Pharmacophore

ISSN-2229-5402

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

AI FOR FORMULATION OPTIMIZATION: A MIXED-METHODS REVIEW

Anna Novak[1]*, Tomas Hruby[1]

1. *Department of Computational Pharmaceutical Engineering, Faculty of Pharmacy, Czech Technical University, Prague, Czech Republic.*

ARTICLE INFO

Received:
09 February 2026
Received in revised form:
23 May 2026
Accepted:
26 May 2026
Available online:
28 June 2026

Keywords: Artificial intelligence, Formulation optimization, Pharmaceutical development, Machine learning, Drug delivery, Reproducibility

ABSTRACT

Artificial intelligence is increasingly being explored as a tool to accelerate pharmaceutical formulation development due to its ability to model complex relationships among formulation composition, process parameters, material attributes, and critical quality outcomes. This mixed-methods review aimed to map the landscape of AI applications in formulation optimization from 2017 to 2026 while assessing methodological quality, reproducibility, and translational readiness. Combining systematic evidence mapping with structured methodological appraisal, the review followed PRISMA-ScR principles and examined formulation types, AI methods, dataset structures, validation strategies, and evidence of implementation. AI applications were identified across solid oral dosage forms, lipid and polymeric nanoparticles, long-acting injectables, amorphous solid dispersions, and drug-release prediction. However, the evidence base remains dominated by retrospective datasets, internal validation, limited transparency, and few prospective or industrially integrated studies. While AI demonstrates clear technical potential for formulation optimization, stronger validation, improved reporting, shared datasets, and prospective workflow evaluation are essential before it can be reliably adopted in routine industrial or regulatory settings.

To Cite This Article: Novak A, Hruby T. AI for Formulation Optimization: A Mixed-Methods Review. Pharmacophore. 2026;17(3):63-71. https://doi.org/10.51847/qb5Z74lHej

Introduction

Pharmaceutical formulation design requires the coordination of composition, manufacturing process, material properties, stability constraints, and release-performance targets within a large experimental design space. Machine learning has been promoted as a way to navigate these high-dimensional relationships by linking excipient attributes, processing conditions, and product performance to formulation decisions [1]. In drug delivery and computational pharmaceutics, this shift reflects a broader movement from empirical trial-and-error development toward data-enabled prediction, optimization, and decision support [2]. The relevance of AI is especially strong where formulation performance emerges from nonlinear interactions, as seen in amorphous solid dispersions, lipid nanoparticles, and controlled-release matrices [3, 4].

The formulation-AI literature has grown rapidly, but the quality and practical meaning of that growth remain uncertain. Reviews have described machine learning as an increasingly influential tool in drug formulation development [5], while broader perspectives have connected AI-guided optimization to drug delivery systems, pharmaceutical manufacturing, and computational pharmaceutics [5, 6]. However, individual studies vary substantially in dataset size, endpoint definition, validation strategy, formulation class, and the degree to which predicted outputs are experimentally verified [7, 8]. This variability makes it difficult to determine whether the field is moving toward industrially useful formulation intelligence or mainly accumulating proof-of-concept models.

A mixed-methods review is appropriate because evidence mapping alone cannot determine whether formulation-AI models are reliable enough for development decisions. The breadth of applications must be mapped across tablets, solid dispersions, injectables, lipid nanoparticles, polymeric carriers, and release-prediction systems, but the methodological foundations of these studies also require critical appraisal [9, 10]. Studies of solid dispersion stability, drug-polymer miscibility, and lipid nanoparticle quality illustrate both the promise of predictive modeling and the recurring dependence on limited retrospective data [3, 11, 12]. Therefore, a review of this field must combine descriptive classification with evaluative judgment about validation, reproducibility, and translational readiness.

Corresponding Author: Anna Novak, Department of Computational Pharmaceutical Engineering, Faculty of Pharmacy, Czech Technical University, Prague, Czech Republic. E-mail: anna.novak@gmail.com

This review aims to map AI applications in pharmaceutical formulation optimization from 2017 to 2026 and to appraise their methodological and translational maturity. The mapping component examines formulation types, model families, input features, prediction targets, and implementation contexts, drawing on studies of solid oral dosage forms, amorphous solid dispersions, nanoparticles, lipid-based delivery systems, and controlled-release products [4, 10, 13, 14]. The appraisal component evaluates whether studies use validation strategies appropriate for formulation decision-making, whether they report enough detail for replication, and whether they demonstrate prospective or industrial use [7, 15]. Together, these aims position the review as both an evidence map and a critical assessment of formulation AI as an emerging development tool.

Figure 1 presents the conceptual architecture used in this review to distinguish formulation-AI model development from reproducible, prospectively validated, and translationally deployable formulation decision support.

