RESEARCH PAPER

AI-Driven Drug Delivery Systems for Herbal Therapeutics: A New Era in Phytopharmaceutical Research

Dr Dhaval Rajde[1*], Dr Anant Kumar Meena[1], Dr Sumegha[1], Dr Jahanvee Gautam[1], Dr Sweety Singh[1]

[1]PG Scholar, PG department of Agadtantra, Post Graduate Institute of Ayurveda, DSRRAU, Jodhpur
Email: dhaval3r.dr@gmail.com
[1]PG Scholar, PG department of Dravyaguna Vigyan, Post Graduate Institute of Ayurveda, DSRRAU, Jodhpur
Email: anantmeenabhu@gmail.com
[1]PG Scholar, PG department of Dravyaguna Vigyan, Post Graduate Institute of Ayurveda, DSRRAU, Jodhpur
Email: meghaarora0406@gmail.com
[1]PG Scholar, PG department of Dravyaguna Vigyan, Post Graduate Institute of Ayurveda, DSRRAU, Jodhpur
Email: jahanveegautam1998@gmail.com
[1]PG Scholar, PG department of Ras Shastra & Bhaisajya Kalpana, Post Graduate Institute of Ayurveda, DSRRAU, Jodhpur
Email: sweetysingh21699@gmail.com

Corresponding author: Dr Dhaval Rajde, Email: dhaval3r.dr@gmail.com

ABSTRACT

Herbal therapeutics offer substantial pharmacological potential, yet their translation into clinically reliable phytopharmaceutical products remains constrained by poor aqueous solubility, low and variable bioavailability, chemical instability, batch-to-batch phytochemical variation, limited pharmacokinetic predictability and insufficient site-specific delivery. This structured narrative review examines how artificial intelligence (AI) can guide the rational development of herbal drug delivery systems. Peer-reviewed literature was identified through targeted, iterative searches of PubMed and major publisher platforms, supplemented by backward reference tracking and authoritative regulatory documents. Searches covered January 2010 to 20 June 2026; earlier landmark studies were retained when foundational to formulation optimization, controlled release or phytochemical delivery. Evidence was organized across four linked methodological layers: machine-learning and deep-learning approaches for phytochemical prioritization; prediction of physicochemical, pharmacokinetic, toxicity and drug-carrier interaction profiles; AI-assisted optimization of phytosomes, liposomes, polymeric nanoparticles, nanocrystals, lipid carriers, nanoemulsions and hydrogels; and future adaptive systems supported by explainable AI, digital twins, reinforcement learning and digital health data. The synthesis shows that AI can reduce experimental burden, expose nonlinear formulation relationships and support multi-objective decisions involving particle attributes, loading, release, stability, manufacturability and safety. However, direct studies integrating AI with herbal delivery remain limited; much of the current rationale is built by connecting mature evidence from natural-product delivery with rapidly advancing AI-enabled pharmaceutical formulation. A proposed phytochemical-carrier-patient digital thread is therefore introduced to preserve traceability from botanical batch identity to product quality, biological disposition and patient context. Translation will require standardized phytochemical datasets, external and prospective validation, uncertainty reporting, explainable and change-controlled models, quality-by-design integration and clinically meaningful performance benchmarks. AI-driven herbal drug delivery should consequently be viewed not as an automated replacement for experimentation, but as a transparent decision-support system capable of moving phytopharmaceutical development from empirical screening toward predictive, adaptive and precision-oriented design.

Keywords: AI-assisted formulation design, adaptive drug delivery, computational pharmaceutics, controlled drug release, deep learning, digital health, digital twin technology, explainable artificial intelligence

How to cite this article: Rajde D, Meena AK, Sumegha, Gautam J, Singh S. AI-Driven Drug Delivery Systems for Herbal Therapeutics: A New Era in Phytopharmaceutical Research. Int J Drug Deliv Technol. 2026;16(71s): 1419-1434. DOI: 10.25258/ijddt.16.71s.167

ABBREVIATIONS

ADMET, absorption, distribution, metabolism, excretion and toxicity; AI, artificial intelligence; ANN, artificial neural network; API, active pharmaceutical ingredient; CQA, critical quality attribute; DoE, design of experiments; FAIR, findable, accessible, interoperable and reusable; GMLP, good machine-learning practice; ML, machine learning; NLC, nanostructured lipid carrier; PBPK, physiologically based pharmacokinetic; PDI, polydispersity index; QbD, quality by design; QSPR, quantitative structure-property relationship; QTPP, quality target product profile; RL, reinforcement learning; SHAP, Shapley additive explanations; SLN, solid lipid nanoparticle; XAI, explainable artificial intelligence.

1. INTRODUCTION

Natural products remain a major source of pharmacologically active scaffolds, therapeutic leads and clinically used medicines. Modern natural-product research has expanded beyond single-compound isolation to include dereplication, network pharmacology, metabolomics and data-driven prioritization, while the historical contribution of natural products to drug discovery remains substantial [1,2]. Herbal medicines occupy a particularly challenging position within this landscape because the therapeutic material may be a purified phytoconstituent, a standardized fraction or a multicomponent extract whose activity depends on composition, processing and biological context. Their widespread use is accompanied by well-recognized concerns regarding variable quality, contamination, interactions, inconsistent pharmacokinetics and incomplete safety surveillance [3].

The scientific modernization of herbal medicine therefore requires more than demonstrating biological activity. It requires reproducible botanical identity, chemical characterization, pharmacological plausibility, dose justification and a dosage form capable of delivering relevant constituents to the intended site at a controlled exposure. Contemporary analytical and molecular approaches have begun to demystify traditional preparations, but translation is still limited when poorly soluble, unstable or extensively metabolized phytochemicals are administered in conventional formulations [4]. Artificial intelligence is increasingly important in natural-product discovery because it can organize high-dimensional chemical and biological information, prioritize structures, infer targets and guide experiments [5]. Yet discovery is only one part of the translational chain. A promising molecule or extract can still fail if its formulation cannot achieve consistent product quality and exposure.

Quality control is unusually consequential for phytopharmaceuticals. Species substitution, geographical origin, harvesting conditions, extraction method, storage and degradation can alter the chemical fingerprint before formulation begins. Robust quality systems consequently depend on authenticated raw materials, validated analytical methods, marker profiles and control of process variability [6]. Drug delivery introduces a second layer of complexity: excipients and carrier structures can change solubilization, intestinal transport, metabolism, tissue distribution, release kinetics and toxicity. Reviews of herbal delivery have described liposomes, phytosomes, nanoparticles, nanoemulsions, micelles and related systems as strategies to improve solubility, stability, bioavailability and targeting [7,8]. More recent assessments emphasize both opportunity and persistent barriers, including limited scale-up evidence, heterogeneity in characterization and uncertain regulatory pathways [9]. Bibliometric growth in herbal nanotechnology further demonstrates rapid expansion of the field, but publication volume alone does not establish product readiness [10].

The central proposition of this review is that AI should be used to connect these otherwise fragmented layers. Instead of optimizing particle size or entrapment efficiency in isolation, an AI-assisted program should link botanical batch information, phytochemical composition, carrier and process variables, product critical quality attributes, biological disposition and the characteristics of the intended patient population. This integrated view is termed the phytochemical-carrier-patient digital thread. It is designed to answer a practical question: given a defined herbal cargo, route, disease context and manufacturing constraint, which formulation is most likely to provide reproducible exposure, safety and therapeutic performance—and how certain is that prediction?

Accordingly, this article reviews the enabling evidence for AI-assisted formulation design, smart nanocarriers, predictive pharmacokinetics, XAI, digital twins and adaptive delivery. It gives particular attention to the boundary between evidence and anticipation. Direct AI-driven development of herbal delivery systems is still sparse. The current field is better understood as a convergence of two comparatively mature domains: advanced delivery of phytochemicals and AI-enabled optimization of pharmaceutical formulations. The review therefore distinguishes (i) direct evidence in herbal formulation, (ii) adjacent evidence from non-herbal formulation and nanomedicine, and (iii) plausible but not yet routine applications such as patient-specific digital twins and RL-controlled therapy.

2. REVIEW METHODOLOGY AND EVIDENCE FRAMEWORK

A structured narrative-review approach was selected because the literature spans heterogeneous evidence types—including natural-product chemistry, formulation science, nanomedicine, machine learning, pharmacokinetics, digital health and regulatory science—and because direct comparative studies suitable for meta-analysis are not yet available. Reporting was informed by the SANRA principles for narrative reviews, particularly justification of importance, statement of aims, description of literature searching, scientific reasoning and appropriate presentation of endpoint data [11]. This article is not presented as a registered systematic review and does not claim exhaustive retrieval.

Targeted and iterative searches were conducted in PubMed and across major publisher platforms, with backward reference tracking from relevant reviews and primary studies. The main time window was January 2010 to 20 June 2026. Earlier studies were retained when they established foundational concepts in neural-network formulation optimization, controlled drug release, nanocarrier engineering or the pharmacokinetic limitations of representative phytochemicals. Search concepts combined terms related to artificial intelligence, machine learning, deep learning, predictive modeling, neural networks, explainable AI, digital twins and reinforcement learning with terms related to herbal medicines, phytopharmaceuticals, phytochemicals, formulation optimization, drug delivery, nanocarriers, bioavailability, controlled release, targeted delivery and pharmacokinetics. Authoritative documents from the US Food and Drug Administration (FDA), European Medicines Agency (EMA) and International Council for Harmonisation (ICH) were used for the regulatory and quality discussion.

English-language peer-reviewed reviews, mechanistic studies, formulation studies and high-relevance methodological papers were prioritized. Studies were included when they provided evidence for at least one of four functions: prediction of phytochemical or drug properties; formulation or nanocarrier optimization; evaluation of delivery-related pharmacokinetics or biological performance; or governance of AI models intended to support pharmaceutical decisions. Commercial claims without peer-reviewed support, duplicative summaries and studies lacking meaningful relevance to formulation or delivery were not used as core evidence. Because reported endpoints and systems were highly heterogeneous, results were synthesized conceptually rather than statistically.

