AI-Assisted Drug Discovery from Phytoconstituents: Predicting Drug–Target Interactions

Muskan Tomar

ABSTRACT

Artificial intelligence (AI) has emerged as a powerful computational tool to support early-stage drug discovery by facilitating the analysis and prioritization of bioactive compounds. In the context of plant-derived phytoconstituents, AI-based methods offer significant potential to assist drug–target interaction (DTI) prediction, virtual screening, and pharmacological prioritization. This narrative review summarizes recent advances in AI-assisted approaches applied to phytochemical-based drug discovery, with a particular focus on machine learning, deep learning, and graph-based models used for DTI prediction. The review also discusses the role of AI in supporting in-silico assessment of absorption, distribution, metabolism, excretion, and toxicity (ADME/Tox) properties, as well as its integration with molecular docking- and pharmacology-oriented evaluation workflows. A structured literature search was conducted using PubMed, Scopus, Web of Science, and Google Scholar, covering publications from 2015 to 2025, with the final literature search completed on 15 December 2025 and employing keywords related to artificial intelligence, drug–target interactions, phytochemicals, and natural product drug discovery. Rather than reporting new experimental findings, this review critically analyzes existing computational strategies, databases, and workflows, highlighting their strengths, limitations, and translational relevance. Overall, AI-assisted DTI prediction is presented as a complementary approach that can guide experimental pharmacology and accelerate the rational development of phytoconstituent-based therapeutics, provided that computational predictions are supported by subsequent in-vitro and in-vivo validation. This review is narrative in scope and is based exclusively on previously published studies, without generating new experimental or computational data.

Keywords: Artificial Intelligence, Phytoconstituents, Drug–target interaction, Machine learning, Deep learning, Graph neural networks, Natural product drug discovery, Molecular docking, ADME/Toxicity prediction, Pharmacology-oriented drug discovery

Department of Pharmacology, Indore Mahavidyalaya, Indore, Madhya Pradesh, India

Correspondence to: Muskan Tomar, muskantomar808@gmail.com

Additional material is published online only. To view please visit the journal online.

Cite this as: Muskan Tomar. AI-Assisted Drug Discovery from Phytoconstituents: Predicting Drug–Target Interactions. Premier Journal of Science 2025;17:100262

DOI: https://doi.org/10.70389/PJS.100262

Peer Review:

Received: 4 December 2025

Last revised: 25 January 2026

Accepted: 26 January 2026

Version accepted: 6

Published: 16 February 2026

Ethical approval: This study did not involve human participants or sensitive data; therefore, ethical approval was not required

Consent: N/a

Funding: No industry funding

Conflicts of interest: N/a

Author contribution: Muskan Tomar – Conceptualization, Writing – original draft, review and editing

Introduction

Natural products, especially plant-derived phytoconstituents, have historically served as a cornerstone in pharmacology and drug discovery due to their remarkable chemical diversity and wide spectrum of biological activities.[1] Classical pharmacotherapy owes many of its most effective therapeutics, including anticancer, antimicrobial, and cardioprotective agents, to compounds isolated from nature's vast repertoire of secondary metabolites. Despite their promise, traditional natural product drug discovery faces significant challenges, including labour-intensive experimental screening, structural complexity, limited scalability, and difficulty in systematically identifying molecular targets relevant to pharmacological effects.[2] In recent years, artificial intelligence (AI) has emerged as a transformative force in early-stage drug discovery by enabling rapid analysis and prediction of drug–target interactions (DTIs) that would be infeasible through conventional methods alone. AI approaches, ranging from machine learning (ML) and deep learning (DL) to advanced architectures like graph neural networks (GNNs), are capable of extracting meaningful patterns from high-dimensional chemical and biological data, accelerating virtual screening, and improving the prioritization of bioactive phytochemicals for pharmacological evaluation.[3] These computational methodologies complement classical pharmacological paradigms by offering predictive insights into ADME (absorption, distribution, metabolism, excretion), toxicity, and off-target risks before costly in-vitro and in-vivo testing begins. As a result, AI-assisted DTI prediction is reshaping how researchers identify and optimize novel drug candidates from nature, bridging the gap between large phytochemical libraries and actionable pharmacological outcomes.[4] An overview of the AI-assisted framework for phytochemical drug discovery is illustrated in Figure 1.

