Integrating Pharmacometrics and Machine Learning for Enhanced Clinical Trial Outcomes

MARIA CACCIAPUOTI, MAURO D'ARCO, ALESSANDRO MINELLA
Department of Information Technology and Electrical Engineering (DIETI)
University of Naples Federico II
Via Claudio, 21, 80125 Napoli
ITALY

Abstract: - A novel computational framework that integrates the mechanistic rigor of Pharmacometrics (PMx) with the predictive power of Multimodal Machine Learning (MML) is conceptually formulated and discussed. The framework is designed to address critical developmental bottlenecks in the drug development sector, which is already witnessing a paradigm shift driven by the rise of Antibody-Drug Conjugates (ADCs)—a class that, according to recent pipeline analyses, accounts for an expanding share of innovative oncology trials. Unfortunately, despite the therapeutic promise of ADCs and other targeted modalities, the industry faces a productivity paradox, characterized by high attrition rates in transitional clinical phases and a median development timeline of seven years. By leveraging transfer learning strategies, we delineate how the synergistic adoption of PMx and MML can conceptually optimize dose selection, inform metastatic risk assessment through genomic-imaging integration, and ultimately accelerate the transition from bench to bedside. This is conceptualized through the definition of methodologies capable of transforming noisy early-phase clinical data into actionable insights, thus bridging the gap between Physiologically Based Pharmacokinetic (PBPK) modeling and deep learning-driven outcomes. Without advancing premature empirical claims, we establish a structured methodological roadmap for a scalable, AI-enhanced clinical trial infrastructure, positioning hybrid modeling as a viable architectural paradigm for the next generation of precision oncology.

Key-Words: - Antibody-Drug Conjugates, Pharmacometrics, Machine Learning, Transfer Learning, Clinical Trial Optimization, Hybrid Modeling, Precision Oncology

Received: May 14, 2026. Revised: July 16, 2026. Accepted: August 18, 2026. Published: November 16, 2026.

1 Introduction

Clinical pharmacology has evolved into a multidisciplinary field closely integrated with data science. This integration, however, is the result of a long-standing methodological evolution that dates back to the 1950s, when clinical pharmacologists first introduced compartmental mathematical models to describe drug concentration–time profiles in patients.

The integration of clinical pharmacology with mathematical modeling, namely pharmacokinetic (PK) modeling, was further advanced through the introduction of nonlinear mixed-effects models in the 1970s by Sheiner [1] and Beal [2]. These approaches enabled the joint modeling of individual pharmacokinetic profiles and inter-individual variability within a population framework. Pharmacometrics (PMx) can thus be defined as the discipline encompassing pharmacokinetic/pharmacodynamic (PK/PD) modeling and quantitative data analysis. Traditional PMx models are typically formulated using systems of differential equations and are therefore grounded in mechanistic knowledge based on biological and pharmacological principles.

Although these methodologies now constitute a cornerstone of quantitative clinical pharmacology and data-driven drug development [3], [4], their early applications were primarily limited to the quantitative description of PK/PD relationships, particularly focusing on drug concentration-time profiles and dose-response dynamics. At this stage, the contribution of PMx to decision-making processes remained relatively limited, with only marginal impact on clinical trial design and development strategies. Subsequently, the application of modeling and simulation (M&S) progressively expanded to critical aspects of experimental design, most notably dose selection across different phases of clinical development. This shift has driven increasing interest from both the pharmaceutical industry and regulatory agencies in the potential of quantitative modeling as a decision-support tool throughout the drug development lifecycle. However, the high structural complexity of PMx models leads to significant computational costs, especially in large-scale population analyses [5]. A critical bottleneck is covariate selection, which remains a resource-intensive process often limited by linear assumptions that cannot capture the

high-dimensional, non-linear interactions inherent in modern data.

In this context, the integration of Machine

Learning (ML) represents a strategic evolution to address these challenges [6]. ML methodologies offer powerful tools for extracting latent patterns from multidimensional datasets, enabling deeper insights into treatment outcomes that traditional PK/PD models might miss.

