Anticancer Phytochemicals in the Age of AI: Target Vulnerability, Tumor Pathway Dependency, Molecular Selectivity, and Safety-Aware Candidate Prioritization

Muhammad Islam[1]*, Nasrin Akhter[1], Shamsul Alam[2]

[1]Department of AI for Ayurvedic and Unani Medicine Informatics, Faculty of Pharmacy, University of Dhaka, Dhaka, Bangladesh.
[2]Department of Machine Learning for Herbal Drug Interactions, Faculty of Pharmacy, Rajshahi University, Rajshahi, Bangladesh.

ABSTRACT

Anticancer phytochemicals occupy an important but methodologically challenging position in contemporary cancer pharmacology because plant-derived compounds may show tumor-relevant biological activity while also carrying uncertainty regarding target engagement, pathway context, selectivity, safety, exposure, and translational readiness. This Original Conceptual Framework Article proposes an AI-assisted framework for prioritizing anticancer phytochemical candidates as research hypotheses rather than therapeutic recommendations. The framework integrates phytochemical identity, compound class, natural product source, chemical characterization, cancer-context definition, target vulnerability, target confidence, oncogenic pathway relevance, tumor pathway dependency, molecular-signature evidence, molecular selectivity, tumor-normal contrast, ADMET evidence, toxicity signals, off-target liability, exposure plausibility, evidence provenance, evidence grading, uncertainty, and validation gaps. Its central logic is that AI can assist chemical annotation, target prediction, pathway inference, molecular-signature interpretation, safety-aware filtering, and candidate ranking only when outputs are interpreted through tumor biology and experimental validation needs. The proposed framework distinguishes AI-predicted mechanisms from experimentally supported target evidence, cytotoxicity from tumor-selective anticancer relevance, and preclinical evidence from clinical claims. The main contribution is a structured prioritization logic that links computational screening to oncology-informed interpretation, safety constraints, and translational boundaries. The framework is intended to support responsible research prioritization, clearer reporting, and validation planning without implying clinical anticancer efficacy.

Key Words: *Anticancer phytochemicals, Artificial intelligence, Target vulnerability, Tumor pathway dependency, Molecular selectivity, Safety-aware screening*

eIJPPR 2025; 15(6):59-70

HOW TO CITE THIS ARTICLE: Islam M, Akhter N, Alam S. Anticancer Phytochemicals in the Age of AI: Target Vulnerability, Tumor Pathway Dependency, Molecular Selectivity, and Safety-Aware Candidate Prioritization. Int J Pharm Phytopharmacol Res. 2025;15(6):59-70. https://doi.org/10.51847/KpnW8CPaFF

INTRODUCTION

Natural products remain important starting points for drug discovery because they provide chemically diverse structures and biologically active scaffolds that can inspire pharmacological investigation, including oncology-oriented research. However, the relevance of a natural product source or phytochemical identity to cancer pharmacology cannot be assumed from origin alone; anticancer interpretation requires evidence that connects chemical identity to tumor-relevant biology, target

Corresponding author: Muhammad Islam
Address: Department of AI for Ayurvedic and Unani Medicine Informatics, Faculty of Pharmacy, University of Dhaka, Dhaka, Bangladesh.
E-mail: ✉ muhammad.islam@du.ac.bd
Received: 20 August 2025; **Revised:** 28 November 2025; **Accepted:** 03 December 2025

This is an **open access** journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

engagement, pathway context, selectivity, safety, and validation readiness [1].

Contemporary natural product discovery increasingly requires integration of chemistry, biology, pharmacology, and translational interpretation. For anticancer phytochemicals, this means that compound annotation, source information, bioactivity evidence, target hypotheses, and developability considerations should be examined together rather than treated as isolated evidence fragments [2]. Such integration is particularly important because tumor biology is context-dependent, and a compound that appears active in one experimental setting may not support a broader anticancer claim.

Machine learning and related AI methods can assist drug discovery by organizing chemical, biological, and pharmacological data into predictive or prioritization workflows. In anticancer phytochemical research, these tools may support target prediction, activity modeling, pathway inference, molecular-signature analysis, candidate triage, and evidence integration, but their outputs remain dependent on data quality, model assumptions, and validation design [3].

The aim of this article is to develop an original AI-assisted conceptual framework for prioritizing anticancer phytochemicals while preserving clear evidence boundaries. The framework is designed to treat AI-predicted anticancer potential as a research hypothesis that requires tumor-context validation, target and pathway evidence, molecular selectivity assessment, safety-aware screening, exposure interpretation, and translational caution before any stronger claim can be justified [4].

Anticancer phytochemicals and AI discovery

Anticancer phytochemicals can be understood as plant-derived compounds or compound classes investigated for tumor-relevant biological effects, including effects on proliferation, survival signaling, cell-cycle regulation, apoptosis-related processes, migration-related biology, or immune-relevant mechanisms where evidence is available. Deep learning can contribute to this research space by representing molecular structures, biological signatures, and pharmacological patterns in ways that support discovery prioritization, although such models do not themselves establish anticancer efficacy [5].

AI-assisted natural product discovery is especially relevant because phytochemicals often present challenges in chemical diversity, annotation, stereochemical complexity, dereplication, source variability, and bioactivity attribution. AI can support chemical-space exploration, target hypothesis generation, similarity-based inference, and prioritization of compounds for follow-up, but these functions must remain connected to reliable chemical characterization and transparent evidence provenance [6].