Evidence was interpreted in three tiers. Tier 1 comprised direct evidence in herbal or phytochemical delivery. Tier 2 comprised enabling evidence from AI-directed pharmaceutical formulations, drug nanoparticles and nanomedicine platforms that could reasonably transfer to herbal systems. Tier 3 comprised prospective applications—especially digital twins, digital-health feedback and reinforcement learning—that are technically plausible but require substantial preclinical, clinical and regulatory validation. This tiering is used throughout the review to prevent extrapolation from being misrepresented as established clinical practice.

3. WHY HERBAL THERAPEUTICS REQUIRE ADVANCED DELIVERY SYSTEMS

3.1 PHYSICOCHEMICAL AND BIOPHARMACEUTICAL BARRIERS

Many pharmacologically interesting phytochemicals are lipophilic, crystalline, chemically labile or poorly permeable. Low aqueous solubility can make dissolution the rate-limiting step for oral absorption, while degradation by light, oxygen, pH or enzymes can reduce the amount reaching systemic circulation. Polyphenols may undergo rapid conjugation; alkaloids may experience transporter-mediated efflux or extensive first-pass elimination; volatile constituents may be lost during processing; and multi-component extracts may contain compounds with divergent solubility and stability requirements. A single carrier must therefore often manage competing functions: protect the cargo, improve dispersion, sustain or trigger release, preserve activity and remain manufacturable.

The delivery problem is not fully described by a mean bioavailability value. Variability among batches and patients is equally important. A formulation that improves average exposure but amplifies between-batch variability may not represent a meaningful advance. Useful endpoints should therefore include dispersion stability, dissolution and release under biorelevant conditions, permeability, metabolic stability, food effects, exposure variability, tissue distribution and the relationship between exposure and pharmacodynamic response. AI is valuable when these endpoints depend on nonlinear interactions among multiple excipients, process settings and cargo descriptors.

3.2 COMPOSITIONAL VARIATION AND STANDARDIZATION

A purified phytoconstituent can be represented by a conventional molecular descriptor set, but a standardized extract requires a richer representation. Minimum inputs may include botanical taxonomy, plant part, geographic origin, extraction solvent, chromatographic fingerprint, marker concentrations, total phenolic or alkaloid content, residual solvent and water content. Spectroscopic or chromatographic embeddings may be more informative than a small number of marker compounds when minor constituents influence solubility, membrane transport or colloidal stability. This creates an opportunity for multimodal learning, but only when analytical data and metadata are consistently generated.

The formulation itself can modify the apparent composition. Selective loading into a lipid phase, adsorption to a polymer, degradation during homogenization or differential release may enrich some constituents and deplete others. Consequently, the relevant product identity is not merely the starting extract fingerprint; it is the fingerprint of the delivered fraction over time. A credible AI workflow should model this transformation and use mass balance, recovery and stability data to preserve traceability.

3.3 PHARMACOKINETIC, TARGETING AND SAFETY BARRIERS

Low oral exposure may arise from poor dissolution, low permeability, metabolism in the gut wall or liver, active efflux, binding to luminal contents or rapid systemic clearance. These mechanisms require different interventions. Increasing dissolution cannot overcome complete metabolic extraction, and permeation enhancement may increase exposure to both active and toxic constituents. Targeted delivery introduces further biological barriers: protein adsorption, mucosal clearance, endosomal trapping, mononuclear phagocyte uptake and heterogeneous disease vasculature. The same carrier attributes that improve cellular uptake in vitro can reduce circulation time or increase off-target accumulation in vivo.

Safety assessment must also be formulation-specific. Nanocarriers can change the tissue exposure of compounds that were previously poorly absorbed. Excipients, degradation products, surface charge and repeated dosing may create toxicological risks not predicted from the free phytochemical. For multicomponent products, herb-drug interactions and effects on metabolic enzymes or transporters remain important. Intelligent delivery must therefore optimize benefit and risk jointly rather than treating bioavailability enhancement as an unconditional objective.

Table 1. Major barriers in herbal drug delivery and the corresponding AI-addressable formulation questions.

Barrier	Therapeutic consequence	Minimum useful data	AI-addressable task
Poor aqueous solubility and slow dissolution	Low or variable oral exposure; high dose requirement	Molecular descriptors, solid state, pH-solubility, dissolution profiles, excipient properties	Predict solubility and excipient compatibility; select carrier class; optimize composition and processing.
Chemical or enzymatic instability	Loss of potency during manufacture, storage or transit	Degradation kinetics, temperature/light/humidity history, antioxidant and packaging variables	Model shelf-life and degradation risk; identify protective formulation and process conditions.
Batch-to-batch phytochemical variation	Inconsistent loading, release, pharmacology and safety	Authenticated source metadata, chromatographic/spectroscopic fingerprints, marker concentrations	Detect atypical batches; predict formulation performance from chemical fingerprint; support adaptive process settings.
Low permeability, efflux or first-pass metabolism	Exposure remains low despite dissolution improvement	Permeability, transporter, microsomal/hepatocyte, metabolite and pharmacokinetic data	Classify limiting mechanism; predict ADMET; prioritize lymphatic, mucosal, prodrug or local-delivery approaches.
Complex drug-carrier interactions	Unexpected loading, aggregation or burst release	Carrier chemistry, surface properties, interaction energies, particle attributes and release profiles	Learn nonlinear composition-property relationships and identify robust design spaces.
Need for targeted or controlled release	Off-target toxicity, short duration or poor tissue concentration	Ligand density, receptor expression, release under relevant stimuli, biodistribution and response	Optimize targeting and release jointly while accounting for uncertainty and biological variability.
Scale-up and manufacturing variability	Laboratory performance is not reproduced at pilot or commercial scale	Equipment geometry, energy input, mixing, temperature, in-process controls and CQAs	Develop scale-aware models, soft sensors and real-time release-support tools.
Patient heterogeneity and herb-drug interactions	Variable response and preventable safety events	Demographics, organ function, genotype/phenotype, co-medications, biomarkers and adherence	Support exposure-response stratification and supervised personalization within validated boundaries.

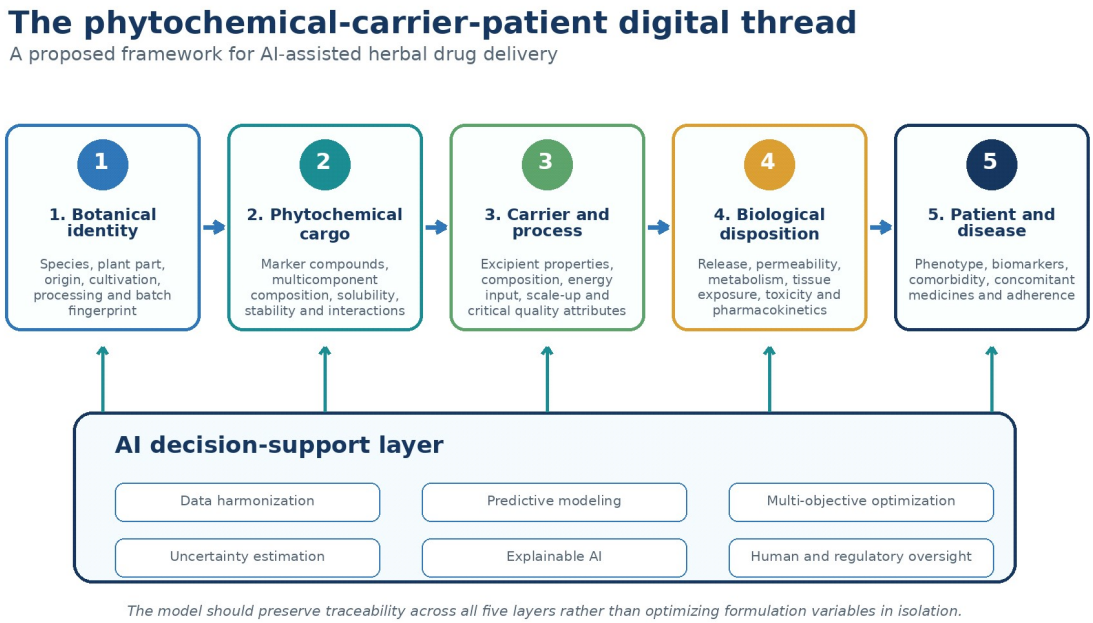

Figure 1. The proposed phytochemical-carrier-patient digital thread. AI becomes most useful when botanical identity, phytochemical composition, carrier/process variables, biological disposition and patient context remain linked and auditable. The framework is conceptual and intended to guide study design.

4. ARTIFICIAL-INTELLIGENCE TOOLBOX FOR COMPUTATIONAL PHARMACEUTICS

4.1 FROM PREDICTION TO PHARMACEUTICAL DECISION SUPPORT

AI in drug delivery should be defined by the decision it supports, not by the novelty of the algorithm. In formulation development, useful decisions include choosing a carrier family, selecting excipients, setting component ratios, predicting product attributes, estimating stability, prioritizing experiments and identifying a robust operating region. Machine-learning-directed formulation has emerged as a coherent field because algorithms can learn complex relationships between molecular, material and process descriptors and experimentally measured outputs [12]. Broad reviews of pharmaceutical technology similarly describe applications in preformulation, dosage-form design, manufacturing, quality control and personalized delivery [13–15].

Machine learning spans linear and nonlinear regression, tree ensembles, support-vector methods, Gaussian processes and neural networks. Deep learning is advantageous when inputs are high-dimensional—such as molecular graphs, spectra, microscopy images or complex release curves—but typically requires more data and stronger regularization. In drug discovery and development, ML has demonstrated value in property prediction, target identification, trial design and pharmacovigilance, while also exposing recurrent risks involving bias, data leakage and limited prospective validation [16,17]. For herbal delivery, the immediate opportunity is not to deploy the largest model; it is to match model complexity to the size, diversity and intended use of the formulation dataset.

4.2 DATA REPRESENTATION ACROSS THE DIGITAL THREAD

A model can only learn what is represented. For a purified phytochemical, inputs may include molecular weight, hydrogen-bonding features, ionization, logP/logD, topological polar surface area, aromaticity, flexibility, melting point, crystal descriptors and degradation liabilities. For extracts, chemical fingerprints, peak tables or learned spectral embeddings are required alongside source and process metadata. Carrier descriptors may include lipid transition temperature, polymer molecular weight, hydrophilic-lipophilic balance, charge, pKa, degradation rate, ligand density and excipient compatibility. Process descriptors include mixing order, shear, temperature, solvent removal, flow rate, energy density, drying and sterilization.