In recent years, the Scientific Machine Learning (SciML) paradigm has emerged as a significant methodological shift. Unlike "black-box" AI, SciML embeds data-driven adaptive capacity directly into mechanistic frameworks, effectively using neural networks to approximate unknown terms within systems of differential equations. By combining the interpretative structure of PMx with the flexibility of ML, the trade-off between physiological consistency and predictive accuracy is reconciled, providing a robust framework for complex tasks such as metastasis risk prediction [7].

At the intersection of pharmacometrics and machine learning, diverse methodological strategies have been proposed to support drug development from discovery through clinical and regulatory evaluation [8]. These include both model-informed approaches, where ML is employed to refine PK/PD models, and purely data-driven predictive methodologies, reflecting a flexible and rapidly evolving methodological landscape. Nevertheless, important structural limitations persist. While traditional mechanistic models ensure interpretability and physiological consistency, they inevitably require the simplification of complex biological phenomena and may lose accuracy when underlying mechanisms are not fully understood. Conversely, ML approaches, despite their flexibility and ability to capture complex nonlinear relationships, typically require large volumes of well-curated and consistently labeled data, making them less effective in clinical settings characterized by limited, noisy, or partially observed data.

A suitable trade-off between interpretability, as addressed by mechanistic models, and predictive performance, as provided by ML approaches, is particularly evident in the study of tumor progression, and especially in the prediction of metastasis risk. The metastatic process is governed by complex and only partially understood biological dynamics, which are difficult to observe in early stages and often not directly measurable in available clinical datasets. As a result, neither purely mechanistic models nor exclusively data-driven approaches are sufficient to fully capture its complexity. Given these considerations, it can be opportune to explore hybrid frameworks that combine the interpretative structure of pharmacometric models with the adaptive capacity of machine learning.

Rather than presenting a completed software

implementation or experimental trial validation, the primary contribution of this perspective paper is to establish a conceptual and methodological framework. We explicitly delineate how physiologically based pharmacokinetic (PBPK) layers can be structurally interfaced with multimodal deep learning architectures via domain-adapted transfer learning. By clarifying the operational tiers, data streams, and mechanistic coupling, this work provides a conceptual blueprint for next-generation clinical decision-support systems in oncology.

The paper is structured as follows: Section 2 discusses the main issues in clinical trials paying attention to the significant case of Antibody-Drug Conjugates (ADCs); Section 3 details the transfer learning strategy and its operational mapping across source and target domains; Section 4 discusses the perspectives toward scalable multi-modal architectures for oncology decision support, detailing the multi-tier hybrid architecture and data interfaces; finally, Section 5 provides concluding remarks distinguishing established literature findings from proposed conceptual directions.

2 Issues in clinical trials

Modern therapeutic strategies are increasingly converging toward precision medicine. The most emblematic example is that of Antibody-Drug Conjugates (ADCs), which have emerged as one of the most promising therapeutic platforms, accounting for approximately 35% of oncology clinical trials initiated in 2024 among targeted biotherapeutic pipelines [9]. Structurally, ADCs consist of a monoclonal antibody linked to a cytotoxic payload through a chemical linker, thereby enabling the selective delivery of chemotherapy to tumor cells [9].

From a clinical pharmacology standpoint, ADCs exhibit multifaceted disposition dynamics that present distinct challenges for conventional empirical modeling. An ADC in circulation behaves as a dynamic mixture characterized by its Drug-to-Antibody Ratio (DAR). Over time, systemic deconjugation gradually reduces the average DAR, releasing unconjugated cytotoxic payload that can lead to severe off-target toxicities. Concurrently, ADC clearance is governed both by non-specific proteolytic elimination and Target-Mediated Drug Disposition (TMDD), where high-affinity binding to target tumor antigens induces receptor internalization and saturation. Furthermore, tumor cell eradication frequently

depends on the "bystander killing effect," wherein uncharged, cleaved payload diffuses across cell membranes to eliminate adjacent antigen-negative malignant cells. Capturing these interdependent phenomena—intact conjugate disposition, payload deconjugation kinetics,

target-mediated clearance, and microenvironmental diffusion—requires mechanistic pharmacometric structures that standard statistical models cannot adequately represent.

