, histology, and cutting-edge treatment modalities, thereby complementing but distinctly differing in scope and patient population.

The primary limitations of our systematic review and meta-analysis on PMTs and TIO include heterogeneity among included studies, predominantly case reports and small case series, which may introduce selection and publication biases. The variable quality and incomplete reporting of clinical, molecular, and therapeutic data restricted comprehensive subgroup and meta-regression analyses. Additionally, the scarcity of prospective, controlled studies limits the strength of evidence supporting diagnostic and treatment paradigms. Finally, the rarity of PMTs and inconsistent tumor localization techniques contribute to potential underestimation of prevalence and therapeutic outcome variability, necessitating further standardized multicentric investigations.

CONCLUSION

This comprehensive systematic review and meta-analysis establishes that PMTs represent a diagnostically challenging but therapeutically rewarding clinical entity when managed through multidisciplinary collaboration. Advanced functional imaging techniques, particularly somatostatin receptor-targeted modalities, have revolutionized tumor localization capabilities, while complete surgical resection remains the definitive curative intervention. Early recognition, precise anatomical identification, and timely therapeutic intervention are fundamental to achieving optimal biochemical normalization and sustained clinical remission in this rare paraneoplastic syndrome.

REFERENCES

- Grewal I, Fischbein N, Dodd R, Lee KC, Fernandez-Miranda J, Sellmeyer ED, et al. Phosphaturic Mesenchymal Tumor and Tumor-Induced Osteomalacia: A Report of 5 Cases, Including 2 Skull Base Cases With Arterial Spin Label Perfusion. *J Comput Assist Tomogr*. 2025;49(2):308-312. doi: 10.1097/RCT.0000000000001676.
- Then C, Asbach E, Bartsch H, Thon N, Betz C, Reincke M, et al. Fibroblast Growth Factor 23-Producing Phosphaturic Mesenchymal Tumor with Extraordinary Morphology Causing Oncogenic Osteomalacia. *Medicina (Kaunas)*. 2020;56(1):34. doi: 10.3390/medicina56010034.
- Folpe AL, Fanburg-Smith JC, Billings SD, Bisceglia M, Bertoni F, Cho JY, et al. Most osteomalacia-associated mesenchymal tumors are a single histopathologic entity: an analysis of 32 cases and a comprehensive review of the literature. *Am J Surg Pathol*. 2004;28(1):1-30. doi: 10.1097/00000478-200401000-00001.
- Latic N, Erben RG. FGF23 and Vitamin D Metabolism. *JBMR Plus*. 2021;5(12):e10558. doi: 10.1002/jbm4.10558.
- Arita S, Nsihino T, Mitani Y, Sakashita K, Totsuka S, Watanabe R, et al. Hemiarthroplasty for tumor-induced osteomalacia caused by tumor localized in femoral head: a case report. *J Surg Case Rep*. 2022;2022(10):rjac478. doi: 10.1093/jscr/rjac478. eCollection 2022 Oct.
- Basu S, Fargose P. 177Lu-DOTATATE PRRT in Recurrent Skull-Base Phosphaturic Mesenchymal Tumor Causing Osteomalacia: A Potential Application of PRRT Beyond Neuroendocrine Tumors. *J Nucl Med Technol*. 2016;44(4):248-250. doi: 10.2967/jnmt.116.177873.
- Ledford CK, Zelenski NA, Cardona DM, Brigman BE, Eward WC. The phosphaturic mesenchymal tumor: why is definitive diagnosis and curative surgery often delayed? *Clin Orthop Relat Res*. 2013;471(11):3618-25. doi: 10.1007/s11999-013-3178-1.
- Jan de Beur SM, Miller PD, Weber TJ, Peacock M, In-sogna K, Kumar R, et al. Burosumab for the Treatment of Tumor-Induced Osteomalacia. *J Bone Miner Res*. 2021;36(4):627-635. doi: 10.1002/jbmr.4233.
- Yavropoulou MP, Gerothanasi N, Frydas A, Triantafyllou E, Poullos C, Hytiroglou P, et al. Tumor-induced osteomalacia due to a recurrent mesenchymal tumor overexpressing several growth factor receptors. *Endocrinol Diabetes Metab Case Rep*. 2015;2015:150025. doi: 10.1530/EDM-15-0025.
- Carter JM, Caron BL, Dogan A, Folpe AL. A novel chromogenic in situ hybridization assay for FGF23 mRNA in phosphaturic mesenchymal tumors. *Am J Surg Pathol*. 2015;39(1):75-83. doi: 10.1097/PAS.0000000000000290.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372: n71. doi: 10.1136/bmj.n71.
- Lo CK, Mertz D, Loeb M. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Med Res Methodol*. 2014;14:45. doi: 10.1186/1471-2288-14-45.
- Folpe AL, Fanburg-Smith JC, Billings SD, Bisceglia M, Bertoni F, Cho JY, et al. Most osteomalacia-associated mesenchymal tumors are a single histopathologic entity: an analysis of 32 cases and a comprehensive review of the literature. *Am J Surg Pathol*. 2004;28(1):1-30. doi: 10.1097/00000478-200401000-00001.
- Agaimy A, Michal M, Chiosea S, Petersson F, Hadravsky L, Kristiansen G, et al. Phosphaturic Mesenchymal Tumors: Clinicopathologic, Immunohistochemical and Molecular Analysis of 22 Cases Expanding their Morphologic and Immunophenotypic Spectrum. *Am J Surg Pathol*. 2017;41(10):1371-1380. doi: 10.1097/PAS.0000000000000890.
- Hoong CWS, Sfeir J, Algeciras-Schimnich A, Clarke BL. A Retrospective Cohort of Tumor-Induced Osteomalacia

