Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

AI FOR NANOMEDICINE DESIGN: A MIXED-METHODS REVIEW OF MODELS AND TRANSLATION

Victor Santos[1]*, Rafael Costa[1], Bruno Teixeira[2]

1. ***Department of AI Drug Discovery Systems, Faculty of Pharmacy, University of Sao Paulo, Sao Paulo, Brazil.***
2. ***Department of Computational Pharmacology, Faculty of Engineering, University of Campinas, Campinas, Brazil.***

ARTICLE INFO

Received:
07 September 2025
Received in revised form:
04 December 2025
Accepted:
09 December 2025
Available online:
28 December 2025

Keywords: Artificial intelligence, Nanomedicine, Machine learning, Lipid nanoparticles, Drug delivery, Nano-QSAR

ABSTRACT

Artificial intelligence offers transformative potential for nanomedicine design because nanoparticle performance emerges from complex interactions among composition, structure, processing, and biological context. Yet the translational impact of these models remains uncertain. This mixed-methods review maps the evidence on AI models for nanomedicine design and appraises their methodological quality and translational readiness. It focuses on nanoparticle formulation, drug delivery systems, nanocarrier optimization, biodistribution, and nano-toxicity prediction. The review combines systematic literature retrieval, descriptive evidence mapping, and critical methodological appraisal. Extracted domains included model type, nanoparticle class, prediction task, dataset source, validation design, reproducibility, interpretability, and proximity to in-vitro, in-vivo, or clinical translation. AI models have been applied across lipid, polymeric, inorganic, dendrimer-like, hybrid, and drug nanoparticle systems. However, methodological rigour is inconsistent, with frequent reliance on small datasets, internal validation, limited code or data sharing, and rare translational case studies. The field has demonstrated convincing proof-of-concept for AI-assisted nanomedicine design, but a substantial gap remains between computational prediction and clinically actionable formulation development. Progress will depend on prospective validation, standardized datasets, reproducible reporting, and stronger integration between computational and experimental teams.

To Cite This Article: Santos V, Costa R, Teixeira B. AI for Nanomedicine Design: A Mixed-Methods Review of Models and Translation. Pharmacophore. 2025;16(6):45-55. https://doi.org/10.51847/efaBcWafDc

Introduction

Nanomedicine design requires simultaneous optimization of particle size, polydispersity, surface charge, composition, morphology, drug loading, release kinetics, stability, uptake, biodistribution, and safety. Machine learning is attractive in this setting because computational models can learn patterns from high-dimensional formulation and biological data that are difficult to capture through one-variable-at-a-time experimentation [1]. Early demonstrations of computationally guided self-assembling drug nanoparticles showed how AI could narrow experimental search spaces while preserving experimental confirmation as the decisive test of utility [1]. More recent lipid nanoparticle studies extend this logic to nucleic-acid delivery, where formulation variables and biological readouts are tightly coupled [2, 3].

The literature expanded rapidly after 2017 as nanoinformatics, nano-QSAR, and formulation-focused machine learning converged with larger experimental datasets and more accessible modelling tools. Reviews on machine-learning-directed drug formulation and nanomedicine data curation describe a shift from isolated predictive studies toward more integrated computational workflows [4, 5]. Public-facing resources and FAIRification efforts, including eNanoMapper-linked workflows and nanoinformatics platforms, have also begun to make nanoparticle datasets more reusable [6, 7]. This growth has encouraged more ambitious models, including transformer-based formulation design and combinatorial ionizable lipid discovery [8, 9].

Despite this expansion, it remains unclear whether AI nanomedicine models are methodologically robust enough to guide translational decision-making. Critical reviews in computational nanotoxicology warn that many models are constrained by small datasets, inconsistent assay conditions, heterogeneous descriptors, and limited external validation [10, 11]. Formulation-

Corresponding Author: Victor Santos; Department of AI Drug Discovery Systems, Faculty of Pharmacy, University of Sao Paulo, Sao Paulo, Brazil. E-mail: victor.santos@gmail.com.

focused reviews similarly note that reported accuracy often exceeds demonstrated real-world utility when models are not prospectively tested across laboratories or nanoparticle classes [12, 13]. This uncertainty justifies a mixed-methods synthesis that maps what has been done while appraising how well it was done.

The objective of this review is to map AI applications in nanomedicine design from 2017 to 2025 and critically appraise their validation rigour, reproducibility, interpretability, and translational readiness. The mapping component captures model types, nanoparticle systems, prediction tasks, input features, and dataset sources, while the appraisal component evaluates whether reported methods support generalisable and clinically meaningful claims [14]. Particular attention is given to studies linking computational prediction with experimental synthesis, in-vitro testing, in-vivo biodistribution, or translational discussion [15, 16]. By integrating these strands, the review aims to distinguish proof-of-concept modelling from evidence that could plausibly support clinical nanomedicine development.

Materials and Methods

Mixed-Methods Design

This review used a mixed-methods design combining descriptive evidence mapping with qualitative methodological appraisal. The mapping component treated each study as a unit of evidence and summarized model family, nanoparticle class, task type, dataset origin, and experimental validation status, following the logic used in broad formulation and nanomedicine AI reviews [4, 5]. The appraisal component examined how convincingly each study supported its claims, including whether validation, reproducibility, and interpretability were adequate for the intended application [10, 11]. This dual approach was selected because AI nanomedicine is not only a question of what models exist, but also whether those models are sufficiently reliable to influence experimental or translational decisions.

Search Strategy and Databases

The search strategy targeted peer-reviewed publications from 2017 to 2025 on artificial intelligence, machine learning, deep learning, nano-QSAR, nanoinformatics, nanoparticle formulation, nanocarrier optimization, drug delivery, biodistribution, cellular uptake, and nanoparticle toxicity. Search strings combined terms such as “machine learning nanomedicine design nanoparticle optimization,” “deep learning lipid nanoparticle drug delivery prediction,” “nanoparticle cellular uptake machine learning,” “nano-QSAR machine learning,” and “nanoparticle toxicity prediction machine learning,” reflecting the terminology used across formulation, delivery, and safety studies [2, 17, 18]. Searches emphasized journals in drug delivery, nanomedicine, pharmaceutical sciences, toxicology, and nanotechnology while allowing inclusion of relevant interdisciplinary studies. Reviews and perspectives were included when they provided conceptual, methodological, or translational context for the appraisal [12, 14].