Natural product source information can enrich candidate interpretation when biosynthetic context, compound class, and chemical annotation are clearly recorded. Nevertheless, source and biosynthetic context should be treated as discovery metadata rather than oncology evidence, because they do not establish tumor relevance, target vulnerability, pathway dependency, or tumor-normal selectivity by themselves [7].

A responsible AI-assisted anticancer phytochemical workflow therefore requires a layered structure in which predictive outputs are interpreted alongside experimental and translational evidence. Chemical annotation, predicted targets, pathway links, safety estimates, and candidate-ranking signals can support prioritization only when uncertainty, evidence strength, and validation requirements are explicitly retained [8].

Target vulnerability and pathway dependency

Target vulnerability refers to the degree to which a molecular target is functionally important in a defined tumor context. In anticancer phytochemical prioritization, a predicted or experimentally proposed target should not be considered meaningful solely because a compound may interact with it; target relevance should be strengthened by functional evidence showing that the target contributes to tumor survival, proliferation, adaptation, or another cancer-relevant phenotype in the relevant context [9].

Tumor pathway dependency adds a second interpretive layer by asking whether a tumor context depends on a pathway or network state that the candidate is hypothesized to modulate. AI-derived dependency maps can help prioritize context-specific vulnerability hypotheses, but they require careful biological interpretation because predicted dependency does not automatically establish response to a phytochemical candidate [10]. Synthetic-lethality prediction can further support vulnerability hypothesis generation, especially where tumor-specific genetic or pathway relationships are implicated, but such predictions require experimental confirmation before mechanistic or translational claims are made [11].

Oncogenic pathway relevance should be distinguished from pathway alteration alone. A pathway may be altered, activated, suppressed, or rewired in cancer datasets without being therapeutically actionable in a given tumor context, and candidate prioritization should therefore examine the direction, context, and functional relevance of pathway evidence [12]. Cancer systems biology also cautions against reducing anticancer interpretation to a single cytotoxic phenotype, because tumor behavior emerges from interacting hallmarks, adaptive states, microenvironmental influences, and context-dependent vulnerabilities [13].

ISSN (Online) 2249-6084 (Print) 2250-1029

Target confidence should integrate multiple evidence streams, including predicted target links, disease association, functional dependency evidence, pathway relevance, molecular signatures, and, where available, experimental target identification. Evidence-integration platforms support this principle by showing that target prioritization is strengthened when diverse evidence types are made explicit rather than collapsed into a single unsupported claim [14]. Dependency-map resources can support targetable vulnerability hypotheses, but the framework proposed here treats such evidence as one component of tumor-context reasoning rather than proof that a phytochemical candidate has validated anticancer activity [15].

Table 1 summarises target vulnerability and tumor pathway dependency features relevant to AI-assisted anticancer phytochemical prioritization.

Table 1. Target Vulnerability and Tumor Pathway Dependency Features for AI-Assisted Anticancer Phytochemical Prioritization: Tumor Context, Target Confidence, Oncogenic Pathway Relevance, Molecular Signatures, Selectivity Needs, Interpretation Boundaries, and Validation Requirements

Framework feature	Core oncology question	Required evidence	Potential prioritization signal	Risk of overinterpretation	Selectivity or safety requirement	Validation need	Framework role
Cancer or tumor context definition	Which tumor context is being addressed?	Cancer type, subtype, model context, molecular background, and evidence source	Candidate hypothesis is tied to a defined tumor setting	Treating broad cancer language as context-specific evidence	Avoid generalizing across tumor types without support	Validate in context-relevant tumor models	Anchors all downstream prioritization logic
Tumor molecular context	What molecular features define the tumor setting?	Genomic, transcriptomic, pathway, or dependency evidence	Candidate aligns with tumor-relevant molecular features	Assuming molecular alteration predicts response	Require comparison with non-dependent or normal contexts	Use molecularly annotated models	Links candidate logic to tumor biology
Target vulnerability	Is the target functionally important in the tumor context?	Functional dependency, perturbation, or orthogonal vulnerability evidence	Target appears necessary for tumor-relevant phenotype	Treating target presence as vulnerability	Assess normal-cell dependency and toxicity risk	Functional perturbation and rescue studies	Separates target prediction from target relevance
Target confidence	How reliable is the compound-target relationship?	Experimental target engagement, chemical proteomics, binding data, or high-confidence prediction	Multiple evidence streams support the same target hypothesis	Treating a single predicted target as confirmed	Screen for off-target and polypharmacology risks	Orthogonal target validation	Grades predicted versus supported mechanisms
Oncogenic pathway relevance	Is the target linked to cancer-relevant signaling?	Pathway alteration, signaling, or systems-biology evidence	Candidate target lies within a tumor-relevant pathway	Assuming pathway association means therapeutic value	Consider normal tissue function of the pathway	Pathway perturbation assays	Provides pathway-level plausibility
Tumor pathway dependency	Does the tumor rely on the implicated pathway?	Dependency screens, perturbation data, or pathway activity signatures	Strong dependency signal in a defined tumor context	Equating pathway alteration with dependency	Require tumor-normal contrast	Test pathway modulation in context-matched models	Distinguishes alteration from functional dependency
Pathway alteration context	Is the pathway activated, suppressed, mutated, or rewired?	Genomic, transcriptomic, proteomic, or phosphoproteomic evidence	Alteration is consistent with candidate mechanism hypothesis	Inferring response direction from alteration alone	Evaluate pathway role in healthy cells	Confirm pathway modulation after candidate exposure	Clarifies pathway interpretation