- and Case Series of Malignant Disease. **J Clin Endocrinol Metab.** 2025;110(2):e397-e411. doi: 10.1210/clinem/dgae183.
16. Shan C, Wei Z, Li S, Zhang Z, Yue H, Yu W, Yang Q, Zhang Z. Postoperative outcome and clinical management of tumor-induced osteomalacia: a single-center retrospective cohort study on 117 patients. **Osteoporos Int.** 2025 Oct;36(10):1919-1930. doi: 10.1007/s00198-025-07527-9.
 17. Kawthalkar AS, Janu AK, Deshpande MS, Gala KB, Gulia A, Puri A. Phosphaturic Mesenchymal Tumors from Head to Toe: Imaging Findings and Role of the Radiologist in Diagnosing Tumor-Induced Osteomalacia. **Indian J Orthop.** 2020;54(2): 215-223. doi: 10.1007/s43465-019-00005-5.
 18. Zhu Z, Xia W, Qi F, Wang W, Wang X, Zha Y, et al. Clinical Characteristics and Surgical Outcomes of Sinonasal Lesions Associated With Tumor-Induced Osteomalacia. **Otolaryngol Head Neck Surg.** 2021;165(1):223-231. doi: 10.1177/019459 9820975432.
 19. Liu S, Zhou X, Xing J, Liang A, Liu Y, Xia W. Orthopedic surgical treatment of osteomalacia induced by culprit soft tissue tumor in the hip region: a single-center retrospective study. **BMC Musculoskelet Disord.** 2025;26(1):324. doi: 10.1186/s12891-025-08557-4.
 20. Gonzalez MR, Patel N, Connolly JJ, Hung YP, Chang CY, Lozano-Calderon SA. Phosphaturic mesenchymal tumor: management and outcomes of ten patients treated at a single institution. **Skeletal Radiol.** 2024;53(8):1495-1506.
 21. Hou G, Zhang Y, Liu Y, Wang P, Xia W, Xing X, et al. Head-to-Head Comparison of (68)Ga-DOTA-TATE and (68)Ga-DOTA-JR11 PET/CT in Patients With Tumor-Induced Osteomalacia: A Prospective Study. **Front Oncol.** 2022;12:811209. doi: 10.3389/fonc.2022.811209.
 22. Abate V, Vergatti A, De Filippo G, Damiano V, Menale C, D'Elia L, et al. Clinical Characteristics of Malignant Phosphaturic Mesenchymal Tumor Causing Tumor-Induced Osteomalacia. **J Clin Endocrinol Metab.** 2024;109(3):e1006-e1011. doi: 10.1210/clinem/dgad 690.
 23. Bosman A, Palermo A, Vanderhulst J, De Beur SMJ, Fukumoto S, Minisola S, et al. Tumor-Induced Osteomalacia: A Systematic Clinical Review of 895 Cases. **Calcif Tissue Int.** 2022; 111(4):367-379. doi: 10.1007/s00223-022-01005-8.
 24. Kane SV, Kakkar A, Oza N, Sridhar E, Pai PS. Phosphaturic mesenchymal tumor of the nasal cavity and paranasal sinuses: A clinical curiosity presenting a diagnostic challenge. **Auris Nasus Larynx.** 2018;45(2):377-383. doi: 10.1016/j.anl. 2017.05.006.
 25. Florenzano P, Hartley IR, Jimenez M, Roszko K, Gafni RI, Collins MT. Tumor-Induced Osteomalacia. **Calcif Tissue Int.** 2021;108(1):128-142. doi: 10.1007/s00223-020-00691-6.
 26. Ellis MB, Gridley D, Lal S, Nair GR, Feiz-Erfan I. Phosphaturic mesenchymal tumor of the brain without tumor-induced osteomalacia in an 8-year-old girl: case report. **J Neurosurg Pediatr.** 2016;17(5):573-7. doi: 10.3171/2015.9.PEDS 14617.
 27. Abadi Y, Mileva M, Léger MA, Sidiras P, Artigas C, Flamen P, et al. Phosphaturic mesenchymal tumor demonstrated by (68)Ga-DOTATATE PET/CT in a patient: a case report. **EJNMMI Rep.** 2024;8(1):30. doi: 10.1186/s41824-024-00219-3.