Inclusion and Exclusion Criteria

Studies were eligible if they developed, evaluated, or critically reviewed AI or machine-learning models applied to nanomedicine design, nanoparticle formulation, drug delivery prediction, biodistribution, cellular uptake, release, or nano-toxicity. Eligible primary studies included supervised learning, deep learning, transformer-based models, nano-QSAR, perturbation-theory machine learning, and platform-oriented approaches, provided that nanoparticle-related inputs and outputs were central to the study [8, 19, 20]. Studies were excluded if they were purely experimental without a computational modelling component, used generic materials informatics without nanomedicine relevance, or lacked sufficient methodological detail to determine the model task. English-language peer-reviewed publications from 2017 through 2025 were included to align the evidence map with the recent AI-driven expansion of the field.

Data Extraction and Evidence Mapping

For each included study, extracted variables included publication year, nanoparticle type, model family, input features, prediction endpoint, dataset size or source, validation strategy, interpretability approach, and evidence of code or data availability. The extraction framework was designed to capture formulation features such as lipid or polymer composition, molar ratios, process parameters, surface properties, zeta potential, drug identity, and molecular descriptors, which are commonly used in lipid nanoparticle, polymeric nanoparticle, and nano-QSAR studies [3, 21, 22]. Translational variables included whether the study proceeded from in-silico prediction to experimental synthesis, in-vitro testing, in-vivo biodistribution, or discussion of scale-up and regulatory relevance [15, 16]. This structure enabled both breadth-oriented mapping and critical comparison of methodological maturity across application areas.

Methodological Appraisal Criteria

Methodological appraisal used predefined criteria covering validation rigour, reproducibility, model complexity, interpretability, reporting completeness, and translational readiness. Validation was judged stronger when studies used independent test sets, external datasets, prospective experimental confirmation, or cross-context generalisation rather than only internal cross-validation [1, 16]. Reproducibility was judged by whether code, data, hyperparameters, preprocessing details, and model-selection procedures were sufficiently available for independent verification, consistent with concerns raised in nanoinformatics and nanotoxicology reviews [7, 11]. Translational readiness was assessed by whether predicted endpoints

were biologically or clinically meaningful and whether models were connected to synthesis, characterization, in-vivo testing, manufacturing constraints, or regulatory decision-making [13, 23].

Results and Discussion

Evidence Mapping

Study Selection and General Characteristics

The included evidence covered a broad but uneven landscape of AI for nanomedicine design, with studies spanning self-assembling drug nanoparticles, lipid nanoparticles, polymeric vectors, inorganic nanomaterials, and toxicity-focused nanomaterial datasets. Publication density increased after 2020, coinciding with greater interest in mRNA delivery, formulation automation, and data-driven nanotoxicology [2, 3, 24]. Lipid nanoparticles were especially prominent in recent work, including models for mRNA delivery, plasmid DNA transfection, and ionizable lipid discovery [9, 21]. In contrast, inorganic nanoparticle and mixture-toxicity studies were more often framed around safety prediction than therapeutic formulation optimization [25, 26].

Types of AI Models Used

The evidence map showed that classical machine-learning algorithms remain common, including random forest, support vector machines, gradient boosting, and ensemble methods, especially in nano-QSAR and toxicity prediction studies [18, 20]. Deep-learning models appeared increasingly in lipid nanoparticle and platform-oriented studies, where larger combinatorial spaces and complex feature interactions motivate neural architectures [2, 3]. Transformer-based neural networks and deep learning-powered optimization platforms represent the newest modelling direction, suggesting a shift from endpoint prediction toward generative or recommendation-style formulation design [8]. However, the field still relies heavily on tabular supervised learning, and model sophistication often exceeds the size and diversity of the available datasets.

Nanoparticle Design Tasks Addressed

The most common prediction tasks involved formulation properties, biological delivery, and safety endpoints rather than complete therapeutic performance. Studies addressed nanoparticle size, encapsulation or delivery efficiency, transfection, cellular uptake, biodistribution, tumor delivery, brain delivery, drug release, and cytotoxicity across diverse particle systems [15, 17, 27]. Perturbation-theory machine learning was applied to coated nanoparticle drug-release systems, illustrating how mechanistic formulation questions can be reframed as predictive modelling tasks [11]. Toxicity studies were particularly numerous, reflecting both regulatory interest and the availability of structured nanotoxicology datasets [20, 28].
Table 1 organizes the AI nanomedicine evidence base by nanoparticle system, modeling strategy, prediction task, translational value, and dominant methodological vulnerability.

Table 1. Evidence Architecture of AI Models for Nanomedicine Design across Nanoparticle Systems and Prediction Tasks