ISSN (Online) 2249-6084 (Print) 2250-1029

Molecular-signature alignment	Does candidate-associated biology align with tumor-relevant signatures?	Transcriptomic, proteomic, or molecular-signature evidence	Signature supports a plausible tumor-relevant hypothesis	Treating signature similarity as mechanism proof	Exclude nonspecific stress or toxicity signatures	Validate signature changes experimentally	Links AI signature logic to mechanism hypotheses
Tumor-normal contrast	Is the candidate hypothesis more tumor-relevant than normal-cell-relevant?	Tumor-normal viability, molecular, or functional comparisons	Evidence suggests a potential tumor-selective window	Treating cytotoxicity as selectivity	Require normal-cell or tissue-relevant comparison	Test selectivity across tumor and normal models	Controls overstatement of anticancer selectivity
Resistance relevance where supported	Does the target or pathway relate to adaptive escape or resistance?	Resistance models, pathway rewiring evidence, or supported resistance mechanisms	Candidate hypothesis addresses resistance-relevant biology	Assuming combination benefit without resistance data	Assess combination toxicity and antagonism	Validate in resistant and non-resistant models	Adds an optional advanced prioritization layer
Mechanistic plausibility	Do target, pathway, signature, and phenotype evidence cohere?	Concordance across target, pathway, omics, and phenotype evidence	Multiple layers support the same hypothesis	Building causality from association alone	Include safety and off-target review	Orthogonal mechanistic validation	Integrates evidence into cautious hypotheses
Validation gap	What evidence is missing before stronger claims can be made?	Explicit inventory of absent target, selectivity, ADMET, or clinical evidence	Clear next validation step can be defined	Hiding uncertainty in ranking language	Require safety and selectivity testing before translational claims	Stage-specific validation planning	Prevents premature therapeutic interpretation

Molecular selectivity and safety-aware screening

Molecular selectivity is a central constraint in anticancer phytochemical prioritization because a candidate may show a tumor-relevant molecular signature without demonstrating selective action on tumor biology. AI-based interpretation of gene-expression signatures can help connect candidate molecules to disease-associated transcriptional states, but such alignment should be read as a hypothesis-generating signal rather than proof of tumor-selective mechanism or clinical relevance [16].

Predicted target profiles can help identify plausible intended targets and possible off-target liabilities, but they should be interpreted as confidence-weighted hypotheses. Target-prediction tools can assist early triage of small molecules, including phytochemical candidates, yet predicted target lists require orthogonal evidence before they are described as supported mechanisms [17]. Exposure plausibility is equally important, because a candidate with an attractive target or pathway hypothesis may have limited prioritization value if predicted pharmacokinetic properties, drug-likeness features, or medicinal chemistry constraints make biologically meaningful exposure unlikely [18].

Safety-aware screening should treat toxicity and ADMET evidence as core prioritization dimensions rather than late-stage annotations. Computational toxicity prediction can flag chemical risk signals and support early caution, but it cannot substitute for experimental safety assessment or justify claims that a natural compound is safe in an oncology context [19]. Integrated ADMET prediction can assist early evaluation of absorption, distribution, metabolism, excretion, and toxicity-related concerns, but predicted properties must remain part of a broader evidence-grading process that includes uncertainty and validation needs [20].

Off-target liability is particularly relevant for phytochemical candidates because polypharmacology may contribute to either mechanistic breadth or misleading nonspecific activity. Computational off-target prediction can help identify unintended interactions that may affect selectivity, toxicity, or mechanism interpretation [21]. Screening strategies for off-target liabilities further support

the need to treat safety-aware filtering as a required layer in candidate prioritization rather than an optional confirmatory step after ranking [22].

Figure 1 illustrates how target vulnerability, tumor pathway dependency, molecular selectivity, and safety filters can be connected during AI-assisted anticancer phytochemical screening.

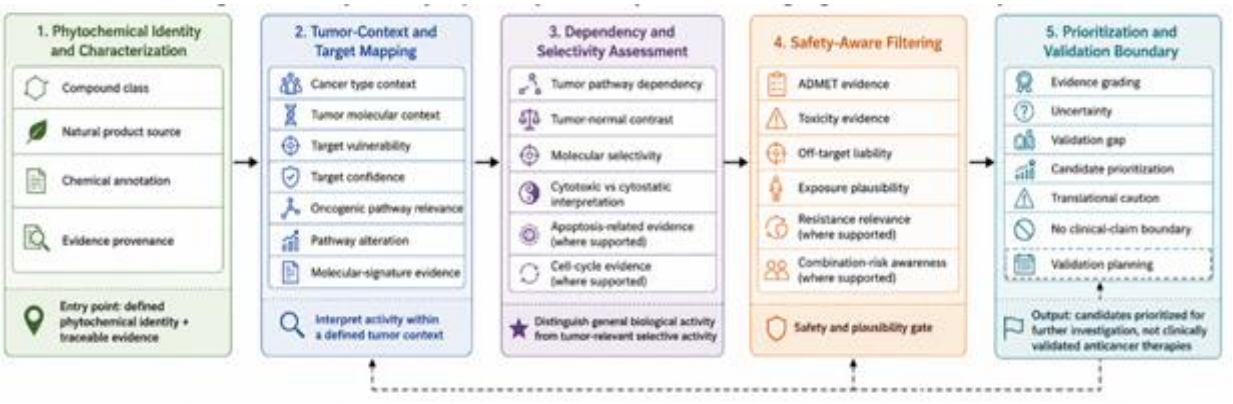

Figure 1. Target Vulnerability, Pathway Dependency, and Safety-Aware Screening Logic for Anticancer Phytochemicals

Alt-text description

A conceptual screening logic diagram showing anticancer phytochemical candidates evaluated through cancer context, target vulnerability, pathway dependency, molecular signatures, selectivity, tumor-normal contrast, ADMET evidence, toxicity signals, off-target risk, uncertainty, and validation needs.