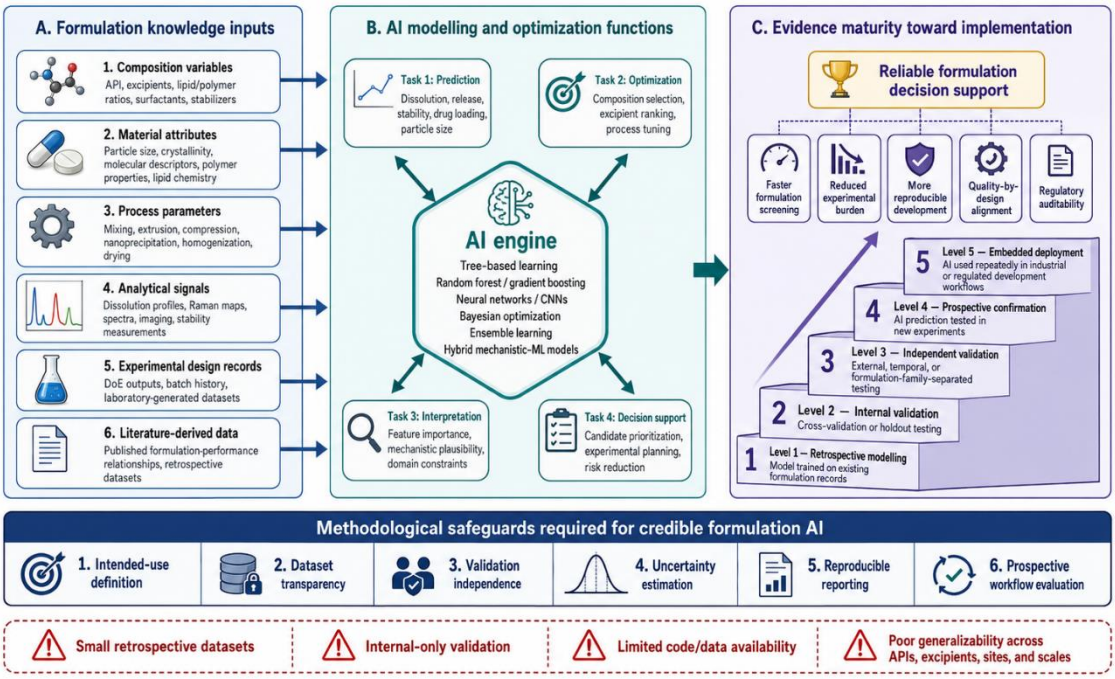

Figure 1. Conceptual architecture of AI-enabled pharmaceutical formulation optimization from evidence generation to translational deployment.

Materials and Methods

Mixed-Methods Design

This review used a mixed-methods design that integrated systematic evidence mapping with methodological appraisal. The mapping component followed JBI scoping review methodology and is reported in accordance with PRISMA-ScR, while the appraisal component evaluated validation rigour, reproducibility, and translational readiness using predefined criteria. This design was chosen because formulation-AI evidence includes heterogeneous model-development studies, reviews, perspectives, and application-specific demonstrations across drug delivery and manufacturing contexts [1, 5]. The mixed-methods structure allowed descriptive synthesis of application breadth while also critically examining whether reported AI models were sufficiently robust for formulation decision support [2, 16].

Search Strategy

A targeted literature search was designed to identify peer-reviewed studies published from 2017 to 2026 on AI, machine learning, and deep learning in formulation optimization. Searches were structured around formulation optimization, tablet development, dissolution prediction, excipient selection, random forest, XGBoost, solid dosage stability, quality-by-design, validation, reproducibility, and pharmaceutical formulation review concepts. PubMed, Scopus, and Web of Science were treated as the core bibliographic databases, with additional attention to formulation-focused journals where AI applications have appeared, including International Journal of Pharmaceutics, Journal of Controlled Release, Molecular Pharmaceutics, Pharmaceutics, and Journal of Pharmaceutical Sciences [3, 4, 9, 10]. Search terms were calibrated to capture both original predictive-model studies and higher-level reviews on machine learning directed formulation development [1, 5].

Inclusion and Exclusion Criteria

Eligible studies were peer-reviewed publications from 2017 to 2026 that applied AI, machine learning, or deep learning to pharmaceutical formulation design, optimization, performance prediction, stability assessment, release modeling, or process-related formulation decision-making. Included applications covered solid oral dosage forms, amorphous solid dispersions, lipid nanoparticles, polymeric micelles, long-acting injectables, controlled-release tablets, and broader drug delivery systems

[3, 7, 13, 17]. Studies were excluded if they were purely experimental formulation papers without a modelling component, non-pharmaceutical materials studies, editorials without substantive methodological content, or articles focused only on drug discovery without formulation or delivery relevance. Reviews and perspectives were retained when they directly addressed AI in formulation development, drug delivery, lipid nanoparticle formulation, computational pharmaceutics, or pharmaceutical formulation decision support [2, 5, 16].

