

**DESIGN OF LIGANDS MULTI-
TARGET DERIVATIVES
FROM THE QUERCETIN
GUIDED PUT ARTIFICIAL
INTELLIGENCE FOR
ALZHEIMER'S AND
PARKINSON'S DISEASES:
MAPPING
POLYPHARMACOLOGY,
GENERATIVE
OPTIMIZATION, AND CO-
FOLDING THROUGH DEEP
LEARNING**

DESIGN DE LIGANTES MULTIALVO DERIVADOS DA QUERCETINA COM O
AUXÍLIO DE INTELIGÊNCIA ARTIFICIAL PARA AS DOENÇAS DE ALZHEIMER
E PARKINSON: MAPEAMENTO DA POLIFARMACOLOGIA, OTIMIZAÇÃO
GENERATIVA E CO-ENOVELAMENTO VIA APRENDIZADO PROFUNDO

Josué Caleb Mendes da Silva de Assumpção¹

Maria Tereza de Assunção Coelho²

Emília Nazaré do Nascimento Melo³

Delvek da Silva Venceslau de Sousa⁴

Miriã Cristina Mendes da Silva⁵

Alexandra Carla Garcia Mendes⁶

ABSTRACT

Neurodegenerative disorders, such as Alzheimer's (AD) and Parkinson's (PD) diseases, feature complex and multifactorial etiologies. Single-target therapies have largely failed to modify the clinical course of these conditions, driving the development of multi-target-directed ligands (MTDLs). Quercetin, a dietary flavonol with documented neuroprotective properties, represents an attractive yet suboptimal starting point due to its modest potency, poor blood-brain barrier (BBB) permeation, and propensity for assay interference (PAINS). This study reports an end-to-end, fully reproducible Artificial Intelligence (AI) workflow designed to overcome these limitations. The approach involves: (i) mapping quercetin's polypharmacology across eight validated AD/PD targets using Quantitative Structure-Activity Relationship (QSAR) models trained on 56,328 curated ChemBL bioactivities (40,207 unique compounds); (ii) profiling ADMET and BBB permeation properties using dedicated machine learning algorithms; (iii) systematically contextualizing quercetin within a compound-target-disease network; and (iv) designing and prioritizing novel quercetin-derived MTDLs through multi-objective optimization, validated by deep Graph Isomorphism Networks (GIN) and protein-ligand co-folding via Boltz-2. The QSAR models, rigorously validated through scaffold splitting, demonstrated high accuracy (ROC-AUC 0.86–0.95; mean 0.91) and successfully recovered approved reference Drugs. Calculations confirmed quercetin as a genuine but weak MTDL with an unfavorable central nervous system profile (predicted BBB probability of 0.29; topological polar surface area of 131 Å²; and an active PAINS alert). From a focused library of 783 generated derivatives, 108 were predicted to outperform the parent compound. The lead candidate improved multi-objective desirability by 142%, significantly enhancing predicted multi-target potency, BBB

permeation, and drug-likeness, while successfully removing the structural alert. However, structure-based co-folding predictions did not fully reproduce this binding affinity gain for the lead compound, exposing a notable informative tension between ligand-based and structure-based computational methods. Overall, this workflow provides innovative, synthetically accessible quercetin-derived MTDL candidates for neurodegeneration, highlighting the necessity of subsequent experimental validation.

Keywords: Quercetin; Multi-Target Ligands; Neurodegeneration; Machine Learning; Generative Drug Design; Blood-Brain Barrier.

RESUMO

Doenças neurodegenerativas, como a doença de Alzheimer (DA) e a doença de Parkinson (DP), são multifatoriais, e fármacos de alvo único têm falhado, em grande parte, em modificar seu curso. Isso motiva o desenvolvimento de ligantes direcionados a múltiplos alvos (MTDLs). A quercetina, um flavonol da dieta com neuroproteção documentada, é um ponto de partida atraente, porém subótimo, devido à baixa potência, à permeação deficiente na barreira hematoencefálica (BHE) e ao potencial de interferência em ensaios. Relatamos um fluxo de trabalho de inteligência artificial (IA) de ponta a ponta e totalmente reproduzível que: (i) mapeia a polifarmacologia da quercetina em relação a oito alvos validados da DA/DP, utilizando modelos de relação quantitativa estrutura-atividade (QSAR) treinados com 56.328 bioatividades curadas do banco de dados ChEMBL (40.207 compostos únicos); (ii) traça o perfil das propriedades ADMET (absorção, distribuição, metabolismo, excreção e toxicidade) e da permeação na BHE com modelos dedicados de aprendizado de máquina; (iii) contextualiza a quercetina em uma rede composto-alvo-doença; (iv) projeta e prioriza novos MTDLs derivados da quercetina por meio de

otimização multiobjetivo, com validação por aprendizado profundo (redes neurais em grafos) e co-folding com o Boltz-2. Os modelos QSAR, validados por divisão de scaffold (esqueleto molecular), apresentaram boa acurácia (ROC-AUC 0,86-0,95; média de 0,91) e identificaram corretamente fármacos de referência aprovados. A quercetina foi confirmada como um ligante multialvo genuíno, porém fraco, com perfil desfavorável para o sistema nervoso central (probabilidade prevista de permeação na BHE de 0,29; área de superfície polar topológica de 131 Å²; alerta de interferência generalizada em ensaios). A partir de uma biblioteca direcionada de 783 derivados, previu-se que 108 seriam superiores à quercetina; o principal candidato apresentou um aumento de 142% na desejabilidade multialvo, elevando simultaneamente a potência multialvo prevista, a permeação na BHE e a semelhança com fármacos (drug-likeness), além de eliminar o alerta de interferência. No entanto, as previsões de co-folding estrutural não reproduziram totalmente o ganho de afinidade do composto líder, evidenciando uma divergência informativa relevante entre métodos baseados em ligantes e métodos baseados em estrutura. Em suma, este fluxo de trabalho computacional fornece candidatos a MTDLs (ligantes multialvo) derivados da quercetina, inovadores e de síntese acessível, para o tratamento da neurodegeneração, ressaltando a necessidade de validação experimental subsequente.

Palavras-chave: Quercetina; Ligantes multialvo; Neurodegeneração; Aprendizado de máquina; Design molecular generativo; Barreira hematoencefálica.