This figure illustrates the integrated workflow for AI-assisted phytochemical drug discovery. Medicinal plants serve as the primary source of phytoconstituents, which are curated into phytochemical libraries and processed through AI models, including ML, DL, and GNNs. These models enable rapid prediction of DTIs, followed by in-silico ADME and toxicity filtering to prioritize drug-like and safe candidates. The shortlisted phytochemicals are subsequently subjected to pharmacological validation through in-vitro and in-vivo studies, facilitating efficient and translational natural product-based drug discovery.

Literature Search Strategy and Methodology

A comprehensive and systematic literature search was conducted to identify relevant studies focusing on AI-based DTI prediction involving phytochemicals and natural products. Major scientific databases, including PubMed, Scopus, Web of Science, and Google Scholar, were searched. The search strategy employed combinations of keywords such as "artificial intelligence", "machine learning", "deep learning", "graph neural networks", "drug–target interaction", "phytochemicals", "natural products", and "AI-assisted drug discovery".[5] Boolean operators (AND/OR) were used to refine the search. The search was restricted to articles published in English between 2015 and 2025, reflecting the rapid evolution of AI methodologies in

Guarantor: Muskan Tomar
Provenance and peer-review: Unsolicited and externally peer-reviewed
Data availability statement: N/a

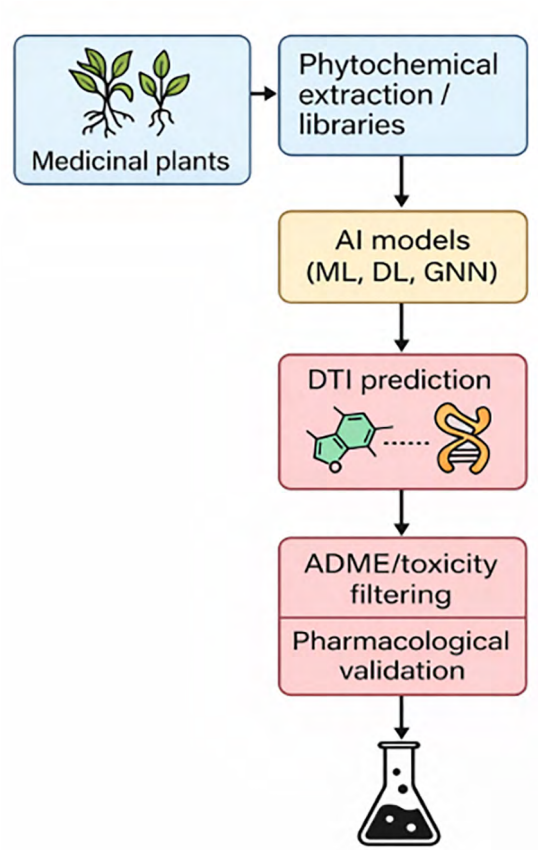

Fig 1 | AI-assisted phytochemical drug discovery framework (The workflow is based on literature published between 2015 and 2025, with the final search conducted on 15 December 2025.)

drug discovery. Reference lists of selected articles were also screened to identify additional relevant studies. This transparent search strategy ensured comprehensive coverage of recent and relevant advancements in AI-assisted phytochemical DTI prediction.[6] The literature search was conducted using the following search strings in various combinations: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "graph neural network") AND ("drug–target interaction" OR "DTI prediction") AND ("phytochemicals" OR "natural products" OR "plant-derived compounds").

The final literature search was completed on 15 December 2025. This manuscript is a narrative review and does not present any original experimental, computational, or proprietary analyses. All methodological descriptions, performance metrics, docking results, and simulation examples discussed are derived from and attributed to previously published studies.