Outputs should reflect the intended decision. Frequently used formulation outputs include particle size, PDI, zeta potential, drug loading, encapsulation efficiency, yield, morphology, release parameters, storage stability, permeability, cytotoxicity and bioavailability. Multi-output models are especially relevant because these endpoints are coupled. For example, increasing lipid content may improve loading but enlarge particles, change digestion and complicate scale-up. A single-objective model can therefore recommend an apparently optimal but practically unusable formulation.

4.3 PROPERTY PREDICTION AND ADMET TRIAGE

Open computational platforms illustrate how physicochemical and ADMET prediction can inform early decisions. SwissADME integrates descriptors and models relevant to solubility, lipophilicity, pharmacokinetics and drug-likeness [18]; admetSAR 2.0 and ADMETlab 2.0 provide broader computational prediction of absorption, distribution, metabolism, excretion and toxicity endpoints [19,20]. Classical models such as ESOL show that interpretable molecular descriptors can support aqueous-solubility estimates [21], while the widely used Lipinski framework illustrates the value—and limitations—of simple property rules in anticipating oral exposure [22]. Deep neural networks and large benchmarking studies have demonstrated improved performance for some QSAR tasks, but performance is dataset- and endpoint-dependent [23,24].

For herbal therapeutics, these tools are best used for triage and hypothesis generation. Predictions for molecules outside a model's applicability domain, metabolites, ionizable polyphenols or complex mixtures may be unreliable. An AI-assisted program should retain the predicted value, uncertainty, model version and relevant experimental confirmation. When a delivery system changes the absorbed species, disposition model inputs must also be updated rather than assuming the pharmacokinetics of the unformulated compound.

Table 2. AI methods relevant to herbal formulation and drug-delivery development.

Method	Typical inputs	Principal use	Main limitation or safeguard
Linear / regularized regression	Curated formulation and process variables	Transparent baseline models; effect estimation; small datasets	May miss nonlinear interactions; retain as a benchmark before complex modeling.
Decision trees and ensemble methods	Mixed numerical and categorical descriptors	Property prediction, ranking, nonlinear interactions and feature importance	Can overfit small datasets; use nested validation and test on new batches.
Gaussian-process regression	Small to medium continuous design spaces	Prediction with uncertainty; Bayesian optimization and active learning	Kernel assumptions and scaling; verify calibrated uncertainty.
Artificial neural networks	Formulation composition, process settings and product attributes	Nonlinear mapping, multi-output prediction and release-profile modeling [25–29]	Data-hungry and sensitive to tuning; compare with simpler models.
Deep learning / graph neural networks	Molecular graphs, spectra, images or large structured datasets	Learned representations for property, interaction and image-based quality prediction	Limited interpretability and transfer outside training chemistry.
Unsupervised learning	Chemical fingerprints, spectra, batches and product profiles	Batch clustering, anomaly detection, material-space mapping and formulation families	Clusters are not automatically mechanistic or clinically meaningful.
Bayesian optimization / active learning	Predictive model plus objective and uncertainty	Select the next most informative or promising experiment; reduce screening burden	Objective function must include safety, robustness and manufacturability.
Mechanistic-ML hybrid models	Mass balance, release, PBPK or process equations plus data	Combine scientific constraints with residual learning and improve extrapolation	Requires careful identifiability, unit consistency and model governance.
Explainable AI	Trained model, features and predictions	Global and local feature attribution, error analysis and decision communication	Post-hoc explanations do not prove causality; test proposed mechanisms experimentally.
Reinforcement learning	State, action, reward and transition model	Sequential experimental design or adaptive control within constrained systems	Unsafe exploration and reward misspecification; require hard constraints and oversight.

5. AI-ASSISTED FORMULATION DESIGN: A PRACTICAL DEVELOPMENT PATHWAY

5.1 LESSONS FROM EARLY NEURAL-NETWORK FORMULATION STUDIES

The use of AI in formulation is not new. Early work showed that artificial neural networks could approximate nonlinear formulation-response relationships and support optimization [25]. Applications to transdermal systems demonstrated the possibility of linking formulation variables with release or permeation outcomes [26]. Subsequent studies used neural networks to optimize chitosan microparticles, sustained-release matrix tablets and controlled-release systems [27–29]. These studies established a durable principle: formulation development can be treated as a supervised learning problem when inputs and outputs are generated under controlled conditions.

The limitations of early studies are equally instructive. Datasets were often small, formulation spaces narrow and validation largely internal. A model may interpolate accurately within a designed experiment yet fail when the botanical batch, excipient grade, equipment or scale changes. Modern workflows should retain the efficiency of neural-network optimization while adding external validation, uncertainty estimates, prospective experiment selection and explicit applicability domains.

5.2 HIGH-THROUGHPUT AND MACHINE-LEARNING-GUIDED NANOMEDICINE DESIGN

High-throughput experimentation and ML now permit broader exploration of material combinations. Quantitative prediction of self-assembly has been used to identify targeted nanomedicines from chemical and biological design variables [30]. High-throughput screening combined with ML has mapped nanomedicine design spaces and linked composition to biological performance [31]. Computationally guided screening has also accelerated the discovery of self-assembling drug nanoparticles [32]. For long-acting injectables, ML models have been used to relate drug and polymer descriptors to release behavior and formulation performance [33]. More recent work has formalized ML-guided high-throughput nanoparticle design as an iterative platform rather than a one-off prediction exercise [34].

These studies are primarily enabling evidence rather than direct demonstrations in herbal products. Nevertheless, the transferable methodological elements are strong: standardized high-throughput preparation, automated characterization, material descriptors, prospective prediction and experimental confirmation. Herbal delivery programs can adopt the same architecture while adding batch fingerprints, multi-analyte loading and stability endpoints.

5.3 DEFINING THE TARGET AND THE OBJECTIVE FUNCTION

The first step is a formulation question that can be answered. A quality target product profile should define the intended route, dosage form, release pattern, strength, stability, target population and key safety or usability constraints. The corresponding objective function should not merely maximize encapsulation efficiency. A clinically relevant objective may seek adequate exposure with low variability, acceptable particle attributes, controlled release, minimal cytotoxicity, long-term stability, feasible cost and resilience to raw-material variability.

Multi-objective optimization should represent trade-offs through a Pareto frontier rather than collapsing every endpoint into an arbitrary weighted score. Weighting may still be used, but weights should be justified by product risk and clinical need. Hard constraints are appropriate for unacceptable toxicity, particle-size limits, residual solvent, dose volume or manufacturability. This prevents the algorithm from trading safety for marginal performance.

5.4 DATASET CONSTRUCTION AND FEATURE ENGINEERING

A useful dataset combines planned variation with realistic sources of variability. DoE remains valuable because it produces balanced, information-rich observations and identifies interactions. Historical data can enlarge the sample, but uncontrolled changes in methods, instruments or materials may create hidden batch effects. Each record should therefore include metadata for botanical batch, excipient supplier and grade, equipment, analytical method, operator or site where relevant, and time between manufacture and measurement.

Feature engineering should be scientifically grounded. Ratios, molar fractions, hydrophilic-lipophilic balance, drug-to-carrier affinity, energy per unit volume, supersaturation indices, dimensionless release parameters and stability slopes may be more transferable than raw recipe quantities. For mixtures, dimension reduction should preserve chemically meaningful variation. Spectral embeddings can be useful, but model interpretation should be related back to identifiable peaks or chemical regions when possible.

5.5 MODEL DEVELOPMENT, VALIDATION AND APPLICABILITY

Data splitting is a major determinant of credibility. Random splitting can overestimate performance when closely related formulations, replicate measurements or the same botanical batch appear in training and test sets. More demanding splits—by batch, scaffold, excipient family, time or site—better approximate future use. Hyperparameter tuning should be nested within cross-validation, with a final untouched test set used once. Metrics should match the decision: mean error alone may hide clinically important tail errors, while classification accuracy may be misleading for imbalanced failure outcomes.

Uncertainty should accompany every recommendation. Gaussian processes, ensembles, conformal prediction or calibrated predictive intervals can identify regions where the model is unsure. The applicability domain should specify the ranges and combinations of phytochemical, material and process variables represented in development. Recommendations outside that domain should trigger additional experiments rather than confident automation.

5.6 ACTIVE LEARNING AND THE MODEL-EXPERIMENT LOOP

Active learning selects experiments expected to be most informative; Bayesian optimization balances exploration of uncertain regions with exploitation of predicted high-performing regions. In herbal formulation, this is valuable when each experiment requires costly phytochemical analysis, stability testing or animal work. The algorithm can propose a small set of diverse candidates, but experimental scientists must screen proposals for feasibility and mechanistic plausibility. Failed formulations are informative and should be retained rather than removed from the dataset.

The end point of an AI cycle is not a predicted optimum but an independently manufactured and characterized product. Confirmatory batches should test robustness to reasonable variation and, where possible, be prepared by a different operator, scale or site. New observations update the model under version control. The closed-loop architecture shown in Figure 2 captures this principle: the model proposes, experiments adjudicate and governance controls each update.

