Regulatory Science for AI-Discovered Phytopharmaceuticals: Evidence Standards, Model Documentation, Safety Evaluation, and Approval Pathways

Yifan Zhao[1]*, Liang Chen[1], Ahmad Fauzi[2]

[1]Department of AI for Traditional Chinese Herbal Formula Synergy, School of Pharmacy, Fudan University, Shanghai, China.

[2]Department of Computational Pharmacognosy for Tropical Fruits, Faculty of Pharmacy, University of Malaya, Kuala Lumpur, Malaysia.

ABSTRACT

Artificial intelligence is increasingly used to identify, rank, and characterize natural-product candidates, yet AI-originated predictions do not resolve the evidentiary challenges associated with complex phytopharmaceutical materials. This article proposes a regulatory science framework for AI-discovered phytopharmaceuticals that integrates two complementary accountability tracks: phytopharmaceutical evidence for identity, quality, composition, consistency, safety, efficacy, and clinical relevance, and AI evidence for data provenance, documentation, transparency, validation, reproducibility, bias control, and lifecycle governance. The framework approach organizes candidate identity and classification, botanical source authentication, phytochemical profiling, batch consistency, quality control, model records, evidence mapping, mechanistic assessment, safety planning, clinical evidence planning where relevant, regulatory translation, risk management, pharmacovigilance readiness, and model-update oversight. Particular emphasis is placed on distinguishing computational prioritization from experimentally supported mechanism, predictive safety from empirical safety evaluation, regulatory alignment from regulatory approval, and documentation readiness from submission readiness. Approval-pathway considerations are therefore treated as structured questions for product-specific scientific assessment rather than as legal recommendations or predetermined routes. The framework also identifies policy priorities for transparent reporting, interdisciplinary review, evidence traceability, terminology harmonization, and continuing monitoring. Its main contribution is a unified regulatory science architecture that connects AI model accountability with the established quality and safety demands of phytopharmaceutical development while preserving clear boundaries against approval claims, reduced-evidence assumptions, autonomous decision-making, and unsupported clinical or regulatory conclusions.

Key Words: *Regulatory science, AI-discovered phytopharmaceuticals, Evidence standards, Model documentation, Safety evaluation, Approval pathways*

eIJPPR 2026; 16(3):20-37

HOW TO CITE THIS ARTICLE: Zhao Y, Chen L, Fauzi A. Regulatory Science for AI-Discovered Phytopharmaceuticals: Evidence Standards, Model Documentation, Safety Evaluation, and Approval Pathways. Int J Pharm Phytopharmacol Res. 2026;16(3):20-37. https://doi.org/10.51847/OxtpDksTGP

INTRODUCTION

Natural products continue to provide chemically and biologically diverse starting points for therapeutic discovery, including isolated compounds, enriched fractions, standardized extracts, and complex preparations. Their development nevertheless presents challenges that differ from those associated with structurally defined single-molecule candidates. Biological source variation, environmental influences, extraction and processing

Corresponding author: Yifan Zhao
Address: Department of AI for Traditional Chinese Herbal Formula Synergy, School of Pharmacy, Fudan University, Shanghai, China.
E-mail: ✉ yifan.zhao@fudan.edu.cn
Received: 10 February 2026; **Revised:** 02 May 2026; **Accepted:** 06 May 2026

conditions, multi-constituent composition, uncertain active principles, and variable analytical characterization can affect the reproducibility and interpretation of pharmacological findings. Contemporary natural-product discovery therefore depends not only on identifying promising biological activity but also on connecting that activity to authenticated materials, chemically characterized preparations, experimentally supported mechanisms, feasible manufacturing controls, and evidence capable of supporting responsible translation [1]. Machine learning and related AI methods are now applied across target identification, molecular representation, virtual screening, activity prediction, de novo design, pharmacokinetic assessment, toxicity prediction, formulation research, clinical development, and manufacturing. These applications can expand the search space, integrate heterogeneous data, and support prioritization of experiments, but their reliability remains conditional on the provenance and fitness of the data, the appropriateness of the modelling task, and the quality of validation. A high-performing model within a development dataset may fail when confronted with new chemical classes, different assay conditions, underrepresented botanical sources, changed manufacturing materials, or clinical contexts that were absent from training. AI discovery should consequently be treated as an evidence-generating and decision-support process rather than as independent proof of biological mechanism, product quality, safety, therapeutic efficacy, or translational readiness [2].

The regulatory science challenge is broader than determining whether an algorithm has acceptable predictive accuracy. Pharmaceutical AI introduces questions concerning data lineage, model purpose, training and test separation, reproducibility, explainability, uncertainty, bias, human responsibility, version control, and continuing oversight. These questions must be connected to the pharmaceutical evidence generated for the candidate rather than assessed as an isolated software exercise. Regulatory-affairs scholarship has accordingly emphasized that AI-enabled pharmaceutical activities require transparent records, accountable review, controlled data practices, and integration with established development functions [3]. For AI-discovered phytopharmaceuticals, this integration is especially important because uncertainty can arise simultaneously from the model, the underlying bioactivity evidence, the botanical material, the composition of the preparation, and the relationship between experimental and intended-use exposures.

AI discovery does not reduce the need to establish botanical identity, phytochemical composition, manufacturing consistency, quality controls, nonclinical safety, clinical evidence where relevant, and product-specific regulatory assessment. Peer-reviewed botanical-development literature demonstrates that coherent quality, nonclinical, and clinical evidence remains central when complex botanical products are evaluated [4]. The objective of this article is therefore to propose an original regulatory science framework that links phytopharmaceutical identity and quality evidence with AI model documentation, mechanistic and safety validation, clinical evidence planning, regulatory translation, pharmacovigilance readiness, lifecycle monitoring, expert oversight, and explicit claim boundaries. The framework is intended to support evidence organization and validation planning. It is not legal advice, an official regulatory guideline, a submission template, an approval pathway, or a determination that any candidate satisfies formal regulatory requirements.

Regulatory science for AI-discovered phytopharmaceuticals

For the purposes of this article, regulatory science for AI-discovered phytopharmaceuticals refers to the systematic organization, evaluation, and communication of scientific evidence needed to understand the identity, quality, biological rationale, safety, clinical relevance, and development status of a phytopharmaceutical candidate whose discovery or prioritization involved AI. Modern natural-product development already combines chemical analysis, biological screening, computational methods, mechanistic investigation, and translational experimentation [5]. The addition of AI creates another evidentiary layer that must be assessed alongside, rather than substituted for, the product evidence. Regulatory science in this setting asks whether the material is adequately defined, whether the model and its data are documented, whether predictions generalize, whether experimental findings support the proposed claims, and whether uncertainties remain visible to responsible reviewers.

This interpretation must be distinguished from regulatory policy, official guidance, legal classification, market authorization, and submission preparation. Regulatory science may identify evidence gaps, organize development questions, improve traceability, and support communication with relevant experts, but it cannot confer approval or determine the legally applicable pathway. Botanical preparations may differ in composition, manufacturing, intended claims, and degree of purification, while product categories and evidentiary conventions may vary between jurisdictions. International botanical-development literature illustrates that these differences complicate attempts to formulate a single universal pathway [6]. Product classification should

therefore be treated as a documented and revisable scientific-development question whose formal resolution requires product-specific assessment within the applicable system.