1. INTRODUCTION

Alzheimer's disease (AD) and Parkinson's disease (PD) represent the two most prevalent neurodegenerative disorders in the world,

posing a rapidly growing global socioeconomic and health challenge (Scheltens *et al.*, 2021; Bloem *et al.*, 2021). The pathophysiology of these conditions is intrinsically complex and multifactorial, characterized by a cascade of simultaneous events that includes cholinergic deficit, misfolding and subsequent aggregation of proteins (such as β -amyloid, tau, and α -synuclein), as well as monoaminergic dysregulation, oxidative stress, and chronic neuroinflammation (Scheltens *et al.*, 2021; Bloem *et al.*, 2021).

Given this systemic complexity, the traditional drug development approach based on the "one drug, one target" paradigm has proven ineffective, accumulating successive failures in modifying the clinical course of these diseases (Grewal *et al.*, 2021; Scheltens *et al.*, 2021). In response, medical science has been strategically adopting the concept of multi-target-directed ligands (MTDLs). This approach consists of the deliberate design of a single chemical entity capable of simultaneously interacting with multiple crucial biological targets, promoting a synergistic and balanced therapeutic effect (Grewal *et al.*, 2021; Zhang *et al.*, 2025).

In this scenario, natural products, especially flavonoids, stand out as privileged sources of structures with polypharmacological and neuroprotective properties (Grewal *et al.*, 2021). Quercetin (3,3',4',5,7-pentahydroxyflavone) is one of the most studied dietary flavonols and exhibits antioxidant, anti-inflammatory, anti-amyloidogenic, and enzyme-modulating activities relevant to Alzheimer's disease (AD) and Parkinson's disease (PD) (Grewal *et al.*, 2021). Despite its potential, the clinical application of native quercetin faces severe pharmacokinetic limitations: its potency is modest, it undergoes extensive phase II metabolism, has very low oral bioavailability, and exhibits poor permeation across the blood-brain barrier (BBB).

Furthermore, the presence of a catechol group in its structure causes it to act as a pan-interference assay (PAINS) compound, functioning more as a chemical model for optimization than as a ready-made drug (Grewal *et al.*, 2021).

Quercetin stands out as a strategic molecular scaffold in drug discovery due to its rigidly defined chemical architecture, based on the flavonol core (C6–C3–C6), which favors stable interactions with different biological targets. The presence of multiple hydroxyl groups in specific positions confers a high capacity for hydrogen bonding and π – π interactions, allowing the simultaneous recognition of distinct active sites, an essential characteristic for multi-target approaches. Although it presents pharmacokinetic limitations, such as low bioavailability, its structure is highly amenable to chemical functionalization, enabling rational modifications aimed at optimizing ADMET properties without loss of the central pharmacophore. Thus, quercetin acts as a versatile lead scaffold, serving as a platform for the design of derivatives with a polypharmacological profile and enhanced neuroprotective potential in neurodegenerative diseases (Zhang *et al.*, 2024; Liu *et al.*, 2024).

In this context, the selection of eight molecular targets—acetylcholinesterase (AChE), butyrylcholinesterase (BChE), β -secretase 1 (BACE1), monoamine oxidases A and B (MAO-A and MAO-B), glycogen synthase kinase-3 β (GSK-3 β), adenosine receptor A2A (ADORA2A), and catechol -O- methyltransferase (COMT)—reflects their integrated participation in critical axes of the pathophysiology of Alzheimer's and Parkinson's diseases, including cholinergic neurotransmission, monoaminergic metabolism, protein aggregation, and pathways of oxidative stress and

neuroinflammation. The simultaneous modulation of these systems by multi-target ligands constitutes a promising strategy to intervene more comprehensively in the complex neurodegenerative network (Knopman). *et al.*, 2024).

In parallel, drug discovery has evolved from approaches based on isolated virtual screenings to integrated computational workflows, in which chemoinformatics, molecular modeling, machine learning, and Artificial Intelligence are employed in a complementary way to support the rational design of new chemical entities. This paradigm is represented by the Design–Make–Test–Analyze (DMTA) cycle, in which information obtained at each stage continuously feeds back into the molecular optimization process, allowing for more efficient exploration of the chemical space, reducing experimental cycles, and accelerating the identification of candidates with more promising pharmacological and pharmacokinetic properties (Vora). *et al.*, 2024; Zhang *et al.*, 2025; Paul *et al.*, 2024).

In this context, recent advances in Artificial Intelligence have significantly expanded the predictive capacity of these computational flows. Technologies such as graph neural networks (GNNs), generative models, and co-folding tools make it possible to predict, with high accuracy, physicochemical properties, pharmacokinetic profiles, and binding affinities of new molecules, overcoming limitations of conventional computational screening strategies (Zhang *et al.*, 2025; Passaro *et al.*, 2025).

However, despite these advances, most studies involving quercetin in the area of neurodegeneration remain limited to the computational characterization of the original molecule, mainly using molecular docking and descriptive analyses. Studies

employing integrated Artificial Intelligence strategies for the rational planning and optimization of new quercetin derivatives with multi-target therapeutic potential for Alzheimer's and Parkinson's diseases remain scarce (Sabarathinan, 2024; Kaur ; Kulkarni ; Wairkar, 2024; Zhang *et al.*, 2025; Passaro *et al.*, 2025).

Given this gap, an integrated, reproducible computational pipeline based on Artificial Intelligence and chemoinformatics was established for the rational design of multi-target ligands derived from quercetin. The strategy integrates predictive modeling, pharmacokinetic property assessment, polypharmacology analysis, and generative molecular design, allowing the characterization of the multi-target profile of quercetin against therapeutic targets relevant to Alzheimer's and Parkinson's diseases, the identification of its main structural and pharmacokinetic limitations, and the prioritization of new derivatives with more favorable predicted pharmacological and pharmacokinetic properties.

In this context, the objective of this study was to establish and apply an Artificial Intelligence-based computational pipeline to identify and prioritize quercetin derivatives with multi-target potential for Alzheimer's and Parkinson's diseases.

2. METHODOLOGY

This study consists of a computational (*in silico*) research, based on molecular modeling, chemoinformatics, and Artificial Intelligence techniques applied to the rational design of drugs. The methodological flow was structured according to the **Design–Make–Test–Analyze (DMTA) cycle, fully adapted to the computational environment, allowing the *in silico*** planning, optimization, and

evaluation of molecular derivatives without the need for experimental synthesis steps or biological validation.