Candidate prioritization logic

Candidate prioritization in anticancer phytochemical discovery should be treated as an evidence-aware process rather than a simple ranking exercise. Omics-guided natural product workflows can support prioritization by linking chemical annotation, bioactivity evidence, molecular profiles, and evidence provenance, but the resulting candidates should be framed as research hypotheses whose value depends on the strength and coherence of supporting evidence [23].

Graph-based machine learning offers a useful conceptual model for integrating heterogeneous relationships among compounds, targets, pathways, disease contexts, safety features, and validation evidence. In the framework proposed here, graph-style reasoning is valuable because it can preserve relationships across evidence layers, but candidate prioritization should retain uncertainty rather than converting network proximity or inferred association into a claim of anticancer efficacy [24].

Cancer drug-repurposing research illustrates how computational prioritization can connect molecular context, candidate activity, and translational hypotheses, but this logic must be adapted cautiously for phytochemicals. A repurposing-like signal may indicate that a compound deserves further study in a tumor-relevant context, yet it should not imply that the compound is clinically active, clinically safe, or suitable for patient use [25]. Interpretable anticancer sensitivity models further support the need for transparent prioritization, because model outputs become more useful when the features driving candidate selection can be inspected and linked to biological uncertainty [26].

Drug-response and synergy models can inform prioritization by connecting molecular features of cancer models to response hypotheses, but these predictions require careful dataset interpretation, external validation, and selectivity-aware follow-up [27]. Combination or synergy hypotheses are especially sensitive to overinterpretation because apparent complementary activity may also increase toxicity, off-target liability, or context-specific uncertainty [28]. Multi-action molecular design based on differential gene-expression information can support the hypothesis that a candidate may influence more than one tumor-relevant biological process, but such design logic must remain constrained by molecular-signature evidence, safety interpretation, and validation readiness [29].

Table 2 maps safety-aware candidate prioritization criteria for anticancer phytochemical discovery.

Table 2. Safety-Aware Candidate Prioritization Criteria for Anticancer Phytochemicals: Chemical Characterization, Cancer Relevance, Target Vulnerability, Pathway Dependency, Molecular Selectivity, Tumor-Normal Contrast, ADMET Evidence, Toxicity Signals, Evidence Strength, and Validation Readiness

Prioritization criterion	Core question	Required evidence	High-priority signal	Caution signal	Safety or selectivity concern	Validation requirement	Decision boundary
Phytochemical identity	Is the candidate clearly identified?	Compound name, structure, stereochemical detail where relevant, source, and purity information	Structurally defined compound or well-characterized fraction	Ambiguous extract, unclear constituent, or poorly annotated compound	Unknown constituents may drive activity or toxicity	Chemical verification and reproducibility checks	Do not prioritize strongly without identity confidence
Chemical characterization	Is chemical evidence sufficient for reproducible study?	Validated annotation, chromatographic or spectral evidence, and batch traceability	Reproducible characterization supports follow-up	Weak annotation or source variability	Misattribution of bioactivity	Analytical confirmation before mechanistic testing	Candidate remains low-readiness if chemistry is uncertain
Cancer relevance	Is evidence linked to a defined cancer context?	Tumor context, model system, molecular background, and assay type	Context-specific tumor-relevant evidence	Generic anticancer wording without context	Overgeneralized oncology interpretation	Context-matched tumor model testing	No broad anticancer claim from narrow evidence
Target vulnerability	Is the target functionally relevant in the tumor context?	Dependency, perturbation, target engagement, or vulnerability evidence	Target appears functionally important in relevant tumor biology	Target prediction or expression alone	Normal-cell dependency may create toxicity risk	Functional validation and target engagement testing	Target prediction alone is insufficient
Pathway dependency	Does the tumor context rely on the implicated pathway?	Pathway activity, pathway alteration, dependency, or perturbation data	Dependency aligns with candidate mechanism hypothesis	Pathway alteration without functional dependency	Essential normal pathway activity may limit selectivity	Pathway modulation and dependency validation	Pathway link remains hypothesis-level unless validated
Molecular selectivity	Is the candidate plausibly selective for intended molecular effects?	Target profile, off-target screen, binding evidence, or selectivity hypothesis	Consistent intended-target evidence with manageable liabilities	Broad promiscuity or poorly explained polypharmacology	Misleading mechanism assignment or toxicity	Orthogonal target and off-target profiling	Predicted selectivity cannot be treated as proven
Tumor-normal selectivity	Is activity more tumor-relevant than normal-cell-relevant?	Tumor-normal comparator data or tissue-relevant normal-cell evidence	Potential tumor-selective response window	Tumor-cell cytotoxicity without normal comparator	General cytotoxicity may be misread as anticancer selectivity	Matched tumor-normal comparison	Cytotoxicity alone does not establish tumor selectivity
ADMET evidence	Are pharmacokinetic and developability concerns considered?	Absorption, distribution, metabolism, excretion, solubility, permeability, transporter, or stability evidence	Plausible exposure and developability profile	Poor predicted exposure or high property uncertainty	Unrealistic concentration assumptions	Experimental ADME follow-up where appropriate	No translational claim without exposure plausibility
Toxicity evidence	Are toxicity risks actively screened?	Toxicity prediction, normal-cell assays, organ-	No major early risk signal and clear	Toxicity alert or missing safety information	Natural origin may be falsely equated with safety	Experimental safety assessment	Natural compounds are not assumed safe