The landscape of ADCs has recently entered a phase of intense dynamism; according to clinical pipeline reports, between late 2024 and early 2025, 148 regulatory status changes were recorded, including 69 clinical progressions and 31 new assets entering Phase I [10]. As of July 2025, the global pipeline encompasses 2,168 conjugates, with 21% represented by next-generation modalities such as Bispecific ADCs (BsADCs), Dual-Payload ADCs, and Antibody-Degrader Conjugates (DACs). Despite this expansion, with 17% of conventional ADCs in active clinical investigation, the field faces substantial structural challenges. Approximately 40% of cumulative historical assets are currently inactive or discontinued, hampered by low early-stage efficiency, where only approximately 10% of Phase I candidates reach approval, and a median development timeline of 7 years [10].

Detailed transition analyses reveal a significant bottleneck in mid-stage development: the Phase I–II success rate (46%) and Phase II–III rate (24%) lag behind broader oncology and general pharmaceutical averages. Conversely, ADCs reaching Phase III demonstrate a robust 82% approval probability. While the cumulative success rate for ADCs (9%) outperforms oncology as a whole (3.4%), it remains below the 14% industry average across all therapeutic classes combined [11].

This development paradigm imposes substantial financial and temporal burdens on sponsors. The high attrition observed in early-stage development translates directly into capital inefficiency, with estimated investments ranging from $1 to $2.8 billion per approved asset when accounting for capitalized discovery failures [9], [12]. In parallel, the global oncology market is projected to reach $441 billion by 2029, making stagnating research and development productivity a critical structural constraint for the competitiveness of pharmaceutical pipelines [13]. In this scenario, the considerable costs associated with clinical trial execution make investment in advanced decision-support technologies increasingly necessary.

In this context, the integration of Artificial Intelligence (AI) represents much more than a mere technological enhancement, as it is a candidate to become a strategic necessity for effective project de-risking. To bridge this gap, AI must be integrated not as a black-box predictor, but as a synergistic partner to Pharmacometrics (PMx). The solution lies in moving beyond simple input-output mappings toward multi-modal frameworks that incorporate

Physiologically Based Pharmacokinetic (PBPK) layers.

The application of machine learning in medicine primarily focuses on predictive modeling of clinical outcomes and risk stratification in high-complexity settings. Unlike traditional statistical approaches, these methods learn non-linear representations of patient trajectories by integrating heterogeneous and longitudinal data. Machine learning thus transforms historical and real-time clinical data into quantitative prognostic indicators, enabling early detection of adverse events and disease progression patterns [14].

In oncology, predictive modeling has substantially improved metastatic risk assessment through the integration of genomic data, clinical biomarkers, and imaging features. These algorithms process complex datasets to estimate the probability of metastatic events, from lymph node involvement to distant organ dissemination, converting large-scale data into actionable clinical predictions [15]. Such systems address the limitations of classical statistical models in high-dimensional settings. They provide a flexible representation of risk that does not rely on linear assumptions or restricted covariate spaces. Both supervised and unsupervised approaches have shown strong performance in identifying latent patterns associated with tumor progression, leveraging structured data (clinical and laboratory variables) and unstructured data (radiological images and genomic sequences) [16].

To address these limitations in data-scarce and high-dimensional clinical settings, advanced learning paradigms such as transfer learning have emerged as effective strategies. In fact, by adopting transfer learning, as discussed in the subsequent section, it is possible to leverage knowledge from data-rich source domains to mitigate the scarcity of early-phase clinical data. This proposed conceptual approach provides a foundation for transforming noisy, high-dimensional datasets into robust prognostic indicators, such as metastatic risk and overall survival, ultimately converting high-risk pipelines into efficient, data-driven portfolios with a higher probability of regulatory success.