A second distinction concerns the status of AI-derived evidence. AI can support molecular prioritization, identify patterns across high-dimensional datasets, propose targets, predict bioactivity, estimate ADMET properties, and help select experiments. It cannot, through prediction alone, authenticate a botanical source, demonstrate batch consistency, establish target engagement, verify a mechanism, characterize mixture effects, confirm safety, or establish clinical benefit. Reviews of pharmaceutical AI consistently position computational systems as tools that can accelerate and inform development while remaining dependent on reliable data, empirical confirmation, and expert interpretation [7]. The regulatory science task is thus to preserve the connection between each AI output, its intended use, its uncertainty, the experiment undertaken to test it, and the downstream decision influenced by the result.

AI-discovered phytopharmaceuticals consequently require dual evidence accountability. The phytopharmaceutical track concerns botanical source authentication, chemical composition, phytochemical profiling, process definition, batch consistency, quality control, pharmacological evidence, safety, clinical evidence where relevant, and continuing product surveillance. The AI track concerns training-data provenance, model documentation, transparency, explainability, reproducibility, leakage control, bias assessment, internal and external validation, experimental confirmation, version control, and monitoring of model changes. Contemporary phytopharmaceutical scholarship similarly emphasizes that standardization, analytical characterization, quality evidence, safety evaluation, and clinical development remain central to translation [8]. The proposed framework connects the two tracks through evidence traceability, multidisciplinary review, risk management, lifecycle monitoring, and explicit boundaries separating scientific planning from regulatory approval.

Evidence standards and model documentation

Evidence standards for AI-discovered phytopharmaceuticals should begin with a claim-specific question: what conclusion is being proposed, and what evidence would be sufficient to support that conclusion at the relevant stage of development? A model that ranks compounds for laboratory testing requires a different evidentiary threshold from a model used to support dose selection, safety interpretation, manufacturing control, or clinical decision-making. Critical analyses of AI in drug discovery have shown that impressive retrospective results can be disconnected from practical impact when tasks, benchmarks, comparator methods, and intended uses are poorly specified [9]. Data quality is equally consequential. Chemical structures may be duplicated or incorrectly represented, assay labels may combine incompatible conditions, negative results may be selectively absent, botanical samples may be incompletely identified, and endpoint definitions may vary between sources. These limitations can propagate bias, uncertainty, leakage, and weak generalization into the model and its outputs [10]. A regulatory science record should therefore specify the proposed claim, the model's decision context, the origin and meaning of each input and label, the applicability domain, and the evidence needed to test the output empirically.

Model transparency and explainability serve related but non-equivalent functions. Transparency concerns whether reviewers can inspect the intended use, data transformations, architecture, feature construction, training procedure, model version, decision thresholds, uncertainty measures, and known limitations. Explainability concerns whether a model output can be interrogated in a manner that is stable, faithful to the model, and scientifically meaningful. Explainable AI can support hypothesis generation, highlight influential features, and help experts investigate potentially spurious relationships, but an explanation does not independently establish causal mechanism or prediction validity [11]. Structured model documentation provides the broader audit trail needed to understand how a system was developed, evaluated, and intended to be used. Documentation approaches in translational science can improve communication of provenance, intended use, performance conditions, limitations, and reproducibility requirements [12]. For phytopharmaceutical applications, this documentation should also identify how botanical, chemical, assay, omics, literature, and clinical data were linked and whether apparently identical records represent the same physical preparation or different products.

Validation should be layered rather than reduced to a single performance result. Internal validation should use leakage-resistant resampling, prespecified metrics, nested tuning where appropriate, calibration analysis, uncertainty characterization, and complete reporting of variability. External validation should use independently sourced data that represent the intended chemical, biological, botanical, or clinical context. The model should be locked before external evaluation, and differences between development and evaluation datasets should be described. Benchmark design is especially important in molecular modelling because random or structurally permissive splits can allow highly similar compounds or related assay records to appear across training and test sets. Empirical work has

demonstrated that some ligand-based benchmarks may reward memorization instead of transferable generalization [13]. Bias assessment should additionally examine performance across chemical scaffolds, botanical taxa, preparation types, laboratories, assay systems, demographic groups where relevant, and other strata that may reveal systematic failure. Experimental validation should then test model-prioritized claims using appropriate controls, replicates, exposure conditions, reference comparators, and orthogonal methods.

Evidence traceability connects these requirements into a reviewable chain. Each conclusion should be traceable from source material and raw data through curation, model version, output, experimental test, expert interpretation, and development decision. Changes to botanical material, analytical procedures, labels, preprocessing, algorithms, or decision thresholds should be versioned and assessed for their effect on prior findings. Best-practice scholarship in chemical machine learning emphasizes transparent data preparation, representation choices, validation design, reproducibility, and limitations reporting [14]. Taken together, realistic model evaluation and fit-for-purpose chemical and biological data are inseparable conditions for credible pharmaceutical AI evidence [9, 10]. A documented model may still be invalid, and a valid model within one domain may still be unsuitable for another; documentation and validation must therefore be reviewed jointly. **Table 1** summarises evidence standards and model documentation requirements for AI-discovered phytopharmaceuticals.

Table 1. Evidence Standards and Model Documentation Requirements for AI-Discovered Phytopharmaceuticals: Botanical Identity, Phytochemical Profiling, Batch Consistency, Quality Control, Data Provenance, Model Transparency, Explainability, Reproducibility, Bias Assessment, Validation, and Evidence Traceability

Evidence or documentation component	Core regulatory science question	Required evidence or record	Documentation function	Failure risk	Validation requirement	Responsible review need	Framework role
Phytopharmaceutical identity	What precisely is the candidate to which the evidence applies?	Product name; preparation type; plant part; processing or extraction method; composition boundary; intended use	Defines the object of evaluation and prevents evidence transfer between non-equivalent materials	Pharmacological, safety, or clinical findings may be assigned to a materially different product	Identity confirmation using fit-for-purpose botanical and chemical methods	Pharmacognosy, analytical chemistry, manufacturing, pharmacology, and clinical review	Entry gate for classification and evidence mapping
Botanical source authentication	Is the declared biological source authentic and traceable?	Accepted taxonomic designation; voucher information where relevant; origin; supplier; collection conditions; authentication methods; chain of custody	Connects experimental material to a verifiable biological source	Misidentification, substitution, adulteration, or irreproducible sourcing	Orthogonal morphological, microscopic, genetic, or chemical authentication as appropriate	Botanical and quality specialists should assess method suitability and discrepancies	Foundational identity and quality control
Chemical composition	What chemical material was studied or processed by the model?	Constituent inventory; marker compounds; concentration ranges; impurities; degradants; analytical procedures	Defines chemical exposure and supports comparison between studies and batches	Mechanism, dose, or safety findings may be attributed to an undefined mixture	Qualified or validated analytical procedures appropriate to development stage	Analytical, pharmacological, toxicological, and manufacturing review	Links identity to biological and safety evidence