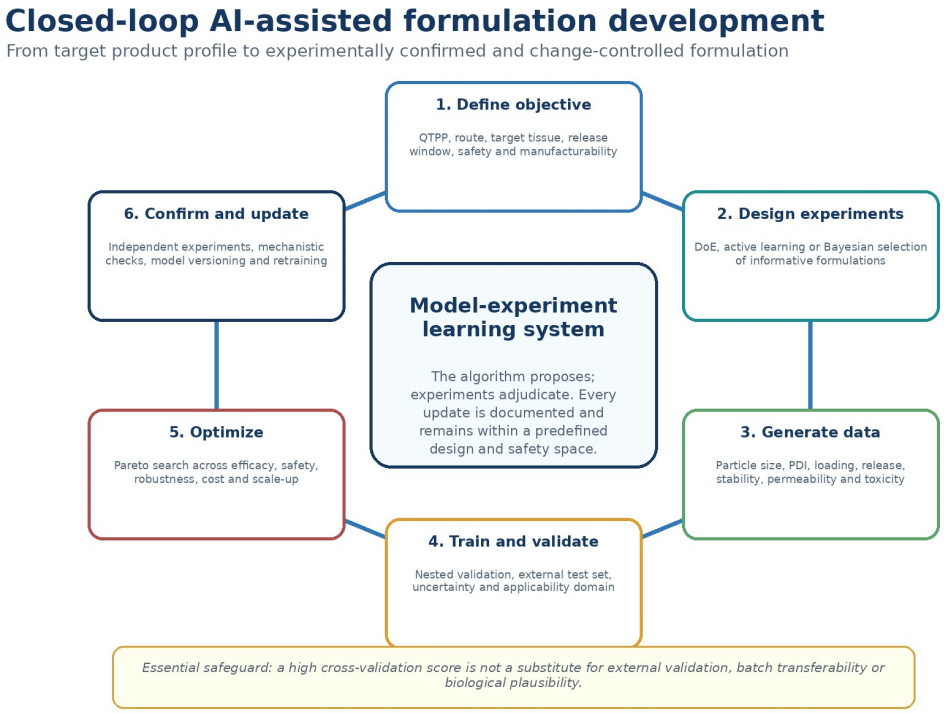

Figure 2. Closed-loop AI-assisted formulation development. A credible workflow combines QTPP definition, informative experimentation, predictive modeling, uncertainty-aware multi-objective optimization and independent confirmation. The loop is repeated under model and data version control.

6. AI-ENABLED SMART NANOCARRIERS FOR HERBAL THERAPEUTICS

Nanocarrier engineering seeks to control the interaction of a therapeutic with biological barriers by manipulating size, shape,

composition, surface chemistry, mechanics and release. Precision nanoparticle design has advanced through an increasingly quantitative understanding of how these attributes influence disposition [35]. General principles include protection of cargo, avoidance or exploitation of biological clearance pathways, controlled interaction with cells and tissues, and release at a therapeutically useful rate [36]. Nanocarriers have become important platforms for targeted therapy [37], and stimuli-responsive systems can couple release to pH, redox conditions, enzymes, temperature or external energy [38].

Translation remains difficult. Analyses of tumor delivery have shown that only a limited fraction of administered nanoparticles may reach solid tumors, underscoring the importance of realistic biodistribution objectives [39]. Clinical translation reviews and surveys of approved nanomedicines show that manufacturability, reproducibility, safety and regulatory clarity are as important as laboratory efficacy [40,41]. Broad reviews of nano-based delivery likewise emphasize the need for scalable preparation and biologically relevant testing [42,43]. These lessons apply directly to herbal nanocarriers, where variability in cargo composition adds another source of uncertainty.

6.1 POLYMERIC NANOPARTICLES, NANOCRYSTALS AND NANOSUSPENSIONS

Biodegradable polymeric nanoparticles can provide cargo protection, sustained release and surface functionalization [44]. For phytochemicals, AI can predict polymer-cargo compatibility, solvent selection, precipitation behavior, particle size, loading and release. The most useful dataset combines molecular and polymer descriptors with process variables such as mixing intensity, phase ratio, addition rate and solvent removal. Mechanistic features—such as predicted interaction energies or polymer glass transition behavior—can improve transferability.

Nanocrystals and nanosuspensions increase the apparent dissolution rate of poorly soluble compounds by reducing particle size and increasing surface area without requiring a large carrier mass. Herbal medicine nanocrystals have been proposed as a versatile strategy for poorly soluble actives and extracts [45]. AI can support stabilizer selection, predict aggregation risk, optimize milling or precipitation parameters and model dissolution under biorelevant conditions. For multicomponent extracts, selective crystallization and differential dissolution must be monitored because the final nanosuspension may not retain the starting chemical ratio.

6.2 LIPOSOMES AND PHYTOSOMES

Liposomes offer aqueous and lipid compartments and can encapsulate cargos with different physicochemical properties. Their classification, preparation and applications are well established [46]. Relevant AI inputs include phospholipid composition, cholesterol fraction, phase-transition temperature, charge, hydration conditions and size-reduction process; outputs include vesicle size, PDI, loading, leakage, oxidation and release. Models should account for storage and biological media because colloidal properties measured immediately after preparation may not predict performance after protein adsorption or digestion.

Phytosomes are phospholipid complexes designed to improve the compatibility and absorption of selected phytochemicals. Their use as topical and systemic delivery platforms has received increasing attention [47]. Unlike a conventional liposome, the phytochemical-phospholipid interaction itself may be central to product identity. Spectroscopy, thermal analysis and molecular simulation can provide features for predicting complex formation, while AI can optimize stoichiometry, solvent, drying and reconstitution. However, a high apparent complexation yield is not sufficient; comparative pharmacokinetics and stability remain necessary.

6.3 SOLID LIPID NANOPARTICLES, NANOSTRUCTURED LIPID CARRIERS AND NANOEMULSIONS

SLNs and NLCs can improve the dispersion of lipophilic phytochemicals, provide controlled release and support oral, dermal or parenteral administration. Formulation variables include solid/liquid lipid composition, surfactant system, phase ratio, temperature history and homogenization energy. AI can learn nonlinear effects on size, crystallinity, loading, expulsion during storage and release. Because lipid digestion controls oral performance, in vitro lipolysis and intestinal models should be included alongside simple buffer release.

Nanoemulsions and self-emulsifying systems are particularly relevant for oral delivery. The composition space is high-dimensional and phase behavior is nonlinear, making it suitable for Bayesian optimization or tree-based modeling. Useful objectives include droplet size after dilution, precipitation risk, digestion, permeability, dose volume and excipient acceptability. AI should not merely reproduce a ternary phase diagram; it should integrate dynamic dispersion and biopharmaceutical endpoints.

6.4 HYDROGELS AND STIMULI-RESPONSIVE SYSTEMS

Hydrogels can localize herbal actives, support mucosal or dermal administration and enable sustained release. Model inputs may include polymer concentration, cross-link density, pH, ionic strength, swelling, rheology and cargo interactions. Release curves can be treated as time-series outputs or summarized through mechanistically relevant parameters. Stimuli-responsive hydrogels add a control layer, but trigger sensitivity must be tested across physiological variability. AI can identify formulations that achieve a target release window while maintaining mechanical integrity and biocompatibility.

6.5 TARGETED AND CONTROLLED DELIVERY

Controlled delivery aims to regulate the rate, duration and location of exposure. Reviews of controlled-release systems emphasize the long-standing contribution of material design, diffusion, degradation and device architecture [48,49]. AI can accelerate this work by predicting release profiles, learning residual deviations from mechanistic models and identifying process parameters that control burst release. Targeting requires additional variables such as ligand density, affinity, orientation, receptor expression and competitive binding. Optimization should use biodistribution and functional response rather than cell uptake alone.

The rapid development of lipid nanoparticles for nucleic-acid delivery illustrates how composition, structure, manufacturing and biological performance can be systematically linked [50]. Herbal delivery is not equivalent to mRNA delivery, but the platform lesson is relevant: standardizable material libraries, high-throughput analytics and rigorous structure-function studies enable data-driven design. The same infrastructure can be adapted to phytochemical cargos provided that botanical and chemical variability is represented explicitly.

Table 3. Delivery platforms, critical quality attributes and high-value AI tasks for herbal therapeutics.

Platform	Critical attributes	High-value AI tasks	Translational caution
Phytosomes / phospholipid complexes	Complex stoichiometry, association, particle properties, oxidation, dissolution and permeability	Predict phytochemical-phospholipid compatibility; optimize solvent and ratio; link analytical signatures to exposure	Complexation evidence should be distinguished from simple physical mixing; human PK is often limited.
Liposomes	Size, lamellarity, PDI, surface charge, loading, leakage, lipid oxidation and release	Select lipids; optimize composition and processing; predict storage leakage and biological-media behavior	Sterility, scale-up, lipid quality and in vivo stability may dominate translation.
Polymeric nanoparticles	Polymer-cargo interaction, size, loading, residual solvent, degradation and release	Material selection; process optimization; hybrid mechanistic-ML release modeling	Polymer grade, solvent removal and scale-dependent mixing require explicit metadata.
Nanocrystals / nanosuspensions	Crystal form, size distribution, surface coverage, aggregation, redispersibility and dissolution	Stabilizer selection; milling/precipitation optimization; predict aggregation and dissolution	Enhanced dissolution may not overcome metabolism; composition of multicomponent cargo may shift.
SLNs / NLCs	Lipid polymorphism, particle size, loading, expulsion, digestion and release	Optimize lipid blend and surfactant; predict storage and lipolysis performance	Apparent loading can decline during lipid recrystallization; biorelevant digestion testing is essential.
Nanoemulsions / self-emulsifying systems	Phase behavior, droplet size after dilution, precipitation, digestion and dose volume	Map composition space; forecast dilution robustness; multi-objective oral performance optimization	High surfactant exposure, food effects and precipitation after dispersion must be addressed.
Hydrogels / local depots	Rheology, gelation, swelling, adhesion, degradation, dose uniformity and release	Predict gelation and release; optimize mechanics and retention; detect batch variability	In vitro release may not represent tissue turnover, enzymes or fluid dynamics.
Ligand-targeted or stimuli-responsive carriers	Ligand density, trigger threshold, biodistribution, off-target uptake and release	Jointly optimize targeting and release; integrate receptor and disease-state data	Cell-uptake gains may not translate to organism-level targeting; complexity increases CMC burden.

7. REPRESENTATIVE PHYTOCONSTITUENTS: EVIDENCE AND AI-ENABLED NEXT STEPS

Representative phytochemicals illustrate both the promise of advanced delivery and the limits of current evidence. Most published studies optimize one platform through conventional experimentation. They provide valuable mechanistic and pharmacokinetic data, but they rarely establish a reusable AI model, external validation or patient-specific adaptation. The cases below should therefore be viewed as data-rich starting points for intelligent formulation rather than examples of mature autonomous development.

7.1 CURCUMIN

Curcumin is a canonical delivery challenge because of poor aqueous solubility, chemical instability, rapid metabolism and low systemic exposure. Its bioavailability limitations have been extensively reviewed [51]. Co-administration with piperine was reported to increase systemic availability, illustrating the importance of metabolic and transport mechanisms in addition to dissolution [52]. A curcumin-phospholipid complex improved pharmacokinetic exposure and pharmacological performance in preclinical work [53]. Nanoparticle encapsulation produced a marked increase in oral bioavailability compared with conventional curcumin preparations in rats [54], and alternative colloidal preparations have also been developed to improve oral exposure [55].