Nanomedicine evidence domain	Typical nanoparticle systems	Dominant AI/model families	Main input feature domains	Common prediction tasks	Translational value of the task	Main methodological vulnerability
Formulation optimization	Lipid nanoparticles, polymeric nanocarriers, drug self-assembling nanoparticles	Random forest, gradient boosting, neural networks, Bayesian or optimization-guided models	Lipid/polymer composition, molar ratios, process conditions, drug identity, formulation descriptors	Particle size, polydispersity, encapsulation efficiency, formulation performance	Can reduce experimental search space and prioritize candidate formulations for synthesis	Internal validation may overstate utility when similar formulations appear across train-test partitions
Nucleic-acid delivery design	Lipid nanoparticles, ionizable lipid libraries, mRNA or plasmid DNA delivery systems	Deep learning, ensemble learning, transformer-based models, recommendation-style optimization	Ionizable lipid structure, helper lipid ratios, cargo type, cell type, transfection readouts	Delivery efficiency, cell-type-preferential transfection, mRNA or plasmid expression	Closest to industrial and clinical relevance because lipid nanoparticles are already linked to translational product development	Dataset scale may still be insufficient for complex architectures unless paired with iterative experimental validation
Cellular uptake and biological interaction prediction	Polymeric, inorganic, lipid, hybrid nanoparticles	Classical ML, ensemble models, nano-QSAR, interpretable feature-importance models	Size, charge, surface chemistry, morphology, protein corona-related variables, cell type, exposure conditions	Cellular uptake, intracellular delivery, biological response	Helps connect physicochemical design variables with biological performance	Assay heterogeneity and biological-context dependence limit generalization
Biodistribution and tissue-targeting prediction	Tumor-targeted nanoparticles, brain-delivery systems, tissue-	Supervised ML, ensemble models, deep-learning variants	Physicochemical properties, route of administration, animal model, dose, surface	Tumor delivery, brain delivery, organ distribution, tissue accumulation	More translationally meaningful than purely physicochemical endpoints because it	External validation across animal models, laboratories, dosing contexts, and

	distribution studies		functionalization, tissue outcome		addresses in-vivo behavior	nanoparticle chemistries is often limited
Drug-release and stability modeling	Coated nanoparticles, polymeric carriers, drug-loaded nanocarriers	Perturbation-theory ML, regression models, ensemble methods	Polymer/coating properties, drug descriptors, release medium, pH, time, temperature, formulation variables	Release kinetics, controlled-release behavior, stability-related outcomes	Supports rational design of release profiles and formulation screening	Mechanistic interpretation may be weak when release conditions are poorly standardized
Nano-toxicity and safety prediction	Inorganic nanoparticles, metal/metal oxide nanoparticles, mixed nanomaterials, therapeutic nanocarriers	Nano-QSAR, random forest, SVM, gradient boosting, toxicity classifiers	Composition, size, surface area, zeta potential, dissolution, exposure dose, assay metadata, cell or organism model	Cytotoxicity, immune response, oxidative stress, hazard classification, mixture toxicity	Important for safety screening and regulatory prioritization	Toxicity datasets are fragmented by assay protocol, exposure context, and incomplete metadata
Nanoinformatics and FAIR data infrastructure	Cross-class nanomaterial datasets and reusable annotation platforms	Data curation workflows, ontology-linked tools, reusable ML pipelines	Metadata standards, descriptor harmonization, assay annotations, dataset provenance, workflow documentation	Dataset reuse, model benchmarking, cross-study comparison, machine-readable annotation	Enables cumulative science and supports standardized benchmarks	Infrastructure does not ensure reproducibility unless individual studies share executable workflows and exact analysis details
Closed-loop or platform-oriented design	Lipid nanoparticles, combinatorial formulation libraries, high-throughput drug-delivery platforms	Active learning, deep learning, transformer-based models, iterative optimization systems	Experimental results from repeated synthesis-testing cycles, multi-objective formulation descriptors, uncertainty estimates	Candidate recommendation, formulation ranking, experimental prioritization, multi-objective optimization	Represents the most promising path from retrospective prediction toward experimentally actionable design	Requires rigorous prospective testing, uncertainty-aware selection, and validation outside the original optimization environment

Input Features and Feature Engineering

Input features varied substantially across studies, but most models used combinations of formulation composition, material descriptors, surface properties, experimental conditions, and biological context. Lipid nanoparticle models typically encoded lipid identity, ionizable lipid structure, molar ratios, helper lipids, and delivery readouts, while plasmid DNA lipid nanoparticle models included design features linked to cell-type-preferential transfection [9, 21]. Cellular uptake and biodistribution studies incorporated physicochemical properties such as size, charge, and surface chemistry alongside experimental or animal-model variables [15, 17]. Nano-QSAR and toxicity models often relied on engineered descriptors, exposure conditions, and assay metadata, making feature harmonization a central determinant of model credibility [18, 22].

Dataset Characteristics and Sources

Dataset size and provenance were highly heterogeneous, ranging from relatively small curated experimental sets to larger combinatorial or literature-derived datasets. Some formulation studies depended on proprietary or newly generated high-throughput datasets, which enabled model development but limited independent reuse when raw data were not fully shared [1, 3]. Public and semi-public nanoinformatics resources, including FAIRified eNanoMapper-linked data and online nanostructure annotation platforms, represent important steps toward a shared data ecosystem [6, 7]. Nevertheless, many datasets remain fragmented by particle type, assay protocol, biological endpoint, and reporting convention, limiting cross-study benchmarking.

Training and Validation Strategies

Training and validation practices were inconsistent across the mapped literature. Many studies used internal cross-validation or random train-test splits, which can inflate performance when similar formulations, closely related nanomaterials, or shared experimental batches appear across partitions [20, 26]. Stronger studies included independent experimental confirmation, such as computationally guided nanoparticle synthesis or in-vivo biodistribution testing after model development [1, 16]. Calibration, uncertainty quantification, temporal validation, and true external laboratory validation were rarely emphasized, even though these practices are essential for assessing whether predictions will generalize beyond the original dataset [11].

Results – Methodological Appraisal

Validation Rigour

Validation rigour was the most important weakness identified in the appraisal. Random splits and internal cross-validation were common in toxicity and formulation prediction, but such designs may overestimate utility when dataset records are correlated through shared nanoparticle cores, similar coatings, repeated assays, or related formulation families [18, 20]. In-vivo distribution models and tumor-delivery prediction studies provided more translationally relevant endpoints, yet still require external validation across animal models, laboratories, and nanoparticle chemistries to support broad claims [15, 16]. The strongest evidence came from studies in which AI predictions were linked to experimental synthesis and testing, because these designs reduce the gap between retrospective modelling and actionable formulation design [1].