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

Phytochemical profiling	Does the analytical profile characterize the complex preparation sufficiently?	Chromatographic, spectrometric, spectroscopic, or other justified fingerprints; reference materials; acceptance logic	Provides a reproducible compositional signature	Compositional drift, substitution, degradation, or loss of relevant constituents may remain undetected	Demonstration of reproducibility, specificity, sensitivity, and discriminatory capacity	Analytical experts should assess whether the profile represents the product adequately	Chemical standardization layer
Batch consistency	Are discovery, nonclinical, clinical, and manufacturing batches comparable?	Comparative profiles; marker ranges; process records; biological comparisons where justified; stability data	Supports bridging of evidence across batches and development stages	Safety or efficacy evidence may not transfer to later material	Evaluation of multiple representative batches and stability-aware comparability	Joint quality, manufacturing, toxicology, and clinical review	Batch-to-evidence traceability gate
Quality control	Are identity, composition, purity, contaminants, strength, and process variability controlled?	Specifications; analytical procedures; release records; stability controls; deviation investigations	Demonstrates controlled production rather than one-time characterization	Variable exposure, contamination, adulteration, or irreproducible study material	Stage-appropriate qualification or validation and longitudinal trend review	Independent quality oversight and escalation of material deviations	Non-negotiable phytopharmaceutical foundation
Training-data provenance	Where did every input, label, and derived feature originate?	Data sources; acquisition dates; inclusion criteria; assay context; curation history; lineage; access conditions	Enables reconstruction, fitness-for-purpose assessment, and duplication detection	Hidden duplication, incompatible assays, circular evidence, or untraceable labels	Sample-level lineage audits and verification against source records	Data-science, domain-science, quality, and governance review	Starting point of the AI evidence dossier
Model documentation	Can an independent reviewer reconstruct the intended use and development process?	Intended use; domain; architecture; features; preprocessing; training procedure; hyperparameters; versions; outputs; limitations	Creates an auditable and version-controlled model record	Irreproducibility, inappropriate reuse, undocumented assumptions, or unverifiable claims	Independent reconstruction or reproducibility testing proportionate to intended use	Multidisciplinary technical and domain review	Central AI accountability mechanism
Model transparency	Are design choices, transformations, assumptions, and output semantics visible?	Model rationale; exclusions; transformations; thresholds; confidence estimates; decision rules	Makes the basis of an output inspectable	Reviewers may confuse implementation choices with biological evidence	Consistency checks between documentation, code, data lineage, and operational behaviour	Technical reviewers and intended users should assess relevance and comprehensibility	Enables appraisal without implying complete interpretability
Explainability	Does the explanation support scientifically	Explanation method; baseline; fidelity	Supports hypothesis generation and	Plausible explanations may obscure invalid	Evaluation of explanation fidelity, robustness,	Domain experts should determine whether	Supportive evidence, not independent validation

	meaningful interrogation?	assessment; stability analysis; sensitivity testing; known limitations	expert investigation	predictions or spurious associations	and domain relevance separately from accuracy	explanations have scientific meaning	
Reproducibility	Can authorized reviewers reproduce the model and reported results?	Versioned code; environment; dependencies; data snapshot; random seeds; preprocessing; evaluation scripts	Tests recoverability of reported analyses	Selective reporting, implementation drift, or inaccessible analytic pathways	Independent reruns with predefined discrepancy tolerances	Technical replication and domain review of reproduced outputs	Auditability and documentation-readiness control
Data leakage control	Were evaluation data isolated from training and model selection?	Split rationale; duplicate checks; structural-similarity analysis; temporal separation where relevant; contamination audit	Protects estimates of generalization	Artificially inflated performance and false translational confidence	Leakage-resistant splitting matched to the intended prediction problem	Independent methodological review before accepting performance evidence	Validation-integrity gate
Bias assessment	Does performance differ across relevant chemical, botanical, experimental, or population groups?	Dataset composition; missingness; subgroup performance; error analysis; mitigation rationale	Identifies systematic weaknesses and underrepresented domains	Poor performance may remain hidden in groups not represented by aggregate metrics	Prespecified stratified analyses and stress testing	Statistical, scientific, safety, and governance review	Representativeness and applicability-domain safeguard
Internal validation	Is performance stable within the development data under controlled resampling?	Cross-validation design; nested tuning; repeated runs; calibration; uncertainty; all prespecified metrics	Estimates development-stage performance and variability	Tuning bias and selective reporting may exaggerate reliability	Nested, leakage-controlled procedures appropriate to dataset structure	Statistical and technical review	Necessary but insufficient first validation layer
External validation	Does the locked model generalize to independently sourced data?	Independent provenance; eligibility rules; distribution comparisons; prespecified analysis; error characterization	Tests transportability beyond development conditions	Internal performance may not transfer to new chemical, botanical, laboratory, or clinical contexts	Independent and sufficiently characterized evaluation representative of intended use	Independent scientific and statistical review	Primary generalization gate
Experimental validation	Are AI-prioritized claims confirmed through appropriate	Prespecified protocols; controls; replicates; reference comparators;	Converts computational hypotheses into product evidence	Predictions may be mistaken for mechanism, safety, efficacy, or quality evidence	Orthogonal assays, replication, and exposure-relevant testing	Experimental, mechanistic, toxicological, and quality review	Boundary between prediction and empirical evidence

	empirical testing?	raw data; deviations; uncertainty			proportionate to the claim		
Evidence traceability	Can every conclusion be followed from source evidence to decision?	Claim-to-data map; version identifiers; analytic lineage; decision records; contradictory findings; sign-offs	Maintains continuity across product and model evidence	Unsupported claims, silent evidence substitution, or loss of version context	Periodic traceability audits and reconciliation after material or model updates	Independent quality and scientific oversight	Cross-cutting ledger connecting both accountability tracks

Safety evaluation requirements

Safety evaluation must be anchored to the identity and composition of the actual phytopharmaceutical candidate. A safety study conducted with an inadequately authenticated plant, an uncharacterized extract, or a batch that is not comparable with the intended material may produce results that cannot be interpreted reliably. Best-practice phytopharmacological research therefore requires authenticated source materials, chemically characterized preparations, appropriate controls, reproducible methods, and transparent reporting [15]. Chemical authentication is also necessary to identify substitution, adulteration, contamination, and compositional inconsistency that can alter exposure and produce misleading toxicity or efficacy findings [16]. These controls are not merely manufacturing concerns: they determine whether the observed safety profile belongs to the product under development. Safety assessment should consequently connect source authentication, phytochemical profiling, process information, batch comparability, contaminant testing, stability, and the materials used in computational, laboratory, animal, and clinical investigations.

Natural origin cannot be treated as a presumption of safety. Herbal and dietary products have been associated with clinically significant adverse effects, including liver injury, while causality assessment is often complicated by incomplete product identification, multi-ingredient composition, variable exposure, concomitant medicines, and uncertain adulteration [17]. Safety planning should therefore address intrinsic toxicity, target-organ effects, safety pharmacology, dose–response relationships, route and duration of exposure, metabolites, susceptible populations, and the possibility that a mixture behaves differently from its isolated constituents. Herb–drug interaction assessment requires particular caution because claims may be based on isolated case reports or mechanistic predictions without adequate temporal, exposure, dechallenge, alternative-cause, or clinical-pharmacology evidence. Critical evaluation of reported interactions demonstrates the importance of structured causality assessment and exposure-relevant mechanistic analysis [18]. AI-generated interaction alerts may help prioritize investigation, but they do not independently establish that a clinically meaningful interaction exists.

ADMET and toxicology models can contribute to early hazard identification, prioritization, and experimental planning, provided that their intended use, applicability domain, training data, uncertainty, and validation are documented. Predictions should be linked to the chemical constituents or mixture representation used by the model and to exposures relevant to the intended product. Empirical evaluation may need to address whole-preparation effects, major constituents, metabolites, impurities, contaminants, batch variability, and interactions among components. Nonclinical evidence should be selected according to the product, route, duration, dose, target population, prior knowledge, and unresolved risks rather than through a universal checklist. Clinical safety evidence, where relevant, should document the administered material, dose, exposure, adverse events, laboratory findings, concomitant products, discontinuations, and follow-up. Pharmacovigilance databases can provide additional evidence about herbal adverse events, although underreporting, inconsistent nomenclature, variable product composition, confounding, and heterogeneous case quality limit causal interpretation [19].