3 Transfer learning strategy

Transfer learning is formally defined as a learning paradigm where knowledge acquired from a source domain and task is distilled and transferred to a target domain characterized by limited data availability (see Figure 1). Within this framework, a model pre-trained on large-scale, high-dimensional datasets learns generalizable feature representations, such as fundamental chemical toxicities or protein-binding patterns, which can then be fine-tuned for specific downstream clinical or pharmacological tasks. This

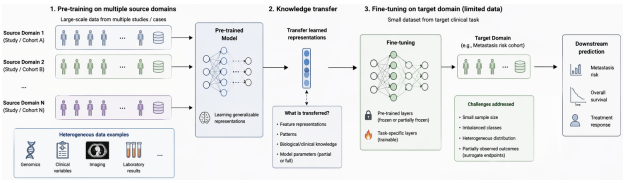


Figure 1: Schematic representation of the conceptual transfer learning workflow. In Phase 1 (Pre-training on multiple source domains), heterogeneous data across multiple cohorts are integrated to learn generalized representations. In Phase 2 (Knowledge transfer), learned features, biological patterns, clinical knowledge, and model parameters are extracted. In Phase 3 (Fine-tuning on target domain), the model is adapted to a data-scarce target trial by freezing pre-trained feature layers and optimizing task-specific layers, directly mitigating small sample size, class imbalance, and sparse observations for downstream predictions.

Source: created by the authors.

approach is particularly vital in biomedical settings where small sample sizes, heterogeneous data distributions, and partially observed outcomes often hinder standard machine learning applications. In the context of clinical trials, transfer learning mitigates the structural limitations of sparse and imbalanced datasets by leveraging latent representations learned from related, data-rich domains. This capability is crucial for metastasis risk modeling, where the target variable is frequently defined indirectly through surrogate endpoints of disease progression.

As schematically detailed in Figure 1, the operational transfer learning architecture is structured into three sequential stages:

1. **Pre-training on Multiple Source Domains (Stage 1):** Large-scale multi-cohort data (Source Domains 1 through N , representing distinct historical studies and patient cohorts A to N) are gathered. As depicted in the heterogeneous data examples box, inputs combine genomics, clinical variables, radiological imaging, and longitudinal laboratory results. Training on this broad distribution enables the network to learn robust, generalizable latent representations.

2. **Knowledge Transfer (Stage 2):** The central question of *"What is transferred?"* is addressed by extracting four distinct informational components: (i) learned feature representations, (ii) latent biological and clinical patterns, (iii) pharmacological domain knowledge, and (iv) model parameters (either partial weights for feature extraction backbones or full multi-task parameter sets).

3. **Fine-Tuning on Target Domain (Stage 3):** The pre-trained architecture is deployed onto a target clinical trial cohort characterized by limited data availability (e.g., an early-phase metastasis risk cohort). In this phase, pre-trained layers are kept frozen or partially frozen to preserve generalizable biological priors, while task-specific downstream layers remain trainable to adapt to target endpoints.

This tiered workflow directly addresses the four structural bottlenecks highlighted in Figure 1: small sample sizes ($N \ll P$), severe class imbalance, heterogeneous multi-site data distributions, and partially observed progression outcomes reliant on surrogate endpoints. The fine-tuned architecture is thereby empowered to deliver reliable downstream predictions for metastatic risk, overall survival trajectories, and individualized treatment response.

Different transfer strategies can be considered, including homogeneous transfer learning, where source and target domains share similar feature spaces, and heterogeneous transfer learning, where knowledge is transferred across different data modalities or representations. These approaches enable the extraction of shared biological patterns and improve both predictive accuracy and model robustness [17].

Within this framework, recent computational breakthroughs have highlighted that specialized pre-trained architectures, specifically Graph Neural Networks (GNNs) for mapping molecular topologies and Transformer-based models for capturing sequential protein dependencies, provide notable predictive advantages in pharmacokinetic and Absorption, Distribution, Metabolism, and Excretion (ADME) prediction tasks [17]. Furthermore, this paradigm has evolved into multimodal frameworks that unify diverse data

streams (including longitudinal clinical trajectories, genomic profiles, and imaging-derived features) within a single predictive architecture [18]. Empirical comparative studies in precision oncology have demonstrated that multimodal integration can yield quantifiable predictive advantages over unimodal baselines—frequently achieving improvements of 5% to 15% in area under the ROC curve (AUC) and concordance index (C -index) for survival and progression endpoints [16], [18]. These findings substantiate the methodological rationale for integrating multi-source clinical data.