Reproducibility and Code/Data Availability

Reproducibility was uneven, and many studies did not provide all elements needed for independent replication. Reviews of machine learning in nanomedicine and nanotoxicology repeatedly identify limited data sharing, inconsistent descriptor definitions, incomplete preprocessing details, and absent code as barriers to cumulative progress [4, 11]. FAIR data workflows and nanoinformatics platforms directly address these weaknesses by improving data structure, metadata capture, and machine readability [6, 7]. However, platform availability alone does not guarantee reproducibility unless individual modelling studies also publish executable workflows, exact train-test splits, and enough experimental metadata to reconstruct the analysis.

Model Complexity and Overfitting Risks

Model complexity was not always justified by dataset scale or diversity. Deep-learning and transformer-based approaches are promising for high-dimensional formulation design, but they carry overfitting risks when the number of experimentally observed formulations is modest relative to the design space [2, 8]. Combinatorial lipid discovery and AGILE-style platforms are more convincing when model complexity is paired with iterative experimental generation and validation of new datapoints [3, 9]. In smaller nano-QSAR and toxicity datasets, simpler baselines may be more appropriate unless authors demonstrate that added complexity improves external generalisation rather than only internal metrics [22, 24].

Interpretability and Domain Relevance

Interpretability was increasingly recognized but inconsistently implemented. Some formulation and delivery studies used feature importance or related explanatory methods to connect predictions with known design principles, such as lipid composition, molar ratios, or physicochemical properties influencing transfection and uptake [17, 21]. Toxicity models also attempted to relate descriptor importance to mechanisms such as dissolution, immune-cell response, and material composition [25, 28]. However, interpretability was often descriptive rather than causal, and few studies tested whether model-derived design rules remained valid in prospective experiments.

Reporting Completeness

Reporting completeness varied across application areas, with stronger studies describing datasets, modelling workflows, validation procedures, and experimental follow-up in enough detail to support critical interpretation. Formulation-oriented reviews argue that minimum reporting should include dataset origin, preprocessing, feature definitions, model architecture, hyperparameter tuning, evaluation metrics, uncertainty, and limitations [12, 13]. Several primary studies provided clear methodological descriptions, but others left ambiguity around split construction, duplicate handling, descriptor calculation, or hyperparameter selection [27, 29]. These reporting gaps make it difficult to distinguish genuinely generalisable models from models that perform well only within narrow experimental contexts.

Comparative Performance

Comparative performance claims were often limited by inadequate baseline selection and lack of statistical testing. Many studies compared multiple algorithms, but superiority claims were less persuasive when differences were small, confidence intervals were absent, or baselines were not tuned with comparable effort [20, 26]. In newer lipid nanoparticle studies, comparisons against alternative modelling strategies and experimental design baselines are particularly important because the practical question is whether AI reduces experimental burden or improves formulation quality, not merely whether one model has a higher retrospective score [3, 8]. Strong comparative evaluation should therefore include robust baselines, repeated splits or external tests, and outcome measures tied to experimental success.

Results – Translational Readiness Assessment

In-Vitro and In-Vivo Validation

Only a minority of studies moved decisively beyond in-silico prediction toward experimentally validated formulation discovery. Computationally guided self-assembling drug nanoparticle design is a prominent example because predictions were connected to high-throughput experimental screening and formulation confirmation [1]. In-vivo relevance was more directly addressed in tumor delivery, tissue distribution, and brain delivery models, which used animal biodistribution or delivery outcomes as endpoints rather than purely physicochemical proxies [15, 16, 27]. Even so, most studies stopped at retrospective prediction, in-vitro characterization, or toxicity classification, leaving prospective and blinded experimental validation underdeveloped [30].

Clinical and Industrial Translation

Clinical and industrial translation remained aspirational rather than established across the mapped evidence. Reviews of artificial intelligence in nanomedicine emphasize that models must eventually account for manufacturability, scale-up, batch variability, regulatory expectations, and clinically meaningful endpoints, not just optimized laboratory readouts [13, 14]. Lipid nanoparticle studies are closest to an industrially relevant trajectory because they address mRNA delivery, ionizable lipid discovery, and formulation design in spaces already connected to clinical product development [3, 8, 9]. However, none of the mapped primary studies demonstrated a fully traceable pathway from AI-designed nanocarrier to clinical candidate, regulatory submission, or human evaluation.

Barriers to Translation

The main barriers to translation were data scarcity, limited external validation, inconsistent reporting, weak reproducibility, and poor alignment between computational endpoints and clinical decision needs. Nanotoxicology reviews highlight that heterogeneous assays, incomplete metadata, and limited harmonization undermine model generalisation, especially when predictions are intended to inform safety assessment [10, 11, 31]. Formulation studies face parallel challenges because models trained within one nanoparticle class or laboratory workflow may not extrapolate to new materials, manufacturing processes, or biological systems [5, 12]. The field therefore needs shared benchmarks, prospective validation, and closer collaboration among computational scientists, formulation scientists, toxicologists, clinicians, industry, and regulators [23].

A Field in Its Early Adolescence

AI-driven nanomedicine design appears to be in an early adolescent phase: energetic, technically ambitious, and rapidly expanding, but not yet governed by stable methodological norms. The field has moved beyond isolated nano-QSAR studies into formulation discovery, lipid nanoparticle optimization, biodistribution modelling, and online nanoinformatics platforms [3, 7, 15]. At the same time, reviews of AI in nanomedicine repeatedly emphasize that proof-of-concept accuracy is not equivalent to translational readiness [14, 23]. The central pattern is therefore one of impressive breadth but uneven maturity, with methodological standards lagging behind modelling ambition.

Figure 1 synthesizes the review findings into an evidence-to-translation landscape showing how AI nanomedicine models progress from nanoparticle data and prediction tasks toward methodological appraisal and translational readiness.