Safety governance should continue throughout development and, where applicable, after broader use. Pharmacovigilance readiness requires product and batch traceability, standardized adverse-event terminology, accessible reporting channels, causality procedures, signal-detection methods, escalation criteria, and clear responsibility for corrective action. Reviews of herbal pharmacovigilance identify persistent challenges related to product identification, reporting quality, surveillance capacity, and risk communication [20]. Risk management should maintain a living record of identified, potential, and uncertain risks, together with mitigation measures, evidence gaps, review triggers, and ownership. Safety interpretation is unreliable when botanical materials are inadequately authenticated, chemically characterized, or

controlled [15, 16]. Uncertainty should be communicated explicitly, and computational predictions, historical use, mechanistic plausibility, or absence of reported events should not be described as proof that a product is safe. **Table 2** maps safety evaluation requirements for AI-discovered phytopharmaceuticals.

Table 2. Safety Evaluation Requirements for AI-Discovered Phytopharmaceuticals: ADMET Evidence, Toxicity Evidence, Safety Pharmacology, Herb–Drug Interaction Risk, Compound-Mixture Safety, Batch Variability, Contaminant and Adulteration Concerns, Nonclinical Safety, Clinical Safety, Pharmacovigilance Readiness, and Risk Management

Safety domain	Core safety question	Required evidence	Evaluation function	Risk if missing	Validation requirement	Regulatory science boundary	Framework role
ADMET evidence	What absorption, distribution, metabolism, excretion, and interaction liabilities are plausible at relevant exposure?	Experimentally grounded ADMET findings; justified computational predictions; applicability domain; uncertainty; constituent and exposure context	Prioritizes liabilities and informs study selection	False reassurance, overlooked metabolites, or inappropriate dose assumptions	Fit-for-purpose benchmark evaluation and empirical confirmation proportionate to the claim	ADMET prediction supports planning but does not establish clinical safety	Early safety-screening layer
Toxicity evidence	What intrinsic, target-organ, genetic, reproductive, immunological, developmental, or other toxicities require investigation?	Appropriate in vitro, in vivo, alternative-method, mechanistic, clinical, and exposure evidence	Characterizes hazards and contributes to risk assessment	Significant toxicity may remain undetected or be attributed to the wrong constituent or product	Appropriate controls, reproducibility, exposure relevance, and product comparability	AI classification or historical use cannot replace necessary empirical safety evaluation	Core hazard-evidence gate
Safety pharmacology	Could the candidate adversely affect vital physiological functions?	Fit-for-purpose cardiovascular, neurological, respiratory, and other justified functional evaluations	Detects acute functional risks not necessarily identified by efficacy studies	Serious functional effects may emerge late in development	Sensitive, controlled, exposure-relevant testing	The required evaluation depends on product characteristics and intended exposure	Functional safety-planning component
Herb–drug interaction risk	Could constituents affect drug metabolism, transport, pharmacodynamics, or exposure?	Composition-aware experimental evidence; mechanistic data; exposure margins; clinical observations; concomitant-use context;	Identifies plausible interaction mechanisms and high-risk combinations	Harmful interactions, therapeutic failure, or incorrect attribution of adverse effects	Predictions require confirmatory and exposure-relevant assessment	A database alert or model output does not independently establish a clinical interaction	Dedicated interaction-risk module

		causality assessment					
Dose or exposure relevance	Were observed effects produced at exposures relevant to intended use?	Dose–response information; constituent and metabolite exposure; route; duration; bioavailability; uncertainty	Connects hazard findings with plausible biological exposure	Laboratory findings may be overinterpreted or dismissed incorrectly	Validated or qualified exposure measurement and justified bridging	Mechanistic plausibility without exposure context does not establish clinical risk magnitude	Exposure anchor for safety interpretation
Compound–mixture safety	Does the whole preparation behave differently from isolated constituents?	Whole-product evidence; constituent-level data where informative; mixture rationale; interaction and matrix-effect assessment	Evaluates additivity, synergy, antagonism, and matrix effects	Safety may be inferred incorrectly from a single constituent	Representative-mixture testing and justified reduction strategies	Evidence for one marker compound cannot automatically substitute for whole-product evidence	Complex-mixture assessment boundary
Contaminant and adulteration concern	Could undeclared chemical, pharmaceutical, microbial, environmental, or processing contaminants alter safety?	Risk-based specifications; supplier controls; validated testing; investigation of unexpected analytical signals	Separates intrinsic product risk from quality failure	Toxicity or apparent efficacy may arise from undeclared materials	Sensitive, specific, and appropriately maintained analytical procedures	Favourable AI predictions do not reduce contamination-control requirements	Quality–safety interface
Batch variability safety concern	Could batch differences alter biological exposure or toxicity?	Comparative chemical and biological profiles; process parameters; marker ranges; stability; representative safety-relevant batches	Determines whether safety findings can be bridged across batches	A safe research batch may not represent clinical or later manufacturing material	Multi-batch and stability-aware comparability assessment	Batch consistency is necessary but does not independently prove safety	Batch-to-safety traceability gate
Nonclinical safety evidence	Is the nonclinical evidence adequate for the product and development question?	Pharmacology; toxicology; toxicokinetics; local tolerance; interaction studies; other	Supports hazard characterization and clinical planning	Human investigation may begin with unresolved or poorly characterized risks	Review of study quality, exposure coverage, product comparability, and	The framework organizes planning and does not prescribe one	Nonclinical safety checkpoint

		risk-driven assessments			reproducibility	universal test package	
Clinical safety evidence where relevant	What adverse effects, dose relationships, interactions, and vulnerable groups emerge in humans?	Prespecified safety outcomes; exposure; adverse-event records; laboratory findings; concomitant medicines; discontinuations; follow-up	Evaluates human tolerability and clinically relevant risk	Predicted or nonclinical safety may be mistaken for human safety	Appropriate study design, monitoring, data integrity, and independent analysis	Clinical evidence needs require formal product-specific assessment	Human safety-evidence layer
Pharmacovigilance readiness	Can adverse events be detected, attributed, evaluated, and communicated during broader use?	Product nomenclature; lot and batch identifiers; reporting channels; coding rules; signal procedures; causality methods; escalation plan	Enables continuing surveillance and investigation	Signals may be missed, pooled across non-equivalent products, or become untraceable	Testing of surveillance processes and continuing data-quality review	Readiness does not imply authorization and cannot replace pre-use safety evidence	Lifecycle surveillance layer
Risk management	How will known, potential, and uncertain risks be monitored and mitigated?	Risk register; mitigation measures; evidence gaps; responsibilities; triggers; review intervals	Converts safety evidence into controlled actions	Known limitations may not influence development or communication	Periodic assessment of mitigation effectiveness and trigger performance	A risk-management plan does not establish that residual risk is acceptable	Cross-stage control and escalation mechanism
Uncertainty communication	What remains unknown, model-dependent, weakly supported, or population-limited?	Confidence intervals; applicability domains; missing-data assessment; assumptions; contradictory evidence; unresolved questions	Prevents certainty inflation and supports proportionate decisions	Preliminary predictions may be communicated as established safety conclusions	Consistency review between uncertainty statements and underlying evidence	Disclosure cannot compensate for missing essential safety evidence	Mandatory output at every safety gate
No safety-claim boundary	Has any computational, historical, mechanistic, or preliminary result been overstated as proof of safety?	Claim audit distinguishing prediction, hypothesis, hazard evidence, risk characterization, clinical evidence, and	Enforces responsible evidentiary language	Unsupported claims that the candidate is safe, non-toxic, or interaction-free	Independent scientific and safety review of external claims	Neither natural origin nor AI discovery establishes safety	Final safety-claim control

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

	surveillance findings

Approval pathways and regulatory translation

Approval-pathway consideration should begin with product definition and evidence status rather than with the assumption that an AI-discovered candidate belongs to a predetermined regulatory category. Entry into clinical development by an AI-originated candidate does not establish efficacy, safety, or approval readiness, because clinical progression remains dependent on the quality of the product, the relevance of the target and mechanism, the adequacy of nonclinical evidence, and the results of appropriately designed human studies [21]. For phytopharmaceuticals, the distinction between a complex botanical preparation, an enriched fraction, a purified constituent, and a structurally defined molecule can alter the scientific questions associated with composition, manufacturing control, dosing, consistency, and clinical evaluation. Comparisons between botanical and single-molecule development approaches illustrate why product identity must precede pathway planning [22]. The proposed approach therefore treats classification as a documented and revisable regulatory science question rather than as a legal determination.