All analyses were conducted using open-source computational tools and libraries, including Python 3, RDKit, scikit-learn, XGBoost, PyTorch Geometric, and Boltz-2, ensuring reproducibility, transparency, and scalability of the workflow. The data used were obtained exclusively from publicly available and freely accessible databases, encompassing structural, pharmacological, and biochemical information from repositories widely used in drug discovery.

Because this is a purely computational study, developed from public data and without the involvement of human participants, experimental animals, or biological material, there was no need to submit it to a Research Ethics Committee, in accordance with current legislation.

2.1. Acquisition And Curation Of Data

Bioactivity data for eight human targets of DA/PD — acetylcholinesterase (AChE), butyrylcholinesterase (BChE), β -secretase 1 (BACE1), monoamine oxidases A and B (MAO-A, MAO-B), glycogen synthase kinase-3 β (GSK-3 β), adenosine receptor A2A (ADORA2A), and catechol -O- methyltransferase (COMT) — were obtained from ChEMBL (Zdrasil). *et al.*, 2024) via their REST API, restricting themselves to records with a defined pChEMBL value (IC50, Ki, Kd, EC50, or Potency). Structures were standardized using RDKit (major fragment, normalization, neutralization) and deduplicated using InChIKey, aggregating replicated measures by the median pChEMBL. Compounds with pChEMBL ≥ 6 were labeled

as active and those with pChEMBL < 5 as inactive; the intermediate zone was reserved from classification. Quercetin and nine approved reference drugs were obtained and standardized identically.

2.2. Representation Molecular And Division Of The Data

The molecules were coded as 2,048-bit Morgan fingerprints (ECFP4) (Rogers; Hahn, 2010) and by physicochemical descriptors. To provide a realistic estimate of generalization, all models were validated by Bemis- Murcko scaffold splitting (Bemis; Murcko, 1996), which avoids scaffold leakage between training and testing; five-fold out-of-fold predictions were pooled for metric estimation. The applicability domain was quantified as the maximum Tanimoto similarity of a query compound to the training assets.

2.3. Modeling QSAR

For each target, classifiers (active/inactive) and regressors (pChEMBL) of random forest (Sharma *et al.*, 2024) and gradient trees. Boosting (Chen; Guestrin, 2016) were trained with class imbalance treatment. Performance was assessed by ROC-AUC, precision- recall AUC (PR-AUC), and Matthews correlation coefficient (MCC) in classification, and by RMSE, R^2 , and Spearman's ρ in regression. The best model per target and outcome was retained for downstream scoring.

2.4. Prediction Of ADMET And From The Barrier Blood-Brain

Drug -likeness and CNS suitability were characterized using RDKit descriptors, Lipinski /Veber/Egan rules, and quantitative drug-likeness assessment (QED) (Bickerton). *et al.*, 2012), a synthetic accessibility score (SA) (Ertl; Schuffenhauer, 2009), an approximation

of the CNS-MPO (Wager *et al.*, 2010) and PAINS structural alerts. A dedicated BBB permeation classifier was trained on the public BBBP (MoleculeNet) dataset (Wu *et al.*, 2018) under the same scaffolding protocol.

2.5. Network Pharmacology

Experimentally measured human targets of quercetin were extracted from ChEMBL (Zdravil. *et al.*, 2024) and crossed with genes associated with AD and PD from the Open Targets platform (Ochoa *et al.*, 2023). Protein-protein interactions between the intersection targets were obtained from STRING (Szklarczyk. *et al.*, 2023), and the enrichment of pathways/processes (KEGG, Reactome, Gene Ontology) was calculated with og (Kolberg *et al.*, 2023).

2.6. Design Generative And Multi-Objective Prioritization

A focused library of quercetin derivatives was generated by two complementary strategies: (i) plausible and synthetically accessible medicinal-chemical transformations of the five hydroxyl groups and aromatic rings (O-alkylation, O-acylation, and aromatic halogenation) and (ii) de novo fragment assembly (BRICS), combining quercetin fragments with fragments of known active inhibitors. Each candidate was scored by a geometric mean desirability integrating predicted multi-target activity (mean probability over the six core enzymatic/kinase targets), BHE probability, QED, and synthetic accessibility, with penalties for PAINS alerts. Non-dominated candidates (Pareto optimums) were identified across the four objectives.

2.7. Network Neural Of Graph

As an independent deep learning comparator, a multitasking graph neural network (graph isomorphism network, GIN) (*et al.*, 2019), implemented in PyTorch Geometric, was trained end-to-end on molecular graphs of the eight-target matrix under scaffold splitting, with masked binary cross-entropy loss (which accommodates missing labels), and used to reassess quercetin, reference drugs, and designed lead candidates.

2.8. Co-Folding Put Apprenticeship Deep And Affinity Of Connection

For priority target-ligand pairs, three-dimensional complexes and binding affinities were predicted with Boltz-2 (Passaro *et al.*, 2025), *a 2025 co-folding model that estimates protein-ligand structure and affinity from sequence and SMILES, at a small fraction of the cost of physical free energy perturbation methods, with accuracy close to AlphaFold 3 class structure prediction (Abramson). et al., 2024).* Canonical UniProt sequences of AChE, MAO-B, and GSK-3 β were used with quercetin, the designed lead candidates, and approved controls.

2.9. Reproducibility

All the code, curated datasets, trained models, and figures are organized into a single project and available to the authors; each step is scripted and deterministic (fixed random seeds).

3. RESULTS

3.1. Models QSAR Accurate And Validated With Honesty

You eight classifiers QSAR validated put scaffold they were evenly robust, with ROC-AUC ranging from 0.855 (MAO-A) to 0.953 (ADORA2A) and a mean of 0.91, and PR-AUC frequently above 0.95 (Table 1, Figure 1). As a positive control, the models correctly classified the approved reference drugs as active against their primary targets (e.g., donepezil-AChE, probability 0.97; rivastigmine-BChE, 0.82; safinamide -MAO-B, 0.94; entacapone -COMT, 0.91), corroborating their reliability. The regression models were moderate, as expected under scaffolding (R^2 0.33-0.63).