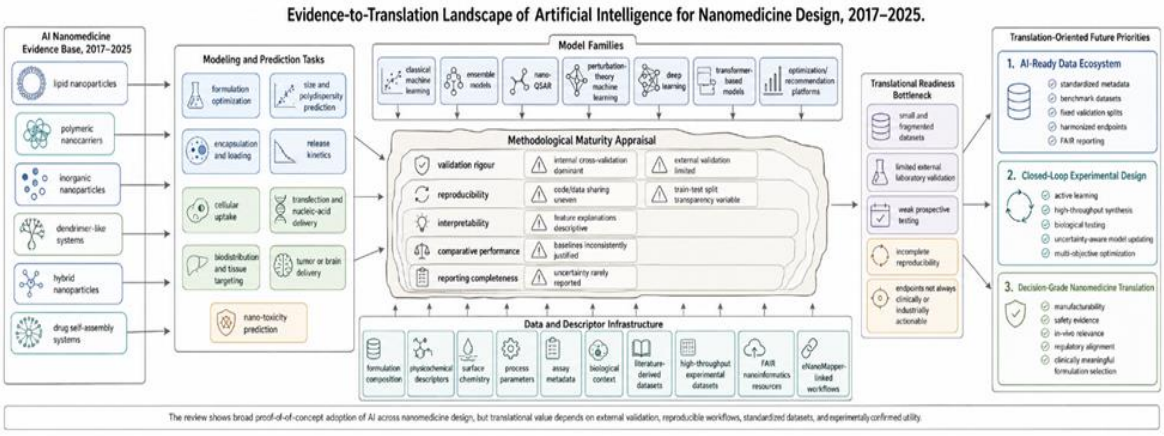

Figure 1. Evidence-to-Translation Landscape of Artificial Intelligence for Nanomedicine Design

The Validation Gap as a Core Weakness

The validation gap is the most persistent threat to the credibility of AI-designed nanomedicine. Internal cross-validation can be useful during development, but it is insufficient when nanoparticle records are correlated by shared composition, similar experimental batches, or repeated assay conditions [12, 20]. Studies predicting nanoparticle delivery to tumors, tissues, and brain provide more clinically relevant endpoints, yet their translational meaning depends on whether models generalize across laboratories, animal models, particle types, and dosing conditions [15, 16, 27]. Prospective experimental confirmation, as demonstrated in computationally guided nanoparticle design, should become the benchmark for claims that AI can guide formulation development [1].

Reproducibility and Open Science

Reproducibility is not a secondary concern in AI nanomedicine; it is a prerequisite for scientific accumulation. When studies omit code, raw data, preprocessing pipelines, descriptor definitions, hyperparameter settings, or train-test splits, independent groups cannot determine whether reported performance reflects robust modelling or dataset-specific artefacts [4, 11]. FAIR-oriented resources and eNanoMapper workflows provide an important infrastructure for reusable nanomaterial data, but their impact depends on routine adoption by authors, reviewers, and journals [6]. Without open workflows, the field risks producing many apparently successful models that cannot be compared, reproduced, or translated.

What Makes a Nanomedicine Model "Translational"?

A translational nanomedicine model should do more than predict a retrospective endpoint with high numerical accuracy. It should operate inside a clearly defined applicability domain, predict properties that matter for formulation performance or clinical development, and generate recommendations that can be synthesized, characterized, scaled, and tested [5, 13]. Lipid nanoparticle models for mRNA delivery and ionizable lipid discovery are closer to this standard because they connect molecular or formulation design with experimentally meaningful delivery outcomes [3, 9]. However, even these studies require stronger evidence on manufacturability, batch robustness, safety, and in-vivo relevance before they can be considered clinically enabling.

Table 2 provides a translational-readiness appraisal framework for distinguishing proof-of-concept AI nanomedicine models from models capable of supporting experimental or regulatory decision-making.

Table 2. Translational Readiness Appraisal Framework for AI-Enabled Nanomedicine Design

Appraisal dimension	Low-readiness pattern	Intermediate-readiness pattern	High-readiness pattern	Evidence needed to support translational claims	Practical implication for future studies
Validation design	Internal cross-validation or random train-test split only	Independent test set or repeated split design within the same dataset	External dataset, prospective experimental confirmation, or independent laboratory replication	Transparent split construction, external test performance, prospective prediction-to-experiment evidence	Claims should be scaled to the validation design; internal validation alone should not justify translational claims
Dataset provenance and representativeness	Small, single-laboratory, poorly described, or literature-mined dataset with limited metadata	Curated dataset with defined inclusion rules and partial metadata harmonization	Multi-source dataset with standardized metadata, assay context, and defined applicability domain	Dataset origin, inclusion criteria, assay protocol, nanoparticle class coverage, missing-data handling	AI nanomedicine studies should define where the model can and cannot be used
Reproducibility	Code, raw data, preprocessing, descriptors, and train-test splits unavailable	Partial sharing of data or model details, but incomplete executable workflow	Public or controlled-access data, code, preprocessing scripts, descriptors, hyperparameters, and exact splits available	Data availability statement, code repository or controlled-access mechanism, workflow documentation	Reproducibility should become a minimum condition for cumulative nanomedicine AI research
Model complexity and overfitting control	Complex model used on small or homogeneous data without sufficient baseline comparison	Multiple models compared, but uncertainty, calibration, or robustness testing limited	Complexity justified by dataset size, external validation, robust baselines, calibration, and uncertainty analysis	Tuned baseline models, repeated experiments, confidence intervals, calibration curves, uncertainty estimates	Deep learning or transformer models should be justified by demonstrable generalization, not novelty alone
Interpretability and domain relevance	Feature importance absent or purely post hoc without biological interpretation	Descriptive explanations linked to known formulation or toxicity variables	Explanations tested against domain knowledge and used to guide prospective formulation or safety decisions	Feature-level explanations, applicability-domain analysis, prospective testing of design rules	Interpretability should help formulation scientists act, not merely decorate model outputs
Endpoint actionability	Endpoint is statistically convenient but weakly linked to formulation or clinical decisions	Endpoint reflects formulation or biological performance but lacks scale-up or safety context	Endpoint supports decision-making about synthesis, delivery, biodistribution, toxicity, manufacturability, or clinical relevance	Justification of endpoint relevance, linkage to experimental or translational decisions	Models should prioritize outcomes that influence real formulation development pathways
Comparative performance	Algorithm superiority claimed from small metric differences without statistical support	Several algorithms compared with common metrics but limited statistical testing	Baselines, statistical uncertainty, external tests, and experimental success measures included	Confidence intervals, significance or robustness testing, experimental burden reduction, success-rate comparison	Comparative claims should address whether AI improves experimental design, not only retrospective scores
Translational integration	Model remains a retrospective computational exercise	Model suggests candidate formulations or mechanisms but lacks full experimental follow-up	Model is embedded in synthesis, characterization, biological testing, manufacturability, or regulatory-evidence pathway	Traceable evidence chain from data to prediction to experiment to decision	The highest-value studies should connect AI outputs to experimentally verified formulation choices

