



Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Human-Relevant Models Need Pharmacological Qualification Before They Can Replace Conventional Translation Pathways

Andrei Popescu^{1*}, Mihai Ionescu¹, Elena Stan², Cristina Radu³

¹Department of Human-Relevant Models and Pharmacological Qualification, Faculty of Pharmacy, University of Bucharest, Bucharest, Romania.

²Department of Translation Pathways and Model Validation, Faculty of Pharmacy, University of Agricultural Sciences Cluj-Napoca, Cluj-Napoca, Romania.

³Department of Preclinical-to-Clinical Translation, Faculty of Pharmacy, Polytechnic University of Bucharest, Bucharest, Romania.

***Email:** andrei.popescu@usamv.ro

ABSTRACT

Human-relevant organoids, tissue models, and microphysiological systems increasingly reproduce biological features that conventional cultures or animal models represent incompletely. Yet biological resemblance is not equivalent to pharmacological qualification. A model can preserve human genetics, architecture, mechanical cues, or disease phenotypes while failing to reproduce clinically relevant exposure, target engagement, response kinetics, toxicity, interindividual variability, or interlaboratory performance. This Translational Perspective argues that replacement decisions should therefore be based on context-specific pharmacological qualification rather than on human origin or physiological complexity alone. Evidence from patient-derived organoids, organ-on-chip systems, quantitative pharmacokinetic models, mechanistic toxicity platforms, and multicentre reproducibility studies shows that strong predictive performance is achievable, but generally within defined diseases, compounds, endpoints, exposure conditions, and experimental architectures. We distinguish biological fidelity, exposure fidelity, mechanistic-readout validity, reproducibility, and human-pharmacology benchmarking as analytically separate requirements. This evidence supports a qualification approach in which the permissible translational claim is constrained by the least-resolved requirement that is necessary for the intended decision. Human-relevant models may consequently replace conventional pathways for narrowly qualified questions while remaining complementary for broader or insufficiently benchmarked decisions.

Keywords: Human-relevant models, Microphysiological systems, Organoids, Organ-on-chip, Translational pharmacology, Model qualification

Received: 25 February 2026; **Revised:** 26 May 2026; **Accepted:** 28 May 2026

INTRODUCTION

New approach methodologies are moving from experimental alternatives toward practical components of pharmaceutical development. Human cellular systems, organoids, engineered tissues, and microphysiological systems can interrogate safety liabilities or biological processes that conventional animal studies represent imperfectly, particularly when human-specific pathways or species differences are central to the question. Their strongest current applications nevertheless remain use-case dependent rather than evidence for technology-wide substitution of established nonclinical pathways [1].

The translational problem therefore changes when human-relevant systems enter a development program. The central question is no longer merely whether a model looks or behaves more like human tissue than an animal or

conventional culture, but whether its outputs can be interpreted quantitatively in relation to human exposure, mechanism, variability, and clinical outcomes. Clinical pharmacology is particularly important because physiologically based pharmacokinetic and quantitative systems pharmacology approaches can place model-derived signals within exposure and response contexts that an isolated in vitro readout cannot provide [2].

Human disease models offer a persuasive response to interspecies limitations and can reproduce aspects of disease architecture with high clinical mimicry. However, model classes differ markedly in cellular composition, maturation, mechanical environment, vascularization, throughput, analytical accessibility, and suitability for particular development questions [3]. The argument developed here is therefore deliberately narrower than a general case for human-based experimentation: human relevance should be treated as an enabling property, whereas replacement requires pharmacological qualification for a declared context of use.

Human relevance is not equivalent to predictive validity

Human pluripotent-stem-cell-derived organoids can recreate selected developmental and disease phenotypes and can support drug screening and safety evaluation, but their translational value depends on properties such as maturation, cellular composition, reproducible production, assay compatibility, and the biological question being tested [4]. A model can therefore demonstrate convincing histology or transcriptomic similarity while remaining unqualified to predict a concentration-response relationship, treatment failure, clinical toxicity, or patient-level benefit. Biological fidelity and predictive validity overlap, but they are not interchangeable constructs.