Search Strings

The following exact search strings were executed:
Database-specific search strategy
PubMed
("artificial intelligence" OR "machine learning" OR "deep learning" OR "graph neural network")
AND ("drug–target interaction" OR "DTI prediction")
AND ("phytochemicals" OR "natural products" OR "plant-derived compounds")
Scopus
TITLE-ABS-KEY ("artificial intelligence" OR "machine learning" OR "deep learning" OR "graph neural network")
AND TITLE-ABS-KEY ("drug–target interaction" OR "DTI prediction")
AND TITLE-ABS-KEY ("phytochemicals" OR "natural products" OR "plant-derived compounds")
Web of Science
TS=("artificial intelligence" OR "machine learning" OR "deep learning" OR "graph neural network")
AND TS=("drug–target interaction" OR "DTI prediction")
AND TS=("phytochemicals" OR "natural products" OR "plant-derived compounds")
Google Scholar
("AI-assisted drug discovery" AND "drug–target interaction" AND phytochemicals)

Inclusion and Exclusion Criteria

Inclusion Criteria:

- Original research articles and review papers focusing on AI, ML, DL, or graph-based models for DTI prediction.
- Studies involving phytochemicals, natural products, or plant-derived compounds.
- Research utilizing validated drug–target databases or benchmark datasets.
- Studies reporting methodological details, model performance, or pharmacological relevance.[7]

Exclusion Criteria:

- Studies limited to classical molecular docking without AI-based prediction.
- Research focusing exclusively on synthetic small-molecule drugs.
- Editorials, conference abstracts, opinions, or non-peer-reviewed articles.
- Studies lacking sufficient methodological transparency or reproducible data.

Records retrieved from all databases were screened for relevance based on title and abstract. Duplicate records were removed before screening. Full-text articles were then assessed for eligibility according to predefined inclusion and exclusion criteria. Only peer-reviewed articles focusing on AI-based DTI prediction involving phytochemicals or natural products were included in the final analysis.

As this study is a narrative review, no formal quantitative risk-of-bias tool was applied. However, the quality of included studies was qualitatively assessed based on publication type, methodological transparency, dataset quality, and relevance to pharmacology-oriented drug discovery. Studies lacking sufficient methodological detail or biological relevance were excluded. The literature search and study selection process is summarized using a PRISMA-style flow diagram (Figure 2).

DOI: https://doi.org/10.70389/PJS.100262 | Premier Journal of Science 2025;XX:100262