Comparison with AI in Other Drug Discovery Areas

Compared with AI in small-molecule drug discovery, AI nanomedicine remains less standardized and more fragmented. Small-molecule modelling benefits from larger shared datasets, established descriptors, common benchmark tasks, and clearer

regulatory experience, whereas nanomedicine datasets are often smaller, assay-dependent, and difficult to harmonize across materials and biological systems [11, 12]. Nanotoxicology has made progress through nano-QSAR and toxicity-prediction models, but even there, mixture effects, dynamic transformations, and incomplete metadata complicate benchmarking [26, 29]. This comparison suggests that nanomedicine AI needs not only better algorithms but also a stronger data culture.

Emerging Opportunities

Emerging opportunities are most compelling where AI is linked to iterative experimentation rather than isolated retrospective prediction. Active learning, multi-objective optimization, and self-driving laboratory concepts could help navigate formulation spaces where particle size, loading, stability, delivery, and toxicity must be optimized simultaneously [1, 5]. Deep learning platforms and transformer-based models suggest a path toward recommendation systems for lipid nanoparticles and related nanocarriers, especially when paired with high-throughput synthesis and biological testing [3, 8]. The next methodological frontier is therefore closed-loop learning in which experimental results continuously update the model and refine the design space.

Limitations of the Mixed-Methods Approach

The mixed-methods approach provides a broad and critical synthesis, but it also has limitations. Evidence mapping necessarily simplifies heterogeneous studies into categories such as model type, nanoparticle class, endpoint, and validation strategy, which may obscure important technical differences between individual workflows [4, 23]. The appraisal is limited to what authors reported, so studies with strong internal practices but incomplete publications may have been judged less favourably than their actual conduct warrants [13]. Conversely, polished reporting cannot substitute for external validation, so the review prioritized methodological evidence over rhetorical claims of translational impact [11].

Strengths and Limitations of this Review

Strengths

A major strength of this review is its combination of systematic evidence mapping with methodological and translational appraisal. Rather than treating AI nanomedicine as a single homogeneous field, the review distinguishes formulation optimization, lipid nanoparticle design, cellular uptake, biodistribution, release prediction, and toxicity modelling as related but methodologically distinct domains [2, 17, 19]. It also integrates primary modelling studies with reviews, perspectives, and data-infrastructure papers, allowing the field's technical outputs to be interpreted alongside its reproducibility and standardization challenges [6, 14]. This structure makes the review useful for researchers designing new models and for stakeholders evaluating whether existing models are ready to guide experiments.

Limitations

This review is limited by its focus on English-language peer-reviewed publications and by the likelihood that some industrial or proprietary AI nanomedicine models remain unpublished. Because commercial formulation platforms, internal pharmaceutical datasets, and confidential regulatory interactions are not consistently visible in the literature, the apparent scarcity of clinical translation may underestimate private-sector activity [13, 23]. The reference set also emphasizes studies with explicit AI, machine learning, deep learning, or nanoinformatics terminology, which may exclude adjacent computational optimization studies that did not use those labels [5]. Finally, methodological appraisal depended on published reporting, which may incompletely capture data provenance, model tuning, and experimental validation practices.

Research Gaps

Standardized Benchmarking

The absence of widely accepted benchmark datasets is a central research gap. Nanomedicine lacks the equivalent of mature benchmark suites that allow models to be compared across common tasks, fixed splits, and transparent evaluation metrics [11, 12]. Current public resources and nanoinformatics platforms are valuable starting points, but they do not yet provide broad standardized benchmarks covering size, encapsulation, release, uptake, biodistribution, immunogenicity, and toxicity across nanoparticle classes [6, 7]. Without shared benchmarks, claims of model superiority will remain difficult to interpret across studies.

Prospective and External Validation

Prospective and external validation remains too rare for a field that aims to guide experimental nanomedicine design. Retrospective performance can identify useful patterns, but the decisive test is whether a model can recommend new formulations that perform as predicted under independent experimental conditions [1]. This is especially important for lipid nanoparticles, polymeric polyplexes, and toxicity prediction, where small changes in composition, cell type, assay conditions, or biological context can substantially alter outcomes [17, 21, 30]. Future studies should therefore include blinded prospective validation whenever feasible and external laboratory replication for models intended to influence translational decisions.

Models for Novel and Complex Nanocarriers

Most current models remain better suited to relatively constrained formulation spaces than to complex, multifunctional, or biomimetic nanocarriers. Lipid nanoparticles have recently attracted sophisticated AI workflows because of their clinical relevance and high-dimensional formulation chemistry [3, 8, 9]. However, targeted nanoparticles, hybrid systems, dendrimer-like carriers, protein corona-aware systems, and stimuli-responsive nanomedicines remain under-represented in standardized modelling studies [14, 23]. Expanding AI to these systems will require richer descriptors, more complete biological metadata, and validation strategies that capture dynamic behaviour in biological environments.

Recommendations

For Researchers

Researchers should design AI nanomedicine studies around generalisation, reproducibility, and experimental actionability from the outset. At minimum, studies should report dataset provenance, preprocessing, descriptor definitions, model-selection procedures, hyperparameters, uncertainty estimates, and exact validation splits [11, 13]. Whenever possible, authors should share code and data or provide controlled-access alternatives when privacy, commercial, or safety constraints prevent full release [6, 7]. The strongest studies will pair prediction with prospective synthesis, characterization, and biological testing, as demonstrated by computationally guided nanoparticle design workflows [1].