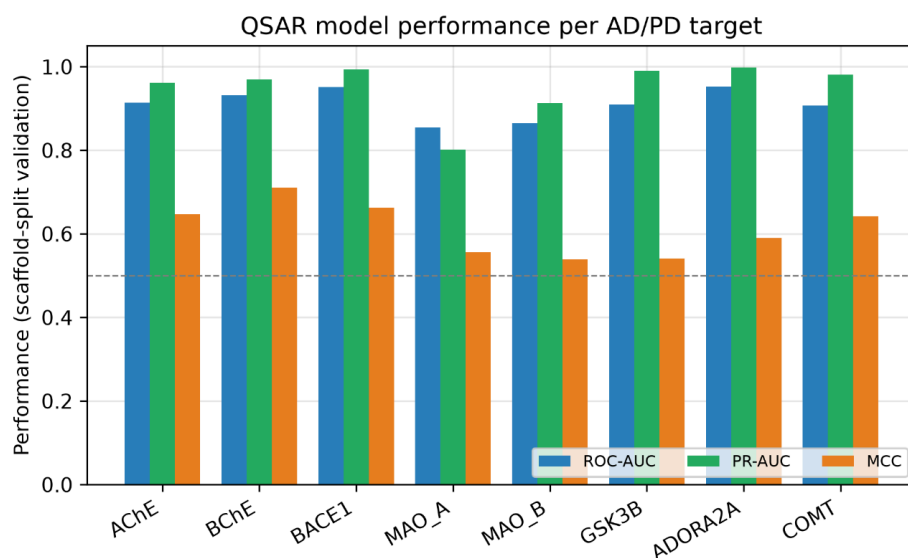
Table 1. Performance, by scaffold division, of the best QSAR classifier for each Alzheimer's/Parkinson's target and the corresponding probability of predicted quercetin activity. AD, disease of Alzheimer's; PD, Parkinson's disease; * outside the strict domain of applicability.

Target	Illness	N	ROC-AUC	PR-AUC	MCC
AChE	FROM THE	4227	0.914	0.962	0.64
BChE	FROM THE	2539	0.932	0.970	0.71
BACE1	FROM THE	8046	0.951	0.994	0.66
GSK3B	FROM THE	4036	0.910	0.990	0.54
MAO- A	DP	2121	0.855	0.801	0.59

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: <https://revistatopicos.com.br/artigos/design-of-ligands-multi-target-derivatives-from-the-quercetin-guided-put-artificial-intelligence-for-alzheimers-and-parkinsons-diseases-mapping-polypharmacology-generative-optimization-and-co-folding-through-deep-learning?noblockage>

Source: The Authors (2026)

Figure 1. Performance, by scaffold division (ROC-AUC, PR-AUC and correlation coefficient) of Matthews), of better classifier QSAR to each target of Alzheimer's/Parkinson's.

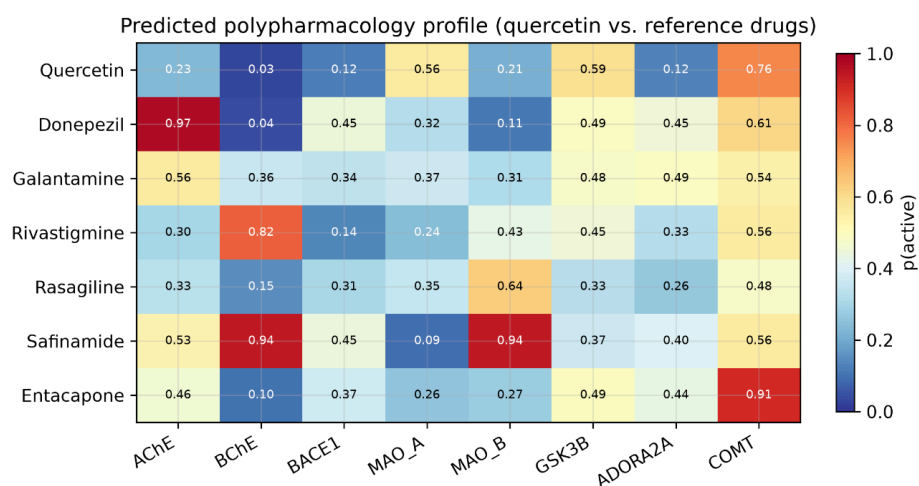


Source: The Authors (2026)

3.2. THE Quercetin And a Ligand Multi-Target Genuine, But Weak

When applied to quercetin, the models predicted a broad, but low to moderate engagement profile (Figure 2). Within the domain of applicability, the highest probabilities of activity they were obtained to MAO-A (0.56) GSK-3 β (0.59) and — as substrate catecholic — COMT (0.76), with weaker predicted activity against AChE (0.23), MAO-B (0.21) and BACE1 (0.12). This pattern is consistent with the known pharmacology of quercetin as a modest and promiscuous modulator, and not as a potent inhibitor, and defines the optimization problem discussed below. A quercetin he has a profile unfavorable to the system nervous central.

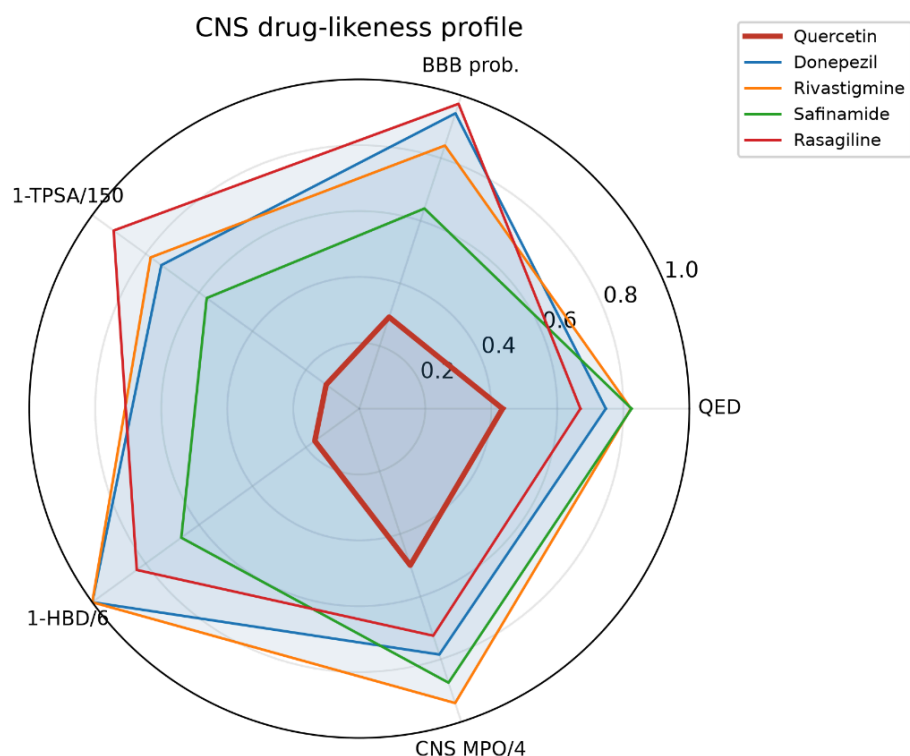
Figure 2. Predicted polypharmacology (probability of activity) of quercetin and reference drugs. approved between you eight targets; drugs reference They are correctly recovered as assets against their primary targets, serving as positive controls.