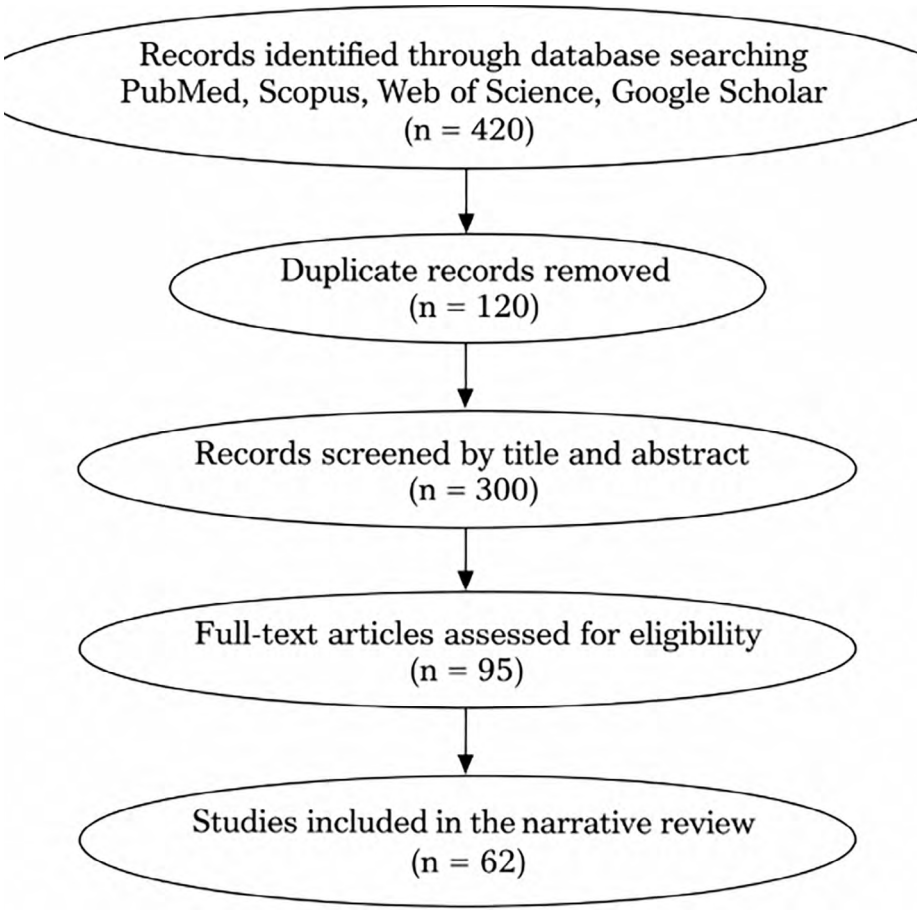

Fig 2 | PRISMA-style flow diagram illustrating the literature search, screening, eligibility assessment, and final inclusion of studies in this narrative review (2015–2025; final search date: 15 December 2025)

Study Screening and Selection Process

All retrieved records were screened by **a single reviewer (Muskan Tomar)** based on title and abstract. Duplicate records were removed before screening. Full-text articles were subsequently assessed for eligibility according to the predefined inclusion and exclusion criteria. As this study is a "Narrative Review", no independent second screener was applied. Any uncertainties regarding study relevance were resolved by careful re-evaluation of the full text and methodological details to ensure consistency and accuracy in study selection.

At the full-text screening stage, studies were excluded primarily due to the absence of AI-based DTI methods, lack of focus on phytochemicals or natural products, relied solely on classical molecular docking approaches, insufficient methodological transparency, or non-peer-reviewed publication status.

Phytoconstituents as Drug Leads

Phytoconstituents, the bioactive secondary metabolites produced by plants, represent one of the most valuable sources of drug leads in pharmacology. Owing to their vast chemical diversity and evolutionary optimization for biological interactions, the plant-derived compounds have historically contributed to the development of numerous clinically approved drugs.[8] Unlike pure synthetic molecules, phytoconstituents often possess complex scaffolds, multiple chiral centers, and diverse functional groups that enable selective and potent interactions with a wide range of molecular targets. From a pharmacological perspective, phytoconstituents exhibit broad therapeutic potential, including anti-inflammatory, antimicrobial, anticancer, neuroprotective, cardioprotective, and metabolic regulatory activities.[9] Many of these compounds act through multi-target mechanisms, offering advantages in complex diseases where modulation of a single target may be insufficient. This intrinsic poly-pharmacology makes phytochemicals attractive candidates for modern drug discovery strategies focused on systems-level therapeutic intervention.[10] Despite their promise, the translation of phytoconstituents into drug candidates has been limited by challenges such as low natural abundance, structural complexity, poor bioavailability, and difficulties in systematic target identification using conventional screening approaches. These limitations have slowed down the experimental validation and reduced the efficiency of traditional pharmacological workflows. Consequently, there is a growing need for computational and AI-assisted strategies to efficiently prioritize phytoconstituents with high drug-likeness and clear pharmacological relevance. Integrating the

AI with phytochemical research offers a rational pathway to unlock the full potential of natural products as drug leads and to streamline their progression from plant sources to clinically relevant therapeutics. Given this chemical and pharmacological complexity of phytoconstituents, computational and AI-based approaches are increasingly required to efficiently identify their molecular targets, which is discussed in the following section.