		risk alerts, or safety literature	uncertainty statement				
Off-target liability	Could unintended interactions explain activity or risk?	Off-target prediction, counterscreening, broad profiling, or liability review	Manageable off-target profile	High-risk liability or unknown profile	Off-target effects may drive toxicity or false mechanism	Off-target validation and counterscreens	Off-target uncertainty lowers prioritization readiness
Exposure plausibility	Are active concentrations biologically plausible?	Concentration-response evidence, solubility, stability, exposure prediction, or pharmacokinetic data	Activity occurs under plausible experimental conditions	Activity only at nonspecific high exposure	Nonspecific stress or assay artifact	Exposure-aware validation	In vitro activity does not imply achievable exposure
Evidence provenance	Is each evidence layer traceable?	Source, dataset, assay, model, computational method, and evidence type	Transparent evidence chain	Opaque dataset or unclear assay provenance	Irreproducible prioritization	Evidence audit and reporting	Opaque evidence cannot support high confidence
Evidence strength	How coherent and reliable is the evidence?	Graded computational, in vitro, preclinical, and clinical evidence where available	Concordant evidence across layers	Single weak evidence stream	Overweighting low-confidence data	Evidence grading and replication	Ranking must reflect evidence quality
Uncertainty	What remains unresolved?	Explicit uncertainty statement and gap inventory	Gaps are defined and testable	Ranking hides missing evidence	False precision in candidate prioritization	Sensitivity analysis and validation planning	Uncertainty must remain visible
Validation readiness	What follow-up is justified?	Defined assays, target tests, pathway studies, selectivity evaluation, safety checks	Clear next validation pathway	No feasible validation plan	Resource use may be poorly directed	Stage-appropriate validation plan	Candidate is prioritized for research, not treatment
Translational relevance	Is there a plausible route to stronger evidence?	Target, pathway, selectivity, safety, exposure, and model relevance	Evidence supports staged translational investigation	Major missing safety or exposure evidence	Premature clinical implication	Preclinical and clinical evidence where appropriate	Translational relevance is not therapeutic proof
No clinical-claim boundary	Are unsupported patient-facing claims avoided?	Explicit separation of prediction, preclinical evidence, and clinical evidence	Candidate framed as a research hypothesis	Language implies treatment, prevention, or clinical efficacy	Medical misinformation risk	Clinical validation before clinical claims	No treatment or prevention recommendation

Proposed conceptual framework

The proposed conceptual framework is designed to connect phytochemical evidence with oncology-informed interpretation and safety-aware prioritization. Its first stage records phytochemical identity, compound class, natural product source, chemical characterization, and evidence provenance. This stage is necessary because anticancer interpretation becomes unstable when chemical identity is uncertain, when source information is incomplete, or when biological activity is attributed to a poorly characterized extract or compound class. Natural products may interact with cancer-relevant signaling pathways, but pathway association must be interpreted as mechanistic plausibility rather than evidence of validated anticancer efficacy [30].

The second stage evaluates target confidence by distinguishing AI-predicted targets from experimentally supported targets. Experimental target-identification approaches, including chemical proteomics, can strengthen mechanistic confidence when they connect a natural product candidate to a molecular target in a biologically relevant system [31]. In the conceptual framework, target confidence is therefore graded across predicted association, orthogonal computational support, biochemical evidence, cellular target engagement, and context-relevant functional evidence.

The third stage links target vulnerability and tumor pathway dependency to molecular selectivity and safety-aware screening. Computational approaches to anticancer natural product discovery can help connect compounds, targets, pathways, and downstream validation priorities, but staged interpretation is required so that target prediction, docking-style inference, pathway mapping, and preclinical evidence do not collapse into premature clinical claims [32]. The framework therefore treats computational prioritization as an upstream filter that must be followed by tumor-context validation, tumor-normal comparison, exposure assessment, toxicity review, and clear separation between preclinical and clinical evidence.

The fourth stage defines translational decision boundaries. Clinical and epidemiological evidence concerning phytochemicals in cancer-related contexts must be interpreted separately from computational predictions, mechanistic studies, and preclinical observations [33]. Accordingly, the framework does not classify a candidate as clinically effective; instead, it classifies the candidate's research readiness according to evidence provenance, target confidence, pathway dependency support, selectivity evidence, safety signals, exposure plausibility, uncertainty, and validation gaps.

The fifth stage integrates source, chemical, mechanistic, and validation evidence into a safety-aware candidate-prioritization decision. Phytochemical source and biosynthetic identity can support reproducibility and discovery interpretation when properly documented [34]. Molecular target evidence can further support mechanistic plausibility, but target claims must be separated into predicted, experimentally supported, functionally validated, and clinically supported categories so that the framework does not overstate anticancer relevance [35].

Table 3 presents the proposed conceptual framework for AI-assisted anticancer phytochemical prioritization.