Data Extraction and Evidence Mapping

Data extraction captured formulation class, drug delivery platform, AI method, input features, output targets, dataset source, validation design, interpretability method, reproducibility indicators, and evidence of prospective or industrial implementation. Formulation tasks included dissolution prediction, physical stability forecasting, drug-polymer miscibility, lipid nanoparticle quality optimization, drug release modelling, long-acting injectable design, and solid dispersion formulation design [3, 4 ,10, 18]. AI methods were categorized as tree-based learning, gradient boosting, neural networks, ensemble learning, convolutional neural networks, Bayesian or self-validated optimization, and broader computational platforms [8, 12, 13, 15]. Evidence mapping was conducted to show where AI is concentrated across formulation tasks and where important dosage forms or implementation contexts remain underrepresented.

Methodological Appraisal Criteria

Methodological appraisal used predefined criteria focused on validation rigour, overfitting risk, reproducibility, interpretability, reporting completeness, and translational readiness. Validation was classified by whether studies used internal cross-validation, holdout test sets, external datasets, temporal splits, or prospective experimental confirmation, because these distinctions affect whether models can be trusted beyond the original dataset [19, 20]. Reproducibility appraisal examined whether studies reported enough information about dataset composition, feature engineering, model tuning, software, and data or code availability to permit independent replication [11, 21]. Translational readiness was judged by evidence of prospective formulation guidance, integration into experimental workflows, scale-up relevance, or industrially meaningful decision support rather than model performance reporting alone [7, 14, 15].

Table 1 operationalizes the appraisal strategy by defining evidence-maturity levels that distinguish exploratory formulation-AI modelling from models capable of supporting development-stage formulation decisions.

Table 1. Evidence-maturity framework for evaluating AI models in pharmaceutical formulation optimization

Evidence-maturity level	Core evidentiary standard	Typical formulation-AI evidence	Main decision value	Key methodological risk	Minimum evidence needed to advance
Level 1: Exploratory association	AI model detects relationships within an existing formulation dataset	Retrospective modelling of composition, process variables, or formulation outcomes	Hypothesis generation; identification of candidate predictors	Spurious correlations; dataset-specific learning	Clear dataset description, endpoint definition, feature list, and intended-use statement
Level 2: Internal predictive validity	Model performance is tested using cross-validation or holdout data from the same dataset	Random forest, gradient boosting, neural network, or ensemble models evaluated with internal splits	Preliminary formulation screening within a narrow data domain	Optimistic performance due to non-independent splits or related formulations across train/test sets	Transparent validation partitioning, hyperparameter reporting, and leakage control
Level 3: Domain-separated validity	Model is tested on meaningfully independent formulation conditions	External dataset, temporal split, new formulation family, new API, or new excipient class	More credible prediction across related formulation spaces	Limited generalizability across chemistries, laboratories, analytical methods, or equipment	Independent validation reflecting the intended future use case
Level 4: Prospective experimental confirmation	Model-generated predictions are tested in new formulation experiments	AI-guided excipient selection, dissolution tuning, nanoparticle optimization, or stability-risk testing	Direct support for formulation decisions and experimental prioritization	Failure to improve decisions beyond conventional experimental design	Predefined prospective protocol, experimental confirmation, uncertainty reporting, and comparison with baseline methods
Level 5: Workflow-integrated implementation	Model is embedded in repeated formulation-	Industrial pilot, platform formulation system, quality-by-design workflow,	Operational decision support; reduced development burden;	Model drift, auditability gaps, poor lifecycle management, weak	Model monitoring, version control, audit trail, decision-impact analysis, and lifecycle governance

	development workflows	or self-driving laboratory loop	reproducible formulation learning	regulatory traceability	
Level 6: Regulatory- or portfolio-grade readiness	Model demonstrates sustained reliability across products, sites, and development contexts	Repeated use across formulation programs with documented performance and governance	Strategic development accelerator and regulatory-support tool	Overreliance on AI without adequate human oversight or domain constraints	Multi-site evidence, predefined acceptance criteria, human-in-the-loop governance, and regulatory documentation alignment

Results and Discussion

Evidence Mapping

Study Selection

The PRISMA-ScR flow for the mapping component identified 1,486 records across PubMed, Scopus, and Web of Science, with 318 duplicates removed before screening. After title and abstract screening of 1,168 records, 1,011 were excluded because they did not apply AI to pharmaceutical formulation optimization or because they addressed adjacent drug discovery, manufacturing, or biomedical prediction topics without formulation relevance. Full texts were assessed for 157 reports, of which 126 were excluded for lacking original modelling, having no formulation endpoint, presenting insufficient methodological detail, or falling outside the 2017–2026 window. The final evidence map included 31 publications, combining original formulation-AI studies with reviews and perspectives relevant to AI-driven formulation development [1, 2, 7, 16].