For Journal Editors

Journal editors should require minimum reporting standards for machine-learning-based nanomedicine studies. Manuscripts should clearly distinguish internal cross-validation from independent external validation and should justify why the chosen validation design matches the claims being made [13, 20]. For deep learning and transformer-based studies, reviewers should ask whether dataset size, regularisation, baseline comparisons, and uncertainty analysis are adequate for the model complexity [2, 8]. Journals can accelerate maturation by requiring data availability statements, code availability statements, transparent split descriptions, and limitations sections that address applicability domain and translational relevance.

For the Community

The nanomedicine AI community should develop open benchmark datasets, blind prediction challenges, and shared reporting templates. Collaborative benchmarks should include representative nanoparticle classes, harmonized metadata, standard endpoints, and predefined train-test splits to reduce selective reporting and enable fair comparison [7, 11]. Such efforts should not focus only on toxicity, although nanotoxicology has provided many of the field's most mature data discussions [10, 31]. Benchmarks for formulation performance, release, uptake, biodistribution, and nucleic-acid delivery would make AI models more useful to drug-delivery scientists and translational teams [21, 27].

For Industry and Regulators

Industry and regulators should begin defining what counts as credible computational evidence for AI-assisted nanomedicine development. Regulatory relevance will require evidence that a model is trained on fit-for-purpose data, validated in an appropriate applicability domain, reproducible, interpretable enough for risk assessment, and connected to experimentally verified formulation attributes [13, 23]. Industrial adoption will also depend on whether AI recommendations can be manufactured reproducibly and integrated into quality-by-design workflows [5, 12]. Early dialogue among model developers, formulation scientists, toxicologists, manufacturers, clinicians, and regulators would reduce the risk that technically impressive models fail at the point of translation.

Implications for Future Research and Translation

Building an AI-Ready Nanomedicine Data Ecosystem

The most important enabling condition for AI-driven nanomedicine is an AI-ready data ecosystem. Such an ecosystem should make nanoparticle composition, synthesis conditions, characterization methods, assay protocols, biological context, and outcome measures findable, accessible, interoperable, and reusable [6]. Online annotation platforms and machine-learning toolkits show how nanostructure metadata and computational workflows can be connected, but broad adoption across laboratories remains incomplete [7]. If the community invests in shared data standards, future models will be better positioned to generalize across particle classes, laboratories, and translational contexts.

From Prediction to Autonomous Experimentation

The field should move from static prediction toward autonomous and semi-autonomous experimentation. In a closed-loop pipeline, AI would propose nanoparticle formulations, robotic or high-throughput systems would synthesize and characterize them, biological assays would generate new outcomes, and the model would update its design strategy [1, 3]. This approach is especially attractive for multi-objective problems such as optimizing delivery, stability, safety, and manufacturability simultaneously [5]. However, closed-loop systems will only be reliable if they include uncertainty-aware model selection, rigorous experimental controls, and prospective validation beyond the initial optimization environment [9].

The Path to Clinical Translation

The path to clinical translation requires a shift from model performance as a computational endpoint to model qualification as part of an evidence package. AI-designed nanomedicines will need traceable data provenance, validated applicability domains, reproducible workflows, experimentally confirmed predictions, safety assessment, manufacturability evidence, and alignment with regulatory expectations [13, 23]. Models predicting toxicity, biodistribution, and delivery should be evaluated not only by statistical accuracy but also by whether they support decisions that reduce risk or improve formulation selection [15, 24]. Translation will therefore depend on integrated evidence chains linking data, models, experiments, manufacturing, and clinical relevance.

Conclusion

AI has been widely applied to nanomedicine design, demonstrating promise across multiple nanoparticle types and prediction tasks. The evidence base now includes formulation optimization, lipid nanoparticle delivery, cellular uptake, biodistribution, drug release, and toxicity prediction. Together, these studies show that computational models can help navigate complex formulation spaces that are difficult to explore experimentally by intuition alone.

However, the methodological quality of many studies is still insufficient to support clinical translation. Validation is often internal, reproducibility is frequently limited, and real-world testing remains rare. As a result, many published models are better interpreted as promising research tools than as decision-ready translational technologies.

Closing the translational gap will require community-driven standardization of datasets, benchmarks, validation designs, and reporting practices. It will also require closer integration between computational scientists and experimental nanomedicine groups. Models should be built with prospective testing, manufacturability, safety, and clinical relevance in mind from the beginning.