This distinction becomes particularly important as organoids and organs-on-chips enter a regulatory environment increasingly receptive to alternative approaches. Technical sophistication and human derivation do not remove questions about reliability, applicability, manufacturing consistency, or the relation between a model output and a development decision [5]. The relevant unit of validation is consequently not “the organoid” or “the organ chip” as a technology class. It is a defined configuration used for a defined pharmacological question.

Patient-derived tumour organoids illustrate why this distinction should be preserved even when results are encouraging. In biliary tract cancer, 61 organoid models were generated from 82 tumours, and organoid drug-screening results were reported as concordant with actual clinical responses in 12 of 13 evaluable matched cases [6]. That result is important because it links an ex vivo pharmacological phenotype to a human treatment outcome. It does not establish the same predictive performance for other malignancies, treatment combinations, sampling sites, or laboratory workflows.

A metastatic colorectal cancer study similarly demonstrated that patient-derived tumour organoids can provide clinically informative predictive testing, while also showing that performance depends on biopsy location, therapeutic context, assay definitions, and the clinical population in which prediction is attempted [7]. These observations support a general qualification principle: predictive validity is not an intrinsic property permanently attached to a model platform. It is a relationship among a model configuration, an intervention, an exposure condition, an endpoint, a patient context, and an external reference.

Organoids and tissue models

The strongest justification for human-derived tissue models arises when conventional systems omit a biological interaction that materially determines pharmacological response. Donor-matched intestinal organoids and tumouroids, for example, have been used with immune components to investigate on-target, off-tumour toxicity from T-cell-engaging bispecific antibodies. The system reproduced clinically relevant epithelial toxicity mechanisms and interindividual differences that depend on the interaction among antigen expression, healthy tissue, tumour tissue, and immune activation [8]. Such evidence supports mechanistic human relevance, but it remains specific to the represented tissue, target, and treatment architecture.

Patient-derived tumour organoids also demonstrate how tissue fidelity and treatment-response fidelity can be evaluated separately. Breast-cancer organoids retained histological and genomic features of their parental tumours, and drug-susceptibility testing was reported to agree with matched clinical response in most evaluated cases [9]. Gastric-cancer organoids have likewise been used for personalized drug screening linked to patient treatment information [10]. These studies strengthen the case for patient-derived models as pharmacological test systems, while leaving disease-specific sampling, culture success, assay turnaround, and treatment representation as boundaries on interpretation.

Earlier gastrointestinal tumour studies provide important proof that clinical response concordance is experimentally achievable. Metastatic gastrointestinal cancer organoids reproduced selected treatment responses observed in patients enrolled in clinical studies [11]. Rectal-cancer tumouroids preserved molecular characteristics

of their source tumours, and ex vivo responses to chemotherapy and radiation correlated with the responses of individual patients; the study established 65 patient-derived cultures spanning primary, metastatic, and recurrent disease [12]. These data support treatment-specific benchmarking rather than an unrestricted claim that an organoid is a general surrogate for the patient.

Pancreatic cancer further exposes the importance of preserving heterogeneity. Patient-derived organoids generated distinct pharmacological response profiles rather than a single stable phenotype shared across patients [13]. For translational qualification, such variation should not automatically be treated as unwanted experimental noise. When technically reproducible, it can represent the biological heterogeneity that a human-relevant model is intended to preserve. The qualification problem is therefore to distinguish biologically meaningful differences from variability introduced by tissue acquisition, culture conditions, analytical methods, or model drift.

Organ-on-chip systems

Organ-on-chip and broader microphysiological systems extend human tissue models by adding controlled interfaces, perfusion, mechanical forces, multicellular organization, and in some systems inter-organ communication. Their value spans mechanistic biology, pharmacology, toxicology, and disease modelling, but the class remains technically heterogeneous [14]. A label such as “organ-on-chip” consequently provides little information about whether a particular implementation is appropriate for absorption, immune toxicity, barrier function, systemic pharmacokinetics, or another development decision.