An AI-enabled curcumin program could combine solid-state descriptors, degradation kinetics, carrier composition, digestion, metabolite formation and pharmacokinetic data. The objective should be a reproducible exposure-response profile rather than the largest fold-increase relative to an exceptionally poor control. External validation should test different curcumin sources and analytical methods, since apparent performance can be confounded by purity, degradation and assay sensitivity.

7.2 BERBERINE

Berberine has very low oral bioavailability and exhibits substantial intestinal first-pass elimination and hepatic distribution in preclinical studies [56]. A formulation that only increases dissolution may therefore have limited effect unless it also addresses permeability, efflux, metabolism or lymphatic transport. AI could classify which mechanism is dominant under a given formulation and use mechanistic-ML models to link in vitro dissolution, intestinal transport and metabolism with systemic and hepatic exposure. Because local gastrointestinal actions may contribute to efficacy, systemic bioavailability should not automatically be treated as the sole target.

7.3 RESVERATROL

Resveratrol formulation research has explored multiple strategies to overcome limited bioavailability, rapid metabolism and stability concerns [57]. The design space includes nanoparticles, lipids, inclusion systems and prodrug-like approaches. A useful model would predict not only parent exposure but also metabolite profiles, tissue distribution and the stability of the trans isomer. Multi-objective optimization is important because aggressive solubilization may not meaningfully improve active exposure if metabolism remains rapid.

7.4 SILYBIN AND SILYMARIN

Phosphatidylcholine complexation has long been used to improve the absorption of silybin. Pharmacokinetic studies of a silybin-phosphatidylcholine complex in healthy volunteers provided early evidence that formulation can alter exposure [58]. This system is well suited to retrospective data integration: phospholipid composition, complexation analytics, dissolution, food effects and human pharmacokinetics could be modeled jointly. The main data challenge is harmonizing older studies with modern analytical and product-characterization standards.

7.5 BOSWELLIC ACIDS

Boswellic acids are lipophilic constituents with challenging oral absorption. A lecithin-based delivery form of Boswellia extract

increased absorption in a clinical pharmacokinetic study [59]. For such extracts, AI should represent the full acid profile and not only a single marker. It could predict differential complexation and absorption across constituents and identify whether the formulation changes the internal ratio of delivered actives. This is particularly important when therapeutic activity may depend on more than one component.

7.6 QUERCETIN AND OTHER POLYPHENOLS

Quercetin, catechins and related polyphenols share recurring problems of solubility, instability, conjugative metabolism and variability in intestinal handling. Although each compound requires specific evidence, the transferable AI opportunity is a platform model trained across a chemically related family. Transfer learning or multitask models could estimate formulation compatibility and identify which observations can be shared across compounds. Such models must preserve compound-specific uncertainty and should not assume that all polyphenols behave equivalently.

Table 4. Representative phytoconstituents and the distinction between established delivery evidence and the next AI-enabled research step.

Phytoconstituent	Established delivery or PK evidence	AI-enabled next step	Current evidence status
Curcumin	Extensive evidence of poor bioavailability; piperine, phospholipid complexes and nanoparticles can alter exposure [51–55].	Integrate degradation, solid state, digestion, metabolites and PK; optimize robust exposure rather than a single in vitro endpoint.	Strong delivery literature; direct AI-guided herbal formulation evidence remains limited.
Berberine	Extensive intestinal first-pass elimination and hepatic distribution help explain very low plasma exposure [56].	Mechanism-classification model linking dissolution, permeability, efflux, metabolism and local versus systemic action.	Strong mechanistic PK evidence; platform selection remains context-dependent.
Resveratrol	Multiple formulation strategies have been proposed to address limited bioavailability and stability [57].	Predict parent and metabolite exposure; include isomer stability and tissue distribution in multi-objective optimization.	Substantial formulation literature, heterogeneous clinical relevance.
Silybin / silymarin	A phosphatidylcholine complex demonstrated altered human pharmacokinetics [58].	Link complexation analytics, formulation variables, food effects and human PK; harmonize historical datasets.	Human PK precedent; modern AI-ready datasets are scarce.
Boswellic acids	A lecithin delivery form increased absorption of boswellic acids [59].	Model the multicomponent acid profile and differential absorption rather than one marker alone.	Clinical PK evidence for a delivery form; limited generalizable formulation models.
Quercetin and related polyphenols	Broad evidence supports solubility, stability and metabolism challenges across the class.	Develop multitask or transfer-learning models across related compounds with compound-specific uncertainty.	Promising platform-level opportunity; requires standardized cross-compound data.

8. PREDICTIVE PHARMACOKINETICS, DIGITAL HEALTH AND PRECISION HERBAL MEDICINE

8.1 CONNECTING FORMULATION ATTRIBUTES TO EXPOSURE

A formulation model becomes clinically relevant only when product attributes are connected to biological exposure and response. The simplest bridge is an empirical relationship between dissolution or release and pharmacokinetic parameters. More informative approaches separate dissolution, permeability, presystemic metabolism, distribution and clearance. Mechanistic release or PBPK components can be combined with ML residual models when equations capture known physiology but systematic discrepancies remain. This hybrid approach is preferable to an unconstrained black box when extrapolation across dose, food state or patient characteristics is required.

For phytopharmaceuticals, the model may need to predict several parent compounds and metabolites simultaneously. Exposure metrics can include maximum concentration, area under the curve, time above a threshold, tissue concentration or local luminal concentration. The relevant metric depends on mechanism of action. A local gastrointestinal effect, for example, may be weakened rather than improved by maximizing systemic absorption. Target product design should therefore begin with a pharmacological hypothesis.

8.2 PRECISION HERBAL MEDICINE

Precision herbal medicine aims to match a standardized product and delivery strategy to a patient or biologically defined subgroup. Potential predictors include age, body size, organ function, inflammatory state, metabolic phenotype, microbiome features, genetic variants, diet, concomitant medicines and adherence. The complexity is substantial because patient variability interacts with botanical and formulation variability. The phytochemical-carrier-patient digital thread addresses this by treating product and patient information as linked inputs rather than separate analyses.

Early applications should focus on bounded decisions: identifying patients at risk of high or low exposure, selecting among validated dosage forms, or recommending additional monitoring. Claims of individualized dosing require prospective evidence, defined update rules and a clinically actionable endpoint. Sensitive attributes and health data also create privacy, fairness and cybersecurity obligations. Precision should not become a label for an opaque model trained on a small convenience sample.

8.3 DIGITAL HEALTH AS A FEEDBACK LAYER

Digital-health tools can supply adherence, symptom, activity, sleep, heart-rate or point-of-care measurements. In chronic conditions, these longitudinal signals could help distinguish formulation failure from non-adherence or disease fluctuation. They may also support exposure-response modeling when conventional clinic visits provide sparse observations. However, consumer-device data can contain missingness, drift and population bias. The measurement process, device version and

data-quality rules should be documented as rigorously as laboratory assays.

The immediate value of digital health is therefore monitoring and hypothesis generation. Closed-loop adaptation should be reserved for validated settings with clear clinical oversight. Herbal products may be perceived as intrinsically safe, but a system that changes dose or release based on noisy signals can create new risk. Human review, predefined action limits and the ability to revert to a safe baseline are essential.

9. EXPLAINABLE AI, UNCERTAINTY AND DATA GOVERNANCE

9.1 WHY EXPLAINABILITY MATTERS

Formulation decisions affect product quality and, ultimately, patient exposure. In high-stakes settings, reliance on a poorly understood black-box model can be inappropriate when an interpretable alternative performs adequately. Rudin argued that interpretable models should be preferred over post-hoc explanations for consequential decisions when possible [60]. In healthcare, explainability is also multidisciplinary: an explanation useful to a formulation scientist may differ from one needed by a regulator, clinician or patient [61].

For herbal delivery, explanations should operate at several levels. Global explanations identify variables that generally influence loading, size, release or stability. Local explanations show why a particular formulation was recommended. Error explanations identify when a model fails by botanical batch, chemistry, carrier family or process scale. Counterfactual explanations can indicate the smallest feasible change expected to move a formulation into an acceptable region. These outputs should support scientific review, not replace it.

9.2 SHAP, LIME AND MECHANISTIC INTERPRETATION

SHAP provides additive feature attributions based on cooperative-game concepts and is widely used to summarize global and local model behavior [62]. LIME approximates a complex prediction locally with an interpretable surrogate [63]. Both can reveal surprising dependencies, but neither establishes causality. Correlated formulation variables can share or redistribute attribution, and explanations may be unstable outside dense regions of the data. Proposed mechanisms—such as a surfactant effect on size or a polymer effect on release—must be tested experimentally.

9.3 UNCERTAINTY, ROBUSTNESS AND DRIFT

A responsible model communicates when it does not know. Predictive intervals, ensemble disagreement and distance from the training domain can inform whether an experimental confirmation is routine or urgent. Robustness analysis should perturb inputs within analytical and manufacturing error ranges. If a small, plausible measurement change causes a large recommendation shift, the model may be unsuitable for decision support.

Model drift can occur when raw-material sources, analytical methods, equipment or patient populations change. Monitoring should compare new input distributions, prediction residuals and outcome rates with the development baseline. Drift thresholds and retraining rules should be predetermined. A retrained model is a new version requiring verification; continuous learning should not mean uncontrolled change.

9.4 FAIR AND AI-READY PHYTOPHARMACEUTICAL DATA

The FAIR principles—findability, accessibility, interoperability and reusability—provide a useful foundation for data stewardship [64]. AI-ready formulation data require more than a spreadsheet of recipes and outcomes. Controlled vocabularies, units, analytical methods, material identifiers, missing-data codes, failed experiments, raw data provenance and links between batches must be retained. Negative results are particularly valuable because they define failure regions and reduce repeated experimentation.

Consortia and shared benchmark datasets could accelerate progress, but commercial formulations and botanical sources create intellectual-property and standardization challenges. Federated or privacy-preserving learning may eventually support cross-site models without pooling sensitive data. Even then, harmonized endpoints and reference materials are necessary; federating inconsistent measurements simply distributes inconsistency.