Major Classes and Chemical Diversities

Phytoconstituents are structurally diverse secondary metabolites produced by plants that serve as a rich reservoir of potential drug leads. Major classes include alkaloids (e.g., morphine, quinine), phenolics and polyphenols (e.g., flavonoids, tannins), terpenoids (e.g., limonoids, diterpenes), glycosides, saponins, anthraquinones, and alkenes, each exhibiting distinctive chemical scaffolds that can interact with a broad range of biological targets. The immense chemical diversity of these compounds arises from evolutionary pressures and biosynthetic complexities, offering unique stereochemistry, functional groups, and bioactive frameworks, often not present in synthetic libraries. Such diversity enhances the likelihood of finding specific interactions with pharmacological targets, making phytochemicals invaluable leads in drug discovery programs.[11,12] The major classes of phytoconstituents, their representative compounds, natural sources, and pharmacological activities are summarized in Table 1.

Known Pharmacological Activities

Phytoconstituents exhibit a wide range of pharmacological activities, including antioxidant, anti-inflammatory, antimicrobial, anticancer, cardioprotective, antidiabetic, neuroprotective, and immunomodulatory effects. Many clinically used drugs are either directly derived from or inspired by these natural products; examples include anticancer agents, analgesics, and antimalarials, highlighting their therapeutic relevance. Their complex structures often engage in multiple molecular targets or biological pathways, contributing to poly-pharmacology and synergy in multi-component extracts.[16–19]

Limitations in Conventional Screening

Despite their promise, conventional pharmacological screening of phytoconstituents faces several limitations. Traditional bioassay-guided fractionation and high-throughput screening are time-consuming, resource-intensive, and often hindered by the low abundance of active compounds in crude extracts. The complex mixture of components and variable composition due to environmental, seasonal, and genetic factors complicate reproducibility and standardization.[20–22] Additionally, isolating pure compounds from plant matrices and identifying their molecular targets through empirical methods alone can be challenging, leading to slow progression from discovery to drug candidate. These practical bottlenecks have historically limited the translation of many promising phytochemicals into clinically viable drugs.[23]

AI Approaches for DTI Prediction

Building on the need for efficient target identification, this section summarizes the key AI approaches used for DTI prediction in phytochemical drug discovery. AI has emerged as a powerful tool for predicting DTIs by enabling rapid and large-scale analysis of chemical and biological data. Unlike conventional experimental approaches, AI-based methods can efficiently handle the high structural diversity of phytoconstituents and the complexity of biological targets. By learning patterns from existing drug–target datasets, AI models facilitate virtual screening, target identification, and binding affinity prediction, thereby accelerating early-stage pharmacological research.[24] In natural product drug discovery, these approaches are particularly valuable for prioritizing bioactive phytochemicals and reducing the time, cost, and experimental burden associated with traditional screening methods. AI-driven DTI prediction typically integrates molecular representations of compounds with protein features to infer interaction likelihoods. ML algorithms utilize predefined molecular descriptors, while advanced DL and graph-based methods automatically learn meaningful representations directly from raw chemical structures and protein sequences.[25] These models not only enhance prediction accuracy, but also support pharmacology-oriented assessments such as poly-pharmacology, off-target effects, and safety profiling. Overall, AI approaches provide a robust computational framework that bridges phytochemical databases with experimental pharmacology, enabling more informed decision-making in natural product-based drug discovery. A comparative overview of AI approaches employed

Table 1 | Major Classes of Phytoconstituents and Their Pharmacological Significances[13–15]