Table 3. Proposed Conceptual Framework for AI-Assisted Anticancer Phytochemical Prioritization: Phytochemical Identity, Target Vulnerability, Tumor Pathway Dependency, Molecular Selectivity, Safety-Aware Screening, Evidence Grading, Candidate Ranking, Validation Planning, and Translational Boundaries

Framework stage	Core purpose	Required input	AI-assisted or conceptual logic	Candidate prioritization output	Caution signal	Validation requirement	Failure risk	Research decision value
Phytochemical identity	Establish what is being evaluated	Compound identity, compound class, natural product source, chemical characterization	AI may assist annotation, dereplication, similarity analysis, and evidence organization	Identity-confidence category	Ambiguous compound, mixture, or extract	Chemical verification and reproducibility check	Misassigned bioactivity	Determines whether prioritization can begin
Evidence provenance	Track where evidence comes from	Assay type, dataset, model, source, computational method, and reporting quality	Evidence is mapped by origin, strength, and uncertainty	Evidence traceability profile	Opaque data source or unclear assay context	Evidence audit	Irreproducible or biased ranking	Improves transparency and reviewability
Cancer-context definition	Specify tumor-relevant setting	Tumor type context, molecular features, model system, and biological question	AI may organize tumor molecular features and candidate-context matching	Context-specific hypothesis	Generic anticancer wording	Context-matched experimental design	Overgeneralized claim	Prevents unsupported pan-cancer interpretation

Target vulnerability assessment	Determine whether implicated target is functionally important	Target prediction, target evidence, dependency evidence, functional evidence	AI may prioritize targets but vulnerability must be evidence-graded	Target-vulnerability confidence level	Target expression or prediction alone	Functional perturbation and target engagement	False target relevance	Identifies biologically meaningful target hypotheses
Tumor pathway dependency assessment	Evaluate whether tumor context relies on implicated pathway	Pathway alteration, dependency data, molecular signatures, pathway activity	AI may infer pathway relevance and dependency hypotheses	Pathway-dependency hypothesis	Pathway alteration without functional dependency	Pathway modulation and dependency testing	Confusing alteration with dependency	Anchors candidate logic in tumor systems biology
Molecular selectivity assessment	Separate intended target effects from broad activity	Target profile, off-target prediction, selectivity evidence	AI may generate selectivity and off-target hypotheses	Molecular selectivity profile	Promiscuity or unknown target spectrum	Orthogonal profiling and counterscreens	Mechanism misassignment	Controls molecular overinterpretation
Tumor-normal selectivity review	Evaluate whether activity is tumor-preferential	Tumor-normal comparator evidence and tissue-relevant normal-cell data	Conceptual filter requires comparator evidence before selectivity language	Tumor-normal contrast category	Tumor-cell cytotoxicity only	Matched tumor-normal testing	Cytotoxicity misread as selectivity	Protects against unsafe anticancer framing
Safety-aware screening	Identify toxicity, ADMET, and off-target concerns	ADMET evidence, toxicity evidence, off-target liability, exposure plausibility	AI may flag risk signals and exposure limitations	Safety-readiness profile	Toxicity alert, poor exposure, or missing safety evidence	Experimental safety and ADME follow-up	Natural-origin safety assumption	Adds safety gate before candidate elevation
Evidence grading	Classify strength and uncertainty	Computational, in vitro, preclinical, clinical, and safety evidence categories	Evidence layers are graded without collapsing into a single unsupported score	Evidence-strength category	Single weak evidence stream	Replication and orthogonal evidence	False precision	Supports transparent candidate comparison
Candidate prioritization	Identify candidates for research follow-up	Identity, target, pathway, selectivity, safety, exposure, and evidence-grade information	AI and conceptual filters support structured triage	Research-priority category	Ranking without uncertainty	Defined validation plan	Premature prioritization	Directs resources toward better-supported hypotheses
Validation planning	Define next evidence step	Target assay, pathway assay, selectivity test, safety check, exposure study	Framework maps gaps to experiments	Validation-readiness plan	No feasible validation route	Stage-appropriate validation	Unactionable hypothesis	Converts prioritization into testable research
Translational decision boundary	Prevent unsupported clinical interpretation	Evidence grade, safety status, exposure plausibility, and	Candidate remains a research hypothesis unless higher	Translational boundary statement	Language implying clinical efficacy	Clinical evidence before clinical claim	Misleading therapeutic claim	Preserves responsible communication

	clinical evidence status	evidence exists

Figure 2 presents the proposed conceptual framework for AI-assisted anticancer phytochemical prioritization across target vulnerability, tumor pathway dependency, molecular selectivity, safety-aware screening, evidence grading, and validation planning.

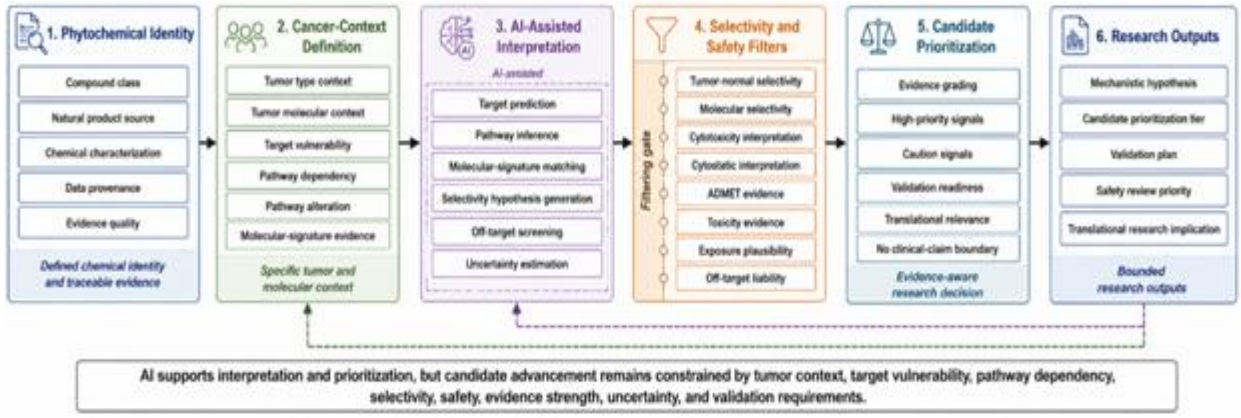

Figure 2. Conceptual Framework for AI-Assisted Anticancer Phytochemical Prioritization