Source: The Authors (2026)

ADMET analysis quantified quercetin's liabilities as a CNS agent: a high topological polar surface area (Figure 3) ($131 \text{ \AA}^2$), five hydrogen bond donors, low QED (0.43), low CNS-MPO score, and a predicted BBB permeation probability of only 0.29—well below approved CNS drugs (0.52-0.97)—along with a PAINS alert arising from the catechol ring B. The BBB classifier itself was accurate (ROC-AUC 0.89). These results precisely identify the properties that an optimized derivative should improve.

Figure 3. Drug-likeness radar for the central nervous system comparing quercetin with approved CNS drugs (QED, BBB probability, topological polar surface area and inversely normalized hydrogen donors, and CNS-MPO score).



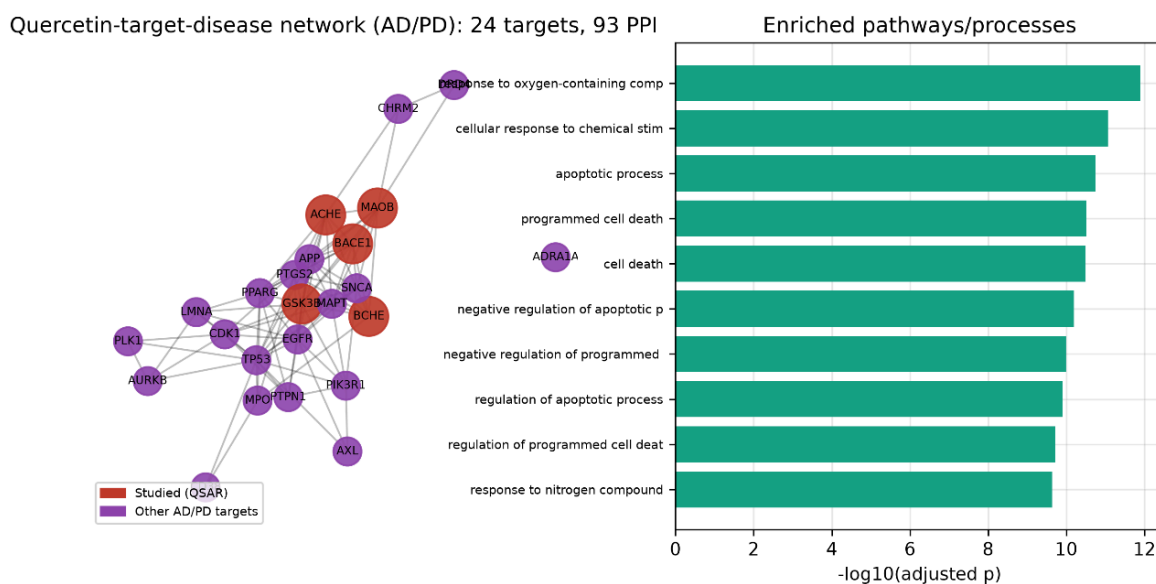
Source: The Authors (2026)

3.3. THE Network Pharmacology League The Quercetin The Programs Of Stress Oxidative And Apoptosis

Quercetin had 151 experimentally measured human targets in ChEMBL, of which 24 overlapped with AD/PD disease genes, including the enzymatic targets studied here (ACHE, BACE1, BCHE, MAOB, GSK3B) and key disease proteins (APP, MAPT/tau, SNCA/ α -synuclein, PTGS2/COX-2, PPARG). The intersection formed a connected network. of interaction protein-protein (93 edges), and the enrichment functional he was dominated by response to oxygen-containing compounds (oxidative stress) and regulation of apoptotic/programmed cell death (adjusted p from g:Profiler $\approx 10^{-11}$ - 10^{-12} ; Figure 4), recovering the canonical neuroprotective mechanism of quercetin directly from the data.

Figure 4. Network pharmacology. Left: quercetin-target-disease network for the quercetin/DA-DP gene intersection (targets

modeled here highlighted). Right: most significantly enriched biological pathways and processes.



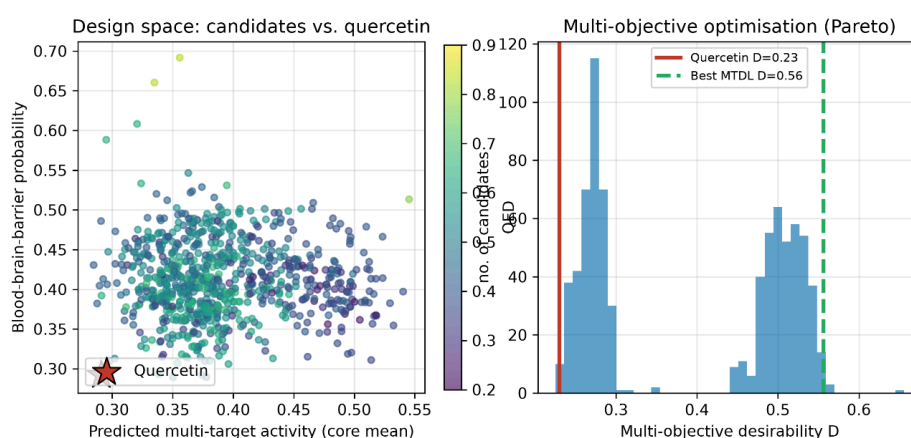
Source: The Authors (2026)