Regulatory translation also requires caution when evidence from one product class or jurisdiction is used to inform another. Comparative analyses of AI-enabled medical technologies show that authorization categories, evaluation practices, and evidence disclosure can differ materially between regulatory systems [23]. These observations cannot be transferred directly to phytopharmaceutical products, but they demonstrate the risk of assuming that technical performance has a uniform regulatory meaning. Authorization may also coexist with limitations in external, prospective, or population-representative validation, as analyses of regulated medical AI have shown [24]. For AI-discovered phytopharmaceuticals, regulatory alignment should consequently mean that evidence is organized transparently, methods are fit for purpose, limitations are disclosed, and product-specific questions are prepared for formal assessment. It should not be interpreted as evidence of likely authorization, regulatory acceptance, legal compliance, or submission sufficiency.

The translation process must extend beyond the initial model and candidate-selection event. Adaptive or retrained AI systems can change when new data, labels, features, thresholds, or deployment conditions are introduced. System-level governance is therefore needed to connect the model, its users, the development environment, update procedures, monitoring activities, and human decisions [25]. Pharmaceutical AI may also influence formulation, manufacturing, quality control, clinical development, and post-market surveillance, making lifecycle documentation necessary across several development functions [26]. Model-update governance should specify which changes require renewed validation, whether previous evidence remains applicable, how version changes are communicated, and when outputs must be restricted or withdrawn from decision processes. Parallel product governance should monitor batch consistency, stability, emerging toxicity, clinical observations, adverse-event reports, and changes in intended use.

Regulatory translation is thus best understood as an evidence-organization process that links candidate classification, phytopharmaceutical quality, AI documentation, mechanistic evidence, safety evaluation, clinical planning, risk management, pharmacovigilance readiness, and lifecycle oversight. Potential pathways may be mapped conditionally to identify evidence dependencies and questions for expert discussion, but the resulting map cannot determine which formal route applies or whether the evidence is sufficient. Submission-readiness language should likewise be restricted to documented gap assessment and organizational preparedness rather than used as a certification claim. **Figure 1** illustrates how evidence standards, model documentation, safety evaluation, quality control, clinical evidence, and lifecycle governance connect to regulatory translation for AI-discovered phytopharmaceuticals.

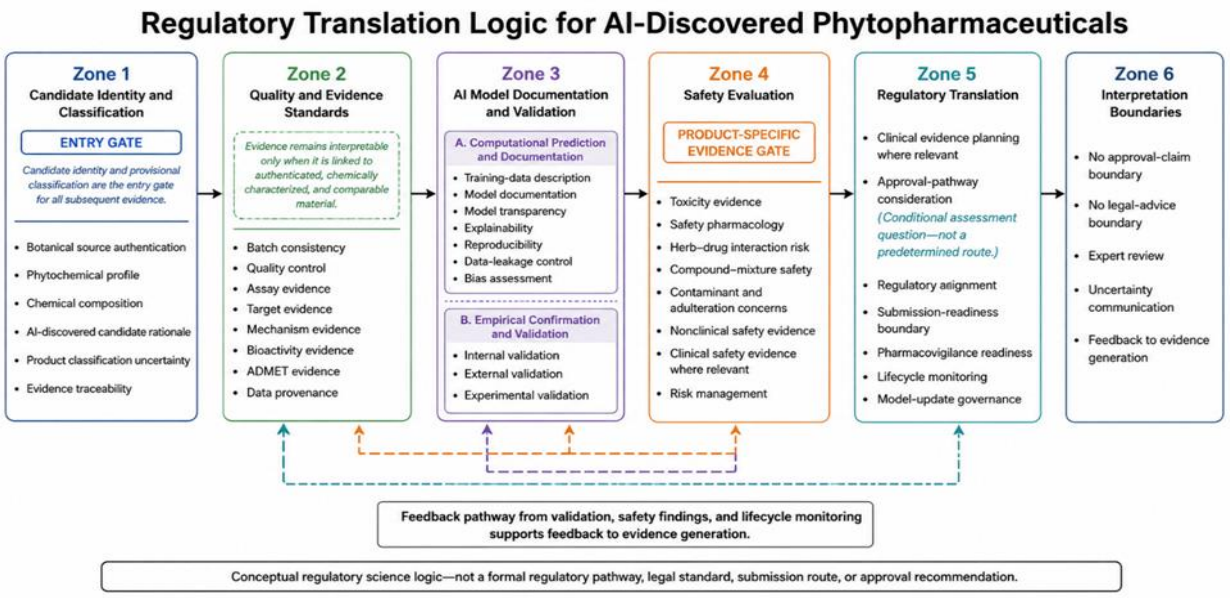

Figure 1. Regulatory Translation Logic for AI-Discovered Phytopharmaceuticals

Alt-text description

A conceptual regulatory translation diagram showing AI-discovered phytopharmaceutical evidence flowing through identity, quality control, model documentation, mechanism evidence, safety evaluation, clinical evidence, pharmacovigilance readiness, lifecycle monitoring, and approval-pathway consideration boundaries.

Proposed regulatory science framework

The proposed framework is a staged, iterative architecture rather than a linear approval algorithm. Its initial stage defines candidate identity and provisional classification by documenting the botanical source, plant part, preparation type, extraction or processing method, chemical composition, intended use, and the precise role played by AI in candidate discovery. This stage is followed by a phytopharmaceutical quality foundation comprising authentication, phytochemical profiling, marker selection, contaminant control, batch consistency, process documentation, and stability-aware comparability. Translational AI scholarship indicates that reliable impact requires clinically or scientifically relevant validation, integration with expert workflows, and continuing evaluation rather than isolated model performance [27]. Accordingly, progression through the framework depends on multidisciplinary review of both product and model evidence.

The AI model-documentation stage records the intended use, training-data provenance, data curation, representations, architecture, preprocessing, model version, hyperparameters, evaluation methods, uncertainty, limitations, and human interventions. Explainability may be included when it supports scientific interrogation, but it is not accepted as a replacement for empirical validation. Critical analysis of explainable AI in healthcare cautions that plausible post hoc explanations can create unwarranted confidence when the underlying model, data, or causal interpretation is weak [28]. The framework therefore separates transparency, explainability, reproducibility, and validity into distinct review questions. Evidence-standard mapping then links each proposed claim—such as candidate prioritization, target association, mechanism, bioactivity, ADMET behaviour, toxicity, quality, or clinical relevance—to the type and strength of evidence needed to support it.