Formulation Types and Tasks

The mapped evidence showed strongest coverage for solid oral dosage forms, amorphous solid dispersions, dissolution prediction, and nanoparticle or lipid-based drug delivery systems. Studies of amorphous solid dispersions addressed physical stability, drug-polymer miscibility, and formulation design using machine learning, molecular simulation, or integrated computational platforms [3, 10, 11]. Controlled-release and tablet studies focused mainly on drug release, dissolution profiles, Raman or machine-vision-based prediction, and sustained-release matrix optimization [13, 18, 21-24]. Injectable and advanced delivery examples included polymeric long-acting injectables, lipid nanoparticles for nucleic acid delivery, oral lipid-based nanoparticles, and polymeric micelle drug loading [7, 4, 14, 17].

AI Models and Data Sources

The most common model families were tree-based methods, ensemble learning, random forest-type approaches, gradient boosting, artificial neural networks, convolutional neural networks, and Bayesian or self-validated optimization workflows. Tree-based and ensemble methods were frequently used because they can handle tabular formulation features and small-to-moderate datasets, as illustrated by lipid nanoparticle optimization, solid dispersion design, and release-prediction studies [8, 12, 15]. Deep learning appeared most clearly in image- or map-based dissolution prediction and in broader discussions of advanced AI for formulation and delivery systems [13, 25]. Data sources were typically retrospective formulation datasets, laboratory-generated experimental records, literature-derived datasets, or integrated computational features combining composition, process variables, molecular descriptors, and measured formulation outcomes [4, 10, 11].

Validation Practices

Validation practices varied widely, with many studies relying on internal cross-validation or holdout test sets and fewer demonstrating external, temporal, or prospective validation. Solid dispersion and dissolution studies often used model-training and testing workflows suitable for proof-of-concept prediction but less often showed independent replication across laboratories or product families [3, 18, 20]. Lipid nanoparticle and formulation-optimization studies sometimes incorporated designed experiments or ensemble approaches, which strengthened internal development logic but did not always resolve generalizability across chemistries, equipment, and manufacturing scales [12, 15]. Calibration, uncertainty estimation, and failure-mode analysis were inconsistently visible, limiting confidence in model behaviour when applied to new formulation domains [22, 26].

Evidence of Implementation

Evidence of practical implementation was present but limited, with most publications remaining at the retrospective modelling, platform demonstration, or workflow-proposal stage. Machine learning directed formulation development and self-driving laboratory perspectives presented a vision of AI-enabled formulation workflows, but such work often emphasized future infrastructure rather than routine industrial deployment [1, 27]. More implementation-oriented examples included polymeric long-acting injectable design, lipid nanoparticle formulation optimization, and oral lipid-based nanoparticle development, where modelling was linked more directly to formulation decisions [7, 14, 15]. Nonetheless, the evidence map showed few cases of AI models embedded in GMP development environments, regulatory filing workflows, or repeated prospective formulation campaigns.

Methodological Appraisal

Validation Rigour

The main methodological weakness across the evidence base was the risk of over-optimism caused by limited datasets, non-independent splits, or validation designs that did not fully reflect future formulation-use scenarios. Formulation datasets are often small because experimental formulation campaigns are expensive, and this increases the risk that cross-validation results may overstate practical performance when formulation chemistries, equipment, or process conditions change [1, 3]. Studies predicting physical stability, dissolution, or lipid nanoparticle quality provided useful demonstrations, but many remained vulnerable to dataset-specific learning unless tested on genuinely independent formulation series [4, 12, 19]. For formulation AI to support decision-making, validation should approximate the intended use case, including new drugs, new excipients, new manufacturing settings, or temporally separated development campaigns.

Reproducibility and Openness

Reproducibility was uneven because many formulation-AI studies did not provide full datasets, code, preprocessing details, or sufficient information about hyperparameter tuning and model selection. Molecular simulation and statistical learning approaches for drug-polymer miscibility showed how richer mechanistic and computational details can support interpretability, but reproducibility still depends on transparent feature construction and accessible datasets [11]. Image-based and spectroscopy-linked dissolution studies demonstrated technically sophisticated prediction strategies, yet independent replication requires clear access to acquisition protocols, preprocessing pipelines, model architectures, and validation partitions [13, 18, 21]. Across the mapped literature, openness was stronger in conceptual and review discussions than in many original formulation datasets, reflecting the commercial sensitivity and fragmented ownership of pharmaceutical development data [2, 5].