4 Toward scalable multi-modal learning frameworks

To clarify the scope of this work, the architecture illustrated in Figure 2 is presented as a proposed conceptual framework rather than a pre-existing, commercially deployed software tool. The purpose of this section is to detail the operational tiers, data interfaces, and hybrid integration mechanisms that define this methodology.

In line with the architectural principles introduced by TrialBench [19], such a system should integrate heterogeneous data sources within a unified predictive architecture. However, in the context of monoclonal antibodies and ADCs, drug representation requires a substantial extension beyond standard approaches based on molecular descriptors or simplified ADME modules. Specifically, this framework incorporates an explicit pharmacometric layer grounded in Physiologically Based Pharmacokinetic (PBPK) and PK/PD models, enabling a mechanistic description of systemic exposure, tissue distribution, target engagement, and dose-response relationships, as illustrated in Figure 2.

As depicted in Figure 2, the framework is organized into five interconnected operational tiers:

Tier 1: Multimodal Data Sources. The input layer aggregates data across two primary avenues: (i) preclinical and Phase I data (including in vitro binding affinities, xenograft profiles, and early dose-escalation observations), and (ii) patient-level multimodal clinical streams. As delineated in Figure 2, these streams encompass Radiology/Imaging (CT, MRI, PET scans), Histology (immunohistochemistry [IHC], H&E staining, and spatial tissue features), Blood tests (longitudinal hematology, biochemistry, circulating tumor DNA [ctDNA], circulating tumor cells [CTCs], and

cytokines), and Clinical/Demographic covariates (age, sex, disease stage, prior treatments, and comorbidity scores).

Tier 2: Modality-Specific Encoding. Rather than pooling raw heterogeneous data directly, dedicated neural encoders extract intermediate feature embeddings tailored to each modality: 3D convolutional or vision-transformer encoders for imaging scans, graph neural architectures for histopathological cell topologies, temporal recurrent encoders for longitudinal blood markers, and dense neural encoders for tabular clinical covariates.

Tier 3: Pharmacometric Layer (Mechanistic Information). Tier 3 constitutes the physiological foundation of the framework. A multi-compartment PBPK model represents drug disposition across key physiological compartments, including lungs, liver, kidneys, muscle, clearance organs, and tumor tissue. Coupled with dynamic PK/PD structures, this layer simulates whole-body kinetics, target-mediated drug disposition (TMDD), dynamic linker deconjugation, and intracellular payload release. As highlighted in Figure 2, the mechanistic layer outputs key pharmacological parameters, including individual clearance rates, steady-state volume of distribution, systemic exposure metrics, and continuous dose-response curves, which serve as mechanistic features alongside data-driven representations.

Tier 4: Multimodal ML Integration (Hybrid Approach). Tier 4 fuses the modality-specific latent embeddings from Tier 2 with the mechanistic pharmacometric features from Tier 3. Within a Multimodal Fusion module, attention-based layers integrate biological exposure constraints with observational patient profiles. The unified representation is then processed by a deep Predictive Layer designed to learn complex, non-linear prognostic patterns while remaining anchored to physiological plausibility.

Tier 5: Outcomes (Decision Support). The predictive layer generates five actionable clinical decision-support outputs (Figure 2):

- **Metastatic Risk Prediction:** Probabilistic risk assessment of organ-specific dissemination.
- **Overall Survival (OS) Prediction:** Continuous time-to-event survival prognostic trajectories.
- **Disease-Free Survival (DFS) Prediction:** Risk modeling for tumor progression and recurrence.
- **Dose Optimization:** Model-informed dose titration balancing efficacy with off-target payload toxicity.
- **Trial Design Support:** Quantitative patient stratification and optimal trial arm allocation.