3.4. THE Design Guided Put AI Generates MTDLs Derivatives From The Quercetin Predicted as Superior To The Original Compound

A library of 783 quercetin derivatives was generated and scored. Removal of the catechol PAINS alert (primarily by O-methylation) eliminated the structural liability in most rational derivatives, and 404 candidates were PAINS-free. Under multi-objective optimization, 108 derivatives were predicted to be superior to quercetin in combined desirability while engaging at least three of the six core targets, and six candidates were Pareto optimal (Figure 5). The leading compound (a 7-O-propyl-4'-O-methyl derivative of quercetin) increased desirability from 0.23 (quercetin) to 0.56—a 142% improvement—while also increasing the activity multi-target prediction, raising the probability of BHE from 0.29 to 0.44, improving the QED from 0.43 to 0.64, remaining readily synthesizable (SA score 2.4) and without PAINS alert (Table 2, Figure 6). Mechanically, O-alkylation masks the catechol (removing the alert and reducing the number of hydrogen bond donors). and the

surface polar, improving like this the permeation from the BHE), while preserve or increases predicted engagement with the target — a coherent and testable structure-activity rationale.

Figure 5. AI-guided design. Left: design space of generated candidates (predicted multi-target activity versus BHE probability, colored by QED), with quercetin highlighted (star). Right: distribution of multi-target desirability, with quercetin and the best multi-target ligand indicated.



Source: The Authors (2026)

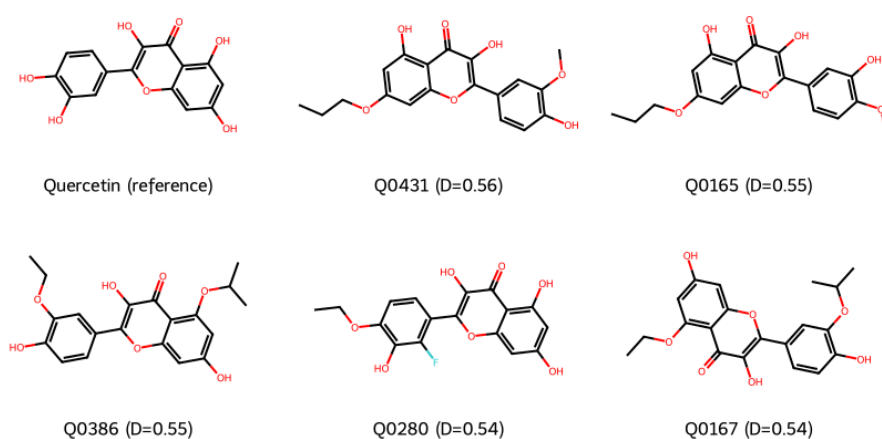
Table 2. Main multi-target ligands derived from quercetin, ordered by multi-target desirability (D), compared with quercetin. n(central) = number of the six central targets engaged ($p \geq 0.5$); BBB = predicted probability of blood-brain barrier; SA = synthetic accessibility.

Compound	Activity	BHE	QED	SA	n(cent)
Quercetin	0.29	0.29	0.43	25	2
Q0431	0.40	0.44	0.64	24	3
Q0165	0.40	0.44	0.64	25	3
Q0386	0.45	0.39	0.63	26	3

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: <https://revistatopicos.com.br/artigos/design-of-ligands-multi-target-derivatives-from-the-quercetin-guided-put-artificial-intelligence-for-alzheimers-and-parkinsons-diseases-mapping-polypharmacology-generative-optimization-and-co-folding-through-deep-learning?noblockage>

Source: The Authors (2026)

Figure 6. Chemical structures of quercetin and the five leading multi-target ligands designed, with their multi-target desirability scores.



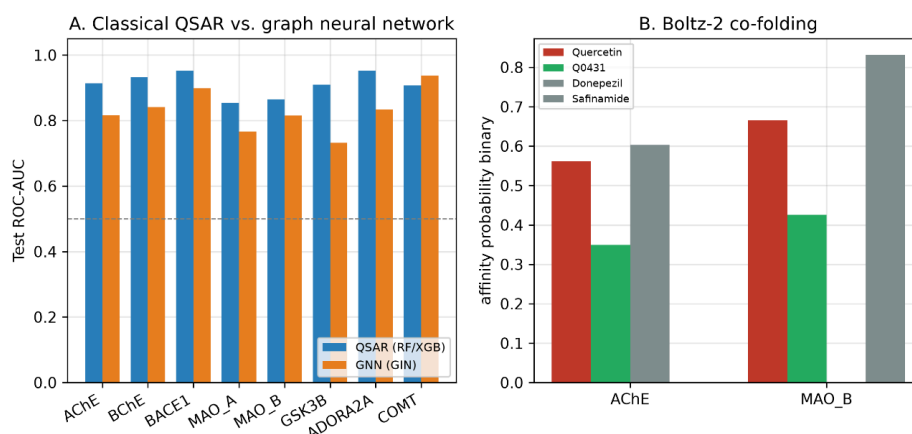
Source: The Authors (2026)

3.5. Validation Put Apprenticeship Deep (Network Neural Of Graph)

An independent multitasking graph neural network, trained end-to-end on molecular graphs under the same scaffolding protocol, corroborated the tree-based models, achieving a test average ROC-AUC of 0.83 across the eight targets (range 0.73-0.94). The graph network replicated the multi-target qualitative classifications of quercetin and the drawn leading candidates obtained with the random forest/ gradient set. boosting, indicating that the

polypharmacology findings are robust to the choice of model family (Figure 7).

Figure 7. Validation by deep learning. (A) ROC-AUC of classical QSAR target test (forest random/XGBoost) versus network neural of graph (graph isomorphism (A) Network). (B) Binding probabilities predicted by Boltz-2 co-folding for priority target-ligand pairs.



Source: The Authors (2026)

3.6. Validation Put Co-Folding And Affinity Of Connection (Boltz-2)

Co-folding with Boltz-2 produced three-dimensional protein-ligand complexes and binding affinity estimates for the priority pairs (Table 3). Against AChE, the leading candidate showed a lower predicted binding probability than quercetin (0.35 vs. 0.56); against MAO-B, the leading candidate showed a lower predicted binding probability. that quercetin (0.43 vs. 0.67). In a way Reassuringly, the approved controls were the ligands with the highest probability of matching their respective targets (donepezil (AChE, 0.60); safinamide (MAO-B, 0.83)), supporting the validity of the structure-based estimates.