Phytoconstituent Class	Representative Compound	Natural Source	Key Pharmacological Activities
Alkaloids	Morphine, quinine, berberine, vincristine	*Papaver somniferum*, *Cinchona* spp., *Berberis* spp.	Analgesic, antimalarial, anticancer, antimicrobial
Flavonoids	Quercetin, kaempferol, catechin, luteolin	Fruits, vegetables, tea, *Ginkgo biloba*	Antioxidant, anti-inflammatory, anticancer, cardioprotective
Phenolic acid	Caffeic acid, ferulic acid, gallic acid	Coffee, cereals, berries	Antioxidant, neuroprotective, anti-inflammatory
Terpenoids	Artemisinin, taxol, limonene	*Artemisia annua*, *Taxus* spp., citrus fruits	Anticancer, antimalarial, anti-inflammatory
Glycosides	Digoxin, salicin	*Digitalis* spp., *Salix* spp.	Cardioprotective, analgesic, anti-inflammatory
Saponins	Ginsenosides, diosgenin	*Panax ginseng*, *Dioscorea* spp.	Immunomodulatory, anticancer, cholesterol-lowering
Tannins	Ellagitannins, proanthocyanidins	Tea, grapes, berries	Antioxidant, antimicrobial, anticancer
Lignans	Secoisolariciresinol, podophyllotoxin	Flaxseed, *Podophyllum* spp.	Anticancer, antiviral, estrogenic modulation

DOI: https://doi.org/10.70389/PJS.100262 | Premier Journal of Science 2025;XX:100262

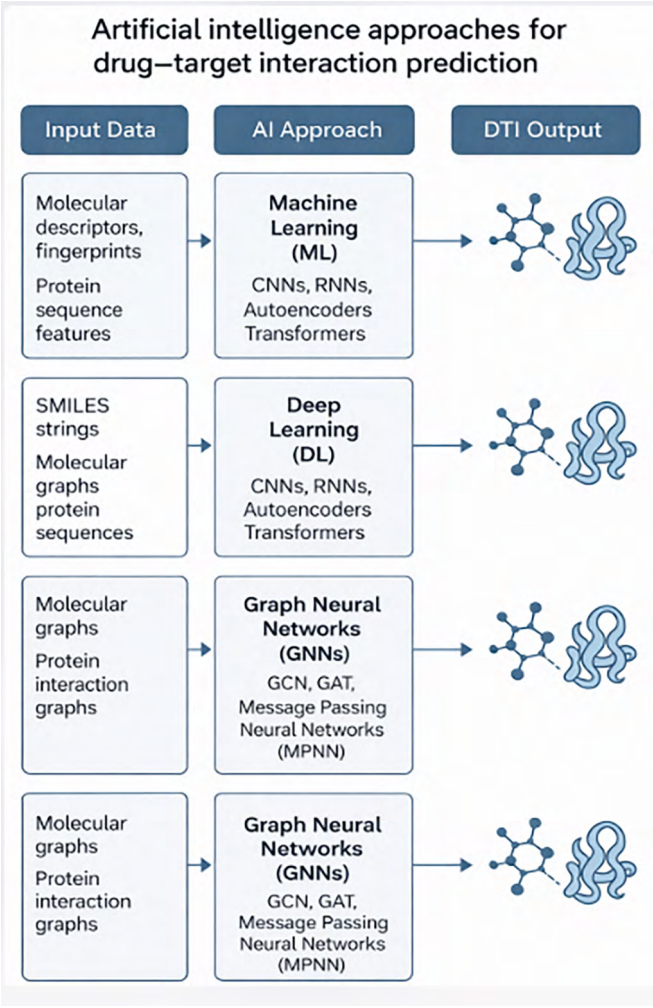

Fig 3 | Artificial intelligence approaches for drug–target interaction prediction

for DTI prediction, along with their strengths and limitations, is presented in Table 2. Different artificial intelligence models employed for DTI prediction are schematically represented in Figure 3. While ML, DL, and GNN models have demonstrated a strong performance in DTI prediction, their applicability to phytochemicals varies considerably. Classical ML models offer better interpretability, but struggle with complex natural product scaffolds. DL and graph-based models capture structural complexity more effectively; however, they are data-intensive and often trained on datasets dominated by synthetic compounds, limiting model portability to underrepresented phytochemicals.