Crucially, as indicated by the dashed feedback pathway in Figure 2, the framework features an *"Iterative learning & updating with new data"* mechanism: incoming trial observations are continuously recycled to dynamically recalibrate

pharmacometric rate constants and update predictive neural weights.

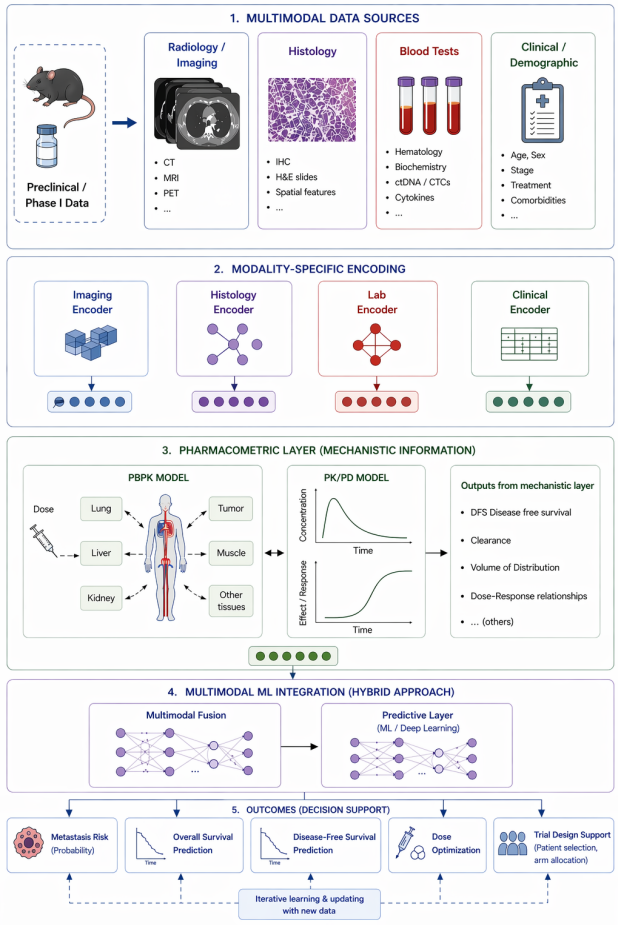


Figure 2: Detailed schematic of the proposed multimodal PBPK–ML conceptual framework for oncology decision support. Tier 1 ingests multimodal data sources alongside preclinical and Phase I kinetics. Tier 2 processes these through modality-specific encoders. Tier 3 incorporates a mechanistic PBPK/PK–PD layer describing multi-organ disposition and dose-response dynamics. Tier 4 integrates mechanistic and empirical representations via multimodal fusion and deep predictive modeling. Tier 5 outputs clinical decision endpoints (metastatic risk, OS, DFS, dose optimization, and trial design), closed by an iterative learning loop.

Source: created by the authors.

From a methodological perspective, training

such a hybrid architecture relies on a composite objective. Rather than optimizing purely on predictive error, the objective balances supervised clinical loss (such as survival hazard estimation or metastatic risk accuracy) with mechanistic fidelity penalties that minimize discrepancies with observed pharmacokinetic concentrations and enforce mass conservation. Furthermore, domain adaptation objectives can be incorporated to minimize distribution shifts between preclinical/historical source domains and target trial cohorts. To protect against unwarranted confidence in small Phase I/II cohorts, uncertainty quantification techniques—such as Bayesian deep ensembles or Monte Carlo dropout—can be implemented, separating data noise (aleatoric uncertainty) from model parameter uncertainty (epistemic uncertainty) to provide trial investigators with explicit confidence boundaries.

Within this setting, preclinical data and early-phase clinical evidence (e.g., Phase I studies) can be jointly leveraged to characterize the kinetic and dynamic profile of a novel therapeutic agent. This information can then be transferred to clinically relevant predictive tasks. The resulting framework does not rely solely on direct outcome prediction; rather, it integrates mechanistic knowledge with observational data to support the estimation of complex endpoints, including metastatic risk, overall survival, and disease-free survival.