Alt-text description

A conceptual framework diagram showing phytochemical identity and cancer context feeding into target vulnerability assessment, tumor pathway dependency assessment, molecular selectivity review, safety-aware screening, evidence grading, candidate prioritization, validation planning, and translational boundaries.

Research implications

For natural product researchers, cancer pharmacologists, computational biologists, and systems pharmacologists, the proposed framework provides a structured way to separate discovery enthusiasm from evidence strength. Reliability and interpretability are particularly important for cancer drug-sensitivity prediction because model outputs can appear precise even when they reflect dataset bias, limited external validity, or insufficient biological explanation [36]. The framework therefore encourages researchers to report why a candidate is prioritized, what evidence supports each layer, and what remains uncertain. For AI model developers and translational oncology researchers, the framework highlights the need for tumor-context-specific datasets, selectivity benchmarks, external validation, exposure-aware modeling, and safety-aware screening. Biological machine-learning workflows require careful protection against data leakage, especially when related samples, overlapping datasets, or improperly separated training and testing structures may inflate apparent performance [37]. In anticancer phytochemical prioritization, leakage-aware validation is essential because overstated model reliability may lead to misleading target, pathway, or candidate-ranking conclusions.

For journal reviewers, funding bodies, and research teams, the framework can function as a reporting and evaluation scaffold. Candidate claims should specify whether evidence concerns chemical identity, predicted target association, experimentally supported target engagement, tumor pathway dependency, molecular signature alignment, tumor-normal selectivity, ADMET plausibility, toxicity risk, off-target liability, validation readiness, or translational boundary. This structure supports clearer peer review and safer research prioritization by preventing AI-derived candidate rankings from being presented as validated anticancer evidence.

CONCLUSION

This article proposed an AI-assisted conceptual framework for anticancer phytochemical prioritization that integrates phytochemical identity, tumor-context definition, target vulnerability, tumor pathway dependency, molecular selectivity, safety-aware screening, evidence grading, candidate prioritization, validation planning, and translational boundaries. The framework positions AI as a tool for organizing and prioritizing research hypotheses, not as evidence of clinical anticancer efficacy. Its central contribution is to connect computational discovery with tumor biology, selectivity assessment, safety interpretation, exposure plausibility, uncertainty reporting, and validation requirements. Responsible anticancer phytochemical discovery requires that predicted targets, pathway links, molecular signatures, and candidate

rankings remain bounded by evidence strength and tested through appropriate experimental and clinical research pathways before any therapeutic claim is considered.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

[1] Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod. 2020;83(3):770-803. doi:10.1021/acs.jnatprod.9b01285

[2] Atanasov AG, Zotchev SB, Dirsch VM, International Natural Product Sciences Taskforce, Supuran CT. Natural products in drug discovery: Advances and opportunities. Nat Rev Drug Discov. 2021;20(3):200-16. doi:10.1038/s41573-020-00114-z

[3] Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-77. doi:10.1038/s41573-019-0024-5

[4] Bender A, Cortes-Ciriano I. Artificial intelligence in drug discovery: What is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. Drug Discov Today. 2021;26(2):511-24. doi:10.1016/j.drudis.2020.12.009

[5] Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23(6):1241-50. doi:10.1016/j.drudis.2018.01.039

[6] Saldívar-González FI, Aldas-Bulos VD, Medina-Franco JL, Plisson F. Natural product drug discovery in the artificial intelligence era. Chem Sci. 2022;13(6):1526-46. doi:10.1039/D1SC04471K

[7] Sahayasheela VJ, Lankadasari MB, Dan VM, Dastager SG, Pandian GN, Sugiyama H. Artificial intelligence in microbial natural product drug discovery: Current and emerging role. Nat Prod Rep. 2022;39(12):2215-30. doi:10.1039/D2NP00035K

[8] Othman ZK, Ahmed MM, Kasimieh O, Musa SS, 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;12:100316. doi:10.1016/j.ibmed.2025.100316

[9] Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. Nature. 2019;568(7753):511-16. doi:10.1038/s41586-019-1103-9

[10] Chiu YC, Zheng S, Wang LJ, Iskra BS, Rao MK, Houghton PJ, et al. Predicting and characterizing a cancer dependency map of tumors with deep learning. Sci Adv. 2021;7(34):eabh1275. doi:10.1126/sciadv.abh1275

[11] Benfatto S, Serçin Ö, Dejure FR, Abdollahi A, Zenke FT, Mardin BR. Uncovering cancer vulnerabilities by machine learning prediction of synthetic lethality. Mol Cancer. 2021;20(1):111. doi:10.1186/s12943-021-01405-8

[12] Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321-37.e10. doi:10.1016/j.cell.2018.03.035

[13] Hanahan D. Hallmarks of cancer: New dimensions. Cancer Discov. 2022;12(1):31-46. doi:10.1158/2159-8290.CD-21-1059