References

1 Nasim N, Sandeep IS, Mohanty S. Plant-derived natural products for drug discovery: Current approaches and prospects. Nucleus. 2022;65(3):399–411. https://doi.org/10.1007/s13237-022-00405-3

2 Simoben CV, Babiaka SB, Moumbock AF, Namba-Nzanguim CT, Eni DB, Medina-Franco JL, et al. Challenges in natural product-based drug discovery assisted with in silico-based methods. RSC Adv. 2023;13(45):31578–94. https://doi.org/10.1039/D3RA06831E

3 Othman ZK, Ahmed MM, Kasimieh O, Musa SS, Francesco Branda F, Cue EG, et al. Artificial intelligence for natural product drug discovery and development: Current landscape, applications, and future directions. Intell-Based Med. 2025;100316. https://doi.org/10.1016/j.ibmed.2025.100316

4 Zhang J, Li H, Zhang Y, Huang J, Ren L, Zhang C, et al. Computational toxicology in drug discovery: Applications of artificial intelligence in ADMET and toxicity prediction. Brief Bioinform. 2025;26(5):bbaf533. https://doi.org/10.1093/bib/bbaf533

5 Bramer WM, De Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: An efficient and complete method to develop literature searches. JMLA. 2018;106(4):531–41. https://doi.org/10.5195/jmla.2018.283

6 Wang Q, Sun B, Yi Y, Velkov T, Shen J, Dai C, et al. Progress of AI-driven drug–target interaction prediction and lead optimization. Int J Mol Sci. 2025;26(20):10037. https://doi.org/10.3390/ijms262010037

7 Gangwal A, Lavecchia A. Artificial intelligence in natural product drug discovery: Current applications and future perspectives. J Med Chem. 2025;68(4):3948–69. https://doi.org/10.1021/acs.jmedchem.4c01257

8 Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. J Nat Prod. 2007;70(3):461–77. https://doi.org/10.1021/np068054v

9 Babalola OO, Bridget K, Oyubu G, Waheed SA, Ajiboye SA, Fakayode AE, et al. Integrating phytochemicals and in silico methods for modern drug discovery: A comprehensive review. Discov Chem. 2025;2(1):297. https://doi.org/10.1007/s44371-025-00373-y

10 Hopkins AL. Network pharmacology: The next paradigm in drug discovery. Nat Chem Biol. 2008;4(11):682–90. https://doi.org/10.1038/nchembio.118

11 Ali S, Khalil AA, Akhtar MS, Amin A, Zaman W. Comprehensive insights into natural bioactive compounds: From chemical diversity and mechanisms to biotechnological innovations and applications. ChemistryOpen. 2025:e202500469. https://doi.org/10.1002/open.202500469

12 Mera IG, Falconí DG, Córdova VM. Secondary metabolites in plants: main classes, phytochemical analysis and pharmacological activities. Rev Bionatura. 2019;4(4). https://doi.org/10.21931/RB/2019.04.04.11

13 Kumar A, Nirmal P, Kumar M, Jose A, Tomar V, Oz E, et al. Major phytochemicals: Recent advances in health benefits and extraction methods. Molecules. 2023;28(2):887. https://doi.org/10.3390/molecules28020887

14 Rodríguez-Negrete EV, Morales-González Á, Madrigal-Santillán EO, Sánchez-Reyes K, Álvarez-González I, Madrigal-Bujaidar E, et al. Phytochemicals and their usefulness in the maintenance of health. Plants. 2024;13(4):523. https://doi.org/10.3390/plants13040523

15 Elshafie HS, Camele I, Mohamed AA. A comprehensive review on the biological, agricultural and pharmaceutical properties of secondary metabolites of plant origin. Int J Mol Sci. 2023;24(4):3266. https://doi.org/10.3390/ijms24043266