10. DIGITAL TWINS, REINFORCEMENT LEARNING AND ADAPTIVE DRUG DELIVERY

10.1 DIGITAL TWIN AS A LAYERED MODEL

A digital twin is a dynamic virtual representation connected to a physical system through data. In personalized medicine, digital-twin concepts have been proposed to integrate biological models and patient-specific information [65], while precision cardiology has provided a detailed vision of how virtual representations might support individualized decisions [66]. In silico clinical trials extend the same logic by using computational populations and disease models to supplement development evidence [67]. Pharmaceutical manufacturing literature has also described digital twins for process understanding, monitoring and control [68].

For herbal therapeutics, a credible digital twin would be layered rather than monolithic. A product twin would represent botanical batch, formulation attributes, release and stability. A physiological twin would represent absorption, metabolism, distribution and relevant disease processes. A patient-state layer would be updated by clinical and digital-health observations. The model would generate predictions with uncertainty and compare them with observed outcomes. This architecture is shown in Figure 3 as a research roadmap, not as a claim of current routine deployment.

10.2 REINFORCEMENT LEARNING FOR SEQUENTIAL DECISIONS

Reinforcement learning optimizes a sequence of actions by learning from rewards. In molecular design, deep RL has been used to propose compounds with desired properties [69,70]. In formulation research, the analogous actions might be selecting the next experiment, changing process settings in a simulator or choosing among validated release options. The reward could combine performance, uncertainty reduction, cost and safety.

Direct application to adaptive herbal dosing is far more demanding. Exploration that is acceptable in simulation may be unsafe in patients, and a reward based on short-term symptoms may produce clinically undesirable policies. Constrained RL, offline evaluation, conservative policy improvement and human authorization are prerequisites. A safer near-term use is experimental design within a validated laboratory space, where the algorithm proposes informative formulations and scientists approve execution.

10.3 ADAPTIVE DELIVERY AND SMART THERAPEUTICS

Adaptive drug delivery refers to systems whose release or dosing responds to measured state. Material-level adaptation can occur through pH-, enzyme-, redox- or temperature-responsive carriers. Algorithm-level adaptation can occur when a controller modifies a permitted parameter based on observed data. Combining the two creates a smart therapeutic, but also couples material, sensor, software and clinical risks. Each component and the integrated system require validation.

For herbal products, the first practical adaptive systems may be local or ex vivo platforms—for example, a hydrogel whose release is selected from a small validated range, or a

manufacturing process adjusted to compensate for a measured batch fingerprint. Fully autonomous patient-level control should remain a long-term objective. The governance layer in Figure 3 includes hard safety limits, clinician and quality-unit oversight, audit trails, cybersecurity and change control.

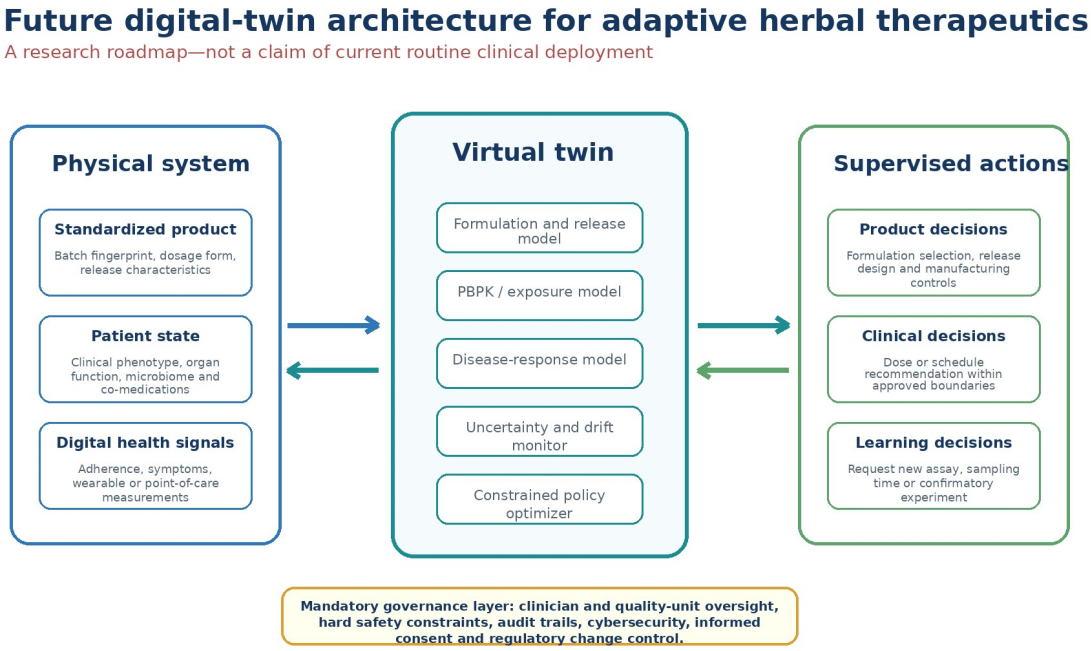

Figure 3. Future digital-twin architecture for adaptive herbal therapeutics. Physical product, patient and digital-health data update a virtual representation containing formulation, exposure and response models. Any action remains constrained by human oversight, safety limits and regulatory change control.

11. TRANSLATION, VALIDATION AND REGULATORY CONSIDERATIONS

11.1 EVIDENCE HIERARCHY AND CONTEXT OF USE

The credibility required of an AI model depends on its context of use. A model that ranks exploratory formulations has lower consequence than one used to set a release specification, waive an experiment or recommend a patient dose. The FDA's 2025 draft guidance proposes a risk-based credibility framework for AI models used to support regulatory decision-making for drugs and biological products [71]. The EMA reflection paper likewise addresses AI across the medicinal-product lifecycle and emphasizes human-centric development, data governance, performance, transparency and lifecycle management [72].

Developers should therefore write a precise context-of-use statement: the user, input data, output, decision, population or product space, and consequence of error. Validation should replicate that context. A model trained on laboratory batches cannot be assumed to support commercial release, and a pharmacokinetic model developed in healthy adults cannot be assumed to personalize therapy in patients with organ impairment.

11.2 INTEGRATION WITH QUALITY BY DESIGN

AI complements rather than replaces pharmaceutical QbD. ICH Q8(R2) frames development around the QTPP, CQAs, material attributes, process parameters, risk assessment and a science-based control strategy [73]. AI can help identify nonlinear relationships, prioritize risk and define a probabilistic design space. Nevertheless, a high-performing model does not eliminate the need for process understanding, validated analytical methods, stability programs or manufacturing controls.

For herbal products, QbD should begin before formulation with botanical-source controls and chemical fingerprints. The model should reveal whether a recommended process setting is compensating for normal raw-material variability or masking an unacceptable batch. Acceptable ranges must be justified by product performance and safety, not only by the model's ability to fit historical data.

11.3 MODEL-INFORMED BIOPHARMACEUTICS AND PHARMACOKINETICS

Regulatory guidances for PBPK analyses emphasize model purpose, data sources, assumptions, verification and reporting [74]. Guidance on biopharmaceutics applications further illustrates how PBPK can support oral product development and manufacturing changes when appropriately qualified [75]. These principles are relevant to AI-enhanced herbal pharmacokinetics. Mechanistic and data-driven components should be separately described, parameters traceable, sensitivity analyses reported and performance evaluated against observations not used for calibration.

For multicomponent products, the model may not need to represent every constituent. It should justify which markers, active moieties or toxicologically relevant components are modeled and why. Unmodeled constituents should be controlled through fingerprint specifications and safety evidence. Claims that a formulation improves bioavailability should identify the measured analyte, matrix, assay, sampling window and clinical significance.

11.4 GOOD AI PRACTICE AND LIFECYCLE MANAGEMENT

In January 2026, FDA and EMA published ten guiding principles of good AI practice in drug development. These emphasize human-centric design, risk-based approaches, standards, clear context of use, multidisciplinary expertise, data governance,

appropriate model development, risk-based performance assessment, lifecycle management and clear essential information [76]. These principles align closely with the needs of AI-driven herbal formulation, where traceability and domain expertise are indispensable.

The EMA/HMA 2025 AI Observatory report, published in June 2026, reflects increasing regulatory attention to guidance, practical applications, collaboration and preparedness for AI in medicines regulation [77]. This evolving environment favors early engagement with regulators when AI outputs will materially support quality, nonclinical, clinical or manufacturing decisions. Documentation should include the model-development plan, data provenance, preprocessing, version control, performance by relevant subgroup or batch, uncertainty, interpretability, cybersecurity and change-management procedures.

11.5 MINIMUM VALIDATION PACKAGE

A minimum translational package should contain: (1) analytically verified and versioned datasets; (2) a predefined context of use and error analysis; (3) comparison with transparent baselines; (4) internal validation that prevents leakage; (5) an external or temporally separated test; (6) prospective confirmation of model-selected formulations; (7) robustness and uncertainty evaluation; (8) mechanistic plausibility checks; (9) documentation of user oversight; and (10) post-deployment or post-transfer monitoring where applicable. For a model affecting product release or clinical decisions, independent review and auditability are essential.

Table 5. Translational roadmap and decision gates for AI-assisted herbal drug delivery.

Development stage	Minimum evidence	Principal risk control	Decision gate
1. Problem formulation	QTPP, intended user, context of use, endpoints and consequence of error	Multidisciplinary review; explicit distinction between exploratory and regulated use	Is the AI output connected to a real, measurable decision?
2. Data readiness	Authenticated botanical material, analytical provenance, material/process metadata, failed experiments	FAIR-aligned schema, units and identifiers; leakage and batch-effect assessment	Are data sufficiently representative and traceable for the intended claim?
3. Model development	Transparent baseline, justified algorithm, predefined metrics and applicability domain	Nested validation, uncertainty estimation, feature review and reproducible code/versioning	Does the model improve the decision rather than only fit the dataset?
4. Prospective formulation confirmation	Independently prepared model-selected candidates and comparator formulations	Blinded or pre-specified testing; retain failures; assess robustness to material and process variation	Is predicted performance reproduced experimentally?
5. Biological and PK validation	Biorelevant release, permeability, safety, PK and exposure-response evidence appropriate to route	Mechanistic checks, relevant controls, analyte-specific assays and uncertainty propagation	Does the delivery improvement produce a meaningful biological or clinical advantage?
6. Scale-up and transfer	Pilot-scale or cross-site batches, equipment and raw-material variability	Scale-aware features, change control, model re-verification and process capability assessment	Does performance remain within specifications after transfer?
7. Regulatory or clinical use	Documented context of use, credibility assessment, human oversight and lifecycle plan [71–77]	Audit trails, cybersecurity, subgroup/batch monitoring, drift thresholds and controlled updates	Is residual model risk acceptable for the consequence of the decision?