[14] Ochoa D, Hercules A, Carmona M, Suveges D, Gonzalez-Uriarte A, Malangone C, et al. Open Targets Platform: Supporting systematic drug-target identification and prioritisation. Nucleic Acids Res. 2021;49(D1):D1302-10. doi:10.1093/nar/gkaa1027

[15] Shimada K, Bachman JA, Muhlich JL, Mitchison TJ. shinyDepMap, a tool to identify targetable cancer genes and their functional connections from Cancer Dependency Map data. Elife. 2021;10:e57116. doi:10.7554/eLife.57116

[16] Méndez-Lucio O, Baillif B, Clevert DA, Rouquié D, Wichard J. De novo generation of hit-like molecules from gene expression signatures using artificial intelligence. Nat Commun. 2020;11(1):10. doi:10.1038/s41467-019-13807-w

[17] Daina A, Michielin O, Zoete V. SwissTargetPrediction: Updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Res. 2019;47(W1):W357-64. doi:10.1093/nar/gkz382

[18] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi:10.1038/srep42717

[19] Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 2018;46(W1):W257-63. doi:10.1093/nar/gky318

[20] Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, et al. ADMETlab 2.0: An integrated online platform for

accurate and comprehensive predictions of ADMET properties. Nucleic Acids Res. 2021;49(W1):W5-14. doi:10.1093/nar/gkab255

[21] Rao MS, Gupta R, Liguori MJ, Hu M, Huang X, Mantena SR, et al. Novel computational approach to predict off-target interactions for small molecules. Front Big Data. 2019;2:25. doi:10.3389/fdata.2019.00025

[22] Van Vleet TR, Liguori MJ, Lynch JJ III, Rao M, Warder S. Screening strategies and methods for better off-target liability prediction and identification of small-molecule pharmaceuticals. SLAS Discov. 2019;24(1):1-24. doi:10.1177/2472555218799713

[23] Wolfender JL, Litaudon M, Touboul D, Queiroz EF. Innovative omics-based approaches for prioritisation and targeted isolation of natural products: New strategies for drug discovery. Nat Prod Rep. 2019;36(6):855-68. doi:10.1039/C9NP00004F

[24] Gaudelet T, Day B, Jamasb AR, Soman J, Regep C, Liu G, et al. Utilizing graph machine learning within drug discovery and development. Brief Bioinform. 2021;22(6):bbab159. doi:10.1093/bib/bbab159

[25] Issa NT, Stathias V, Schürer S, Dakshanamurthy S. Machine and deep learning approaches for cancer drug repurposing. Semin Cancer Biol. 2021;68:132-42. doi:10.1016/j.semcancer.2019.12.011

[26] Cadow J, Born J, Manica M, Oskooei A, Rodríguez Martínez M. PaccMann: A web service for interpretable anticancer compound sensitivity prediction. Nucleic Acids Res. 2020;48(W1):W502-8. doi:10.1093/nar/gkaa327

[27] Baptista D, Ferreira PG, Rocha M. Deep learning for drug response prediction in cancer. Brief Bioinform. 2021;22(1):360-79. doi:10.1093/bib/bbz171

[28] Kuenzi BM, Park J, Fong SH, Sanchez KS, Lee J, Kreisberg JF, et al. Predicting drug response and synergy using a deep learning model of human cancer cells. Cancer Cell. 2020;38(5):672-84.e6. doi:10.1016/j.ccell.2020.09.014

[29] Pravalphruekul N, Piriyajitakonkij M, Phunchongharn P, Piyayotai S. De novo design of molecules with multiaction potential from differential gene expression using variational autoencoder. J Chem Inf Model. 2023;63(13):3999-4011. doi:10.1021/acs.jcim.3c00355

[30] Hashem S, Ali TA, Akhtar S, Nisar S, Sageena G, Ali S, et al. Targeting cancer signaling pathways by natural products: Exploring promising anti-cancer agents. Biomed Pharmacother. 2022;150:113054. doi:10.1016/j.biopha.2022.113054

[31] Bhukta S, Gopinath P, Dandela R. Target identification of anticancer natural products using a chemical proteomics approach. RSC Adv. 2021;11(45):27950-64. doi:10.1039/D1RA04283A

[32] 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. doi:10.3390/biomedicines12010201

[33] Rudzińska A, Juchaniuk P, Oberda J, Wiśniewska J, Wojdan W, Szklener K, et al. Phytochemicals in cancer treatment and cancer prevention: Review on epidemiological data and clinical trials. Nutrients. 2023;15(8):1896. doi:10.3390/nu15081896

[34] Dogra A, Gupta D, Bag S, Ahmed Z. Biosynthesis of anticancer phytochemical compounds and their potential applications in cancer therapy. Front Pharmacol. 2023;14.

[35] Kubczak M, Szustka A, Rogalińska M. Molecular targets of natural compounds with anti-cancer properties. Int J Mol Sci. 2021;22(24):13659. doi:10.3390/ijms222413659

[36] Lenhof K, Eckhart L, Rolli LM, Lenhof HP. Trust me if you can: A survey on reliability and interpretability of machine learning approaches for drug sensitivity prediction in cancer. Brief Bioinform. 2024;25(5):bbae379. doi:10.1093/bib/bbae379

[37] Bernett J, Blumenthal DB, Grimm DG, Haselbeck F, Joeres R, Kalinina OV, et al. Guiding questions to avoid data leakage in biological machine learning applications. Nat Methods. 2024;21(8):1444-53. doi:10.1038/s41592-024-02362-y