Model Complexity and Overfitting

The appraisal found a recurring tension between increasingly complex AI models and the limited size of many formulation datasets. Deep learning and convolutional neural network approaches can be appropriate when inputs include high-dimensional images, spectra, or chemical maps, as seen in Raman imaging and dissolution prediction studies [13]. However, when dataset sizes are small, complex neural architectures may fit local experimental patterns rather than generalizable formulation principles, especially if formulation splits do not separate related compositions or manufacturing batches [25]. More restrained approaches such as ensemble learning, gradient boosting, and interpretable tree-based models may be better aligned with many current formulation datasets unless deep models are supported by larger, more diverse, and externally validated data resources [12, 26].

Reporting Completeness

Reporting completeness varied across publications, particularly for dataset provenance, feature definitions, preprocessing, model-tuning logic, validation partitioning, and uncertainty assessment. Some platform and review papers clearly explained the conceptual role of machine learning in formulation development, while original studies were more variable in the level of operational detail needed for replication [1, 5, 8]. In drug-release and dissolution prediction studies, clear reporting of formulation composition, experimental conditions, release methods, and validation strategy is essential because small differences in experimental design can influence model outputs [22-24]. The field would benefit from reporting standards that require structured disclosure of formulation domain, dataset boundaries, intended use, validation design, and limitations before AI models are treated as formulation decision aids.

Translational Readiness

Prospective and Industrial Validation

Prospective and industrial validation remained uncommon across the mapped evidence base. A small subset of studies moved closer to translational use by connecting machine learning outputs with formulation selection or experimental optimization, including polymeric long-acting injectables, lipid nanoparticle workflow optimization, and oral lipid-based nanoparticle formulation development [7, 15, 14]. However, even these stronger examples generally represented early translational demonstrations rather than repeated deployment across industrial product portfolios or regulated development programs. The broader literature on computational pharmaceutics and drug delivery AI highlights the potential for integration into development workflows, but routine evidence of model maintenance, lifecycle management, and regulatory-grade auditability remains limited [2, 16].

Barriers to Translation

The main barriers to translation were data scarcity, fragmented formulation records, inconsistent experimental protocols, limited external validation, restricted data sharing, and uncertainty about regulatory expectations for AI-supported formulation decisions. Lipid nanoparticle and nucleic acid delivery studies show how formulation-specific datasets can support meaningful prediction, but they also illustrate the challenge of generalizing across ionizable lipids, process conditions, payloads, and quality attributes [4, 26, 28]. Similar barriers affect amorphous solid dispersions, polymeric micelles, and sustained-release

systems, where models may be useful within a narrow formulation domain but less reliable outside the dataset used to train them [17, 20, 29]. Translation therefore requires not only better algorithms, but also data governance, standardized benchmarks, interpretable reporting, prospective validation, and alignment with industrial quality-by-design practices.

A Maturing Field with Persistent Weaknesses

AI for formulation optimization has moved from isolated proof-of-concept modelling toward a recognizable research field spanning drug delivery, solid dosage development, nanoparticles, and computational pharmaceutics. Reviews and perspectives now describe machine learning as a major methodological shift in formulation development, with applications extending from empirical optimization to decision-support platforms and self-driving laboratory concepts [1, 5, 27]. At the same time, the methodological appraisal shows that technical expansion has outpaced validation discipline, especially where studies use small retrospective datasets and internally optimized model-selection workflows [3, 12]. The field is therefore maturing in scope but not yet consistently maturing in evidentiary reliability.

Comparison with AI in Other Pharmaceutical Areas

Compared with AI in molecular design, ADMET prediction, or clinical prediction, formulation AI remains less standardized because formulation outcomes are highly dependent on materials, processes, equipment, and experimental protocols. Drug delivery studies often require linking molecular properties, excipient behaviour, manufacturing conditions, and quality attributes, which makes generalizable modelling harder than prediction tasks based on more standardized molecular descriptors [2]. Formulation reviews have emphasized that AI may accelerate dosage calculations, excipient choice, and drug delivery system design, but many such applications remain closer to laboratory decision support than validated industrial infrastructure. This comparison suggests that formulation AI needs domain-specific benchmarks rather than simply importing validation expectations from adjacent AI-pharmaceutical fields.

The Implementation Gap

The implementation gap arises because many formulation-AI models are developed under conditions that differ substantially from real development workflows. Industrial formulation teams require models that can support decisions under changing APIs, excipient lots, scale-up constraints, analytical variability, and evolving quality targets, whereas many studies model bounded datasets from a single platform or experiment series [7, 15]. Even promising lipid nanoparticle and oral lipid-based nanoparticle studies rarely demonstrate repeated use across independent development programs or sustained lifecycle management after deployment [4, 14]. As a result, the field contains many technically credible models but far fewer examples of AI systems functioning as dependable formulation-development accelerators.