12. RESEARCH GAPS AND FUTURE ROADMAP

12.1 BUILD BENCHMARK DATASETS THAT REFLECT BOTANICAL REALITY

The largest bottleneck is not algorithm availability but interoperable data. Benchmark datasets should include authenticated plant material, extraction and analytical metadata, full formulation recipes, process conditions, critical attributes, failed formulations, stability and biological endpoints. Multi-batch studies are essential. A model developed from one extract batch cannot demonstrate resilience to the variability that defines phytopharmaceutical development.

12.2 MOVE FROM RETROSPECTIVE FITTING TO PROSPECTIVE BENEFIT

Many AI studies report cross-validation metrics without testing whether the model selects better experiments or products. Future work should compare AI-assisted development with conventional DoE or expert selection using prospective endpoints: number of experiments, time, cost, success rate, robustness and biological performance. Prospective benefit is the strongest evidence that AI changes development rather than merely describing completed experiments.

12.3 COUPLE FORMULATION, PHARMACOKINETICS AND PHARMACODYNAMICS

Models should connect product attributes to exposure and response. Release prediction alone is insufficient when metabolism or targeting dominates. Hybrid mechanistic-ML approaches are promising because they retain mass balance and physiological structure while learning complex residual relationships. Multi-analyte models are needed for standardized extracts, with explicit justification for marker selection and handling of unmeasured constituents.

12.4 ESTABLISH EXPLAINABILITY AND UNCERTAINTY AS PERFORMANCE REQUIREMENTS

Future publications should report calibration, predictive intervals, applicability domain, batch- or subgroup-specific errors and stability of explanations. Feature attribution should lead to confirmatory experiments. A model that is slightly less accurate but more stable, interpretable and transferable may be more valuable in pharmaceutical development than a fragile black box with a superior internal score.

12.5 DESIGN FOR MANUFACTURING AND REGULATION FROM THE BEGINNING

Laboratory optimization should include scale-relevant variables and permitted excipients, analytical specifications, sterilization or microbial controls, packaging and stability. Model files, code, data transformations and version histories should be controlled. Early regulatory dialogue is warranted when AI will justify a

design space, specification, manufacturing control or clinical decision.

12.6 USE DIGITAL TWINS AND REINFORCEMENT LEARNING IN BOUNDED STEPS

Near-term digital twins should focus on product and process simulation, or on PK prediction with supervised updating. RL should first be used for laboratory experiment selection or simulation-based process control. Progression toward adaptive patient-level delivery should occur only after conservative offline evaluation, clinically meaningful prospective testing and demonstration that human oversight can reliably detect and prevent unsafe recommendations.

A practical five-year research sequence is therefore: standardize data; validate AI-assisted formulation prospectively; link product attributes to PK; establish multi-site and multi-batch transferability; and only then test bounded adaptive or personalized decisions. This sequence places evidence generation ahead of technological spectacle.

13. LIMITATIONS OF THIS REVIEW

This is a structured narrative review rather than a registered systematic review. Searches were targeted and iterative, and no formal risk-of-bias instrument or quantitative meta-analysis was applied. The literature is heterogeneous in cargo, platform, analytical method, biological model and outcome definition. Direct studies combining AI with herbal drug-delivery development remain limited, so several recommendations are extrapolated from adjacent evidence in pharmaceutical formulation, nanomedicine, natural-product delivery and model-informed development. These extrapolations are explicitly identified as enabling or future-facing rather than established clinical practice. Regulatory documents cited here reflect the status available through 20 June 2026 and may evolve.

14. CONCLUSION

AI-driven herbal drug delivery is best understood as a disciplined integration of computational pharmaceutics, standardized phytopharmaceutical science and experimentally verified delivery design. Its near-term value lies in learning nonlinear formulation relationships, reducing uninformative experimentation, predicting property and stability risks, and optimizing multiple product attributes under manufacturing and safety constraints. Smart nanocarriers—including phytosomes, liposomes, polymeric nanoparticles, nanocrystals, lipid carriers, nanoemulsions and hydrogels—provide rich but complex design spaces in which these methods can be useful.

The field should avoid two errors: treating every fitted model as intelligence and treating future digital twins or adaptive systems as already validated therapy. Direct AI-herbal evidence is still developing. Progress depends on a traceable phytochemical-carrier-patient digital thread, multi-batch datasets, external and prospective validation, uncertainty estimation, interpretable decisions, quality-by-design integration and regulatory lifecycle controls. Under these conditions, AI can shift herbal formulation from empirical trial-and-error toward predictive, explainable and precision-oriented development while preserving the central role of pharmaceutical science and experimental evidence.

Declarations

Author contributions: To be completed according to the target journal's contributorship taxonomy.

Funding: No specific funding information was provided for this manuscript draft.

Conflicts of interest: The authors should declare all relevant financial and non-financial interests before submission.

Ethics approval: Not applicable; this article is a review of published literature and does not report new human or animal experimentation.

Data availability: No new dataset was generated. Sources used in the review are listed in the reference section.

Acknowledgements: To be completed by the authors.

REFERENCE

1. Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 2021;20:200-216. doi:10.1038/s41573-020-00114-z.

2. 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.

3. Ekor M. The growing use of herbal medicines: issues relating to adverse reactions and challenges in monitoring safety. Front Pharmacol. 2014;4:177. doi:10.3389/fphar.2013.00177.

4. Li FS, Weng JK. Demystifying traditional herbal medicine with modern approach. Nat Plants. 2017;3:17109. doi:10.1038/nplants.2017.109.

5. Mullowney MW, Duncan KR, Elsayed SS, et al. Artificial intelligence for natural product drug discovery. Nat Rev Drug Discov. 2023;22(11):895-916. doi:10.1038/s41573-023-00774-7.

6. Wang H, Chen Y, Wang L, Liu Q, Yang S, Wang C. Advancing herbal medicine: enhancing product quality and safety through robust quality control practices. Front Pharmacol. 2023;14:1265178. doi:10.3389/fphar.2023.1265178.

7. Ajazuddin, Saraf S. Applications of novel drug delivery system for herbal formulations. Fitoterapia. 2010;81(7):680-689. doi:10.1016/j.fitote.2010.05.001.

8. Bonifácio BV, da Silva PB, Ramos MAS, Negri KMS, Bauab TM, Chorilli M. Nanotechnology-based drug delivery systems and herbal medicines: a review. Int J Nanomedicine. 2014;9:1-15. doi:10.2147/IJN.S52634.

9. Teja PK, Mithiya J, Kate AS, Bairwa K, Chauthe SK. Herbal nanomedicines: recent advancements, challenges, opportunities and regulatory overview. Phytomedicine. 2022;96:153890. doi:10.1016/j.phymed.2021.153890.

10. Ai S, et al. Collision of herbal medicine and nanotechnology: a bibliometric analysis of herbal nanoparticles from 2004 to 2023. J Nanobiotechnology. 2024;22:140. doi:10.1186/s12951-024-02426-3.

11. Baethge C, Goldbeck-Wood S, Mertens S. SANRA-a scale for the quality assessment of narrative review articles. Res Integr Peer Rev. 2019;4:5. doi:10.1186/s41073-019-0064-8.

12. Bannigan P, Aldeghi M, Bao Z, Häse F, Aspuru-Guzik A, Allen C. Machine learning directed drug formulation development. Adv Drug Deliv Rev. 2021;175:113806. doi:10.1016/j.addr.2021.05.016.

13. Vora LK, Gholap AD, Jetha K, Thakur RRS, Solanki HK, Chavda VP. Artificial intelligence in pharmaceutical technology and drug delivery design. Pharmaceutics. 2023;15(7):1916. doi:10.3390/pharmaceutics15071916.

14. Noorain, Srivastava V, Parveen B, Parveen R. Artificial intelligence in drug formulation and development: applications and future prospects. Curr Drug Metab. 2023;24(9):622-634. doi:10.2174/0113892002265786230921062205.

15. Serrano DR, Luciano FC, Anaya BJ, et al. Artificial intelligence applications in drug discovery and drug delivery: revolutionizing personalized medicine. Pharmaceutics. 2024;16(10):1328. doi:10.3390/pharmaceutics16101328.

16. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-477. doi:10.1038/s41573-019-0024-5.

17. 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.

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. Yang H, Lou C, Sun L, et al. admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. Bioinformatics. 2019;35(6):1067-1069. doi:10.1093/bioinformatics/bty707.

20. Xiong G, Wu Z, Yi J, et al. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Res. 2021;49(W1):W5-W14. doi:10.1093/nar/gkab255.

21. Delaney JS. ESOL: a simple model for aqueous solubility based on molecular topology. J Chem Inf Comput Sci. 2004;44(3):1000-1005. doi:10.1021/ci034243x.

22. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46(1-3):3-26. doi:10.1016/S0169-409X(00)00129-0.

23. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neural nets as a method for quantitative structure-activity relationships. J Chem Inf Model. 2015;55(2):263-274. doi:10.1021/ci500747n.

24. Mayr A, Klambauer G, Unterthiner T, et al. Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. Chem Sci. 2018;9:5441-5451. doi:10.1039/C8SC00148K.

25. Takayama K, Fujikawa M, Nagai T. Artificial neural network as a novel method to optimize pharmaceutical formulations. Pharm Res. 1999;16:1-6. doi:10.1023/A:1011986823850.

26. Takayama K, Takahara J, Fujikawa M, Ichikawa H, Nagai T. Formula optimization based on artificial neural networks in transdermal drug delivery. J Control Release. 1999;62(1-2):161-170. doi:10.1016/S0168-3659(99)00033-4.

27. Leonardi D, Salomón CJ, Lamas MC, Olivieri AC. Development of novel formulations for Chagas disease: optimization of benznidazole chitosan microparticles based on artificial neural networks. Int J Pharm. 2009;367(1-2):140-147. doi:10.1016/j.ijpharm.2008.09.036.