For monoclonal antibodies and next-generation ADCs, system dynamics cannot be adequately represented through a simple input–output mapping. Instead, a structured representation of the underlying biological processes is required [20].

The integration of PBPK modeling with machine learning addresses a critical industry need: combining the high predictive performance of deep learning with the rigorous interpretability of pharmacometrics [21]. This hybrid synergy improves model generalization across diverse patient populations and clinical settings. From a translational standpoint, the long-term objective is a generalizable decision-support platform suitable for industrial-scale applications. Such a platform aims to streamline candidate selection, dose optimization, and trial design through a reusable, scalable architecture, providing a foundational conceptual roadmap for the next generation of oncology drug development.

5 Conclusion

The application of machine learning models, both in purely data-driven settings and within model-informed machine learning paradigms, is expected to play a very important role in drug development. These methodologies provide

predictive tools that support decision-making in complex and high-dimensional clinical contexts.

Established findings in current literature demonstrate that purely data-driven deep learning models are prone to overfitting, lack mechanistic interpretability, and struggle when applied to the sparse, noisy, and non-uniformly sampled cohorts typical of early-phase oncology trials. Conversely, traditional pharmacometric models provide

physiological transparency and mass conservation, but cannot readily scale to high-dimensional multimodal data (such as digital pathology and genomics) or capture complex non-linear progression trajectories without restrictive linear assumptions.

The conceptual contribution of this work lies in establishing a structured architectural roadmap that unifies these paradigms. By interfacing multi-compartment PBPK/PK–PD mechanisms directly with modality-specific neural encoders and transfer learning pipelines, the proposed framework preserves physiological plausibility while capturing high-dimensional patient heterogeneity. The multi-tiered data flow, combined with iterative updating and uncertainty estimation, provides a clear methodological design for future implementations.

Future perspectives and practical translation will require assembling collaborative, multi-institutional clinical trial datasets that combine longitudinal pharmacokinetic measurements with paired imaging and multi-omics profiles. Evaluating this hybrid framework against standalone empirical and mechanistic models across prospective ADC pipelines represents the necessary next step to transition precision oncology from empirical trial-and-error toward AI-guided, physiologically anchored decision support.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the author(s) utilized artificial intelligence (AI) technologies solely for the purpose of linguistic refinement and grammatical correction. Following the application of these tools, the author(s) comprehensively reviewed, verified, and edited the resulting text. The author(s) maintain full accountability and responsibility for the accuracy, integrity, and final content of the publication.

References:

[1] L. B. Sheiner, B. Rosenberg, V. V. Marathe, "Estimation of population characteristics of pharmacokinetic parameters from routine clinical data", J. Pharmacokinet. Biopharmaceut., Vol. 5, 1977, pp. 445–479.

[2] L. B. Sheiner, S. L. Beal, "Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data", J. Pharmacokinet. Biopharmaceut., Vol. 8, 1980, pp. 553–571.

[3] L. Zhang, M. Pfister, B. Meibohm, "Concepts and challenges in quantitative pharmacology and model-based drug development", AAPS Journal, Vol. 10, 2008, pp. 552–559.

[4] G. Koch, J. Schropp, "Mathematical concepts in pharmacokinetics and pharmacodynamics with application to tumor growth", Nonautonomous Dynamical Systems in the Life Sciences, Springer, 2013.

[5] G. Koch, M. Pfister, I. Daunhawer, M. Wilbaux, S. Wellmann, J. E. Vogt, "Pharmacometrics and Machine Learning Partner to Advance Clinical Data Analysis", Clinical Pharmacology & Therapeutics, Vol. 107, 2020, pp. 926–933. doi: 10.1002/cpt.1774

[6] L. Hutchinson, B. Steiert, A. Soubret, J. Wagg, A. Phipps, R. Peck, J. E. Charoin, B. Ribba, "Models and Machines: How Deep Learning Will Take Clinical Pharmacology to the Next Level", CPT: Pharmacometrics & Systems Pharmacology, Vol. 8, 2019, pp. 131–134. doi: 10.1002/psp4.12377