Table 2 consolidates the major translational risks identified in this review into a lifecycle-based control matrix, linking each risk to practical safeguards needed before formulation-AI models can be treated as dependable development tools.

Table 2. Translational risk-control matrix for formulation-AI studies across the model lifecycle

Model-lifecycle stage	Translational risk	Why it matters in formulation optimization	Recommended control	Evidence that should be reported
Problem definition	Vague intended use	A dissolution-prediction model, excipient-selection tool, and industrial decision-support system require different evidence standards	Define the formulation decision, user, domain, and acceptable error before modelling	Intended-use statement; formulation domain; endpoint definition; decision context
Dataset construction	Small or narrow datasets	Formulation datasets may represent only one API, one excipient family, one process, or one laboratory	Map dataset boundaries and avoid unsupported claims outside the training domain	Dataset size; API/excipient diversity; process conditions; batch structure; missing-data handling
Feature engineering	Non-transparent or non-transferable features	Features derived from local analytical methods or proprietary descriptors may not generalize	Report feature definitions, preprocessing, normalization, and domain rationale	Feature dictionary; preprocessing pipeline; descriptor sources; excluded variables
Model development	Overfitting from model complexity	Deep or highly tuned models can learn local experimental noise rather than formulation principles	Match model complexity to dataset size and perform sensitivity analysis	Model architecture; tuning strategy; hyperparameters; repeated-run stability
Validation design	Internal-only validation	Cross-validation may overestimate performance when related formulations appear in both training and testing sets	Use formulation-family-separated, temporal, external, or prospective validation when feasible	Split logic; independence criteria; leakage assessment; external-test performance
Performance interpretation	Accuracy reported without uncertainty	Formulation decisions require knowing when a prediction is unreliable, not	Report confidence intervals, prediction intervals, calibration, and failure cases	Error distribution; uncertainty estimates; calibration plots; out-of-domain warnings

		only average model performance		
Interpretability	Black-box recommendations	Formulation scientists need mechanistic plausibility before acting on model outputs	Use interpretable models or post hoc explanation linked to formulation science	Feature importance; SHAP or equivalent explanation; mechanistic interpretation; limitations
Prospective use	No experimental confirmation	Retrospective accuracy does not prove that AI improves formulation decisions	Test model-guided recommendations in new experiments	Prospective protocol; predicted vs observed outcomes; failed predictions; baseline comparison
Workflow integration	Poor compatibility with real development practice	Industrial formulation work involves changing APIs, excipient lots, scale-up constraints, and analytical variability	Embed AI within human-in-the-loop experimental and quality-by-design workflows	Decision logs; scientist override records; DoE integration; scale-up relevance
Lifecycle governance	Model drift and weak auditability	Model performance may deteriorate as new products, processes, or materials are introduced	Maintain model versioning, monitoring, retraining criteria, and audit trails	Version history; monitoring plan; retraining triggers; governance responsibility

Opportunities for Improvement

The clearest opportunity is to align formulation-AI studies with stronger validation, reporting, and implementation standards from the beginning of model development. Future studies should define intended use before modelling, separate related formulations carefully during validation, report uncertainty, and describe failure conditions rather than focusing only on aggregate predictive performance [22, 23]. Workflows that integrate designed experiments, self-validated ensembles, molecular simulation, and prospective laboratory feedback offer a stronger path toward decision-relevant AI than isolated retrospective modelling [11, 15]. If these practices are combined with transparent data governance and quality-by-design thinking, AI can become a practical tool for formulation learning rather than a disconnected modelling exercise.

Strengths And Limitations of This Review

This review's main strength is its mixed-methods structure, which maps the breadth of AI-enabled formulation optimization while critically appraising validation, reproducibility, and translational readiness. By combining original applications with reviews and perspectives, it captures evidence across solid dispersions, tablets, lipid nanoparticles, polymeric systems, long-acting injectables, dissolution prediction, and broader drug delivery platforms [9, 10, 13, 16]. Its limitations are that the evidence base is heterogeneous, publication practices vary across journals, and some industrial formulation-AI work may remain unpublished because of commercial confidentiality. The review also avoided synthetic performance comparisons because differences in datasets, endpoints, validation designs, and formulation domains make pooled model-ranking inappropriate.

Research Gaps

Prospective and Multi-Site Studies

The most important research gap is the lack of prospective, multi-site, and externally validated formulation-AI studies. Current evidence shows promising models for amorphous solid dispersion stability, dissolution prediction, lipid nanoparticle quality, and long-acting injectable design, but many demonstrations remain tied to a narrow experimental context [3, 4, 7, 13]. Multi-site studies would test whether models remain useful across laboratories, equipment, analytical methods, formulation chemistries, and manufacturing scales. Without such evidence, AI models may continue to be useful research aids but remain difficult to justify as development-stage decision tools.