Table 3. Boltz-2 co-folding predictions for priority target-ligand pairs. Highest probability of binding. Lower affinity values indicate a

stronger predicted link.

Target	Binder	Probability connection	Affinity value
AChE	Quercetin	0.561	0.931
AChE	Q0431	0.349	1,441
AChE	Donepezilla	0.603	0.846
MAO- B	Quercetin	0.666	0.644
MAO- B	Q0431	0.426	1.302
MAO- B	Safinamide	0.831	-0.491

Source: The Authors (2026).

4. DISCUSSION

This study demonstrates that a transparent AI stream can go beyond the now abundant descriptive computational analyses of quercetin to actively design new chemical matter. Three findings deserve highlighting. First, by validating each QSAR model by scaffolding and retrieving approved drugs as positive controls, we provide an honest and uninflated estimate of predictive performance—a frequent weakness of virtual screening reports. Second, the models realistically portray quercetin: a genuine, yet weak and promiscuous, multi-target ligand with an unfavorable CNS profile, which explains the gap between its abundant in vitro neuroprotection and its limited translation. Third, and most importantly, multi-objective generative optimization yielded synthetically accessible derivatives. predicted as simultaneously more powerful between targets, more permeable to the brain and more drug -like than the original compound, with the catechol passive removed.

The underlying chemistry of the leading candidates is deliberately conservative and experimentally accessible: selective O-alkylation of the flavonol hydroxyl groups is well preceded and addresses directly to the two main deficiencies from the quercetin — polarity/number of excessive hydrogen donors (which limit BBB penetration) and redox-active catechol (responsible for PAINS alert and metabolic instability). The fact that an unbiased, machine learning-guided objective has rediscovered this medicinal-chemical logic reinforces confidence in the approach.

It is important to note that structure-based co-folding offered a more cautious and honest perspective. Boltz-2 correctly classified the approved controls as the strongest ligands of their respective targets (donepezil for AChE, safinamide for MAO-B), validating the method; however, it did not corroborate the ligand-based prediction of Enhanced engagement was observed for the lead candidate, which showed a lower predicted binding probability than quercetin against both AChE and MAO-B. This discrepancy between ligand-based (QSAR and graph neural network) and structure-based (co-folding) approaches is itself informative: masking hydroxyl groups that improves predicted BBB permeation and drug-likeness can simultaneously attenuate polar contacts that contribute to binding. Far from being a weakness, this tension delineates a concrete and testable hypothesis and reinforces that the designed molecules are computational candidates whose affinity and selectivity ultimately need to be resolved experimentally.

Our work has clear limitations. All results are computational; the predicted activities and affinities, however well validated internally, must be confirmed by synthesis and biological assay. Some predictions of quercetin (COMT, ADORA2A, BChE) fell outside the

strict domain of applicability and are reported as having lower confidence. The QSAR regression under scaffolding is necessarily moderate. Finally, the designed library emphasizes synthetically obvious transformations; a broader generative exploration and experimental medicinal chemistry are natural extensions. Prospective benchtop validation of the leading MTDL candidates against AChE, MAOs, and GSK-3 β , along with cellular models of neuroprotection and BBB, is the next priority step.

5. CONCLUSION/FINAL CONSIDERATIONS

An end-to-end, reproducible AI workflow mapped the polypharmacology of quercetin. In eight DA/PD targets, their CNS liabilities were quantified, they were placed within a disease network dominated by oxidative stress and apoptosis programs, and novel multi-target quercetin-derived ligands were designed that were predicted to be superior to the source compound in potency, brain permeation, and drug-likeness, while remaining synthetically accessible. The approach illustrates as the apprenticeship of machine modern he can convert one product natural promising, but flawed, in concrete and testable candidates for neurodegeneration, and provides a transferable model to other phytochemicals.

REFERENCES

ABRAMSON, J. et al. **Accurate structure prediction of biomolecular interactions with AlphaFold 3**. Nature, v. 630, n. 8016, p. 493–500, 2024. Available at: <https://doi.org/10.1038/s41586-024-07487-w>
Accessed on: January 12, 2026.

BEMIS, GW; MURCKO, MA **The properties of known drugs. 1. Molecular frameworks**. Journal Journal of Medicinal Chemistry, v.

39, n. 15, p. 2887–2893, 1996. Available at: <https://pubs.acs.org>. Accessed on: March 3, 2026.

BICKERTON, GR et al. **Quantifying the chemical beauty of Drugs.** Nature. Chemistry, v. 4, n. 2, p. 90–98, 2012. Available at: <https://doi.org/10.1038/nchem.1243>. Accessed on: November 27, 2025.

BLOEM, BR; OKUN, MS; KLEIN, C. **Parkinson's disease.** The Lancet, v. 397, n. 10291, p. 2284–2303, 2021. Available at: [https://doi.org/10.1016/S0140-6736\(21\)00218-X](https://doi.org/10.1016/S0140-6736(21)00218-X). Accessed on: February 18, 2026.

CHEN, T.; GUESTRIN, C. **XGBoost: a scalable tree boosting system.** In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016. p. 785–794. Available at: <https://doi.org/10.1145/2939672.2939785>. Accessed on: December 5, 2025.

ERTL, P.; SCHUFFENHAUER, A. **Estimation of synthetic accessibility score of drug -like molecules based on molecular complexity and fragment Contributions** Journal of Cheminformatics, v. 1, n. 1, p. 8, 2009. Available at: <https://doi.org/10.1186/1758-2946-1-8>. Accessed on: January 22, 2026.

GREWAL, AK et al. **Mechanistic insights and perspectives involved in neuroprotective action of quercetin.** Biomedicine & Pharmacotherapy, v. 140, p. 111729, 2021. Available at: <https://doi.org/10.1016/j.biopha.2021.111729>. Accessed on: Apr. 14, 2026.

KAUR, K.; KULKARNI, Y.A.; WAIRKAR, S. **Exploring the potential of quercetin in Alzheimer's disease: pharmacodynamics, pharmacokinetics, and Nanodelivery systems.** Brain Research, v.

1834, p. 148905, 2024. Available at: <https://doi.org/10.1016/j.brainres.2024.148905>. Accessed on: January 9, 2026.