Human organ chips can reproduce selected physiological and therapeutic responses under experimentally controlled conditions [15]. Intestinal chips illustrate the importance of specifying which components are pharmacologically consequential. Whereas conventional intestinal organoids may lack sustained fluid flow, peristalsis-like mechanics, stromal or endothelial interfaces, immune interactions, and long-term living microbiota, microfluidic intestinal systems can incorporate subsets of these features [16]. Their advantage is therefore not complexity itself, but the capacity to represent a missing process when that process governs the question under investigation.

Translation also depends on whether a platform can move beyond specialist laboratories. Materials, device architecture, manufacturing processes, fluid handling, throughput, interoperability, and end-user requirements can all affect dissemination of organ-chip technologies [17]. A pharmacologically convincing prototype may consequently remain unsuitable for routine development if its performance cannot be reproduced at the scale, frequency, and operational robustness required by its intended use.

The resulting cross-scale problem is summarized in **Table 1**. The table does not rank organ-chip technologies; it separates the evidence needed to move from a measured chip-level phenomenon toward a stronger development claim without treating transfer across scales as automatic.

Table 1 is needed because organ-chip measurements acquire different meanings as they cross from engineered microenvironment to tissue function, systemic pharmacology, and development decisions; the transfer boundary between those levels cannot be represented adequately by a single statement of physiological realism.

Table 1. Cross-scale qualification boundaries that determine what an organ-on-chip measurement can support in development

Manuscript-derived decision case	Organ-chip classification basis Relevant model	Immediate decision consequence Transfer boundary Next validation step Permissible development claim Evidence layer Consequence
Scientific anchor	Organ-chip classification basis Relevant model	Immediate decision consequence Transfer boundary Next validation step Permissible development claim Evidence layer Consequence

Barrier or interface response	Organ-chip classification basis Reproduced tissue interface under defined flow or mechanical conditions	Immediate decision consequence Establishes whether a drug or perturbation alters a local tissue function
	Relevant model or patient context Organ, disease state, cellular source	Transfer boundary that remains Local barrier behaviour does not establish systemic exposure or clinical response
		Next validation step Compare with human physiological or clinical barrier data under relevant exposure
		Permissible development claim Mechanistic or local functional relevance
		Evidence layer Tissue function Consequence for pharmacological qualification Qualify the represented interface, not the entire platform
Drug transport or metabolic handling	Organ-chip classification basis Dynamic perfusion with relevant absorptive or metabolic compartment	Immediate decision consequence Supports interrogation of absorption, clearance, or metabolite formation
	Relevant model or patient context Drug properties, route, organ configuration	Transfer boundary that remains Chip concentration-time behaviour requires scaling before human PK interpretation
		Next validation step Benchmark against reference compounds and human PK parameters
		Permissible development claim PK-relevant model performance for defined compounds
		Evidence layer Exposure and PK Consequence for pharmacological qualification Require quantitative exposure translation
Mechanically or immunologically conditioned response	Organ-chip classification basis Controlled flow, deformation, immune interaction, microbiome, or related context	Immediate decision consequence Tests whether missing physiological context changes pharmacological behaviour
	Relevant model or patient context Disease and patient-specific biology	Transfer boundary that remains Added biological complexity does not itself prove improved prediction
		Next validation step Show that the added process changes a human-relevant endpoint reproducibly
		Permissible development claim Context-specific mechanistic relevance
		Evidence layer Mechanism Consequence for pharmacological qualification Include only pharmacologically necessary complexity
Inter-organ exposure-response relationship	Organ-chip classification basis Multiple mature tissues linked through controlled vascular communication	Immediate decision consequence Allows distribution, metabolism, response, and toxicity to be studied together
	Relevant model or patient context Compound and organ network	Transfer boundary that remains Represented organs remain an incomplete approximation of whole-body physiology
		Next validation step Compare system PK/PD with human concentration-time and response data
		Permissible development claim Multi-organ PK/PD relevance within represented physiology
		Evidence layer System pharmacology Consequence for pharmacological qualification Define explicitly which systemic processes are captured
Transfer into routine development	Organ-chip classification basis Reproducible device, workflow, readout, and operational performance	Immediate decision consequence Determines whether evidence can be generated consistently across programs or sites
	Relevant model or patient context Laboratory, operator, platform version	Transfer boundary that remains Successful performance in one laboratory does not establish implementation robustness
		Next validation step Cross-site testing, quality control, and context-specific performance benchmarks
		Permissible development claim Operationally qualified use for a specified decision
		Evidence layer Implementation Consequence for pharmacological qualification Treat deployability as part of qualification, not an afterthought