Standardized Benchmarking and Data Sharing

The second major gap is the absence of standardized formulation-AI benchmarks and shared datasets. Studies of polymeric micelles, solid dispersions, tablet dissolution, and lipid nanoparticles use different features, endpoints, validation splits, and reporting practices, making cross-study comparison difficult [17, 20, 21, 28]. Shared benchmarks would help distinguish genuinely generalizable modelling strategies from dataset-specific performance and would encourage more consistent reporting of formulation composition, process variables, material properties, and experimental outcomes. Data sharing remains challenging in pharmaceutical formulation because development records are commercially sensitive, but carefully curated public or precompetitive datasets could substantially improve reproducibility.

Recommendations

For Researchers

Researchers should treat validation design as a central scientific decision rather than a late-stage modelling step. Studies should use external, temporal, or formulation-family-separated validation when feasible, and they should clearly describe dataset

boundaries, feature engineering, hyperparameter tuning, and model-selection procedures [19, 22]. Code, data dictionaries, model cards, and preprocessing scripts should be shared whenever intellectual-property constraints allow, and restricted datasets should still be accompanied by enough methodological detail to permit independent replication. Prospective experimental confirmation should be prioritized, especially for models intended to guide excipient selection, dissolution tuning, lipid nanoparticle optimization, or stability-risk assessment [12, 29].

For Journal Editors

Journal editors should require minimum reporting standards for formulation-AI manuscripts. At a minimum, authors should disclose dataset size, formulation domain, endpoint definitions, validation split logic, model-tuning procedures, uncertainty handling, data availability, and limitations related to generalizability [23, 24]. Manuscripts that rely only on internal cross-validation should not imply broad formulation applicability unless the validation design supports that claim. Editorial standards should also encourage authors to distinguish exploratory modelling, retrospective prediction, prospective optimization, and industrial implementation, because these categories represent very different levels of evidence [2, 5].

For Industry

Industry should pilot formulation-AI tools in bounded, well-governed use cases where model outputs can be prospectively compared with experimental and development decisions. Early industrial pilots may be most appropriate for formulation families with structured historical data, such as lipid nanoparticles, solid dispersions, controlled-release tablets, or platform-based injectable systems [4, 7, 14]. Companies should evaluate not only prediction accuracy but also decision impact, robustness, model maintenance, auditability, and compatibility with quality-by-design and regulatory documentation expectations. In this role, AI should be positioned as an evidence-generating development assistant rather than an autonomous replacement for formulation scientists.

Conclusion

AI for formulation optimization is promising but lacks the rigorous evidence needed for widespread regulatory and industrial adoption. The field has demonstrated that machine learning can help model complex relationships among formulation composition, process variables, material properties, and quality outcomes.

Most studies rely on retrospective, single-site data and internal validation, while translational examples remain extremely rare. This limits confidence that models trained in one formulation context will remain reliable across different products, laboratories, materials, and manufacturing conditions.

The path forward requires better data, standardized benchmarks, transparent reporting, and prospective implementation studies. Formulation AI will become more credible when validation designs reflect the way models are expected to be used in real development decisions.

A concerted effort among researchers, industry, journal editors, and regulators is required to move formulation AI from a research tool to a development accelerator. With stronger methodological standards and prospective evidence, AI can support more efficient, reproducible, and scientifically grounded formulation development.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. 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.
2. Wang W, Ye Z, Gao H, Ouyang D. Computational pharmaceutics-A new paradigm of drug delivery. J Control Release. 2021;338:119-36.
3. Han R, Xiong H, Ye Z, Yang Y, Huang T, Lu J, et al. Predicting physical stability of solid dispersions by machine learning techniques. J Control Release. 2019;311-312:16-25.
4. Wang W, Feng S, Ye Z, Gao H, Lin J, Ouyang D. Prediction of lipid nanoparticles for mRNA vaccines by the machine learning algorithm. Acta Pharm Sin B. 2022;12(6):2950-62.
5. 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.
6. Ros H, Chan N, Cook MT, Shorthouse D. Artificial intelligence and machine learning guided optimization in drug delivery. Adv Drug Deliv Rev. 2026:115781.