Mechanistic and safety evaluation operate as connected but independent assessment stages. Mechanistic evidence should establish whether predicted targets or pathways are supported by orthogonal experiments, exposure-relevant effects, appropriate controls, reproducibility, and evidence relating to the actual preparation rather than only to an isolated constituent. Safety planning should integrate ADMET information, toxicology, safety pharmacology, herb–drug interaction risk, mixture effects, contaminants, batch variability, nonclinical evidence, and clinical safety evidence where relevant. Human expert oversight is essential because machine-learning systems can produce automation bias, workflow disruption, accountability gaps, and other unintended consequences even when technical performance appears favourable [29]. The framework therefore requires a documented expert-review record, including disagreements, residual uncertainties, rejected claims, and reasons for progression, restriction, or additional evidence generation.

Clinical evidence planning and regulatory translation mapping are conditional stages. The framework does not

assume that every candidate requires an identical clinical programme or that one regulatory category applies across products and jurisdictions. Instead, it records the proposed indication, population, route, formulation, dose rationale, comparability of study material, unresolved risks, and the clinical questions that would require formal product-specific assessment. Lifecycle monitoring begins before any broader deployment or authorization claim and continues through model and product development. Dataset shift can reduce performance when the chemical, biological, laboratory, population, or operational environment differs from development conditions, making predefined monitoring and revalidation triggers necessary [30]. Predictive toxicology evidence likewise requires transparent assumptions, an appropriate applicability domain, uncertainty characterization, and fit-for-purpose empirical validation before it can influence higher-stakes safety conclusions [31].

The final architecture connects risk management, model-update governance, product surveillance, pharmacovigilance readiness, policy feedback, and a no-approval-claim boundary. Pharmaceutical AI must be governed across discovery, development, clinical testing, manufacturing, quality control, and lifecycle use rather than treated as a one-time candidate-generation tool [32]. Expert review remains a cross-cutting requirement because translational integration failures, explanation-related overconfidence, and unintended consequences can persist despite apparently strong technical outputs [27–29]. The framework may generate an identity record, quality dossier, model dossier, claim–evidence map, mechanistic confidence profile, safety plan, conditional clinical evidence roadmap, translation map, risk register, and monitoring plan. These are proposed regulatory science outputs and do not constitute legal classifications, official requirements, regulator-approved procedures, submission certifications, or approval determinations. **Table 3** presents the proposed regulatory science framework for AI-discovered phytopharmaceuticals.

Table 3. Proposed Regulatory Science Framework for AI-Discovered Phytopharmaceuticals: Candidate Identity, Quality Foundation, AI Model Documentation, Evidence Mapping, Mechanistic Evidence, Safety Evaluation, Clinical Evidence Planning, Regulatory Translation, Approval-Pathway Consideration, Lifecycle Monitoring, Policy Feedback, and Claim Boundaries

Framework stage	Core purpose	Required input	Regulatory science function	Potential output	Validation need	Failure risk	Feedback mechanism	Decision boundary
Candidate identity and classification	Define the candidate and the provisional development category	Botanical source; plant part; preparation; processing method; chemical composition; intended use; AI discovery rationale	Prevents evidence from being assigned to an undefined or non-equivalent product	Candidate identity statement and provisional classification map	Independent botanical, chemical, pharmaceutical, and regulatory science review	Misclassification or inappropriate borrowing of evidence from different products	Revise the classification when composition, claims, or manufacturing change	Classification remains provisional and nonbinding
Phytopharmaceutical quality foundation	Establish authenticated and controlled material for evidence generation	Source records; analytical profiles; specifications; process information; contaminants; stability; batch data	Makes pharmacological, safety, and clinical evidence interpretable	Quality evidence dossier and batch-comparability map	Method qualification or validation and representative-batch assessment	Uncontrolled, contaminated, or non-reproducible material progresses through development	Batch trends, deviations, and stability findings update evidence status	Progression should pause when identity or comparability is unresolved
AI model documentation	Record the model, data, intended use, limitations, and versions	Data lineage; curation history; model architecture; preprocessing; code environment; metrics;	Enables reconstruction, review, and version-specific interpretation	Version-controlled model dossier	Reproducibility audit and consistency check against implementation	Opaque, irreproducible, or incorrectly reused model evidence	Every controlled model change updates the dossier	Documentation does not establish validity by itself

		uncertainty; human interventions						
Evidence standard mapping	Match each proposed claim to an appropriate evidence requirement	Intended claim; development stage; decision context; evidence inventory; uncertainty	Prevents prediction, association, or tradition from being represented as proof	Claim–evidence matrix and evidence-gap register	Independent scientific and methodological review	Weak evidence supports high-level claims	New evidence changes claim status and closes or reopens gaps	Unsupported claims remain hypothesis-level
Mechanistic evidence assessment	Evaluate target, pathway, and biological-plausibility claims	Model predictions; assays; target-engagement data; constituent profiles; exposure information; orthogonal experiments	Separates computational hypotheses from experimentally supported mechanisms	Mechanistic confidence profile	Replication, orthogonal methods, controls, and exposure relevance	Spurious target narratives misdirect development	Contradictory or new findings revise confidence	Mechanistic evidence does not establish clinical efficacy
Safety evaluation planning	Organize risk-based safety evidence for the actual product	Quality data; dose and route; ADMET; toxicology; interactions; contaminants; exposure; historical and clinical signals	Coordinates empirical safety assessment and uncertainty management	Safety evidence plan and living risk register	Specialist review of methods, exposure coverage, and material comparability	Important hazards are omitted or evaluated using non-representative material	Findings update formulation, dose, model assumptions, and study priorities	A safety plan is not a safety determination
Clinical evidence planning where relevant	Define human evidence questions for the proposed use	Product definition; nonclinical evidence; mechanism; dose rationale; safety risks; target population; endpoints	Aligns clinical evaluation with the actual product and intended claims	Conditional clinical evidence roadmap	Clinical, statistical, pharmacological, quality, and ethical review	Premature studies, unsuitable endpoints, or non-comparable clinical material	Clinical findings feed back to product, mechanism, and safety assessment	The framework does not authorize clinical investigation
Regulatory translation mapping	Organize evidence for potential formal assessment	Classification map; quality dossier; model dossier; safety status; clinical plans; jurisdictional context	Converts evidence into reviewable questions, dependencies, and gaps	Conditional translation map and consultation-question set	Product-specific regulatory science and expert review	Generic assumptions obscure product- or jurisdiction-specific requirements	Formal feedback, when available, is logged and incorporated	Mapping supports planning rather than formal determination
Approval-pathway consideration	Compare plausible development options without selecting or guaranteeing one	Product type; intended claims; evidence maturity; classification uncertainty; applicable context	Clarifies evidentiary consequences associated with possible classifications	Conditional pathway-options map	Product- and jurisdiction-specific formal assessment outside the framework	A pathway option is misrepresented as an official route	New evidence or authoritative feedback revises the options	No approval route is recommended