KNOPMAN, DS et al. **Alzheimer disease**. Nature Reviews Disease Primers, v. 10, 2024. Available at: <https://doi.org/10.1038/s41572-024-00476-8>. Accessed on: February 28, 2026.

KOLBERG, L. et al. g: **Profiler—interoperable web service for functional enrichment analysis (2023 update)**. Nucleic Acids Research, v. 51, n. W1, p. W207–W212, 2023. Available at: <https://doi.org/10.1093/nar/gkad347>. Accessed on: December 11, 2025.

LIU, H. et al. **Structural modification strategies of flavonoids for drug development: focus on quercetin derivatives**. European Journal of Medicinal Chemistry, 2024. Available at: <https://doi.org/10.1016/j.ejmech.2024.115874>. Accessed on: April 2, 2026.

OCHOA, D. et al. **The next-generation Open Targets Platform**. Nucleic Acids Research, v. 51, n. D1, p. D1353–D1359, 2023. Available at: <https://doi.org/10.1093/nar/gkac1046>. Accessed on: Nov. 30, 2025.

PASSARO, S. et al. **Boltz-2: towards accurate and efficient binding affinity prediction**. bioRxiv, 2025. Available at: <https://doi.org/10.1101/2025.06.14.659707>. Accessed on: July 7, 2026.

PAUL, D. et al. **Artificial intelligence in drugs discovery and development**. Drug Discovery Today, 2024. Available at: <https://doi.org/10.1016/j.drudis.2024.103861>. Accessed on: March 19, 2026.

ROGERS, D.; HAHN, **M. Extended-connectivity fingerprints.** Journal of Chemical Information and Modeling, v. 50, n. 5, p. 742–754, 2010. Available at: <https://doi.org/10.1021/ci100050t>. Accessed on: February 25, 2026.

SABARATHINAM, **S. Unraveling the therapeutic potential of quercetin and quercetin-3-O-glucuronide in Alzheimer's disease through network pharmacology, molecular docking, and dynamic simulations.** Scientific Reports, v. 14, p. 14852, 2024. Available at: <https://doi.org/10.1038/s41598-024-61779-9>. Accessed on: January 6, 2026.

SCHELTENS, P. et al. **Alzheimer's disease.** The Lancet, v. 397, n. 10284, p. 1577–1590, 2021. Available at: [https://doi.org/10.1016/S0140-6736\(20\)32205-4](https://doi.org/10.1016/S0140-6736(20)32205-4). Accessed on: March 21, 2026.

SHARMA, A. et al. **Machine learning in drugs discovery: recent advances and applications.** Briefings in Bioinformatics, 2024. Available at: <https://doi.org/10.1093/bib/bbae123>. Accessed on: April 10, 2026.

SZKLARCZYK, D. et al. **The STRING database in 2023.** Nucleic Acids Research, v. 51, n. D1, p. D638–D646, 2023. Available at: <https://doi.org/10.1093/nar/gkac1000>. Accessed on: Dec. 15, 2025.

VORA, J. et al. **Artificial intelligence in drugs discovery: recent advances and future perspectives.** Drug Discovery Today, 2024. Available at: <https://doi.org/10.1016/j.drudis.2024.103978>. Accessed on: April 4, 2026.

WAGER, TT et al. **Moving beyond rules: the development of a central nervous system multi-parameter Optimization (CNS MPO)**

desirability tool. ACS Chemical Neuroscience, v. 1, n. 6, p. 435–449, 2010. Available at: <https://doi.org/10.1021/cn100008c>. Accessed on: February 13, 2026.

WU, Z. et al. **MoleculeNet: a benchmark for molecular machine learning.** Chemical Science, v. 9, n. 2, p. 513–530, 2018. Available at: <https://doi.org/10.1039/C7SC02664A>. Accessed on: Nov. 26, 2025.

XU, K. et al. **How powerful are graph neural networks? In:** International Conference on Learning Representations (ICLR). 2019. Available at: <https://openreview.net>. Accessed on: January 17, 2026.

ZDRAZIL, B. et al. **The ChEMBL Database in 2023.** Nucleic Acids Research, v. 52, n. D1, p. D1180–D1192, 2024. Available at: <https://doi.org/10.1093/nar/gkad1004>. Accessed on: April 1, 2026.

ZHANG, Q. et al. **Graph neural networks in modern AI- aided drug discovery.** Chemical Reviews, v. 125, n. 20, p. 10001–10103, 2025. Available at: <https://doi.org/10.1021/acs.chemrev.5c00000>. Accessed on: July 8, 2026.

ZHANG, Y. et al. **Quercetin as a multi-target therapeutic candidate for neurodegenerative diseases: advances and challenges.** International Journal of Biological Macromolecules, 2024. Available at: <https://doi.org/10.1016/j.ijbiomac.2024.129876>. Accessed on: March 29, 2026.

¹ University of the Amazon (UNAMA). Bachelor's degree in Biomedicine; Specialist in Clinical and Toxicological Analyses (UNAMA). E-mail: [acesse o artigo original para visualizar o e-mail](#).
ORCID: <https://orcid.org/0009-0008-9491-2575>.

² University of the Amazon (UNAMA). Bachelor's degree in Social Work; Specialist in Environmental Analysis (ICB/UFGA); Multiprofessional Specialization in Primary Care and Family Health (UFGA). E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0009-0002-4214-1111>.

³ Federal Rural University of the Amazon (UFRA). Bachelor's degree in Agronomy; Specialist in Environmental Management (NUMA/UFGA); Master's degree in Food Science and Technology (PPGCTA/UFGA). E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0000-0002-1207-4213>.

⁴ Paulista University (UNIP) Systems Analysis and Development; Postgraduate Program in Electrical Engineering (Applied Computing – Computational Intelligence), Federal University of Pará (UFGA); Plim.AI. E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0000-0002-4542-3926>.

⁵ University of the Amazon (UNAMA). Bachelor's degree in Nutrition; Master's degree in Food Science and Technology (PPGCTA/UFGA). E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0009-0003-7706-2496>

⁶ University of the Amazon (UNAMA). Dentist: Bachelor's degree in Dentistry; Bachelor's degree in Pedagogy (UVA-CE). E-mail: [acesse o artigo original para visualizar o e-mail](#)