28. Mandal U, Gowda V, Ghosh A, Selvan S, Solomon S, Pal TK. Optimization of metformin HCl 500 mg sustained release matrix tablets using artificial neural network based on multilayer perceptrons model. Chem Pharm Bull (Tokyo). 2008;56(2):150-155. doi:10.1248/cpb.56.150.

29. Petrović J, Ibrić S, Betz G, Đurić Z. Optimization of matrix tablets controlled drug release using Elman dynamic neural networks and decision trees. Int J Pharm. 2012;428(1-2):57-67. doi:10.1016/j.ijpharm.2012.02.031.

30. Shamay Y, Shah J, Işık M, et al. Quantitative self-assembly prediction yields targeted nanomedicines. Nat Mater. 2018;17(4):361-368. doi:10.1038/s41563-017-0007-z.

31. Yamankurt G, Berns EJ, Xue A, Lee A, Bagheri N, Mrksich M, Mirkin CA. Exploration of the nanomedicine-design space with high-throughput screening and machine learning. Nat Biomed Eng. 2019;3(4):318-327. doi:10.1038/s41551-019-0351-1.

32. Reker D, et al. Computationally guided high-throughput design of self-assembling drug nanoparticles. Nat Nanotechnol. 2021;16:725-733. doi:10.1038/s41565-021-00870-y.

33. Bannigan P, et al. Machine learning models to accelerate the design of polymeric long-acting injectables. Nat Commun. 2023;14:35. doi:10.1038/s41467-022-35343-w.

34. Ortiz-Perez A, van Tilborg D, van der Meel R, Grisoni F, Albertazzi L. Machine learning-guided high throughput nanoparticle design. Digital Discovery. 2024;3(7):1280-1291. doi:10.1039/D4DD00104D.

35. Mitchell MJ, Billingsley MM, Haley RM, Wechsler ME, Peppas NA, Langer R. Engineering precision nanoparticles for drug delivery. Nat Rev Drug Discov. 2021;20(2):101-124. doi:10.1038/s41573-020-0090-8.

36. Blanco E, Shen H, Ferrari M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. Nat Biotechnol. 2015;33(9):941-951. doi:10.1038/nbt.3330.

37. Peer D, Karp JM, Hong S, Farokhzad OC, Margalit R, Langer R. Nanocarriers as an emerging platform for cancer therapy. Nat Nanotechnol. 2007;2(12):751-760. doi:10.1038/nnano.2007.387.

38. Mura S, Nicolas J, Couvreur P. Stimuli-responsive nanocarriers for drug delivery. Nat Mater. 2013;12(11):991-1003. doi:10.1038/nmat3776.

39. Wilhelm S, Tavares AJ, Dai Q, et al. Analysis of nanoparticle delivery to tumours. Nat Rev Mater. 2016;1:16014. doi:10.1038/natrevmats.2016.14.

40. Anselmo AC, Mitragotri S. Nanoparticles in the clinic: an update. Bioeng Transl Med. 2019;4(3):e10143. doi:10.1002/btm2.10143.

41. Bobo D, Robinson KJ, Islam J, Thurecht KJ, Corrie SR. Nanoparticle-based medicines: a review of FDA-approved materials and clinical trials to date. Pharm Res. 2016;33(10):2373-2387. doi:10.1007/s11095-016-1958-5.

42. Patra JK, Das G, Fraceto LF, et al. Nano based drug delivery systems: recent developments and future prospects. J Nanobiotechnology. 2018;16:71. doi:10.1186/s12951-018-0392-8.

43. Shi J, Kantoff PW, Wooster R, Farokhzad OC. Cancer nanomedicine: progress, challenges and opportunities. Nat Rev Cancer. 2017;17(1):20-37. doi:10.1038/nrc.2016.108.

44. Soppimath KS, Aminabhavi TM, Kulkarni AR, Rudzinski WE. Biodegradable polymeric nanoparticles as drug delivery devices. J Control Release. 2001;70(1-2):1-20. doi:10.1016/S0168-3659(00)00339-4.

45. Guo M, Qin S, Wang S, et al. Herbal medicine nanocrystals: a potential novel therapeutic strategy. Molecules. 2023;28(17):6370. doi:10.3390/molecules28176370.

46. Akbarzadeh A, Rezaei-Sadabady R, Davaran S, et al. Liposome: classification, preparation, and applications. Nanoscale Res Lett. 2013;8:102. doi:10.1186/1556-276X-8-102.

47. Alharbi WS, Almughem FA, Almehmady AM, et al. Phytosomes as an emerging nanotechnology platform for the topical delivery of bioactive phytochemicals. Pharmaceutics. 2021;13(9):1475. doi:10.3390/pharmaceutics13091475.

48. Adepu S, Ramakrishna S. Controlled drug delivery systems: current status and future directions. Molecules. 2021;26(19):5905. doi:10.3390/molecules26195905.

49. Park K. Controlled drug delivery systems: past forward and future back. J Control Release. 2014;190:3-8. doi:10.1016/j.jconrel.2014.03.054.

50. Hou X, Zaks T, Langer R, Dong Y. Lipid nanoparticles for mRNA delivery. Nat Rev Mater. 2021;6(12):1078-1094. doi:10.1038/s41578-021-00358-0.

51. Anand P, Kunnumakkara AB, Newman RA, Aggarwal BB. Bioavailability of curcumin: problems and promises. Mol Pharm. 2007;4(6):807-818. doi:10.1021/mp700113r.

52. Shoba G, Joy D, Joseph T, Majeed M, Rajendran R, Srinivas PS. Influence of piperine on the pharmacokinetics of curcumin in animals and human volunteers. Planta Med. 1998;64(4):353-356. doi:10.1055/s-2006-957450.

53. Maiti K, Mukherjee K, Gantait A, Saha BP, Mukherjee PK. Curcumin-phospholipid complex: preparation, therapeutic evaluation and pharmacokinetic study in rats. Int J Pharm. 2007;330(1-2):155-163. doi:10.1016/j.ijpharm.2006.09.025.

54. Shaikh J, Ankola DD, Beniwal V, Singh D, Kumar MNVR. Nanoparticle encapsulation improves oral bioavailability of curcumin by at least 9-fold when compared to curcumin

administered with piperine. Eur J Pharm Sci. 2009;37(3-4):223-230. doi:10.1016/j.ejps.2009.02.019.

55. Sasaki H, Sunagawa Y, Takahashi K, et al. Innovative preparation of curcumin for improved oral bioavailability. Biol Pharm Bull. 2011;34(5):660-665. doi:10.1248/bpb.34.660.

56. Liu YT, Hao HP, Xie HG, et al. Extensive intestinal first-pass elimination and predominant hepatic distribution of berberine explain its low plasma levels in rats. Drug Metab Dispos. 2010;38(10):1779-1784. doi:10.1124/dmd.110.033936.

57. Amri A, Chaumeil JC, Sfar S, Charrueau C. Administration of resveratrol: what formulation solutions to bioavailability limitations? J Control Release. 2012;158(2):182-193. doi:10.1016/j.jconrel.2011.09.083.

58. Barzaghi N, Crema F, Gatti G, Pifferi G, Perucca E. Pharmacokinetic studies on IdB 1016, a silybin-phosphatidylcholine complex, in healthy human subjects. Eur J Drug Metab Pharmacokinet. 1990;15:333-338. doi:10.1007/BF03190223.

59. Hüsch J, Bohnet J, Fricker G, et al. Enhanced absorption of boswellic acids by a lecithin delivery form (Phytosome) of Boswellia extract. Fitoterapia. 2013;84:89-98. doi:10.1016/j.fitote.2012.10.002.

60. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1:206-215. doi:10.1038/s42256-019-0048-x.

61. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20:310. doi:10.1186/s12911-020-01332-6.

62. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30. 2017:4765-4774.

63. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016:1135-1144. doi:10.1145/2939672.2939778.

64. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018. doi:10.1038/sdata.2016.18.

65. Björnsson B, Borrebaeck C, Elander N, et al. Digital twins to personalize medicine. Genome Med. 2020;12:4. doi:10.1186/s13073-019-0701-3.

66. Corral-Acero J, Margara F, Marciniak M, et al. The digital twin to enable the vision of precision cardiology. Eur Heart J. 2020;41(48):4556-4564. doi:10.1093/eurheartj/ehaa159.

67. Pappalardo F, Russo G, Musuamba Tshinanu F, Viceconti M. In silico clinical trials: concepts and early adoptions. Brief Bioinform. 2019;20(5):1699-1708. doi:10.1093/bib/bby043.

68. Chen Y, Yang O, Sampat C, Bhalode P, Ramachandran R, Ierapetritou M. Digital twins in pharmaceutical and biopharmaceutical manufacturing: a literature review. Processes. 2020;8(9):1088. doi:10.3390/pr8091088.

69. Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design. Sci Adv. 2018;4(7):eaap7885. doi:10.1126/sciadv.aap7885.

70. Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-novo design through deep reinforcement learning. J Cheminform. 2017;9:48. doi:10.1186/s13321-017-0235-x.

71. US Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products: draft guidance for industry and other interested parties. Silver Spring (MD): FDA; January 2025. Accessed 20 June 2026.

72. European Medicines Agency. Reflection paper on the use of artificial intelligence in the medicinal product lifecycle. Amsterdam: EMA; adopted September 2024. Accessed 20 June 2026.

73. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q8(R2): Pharmaceutical Development. Geneva: ICH; 2009.

74. US Food and Drug Administration. Physiologically based pharmacokinetic analyses-format and content: guidance for industry. Silver Spring (MD): FDA; 2018.

75. US Food and Drug Administration. The use of physiologically based pharmacokinetic analyses-biopharmaceutics applications for oral drug product development, manufacturing changes, and controls: guidance for industry. Silver Spring (MD): FDA; 2020.

76. US Food and Drug Administration; European Medicines Agency. Guiding principles of good AI practice in drug development. January 2026. Accessed 20 June 2026.

77. European Medicines Agency; Heads of Medicines Agencies. 2025 AI Observatory report. EMA/67888/2026. Amsterdam: EMA; June 2026. Accessed 20 June 2026. End of manuscript