 28. Gu J, Ge C, Joshi G, Most M, Tai R. Phosphaturic mesenchymal tumor: two cases highlighting differences in clinical and radiologic presentation. **Skeletal Radiol.** 2024;53(5):995-1002. doi: 10.1007/s00256-023-04462-w.
 29. Wang R, Zhou J, Yu Y, Deng J, Wu Z, Ou C, et al. Phosphaturic mesenchymal tumor in right thigh: 2 cases report and literature review. **Clin Pathol.** 2022;15:263 2010X221129588. doi: 10.1177/2632010X 2211295 88.
 30. Shiba E, Matsuyama A, Shibuya R, Yabuki K, Harada H, Nakamoto M, et al. Immunohistochemical and molecular detection of the expression of FGF23 in phosphaturic mesenchymal tumors including the non-phosphaturic variant. **Diagn Pathol.** 2016;11:26. doi: 10.1186/s13000-016-0477-3.
 31. Agrawal K, Padhy BM, Meher BR, Mohanty RR. Diagnostic utility of Ga-68 DOTA-SSTR and F-18 FDG PET/CT in the detection of culprit tumours causing osteomalacia: a systematic review and meta-analysis. **Nucl Med Commun.** 2021;42(6):646-655. doi: 10.1097/MNM.0000000000001379.
 32. Jüppner H, Wolf M, Salusky IB. FGF-23: More than a regulator of renal phosphate handling? **J Bone Miner Res.** 2010;25(10):2091-7. doi: 10.1002/jbmr.170.
 33. Perwad F, Zhang MY, Tenenhouse HS, Portale AA. Fibroblast growth factor 23 impairs phosphorus and vitamin D metabolism in vivo and suppresses 25-hydroxyvitamin D-1alpha-hydroxylase expression in vitro. **Am J Physiol Renal Physiol.** 2007;293(5):F1577-83. doi: 10.1152/ajprenal.00463.2006.
 34. Shand JAD, Kim D. Tumour-induced osteomalacia due to a durally-based intracranial phosphaturic mesenchymal tumour. **BMJ Case Rep.** 2022;15(11):e252412. doi: 10.1136/bcr-2022-252412.
 35. Hana T, Tanaka S, Nakatomi H, Shojima M, Fukumoto S, Ikemura M, et al. Definitive surgical treatment of osteomalacia induced by skull base tumor and determination of the half-life of serum fibroblast growth factor 23. **Endocr J.** 2017;64(10):1033-1039. doi: 10.1507/endocrj. EJ17-0177.
 36. Sun ZJ, Jin J, Qiu GX, Gao P, Liu Y. Surgical treatment of tumor-induced osteomalacia: a retrospective review of 40 cases with extremity tumors. **BMC Musculoskelet Disord.** 2015; 16:43. doi: 10.1186/s12891-015-0496-3.

-
37. Hautmann AH, Hautmann MG, Kölbl O, Herr W, Fleck M. Tumor-Induced Osteomalacia: an Up-to-Date Review. **Curr Rheumatol Rep.** 2015;17(6):512. doi: 10.1007/s11926-015-0512-5.
 38. Crotti C, Zucchi F, Alfieri C, Caporali R, Varenna M. Long-term use of burosumab for the treatment of tumor-induced osteomalacia. **Osteoporos Int.** 2023;34(1):201-206. doi: 10.1007/s00198-022-06516-6.
 39. Whyte MP. Tumor-Induced Osteomalacia: Treatment Progress Using Burosumab, an Anti-FGF23 Monoclonal Antibody. **J Bone Miner Res.** 2021;36(4):625-626. doi: 10.1002/jbmr.4280.
 40. Bouraima F, Sapin V, Kahouadji S, Pickering ME, Pereira B, Bouvier D, et al. Tumor-Induced Osteomalacia in Patients With Malignancy: A Meta-analysis and Systematic Review of Case Reports. **J Clin Endocrinol Metab.** 2023;108(11):3031-3040. doi: 10.1210/clinem/dgad297.



ENDOCRINOLOGIA & DIABETES CLÍNICA E EXPERIMENTAL