The next decade offers rich opportunities for AI to accelerate the development of safer and more effective nanomedicines. Those opportunities will be realized only if the field collectively raises its methodological standards. AI can become a powerful engine for nanomedicine translation, but only when predictive performance is matched by reproducibility, external validation, and experimentally demonstrated utility.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Reker D, Rybakova Y, Kirtane AR, Cao R, Yang JW, Navamajiti N, et al. Computationally guided high-throughput design of self-assembling drug nanoparticles. Nat Nanotechnol. 2021;16(6):725-33.
2. Harrison PJ, Wieslander H, Sabirsh A, Karlsson J, Malmsjö V, Hellander A, et al. Deep-learning models for lipid nanoparticle-based drug delivery. Nanomedicine. 2021;16(13):1097-110.
3. Xu Y, Ma S, Cui H, Chen J, Xu S, Gong F, et al. AGILE platform: a deep learning powered approach to accelerate LNP development for mRNA delivery. Nat Commun. 2024;15(1):6305.
4. Chen C, Yaari Z, Apfelbaum E, Grodzinski P, Shamay Y, Heller DA. Merging data curation and machine learning to improve nanomedicines. Adv Drug Deliv Rev. 2022;183:114172.
5. 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.
6. Kochev N, Jeliazkova N, Paskaleva V, Tancheva G, Iliev L, Ritchie P, et al. Your spreadsheets can be fair: A tool and fairification workflow for the enanomapper database. Nanomaterials. 2020;10(10):1908.
7. Wang T, Russo DP, Demokritou P, Jia X, Huang H, Yang X, et al. An online nanoinformatics platform empowering computational modeling of nanomaterials by nanostructure annotations and machine learning toolkits. Nano Lett. 2024;24(33):10228-36.
8. Chan A, Kirtane AR, Qu QR, Huang X, Woo J, Subramanian DA, et al. Designing lipid nanoparticles using a transformer-based neural network. Nat Nanotechnol. 2025;20(10):1491-501.
9. Li B, Raji IO, Gordon AG, Sun L, Raimondo TM, Oladimeji FA, et al. Accelerating ionizable lipid discovery for mRNA delivery using machine learning and combinatorial chemistry. Nat Mater. 2024;23(7):1002-8.
10. Singh AV, Ansari MH, Rosenkranz D, Maharjan RS, Kriegel FL, Gandhi K, et al. Artificial intelligence and machine learning in computational nanotoxicology: unlocking and empowering nanomedicine. Adv Healthc Mater. 2020;9(17):1901862.
11. Zhou Y, Wang Y, Peijnenburg W, Vijver MG, Balraadjsing S, Dong Z, et al. Application of machine learning in nanotoxicology: a critical review and perspective. Environ Sci Technol. 2024;58(34):14973-93.

12. Bao Z, Bufton J, Hickman RJ, Aspuru-Guzik A, Bannigan P, Allen C. Revolutionizing drug formulation development: The increasing impact of machine learning. Adv Drug Deliv Rev. 2023;202:115108.
13. Dorsey PJ, Lau CL, Chang TC, Doerschuk PC, D'Addio SM. Review of machine learning for lipid nanoparticle formulation and process development. J Pharm Sci. 2024;113(12):3413-33.
14. Serov N, Vinogradov V. Artificial intelligence to bring nanomedicine to life. Adv Drug Deliv Rev. 2022;184:114194.
15. Lin Z, Chou WC, Cheng YH, He C, Monteiro-Riviere NA, Riviere JE. Predicting nanoparticle delivery to tumors using machine learning and artificial intelligence approaches. Int J Nanomedicine. 2022;17:1365-79.
16. Mi K, Chou WC, Chen Q, Yuan L, Kamineni VN, Kuchimanchi Y, et al. Predicting tissue distribution and tumor delivery of nanoparticles in mice using machine learning models. J Control Release. 2024;374:219-29.
17. Loecher A, Bruyns-Haylett M, Ballester PJ, Borros S, Oliva N. A machine learning approach to predict cellular uptake of pBAE polyplexes. Biomater Sci. 2023;11(17):5797-808.
18. Concu R, Kleandrova VV, Speck-Planche A, Cordeiro MN. Probing the toxicity of nanoparticles: a unified in silico machine learning model based on perturbation theory. Nanotoxicology. 2017;11(7):891-906.
19. Santana R, Zuluaga R, Gañán P, Arrasate S, Onieva E, González-Díaz H. Predicting coated-nanoparticle drug release systems with perturbation-theory machine learning (PTML) models. Nanoscale. 2020;12(25):13471-83.
20. Ji Z, Guo W, Wood EL, Liu J, Sakkiah S, Xu X, et al. Machine learning models for predicting cytotoxicity of nanomaterials. Chem Res Toxicol. 2022;35(2):125-39.
21. Cheng L, Zhu Y, Ma J, Aggarwal A, Toh WH, Shin C, et al. Machine learning elucidates design features of plasmid deoxyribonucleic acid lipid nanoparticles for cell type-preferential transfection. ACS Nano. 2024;18(42):28735-47.
22. Na M, Nam SH, Moon K, Kim J. Development of a nano-QSAR model for predicting the toxicity of nano-metal oxide mixtures to Aliivibrio fischeri. Environ Sci Nano. 2023;10(1):325-37.
23. Chou WC, Canchola A, Zhang F, Lin Z. Machine learning and artificial intelligence in nanomedicine. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2025;17(4):e70027.
24. Ahmadi M, Ayyoubzadeh SM, Ghorbani-Bidkorpeh F. Toxicity prediction of nanoparticles using machine learning approaches. Toxicology. 2024;501:153697.
25. Huang Y, Li X, Cao J, Wei X, Li Y, Wang Z, et al. Use of dissociation degree in lysosomes to predict metal oxide nanoparticle toxicity in immune cells: Machine learning boosts nano-safety assessment. Environ Int. 2022;164:107258.
26. Zhang F, Wang Z, Peijnenburg WJ, Vijver MG. Machine learning-driven QSAR models for predicting the mixture toxicity of nanoparticles. Environ Int. 2023;177:108025.
27. Yousfan A, Al Rahwanji MJ, Hanano A, Al-Obaidi H. A comprehensive study on nanoparticle drug delivery to the brain: application of machine learning techniques. Mol Pharm. 2023;21(1):333-45.
28. Shirokii N, Din Y, Petrov I, Seregin Y, Sirotenko S, Razlivina J, et al. Quantitative prediction of inorganic nanomaterial cellular toxicity via machine learning. Small. 2023;19(19):2207106.
29. Khokhlov I, Legashev L, Bolodurina I, Shukhman A, Shoshin D, Kolesnik S. Prediction of dynamic toxicity of nanoparticles using machine learning. Toxics. 2024;12(10):750.
30. Hosni Z, Achour S, Saadi F, Chen Y, Al Qaraghuli M. Machine learning-driven nanoparticle toxicity. Ecotoxicol Environ Saf. 2025;299:118340.
31. Singh AV, Varma M, Laux P, Choudhary S, Datusalia AK, Gupta N, et al. Artificial intelligence and machine learning disciplines with the potential to improve the nanotoxicology and nanomedicine fields: a comprehensive review. Arch Toxicol. 2023;97(4):963-79.