16 Riaz M, Khalid R, Afzal M, Anjum F, Fatima H, Zia S, et al. Phytobioactive compounds as therapeutic agents for human diseases: A review. Food Sci Nutr. 2023;11(6):2500–29. https://doi.org/10.1002/fsn3.3308

17 Chunarkar-Patil P, Kaleem M, Mishra R, Ray S, Ahmad A, Verma D, et al. Anticancer drug discovery based on natural products: from computational approaches to clinical studies. Biomedicines. 2024;12(1):201. https://doi.org/10.3390/biomedicines12010201

18 Ansari P, Reberio AD, Ansari NJ, Kumar S, Khan JT, Chowdhury S, et al. Therapeutic potential of medicinal plants and their phytoconstituents in diabetes, cancer, infections, cardiovascular diseases, inflammation and gastrointestinal disorders. Biomedicines. 2025;13(2):454. https://doi.org/10.3390/biomedicines13020454

19 Bose PA, Sohag MM, Rabbee MF, Zamee TM, Kona J, Elora B, et al. Pharmacological overview of bioactive natural products from *Gynura procumbens* (Lour.) Merr. Plants. 2025;14(17):2714. https://doi.org/10.3390/plants14172714

20 Mashele SS. Phytochemicals as multifunctional agents: Antimicrobial, enzyme inhibitory, and wound-healing potentials in the era of drug resistance. Preprint. https://doi.org/10.20944/preprints202510.1183.v1

21 Altemimi A, Lakhssassi N, Baharlouei A, Watson DG, Lightfoot DA. Phytochemicals: extraction, isolation, and identification of bioactive compounds from plant extracts. Plants. 2017;6(4):42. https://doi.org/10.3390/plants6040042

22 Weller MG. A unifying review of bioassay-guided fractionation, effect-directed analysis and related techniques. Sensors. 2012;12(7):9181–209. https://doi.org/10.3390/s120709181

23 Atanasov AG, Waltenberger B, Pferschy-Wenzig EM, Linder T, Wawrosch C, Uhrin P, et al. Discovery and resupply of pharmacologically active plant-derived natural products: a review. Biotechnol Adv. 2015;33(8):1582–614. https://doi.org/10.1016/j.biotechadv.2015.08.001

24 Liao Q, Zhang Y, Chu Y, Ding Y, Liu Z, Zhao X, et al. Application of artificial intelligence in drug–target interaction prediction: a review. NPJ Biomed Innov. 2025;2(1):1. https://doi.org/10.1038/s44385-024-00003-9

25 Talukder MA, Kazi M, Alazab A. Predicting drug–target interactions using machine learning with improved data balancing and feature engineering. Sci Rep. 2025;15:19495. https://doi.org/10.1038/s41598-025-03932-6

List of Abbreviations

ADME – Absorption, Distribution, Metabolism and Excretion

AI – Artificial Intelligence

BBB – Blood–Brain Barrier

Bo5 – Rule of Five (Lipinski)

CNN – Convolutional Neural Network

CNS – Central Nervous System

CYP – Cytochrome P450

DL – Deep Learning

DTI – Drug–Target Interaction

GAT – Graph Attention Network

GCN – Graph Convolutional Network

GNN – Graph Neural Network

GPCR – G-Protein Coupled Receptor

hERG – Human Ether-à-go-go-Related Gene

IMPAT – Indian Medicinal Plants, Phytochemistry and Therapeutics

k-NN – k-Nearest Neighbours

MD – Molecular Dynamics

ML – Machine Learning

MPNN – Message Passing Neural Network

NPASS – Natural Product Activity and Species Source Database

QSAR – Quantitative Structure–Activity Relationship

RF – Random Forest

RNN – Recurrent Neural Network

SMILES – Simplified Molecular Input Line Entry System

SVM – Support Vector Machine

TCMSP – Traditional Chinese Medicine Systems Pharmacology Database

TPSA – Topological Polar Surface Area

TTD – Therapeutic Target Database

UNPD – Universal Natural Product Database

XAI – Explainable Artificial Intelligence