Risk management	Track quality, safety, model, operational, and communication risks	Evidence gaps; uncertainty; identified hazards; monitoring data; mitigation measures	Makes residual risk, ownership, and escalation logic explicit	Prioritized mitigation and escalation plan	Periodic review of mitigation effectiveness	Known uncertainties become normalized or forgotten	Actions, outcomes, and residual risks are tracked continuously	Risk management does not establish acceptability of residual risk
Lifecycle monitoring	Detect changes in product quality, evidence, context, safety, and model performance	Batch trends; drift indicators; complaints; adverse events; new literature; validation data; use-context information	Maintains evidence validity over time	Monitoring reports and re-evaluation triggers	Prospective thresholds and independent review of triggered events	Evidence becomes stale while claims remain unchanged	Monitoring initiates revalidation, restriction, withdrawal, or evidence regeneration	Validity is not assumed to be permanent
Model update governance	Control retraining, feature changes, threshold changes, and new data incorporation	Change description; retraining data; version history; comparative evaluation; bias analysis; rollback plan	Preserves traceability and determines revalidation needs	Change-impact assessment and controlled release record	Locked comparative validation and applicability-domain review	Silent updates invalidate previous evidence	Update outcomes feed back into model and claim records	Updated outputs require renewed evidentiary assessment
Policy feedback	Translate recurring evidence gaps into proposed research or infrastructure priorities	Aggregated framework experience; unresolved gaps; reporting failures; stakeholder input	Supports learning-system development in regulatory science	Proposed reporting conventions, research priorities, or capacity needs	Conflict-of-interest and representativeness review	Product-specific experience is overgeneralized into policy	Proposals are revised through multidisciplinary consultation	Outputs are proposed priorities, not binding policy
Expert review	Apply accountable multidisciplinary judgment at every major transition	Complete evidence package; uncertainties; model outputs; contradictory findings; decision criteria	Integrates botanical, analytical, computational, pharmacological, safety, clinical, quality, and regulatory perspectives	Signed review record, dissent log, and decision rationale	Reviewer competence, independence, and conflict management	Automation bias or fragmented accountability	Disagreements and new evidence trigger reassessment	Human review cannot substitute for missing evidence
No approval-claim boundary	Prevent framework outputs from being represented as authorization or legal sufficiency	Reports; figures; tables; external claims; terminology; readiness assessments	Maintains the distinction between planning, alignment, and formal regulatory decisions	Claim-boundary statement and terminology-compliance record	Independent final scientific and editorial audit	Misleading claims of compliance, submission readiness, acceptance, or approval	Problematic wording is corrected and traced to its source	The framework cannot confer regulatory approval

Figure 2 presents the proposed regulatory science framework for AI-discovered phytopharmaceuticals, integrating evidence standards, model documentation, safety evaluation, regulatory translation, lifecycle monitoring, policy feedback, and claim boundaries.

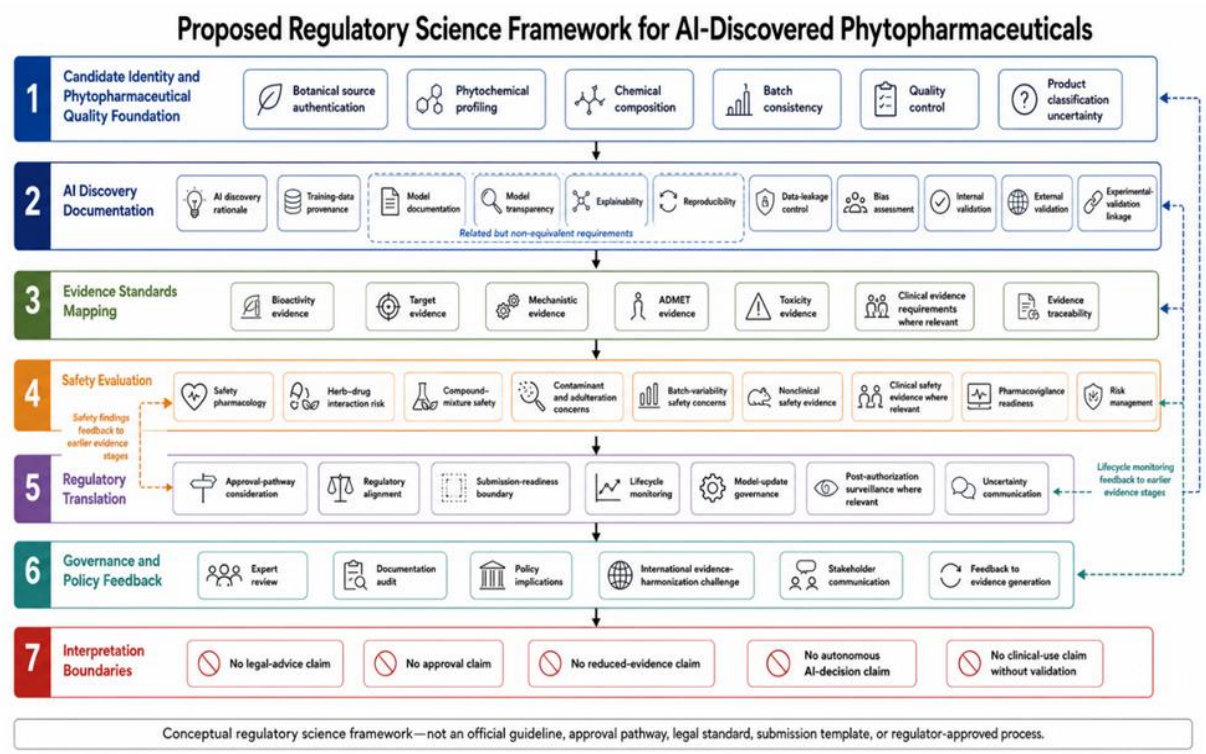

Figure 2. Proposed Regulatory Science Framework for AI-Discovered Phytopharmaceuticals

Alt-text description

A conceptual regulatory science framework showing AI-discovered phytopharmaceutical candidates moving through quality evidence, model documentation, evidence mapping, mechanism assessment, safety evaluation, clinical evidence planning, regulatory translation, lifecycle monitoring, expert oversight, policy feedback, and no-approval-claim boundaries.

Policy implications

For researchers, AI developers, and phytopharmaceutical manufacturers, the framework implies that documentation should begin at candidate conception rather than being reconstructed after favourable results emerge. Botanical identity, material provenance, chemical composition, batch history, training-data lineage, model versions, validation designs, experimental outcomes, and decision records should be retained in interoperable evidence structures. Recent phytopharmaceutical regulatory scholarship highlights the continuing importance of product definition, standardization, quality, safety, efficacy evidence, and implementation capacity when developing distinct herbal-medicine categories [33]. Funders and research institutions could therefore prioritize infrastructure for authenticated materials, reference-quality analytical data, negative and inconclusive findings, external validation resources, and multidisciplinary regulatory science expertise.

Journals, ethics committees, quality teams, and pharmacovigilance systems also influence the reliability of the evidence ecosystem. Journals can require transparent descriptions of botanical materials, chemical profiles, dataset provenance, splitting strategies, applicability domains, uncertainty, model versions, and experimental confirmation. Ethics and governance review should examine whether AI outputs are being used within the evidentiary scope for which they were validated. Quality teams should connect manufacturing changes to prior pharmacological, toxicological, and clinical evidence, while pharmacovigilance systems should preserve product- and batch-level traceability. Analyses of botanical regulatory submissions can reveal recurring quality, nonclinical, clinical, and documentation issues, although those findings remain product-, indication-, and jurisdiction-specific [34]. Such analyses can inform research priorities without functioning as universal submission rules.

Broader policy research should examine harmonized terminology for AI-discovered candidates, phytopharmaceutical preparations, model changes, evidence status, and regulatory alignment. Shared terminology could reduce ambiguity between a predicted association, an experimentally confirmed finding, an evidence gap, a development option, and a formal regulatory decision. International harmonization efforts should focus on transferable scientific records and transparent documentation while respecting differences in product classification and applicable requirements. These implications are proposed research, reporting,

infrastructure, and governance priorities for researchers, manufacturers, journals, funders, safety systems, and regulatory science communities. They are not legal advice, official guidance, mandatory standards, approval recommendations, or claims that one evidence package will be accepted across authorities.