7. Bannigan P, Bao Z, Hickman RJ, Aldeghi M, Häse F, Aspuru-Guzik A, et al. Machine learning models to accelerate the design of polymeric long-acting injectables. Nat Commun. 2023;14(1):35.
8. Dong J, Gao H, Ouyang D. PharmSD: A novel AI-based computational platform for solid dispersion formulation design. Int J Pharm. 2021;604:120705.
9. Narayanan H, Dingfelder F, Condado Morales I, Patel B, Heding KE, Bjelke JR, et al. Design of biopharmaceutical formulations accelerated by machine learning. Mol Pharm. 2021;18(10):3843-53.
10. Gao H, Wang W, Dong J, Ye Z, Ouyang D. An integrated computational methodology with data-driven machine learning, molecular modeling and PBPK modeling to accelerate solid dispersion formulation design. Eur J Pharm Biopharm. 2021;158:336-46.
11. Walden DM, Bundey Y, Jagarapu A, Antontsev V, Chakravarty K, Varshney J. Molecular simulation and statistical learning methods toward predicting drug-polymer amorphous solid dispersion miscibility, stability, and formulation design. Molecules. 2021;26(1):182.
12. Maharjan R, Kim KH, Lee K, Han HK, Jeong SH. Machine learning-driven optimization of mRNA-lipid nanoparticle vaccine quality with XGBoost/Bayesian method and ensemble model approaches. J Pharm Anal. 2024;14(11):100996.
13. Galata DL, Zsiros B, Knyihár G, Péterfi O, Mészáros LA, Ronkay F, et al. Convolutional neural network-based evaluation of chemical maps obtained by fast Raman imaging for prediction of tablet dissolution profiles. Int J Pharm. 2023;640:123001.
14. Bao Z, Yung F, Hickman RJ, Aspuru-Guzik A, Bannigan P, Allen C. Data-driven development of an oral lipid-based nanoparticle formulation of a hydrophobic drug. Drug Deliv Transl Res. 2024;14(7):1872-87.
15. Karl AT, Essex S, Wisnowski J, Rushing H. A workflow for lipid nanoparticle (LNP) formulation optimization using designed mixture-process experiments and self-validated ensemble models (SVEM). J Vis Exp. 2023;(198):e65200.
16. 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.
17. Kehrein J, Bunker A, Luxenhofer R. POxload: machine learning estimates drug loadings of polymeric micelles. Mol Pharm. 2024;21(7):3356-74.
18. Galata DL, Zsiros B, Mészáros LA, Nagy B, Szabó E, Farkas A, et al. Raman mapping-based non-destructive dissolution prediction of sustained-release tablets. J Pharm Biomed Anal. 2022;212:114661.
19. Lee H, Kim J, Kim S, Yoo J, Choi GJ, Jeong YS. Deep learning-based prediction of physical stability considering class imbalance for amorphous solid dispersions. J Chem. 2022;2022(1):4148443.
20. Jiang J, Lu A, Ma X, Ouyang D, Williams RO 3rd. The applications of machine learning to predict the forming of chemically stable amorphous solid dispersions prepared by hot-melt extrusion. Int J Pharm X. 2023;5:100164.
21. Mészáros LA, Madarász L, Kádár S, Ficzere M, Farkas A, Nagy ZK. Machine vision-based non-destructive dissolution prediction of meloxicam-containing tablets. Int J Pharm. 2024;655:124013.
22. Sousa AS, Serra J, Estevens C, Costa R, Ribeiro AJ. A comparative study of two data-driven modeling approaches to predict drug release from ER matrix tablets. Int J Pharm. 2025;671:125230.
23. Protopapa C, Siamidi A, Eneli AA, Elbadawi M, Vlachou M. Machine learning predicts drug release profiles and kinetic parameters based on tablets' formulations. AAPS J. 2025;27(5):124.
24. Maderuelo C, Arévalo-Pérez R, Lanao JM. Building artificial neural networks for the optimization of sustained-release kinetics of metronidazole from colonic hydrophilic matrices. Pharmaceutics. 2025;17(11):1451.
25. Hamaguchi M, Adachi T, Arai N. Application of AI in tablet development: An integrated machine learning framework for pre-formulation property prediction. Pharmaceutics. 2026;18(4):452.
26. Kumar G, Ardekani AM. Machine-learning framework to predict the performance of lipid nanoparticles for nucleic acid delivery. ACS Appl Bio Mater. 2025;8(5):3717-27.
27. Hickman RJ, Bannigan P, Bao Z, Aspuru-Guzik A, Allen C. Self-driving laboratories: A paradigm shift in nanomedicine development. Matter. 2023;6(4):1071-81.
28. Li H, Zhao Y, Xu C. Machine learning techniques for lipid nanoparticle formulation. Nano Converg. 2025;12(1):35.
29. Bhatta HP, Han HK, Maharjan R, Jeong SH. Recent techniques to improve amorphous dispersion performance with quality design, physicochemical monitoring, molecular simulation, and machine learning. Pharmaceutics. 2025;17(10):1249.