[7] C. Rackauckas et al., "Universal differential equations for scientific machine learning", arXiv preprint, 2021. Available: https://arxiv.org/abs/2001.04385

[8] Y. Ryeznik, O. Sverdlov, "Pharmacometrics and machine learning in drug development", Artificial Intelligence for Drug Product Lifecycle Applications, Academic Press, 2025, pp. 99–108. doi: 10.1016/B978-0-323-91819-0.00004-X

[9] IQVIA Institute for Human Data Science, "Global Oncology Trends 2025", IQVIA, May 2025. Available: https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/global-oncology-trends-2025

[10] Beacon (Hanson Wade), "ADC Landscape Review Series: H1 2025", Beacon Intelligence, 2025. Available: https://www.beacon-intelligence.com/landscape-reviews/adc-h1-2026-landscape-review/

[11] Beacon (Hanson Wade), "ADC Success Rates & Time in Phase", Beacon ADC Database, 2025. Available: https://www.beacon-intelligence.com/solutions/adc-database/

[12] O. J. Wouters, M. McKee, J. Luyten, "Estimated Research and Development Investment Needed to Bring a New Medicine to Market", JAMA, Vol. 323, No. 9, 2020, pp. 844–853. doi: 10.1001/jama.2020.1166

[13] Deloitte Centre for Health Solutions, "Measuring the Return from Pharmaceutical Innovation", 15th edition, Deloitte, March 2025. Available: https://www.deloitte.com/us/en/Industries/life-sciences-health-care/articles/measuring-return-from-pharmaceutical-innovation.html

[14] J. Sidey-Gibbons, C. Sidey-Gibbons, "Machine learning in medicine: a practical introduction", BMC Medical Research Methodology, Vol. 19, 2019, Article 64. doi: 10.1186/s12874-019-0681-4

[15] L. Yu, Z. Huang, Z. Xiao, X. Tang, Z. Zeng, X. Tang, W. Ouyang, "Unveiling the best predictive models for early onset metastatic cancer: Insights and innovations (Review)", Oncology Reports, Vol. 51, No. 4, 2024, Article 60. doi: 10.3892/or.2024.8719

[16] G. H. Abbas, E. R. Khouri, O. Thaher, S. Taha, M. Vladimirov, R. J. Oviedo, J. Schmidt, D. Bausch, S. Pouwels, "Predictive modeling for metastasis in oncology: current methods and future directions", Annals of Medicine and Surgery, Vol. 87, No. 6, 2025, pp. 3489–3508. doi: 10.1097/MS9.0000000000003279

[17] W. Guo, Y. Dong, G.-F. Hao, "Transfer learning empowers accurate pharmacokinetics prediction of small samples", Drug Discovery Today, Vol. 29, No. 4, 2024, Article 103946. doi: 10.1016/j.drudis.2024.103946

[18] Khan SN, Danishuddin, Khan MWA, Guarnera L, Akhtar SMF. "Multi-modal AI in precision medicine: integrating genomics, imaging, and EHR data for clinical insights." Front Artif Intell. 2026 Jan 7;8:1743921. doi: 10.3389/frai.2025.1743921. PMID: 41573270; PMCID: PMC12819606.

[19] J. Chen, Y. Hu, M. Cai et al., "TrialBench: Multi-Modal AI-Ready Datasets for Clinical Trial Prediction", Scientific Data, Vol. 12, 2025, Article 1564. doi: 10.1038/s41597-025-05680-8

[20] W.-C. Chou, Z. Lin, "Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling", Toxicological Sciences, Vol. 191, No. 1, 2023, pp. 1–14. doi: 10.1093/toxsci/kfac101

[21] Q. Yao, Y. Lee, T. Zhou, "Pharmacokinetic/pharmacodynamic modeling of therapeutic antibodies in oncology: current advances and future perspectives", Journal of Pharmaceutical Investigation, Vol. 56, 2026, pp. 53–77. doi: 10.1007/s40005-025-00780-4

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflicts of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International , CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 https://creativecommons.org/licenses/by/4.0/deed.en_US