CONCLUSION

This article proposes a regulatory science framework for AI-discovered phytopharmaceuticals that integrates candidate identity, botanical authentication, phytochemical characterization, quality control, AI model documentation, evidence standards, mechanistic assessment, safety evaluation, clinical evidence planning where relevant, regulatory translation, risk management, lifecycle monitoring, expert oversight, policy feedback, and explicit claim boundaries. Its central principle is dual evidence accountability: AI-derived prioritization must be supported by transparent data and validated models, while the candidate itself must satisfy the scientific demands associated with a defined, consistent, safe, and clinically interpretable phytopharmaceutical product. AI can strengthen candidate generation and evidence integration, but it cannot replace product characterization, empirical validation, safety investigation, human judgment, or formal regulatory assessment. The framework is therefore intended to improve evidence organization, traceability, validation planning, and interdisciplinary communication without functioning as an approval pathway, legal standard, submission template, or determination of regulatory readiness.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

[1] Atanasov AG, Zotchev SB, Dirsch VM, 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

[2] 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

[3] Patil RS, Kulkarni SB, Gaikwad VL. Artificial intelligence in pharmaceutical regulatory affairs. Drug Discov Today. 2023;28(9):103700. doi:10.1016/j.drudis.2023.103700

[4] Wu C, Lee SL, Taylor C, Li J, Chan YM, Agarwal R, et al. Scientific and regulatory approach to botanical drug development: A U.S. FDA perspective. J Nat Prod. 2020;83(2):552-62. doi:10.1021/acs.jnatprod.9b00949

[5] Thomford NE, Senthebane DA, Rowe A, Munro D, Seele P, Maroyi A, et al. Natural products for drug discovery in the 21st century: Innovations for novel drug discovery. Int J Mol Sci. 2018;19(6):1578. doi:10.3390/ijms19061578

[6] Ahn K. The worldwide trend of using botanical drugs and strategies for developing global drugs. BMB Rep. 2017;50(3):111-6. doi:10.5483/BMBRep.2017.50.3.221

[7] Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021;26(1):80-93. doi:10.1016/j.drudis.2020.10.010

[8] Tiwari R, Mishra S, Chauhan A, Gulati P, Dhobi M. The dawn till dusk of phytopharmaceuticals. Saudi Pharm J. 2024;32(11):102185. doi:10.1016/j.jsps.2024.102185

[9] Bender A, Cortés-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

[10] Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: What is realistic, what are illusions? Part 2: A discussion of chemical and biological data. Drug Discov Today. 2021;26(4):1040-52. doi:10.1016/j.drudis.2020.11.037

[11] Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nat Mach Intell. 2020;2(10):573-84. doi:10.1038/s42256-020-00236-4

[12] Brereton TA, Malik MM, Lifson M, Greenwood JD, Peterson KJ, Overgaard SM. The role of artificial intelligence model documentation in translational science: Scoping review. Interact J Med Res. 2023;12:e45903. doi:10.2196/45903

[13] Wallach I, Heifets A. Most ligand-based classification benchmarks reward memorization rather than generalization. J Chem Inf Model. 2018;58(5):916-32. doi:10.1021/acs.jcim.7b00403

[14] Artrith N, Butler KT, Coudert FX, Han S, Isayev O, Jain A, et al. Best practices in machine learning for chemistry. Nat Chem. 2021;13(6):505-8. doi:10.1038/s41557-021-00716-z

[15] Heinrich M, Appendino G, Efferth T, Fürst R, Izzo AA, Kayser O, et al. Best practice in research—

overcoming common challenges in phytopharmacological research. J Ethnopharmacol. 2020;246:112230. doi:10.1016/j.jep.2019.112230

[16] Ichim MC, Booker A. Chemical authentication of botanical ingredients: A review of commercial herbal products. Front Pharmacol. 2021;12:666850. doi:10.3389/fphar.2021.666850

[17] Navarro VJ, Khan I, Björnsson E, Seeff LB, Serrano J, Hoofnagle JH. Liver injury from herbal and dietary supplements. Hepatology. 2017;65(1):363-73. doi:10.1002/hep.28813

[18] Awortwe C, Makiwane M, Reuter H, Muller C, Louw J, Rosenkranz B. Critical evaluation of causality assessment of herb–drug interactions in patients. Br J Clin Pharmacol. 2018;84(4):679-93. doi:10.1111/bcp.13490

[19] Kongkaew C, Phan DTA, Janusorn P, Mongkhon P. Estimating adverse events associated with herbal medicines using pharmacovigilance databases: Systematic review and meta-analysis. JMIR Public Health Surveill. 2024;10:e63808. doi:10.2196/63808

[20] Choudhury A, Singh PA, Bajwa N, Dash S, Bisht P. Pharmacovigilance of herbal medicines: Concerns and future prospects. J Ethnopharmacol. 2023;309:116383. doi:10.1016/j.jep.2023.116383

[21] Jayatunga MKP, Ayers M, Bruens L, Jayanth D, Meier C. How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons. Drug Discov Today. 2024;29(6):104009. doi:10.1016/j.drudis.2024.104009

[22] Bonn-Miller MO, ElSohly MA, Loflin MJE, Chandra S, Vandrey R. Cannabis and cannabinoid drug development: Evaluating botanical versus single molecule approaches. Int Rev Psychiatry. 2018;30(3):277-84. doi:10.1080/09540261.2018.1474730

[23] Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): A comparative analysis. Lancet Digit Health. 2021;3(3):e195-203. doi:10.1016/S2589-7500(20)30292-2

[24] Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: Limitations and recommendations from an analysis of FDA approvals. Nat Med. 2021;27(4):582-4. doi:10.1038/s41591-021-01312-x

[25] Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. NPJ Digit Med. 2020;3:53. doi:10.1038/s41746-020-0262-2

[26] Huanbutta K, Burapapadh K, Kraisit P, Sriamornsak P, Ganokratanaa T, Suwanpitak K, et al. The artificial intelligence-driven pharmaceutical industry: A paradigm shift in drug discovery, formulation development, manufacturing, quality control, and post-market surveillance. Eur J Pharm Sci. 2024;203:106938. doi:10.1016/j.ejps.2024.106938

[27] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2

[28] Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11):e745-50. doi:10.1016/S2589-7500(21)00208-9

[29] Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA. 2017;318(6):517-8. doi:10.1001/jama.2017.7797

[30] Subbaswamy A, Saria S. From development to deployment: Dataset shift, causality, and shift-stable models in health AI. Biostatistics. 2020;21(2):345-52. doi:10.1093/biostatistics/kxz041

[31] Hartung T, Rao M, Behl M, Maertens A. AI snake oil? A risk/benefit analysis for toxicology. Front Artif Intell. 2026;9:1779276. doi:10.3389/frai.2026.1779276

[32] Niazi SK. The coming of age of AI/ML in drug discovery, development, clinical testing, and manufacturing: The FDA perspectives. Drug Des Devel Ther. 2023;17:2691-725. doi:10.2147/DDDT.S424991

[33] Jadhav R, Das J, Rajyaguru S, Kyada S, Kale V, Teja PK, et al. Regulations, current development, and future prospects of phytopharmaceuticals, a new class of herbal medicines in India. Discover Pharm Sci. 2025;1:6. doi:10.1007/s44395-025-00006-4

[34] Park JYK, Lee D, Rui L, Gao X, Furness MS, Wu C. Analysis of regulatory botanical submission profile for cancer management from the U.S. FDA perspectives. Ther Innov Regul Sci. 2025;59(5):1129-37. doi:10.1007/s43441-025-00786-y