Exposure and dosing fidelity

Human-relevant biology is pharmacologically uninformative if the tissue experiences an exposure that has no defensible relationship to the human concentration-time profile. For organ-chip applications intended to inform pharmacokinetics, relevant requirements include physiological parameters, PK-specific readouts, reference compounds, and explicit scaling approaches [18]. Nominal medium concentration should therefore not be treated as a substitute for systemic exposure, free drug concentration, tissue distribution, metabolite exposure, or duration above a pharmacologically active concentration.

Quantitative translation is nevertheless feasible in defined systems. Fluidically coupled vascularized organ chips combined with computational scaling have been used to predict human pharmacokinetic behaviour for nicotine and cisplatin [19]. The importance of this study lies less in declaring a general predictive capability than in demonstrating the architecture of a defensible translation: experimentally generated organ-level information was connected to a quantitative model and assessed against human pharmacokinetic behaviour.

Inter-organ communication can also change pharmacodynamic interpretation. A vascularly linked system containing mature heart, liver, bone, and skin tissues maintained tissue phenotypes for four weeks and reproduced aspects of doxorubicin pharmacokinetics, pharmacodynamics, and cardiotoxicity-associated biomarker behaviour [20]. The comparison with isolated or differently linked tissues showed that apparently local pharmacological readouts may depend on upstream distribution, metabolism, or tissue cross-talk. Exposure qualification must therefore ask not only whether the administered concentration is plausible, but whether the model contains the processes needed to generate the pharmacologically relevant exposure at the site of response.

Mechanistic readouts

Human-relevant models frequently generate rich morphological, transcriptomic, proteomic, electrophysiological, metabolic, and imaging outputs. Their availability can create an interpretive hazard: a responsive biomarker may be described as mechanistic even when it is downstream, nonspecific, exposure-dependent, or disconnected from the human decision endpoint. A useful distinction in microphysiological systems is between designed features, which are imposed by model architecture, and emergent features, which arise from cellular and tissue function. Quantifying both against corresponding human physiology provides a more rigorous route to judging fit for purpose [21].

Mechanistic readouts are especially valuable when they resolve a translational discrepancy. Human and animal Liver-Chip systems have reproduced selected species-specific drug toxicities, demonstrating that a human-derived system can help identify when toxicity depends on species biology rather than merely detecting a generic injury signal [22]. Such findings support a mechanism-level interpretation within hepatic toxicity; they do not demonstrate that the same platform predicts every toxicological outcome or that a mechanistic readout can replace exposure and clinical benchmarking.

The distinction between plausible mechanism and qualified decision utility becomes clearest when performance is tested against an external benchmark. In a blinded evaluation involving 870 Liver-Chips and 27 known hepatotoxic or non-toxic small molecules, the two-donor, protein-binding-corrected analysis achieved 87% sensitivity and 100% specificity for the defined drug-induced liver injury context [23]. The significance of this result is its boundedness: drug set, endpoint, analysis procedure, donor configuration, and benchmark were explicit. It therefore illustrates what qualification looks like when a mechanistic platform is converted into a testable development claim.

Figure 1 makes this distinction visually explicit by separating the generation of a biologically credible readout from the successive boundaries that determine whether that readout can support mechanistic, predictive, or development-level inference.

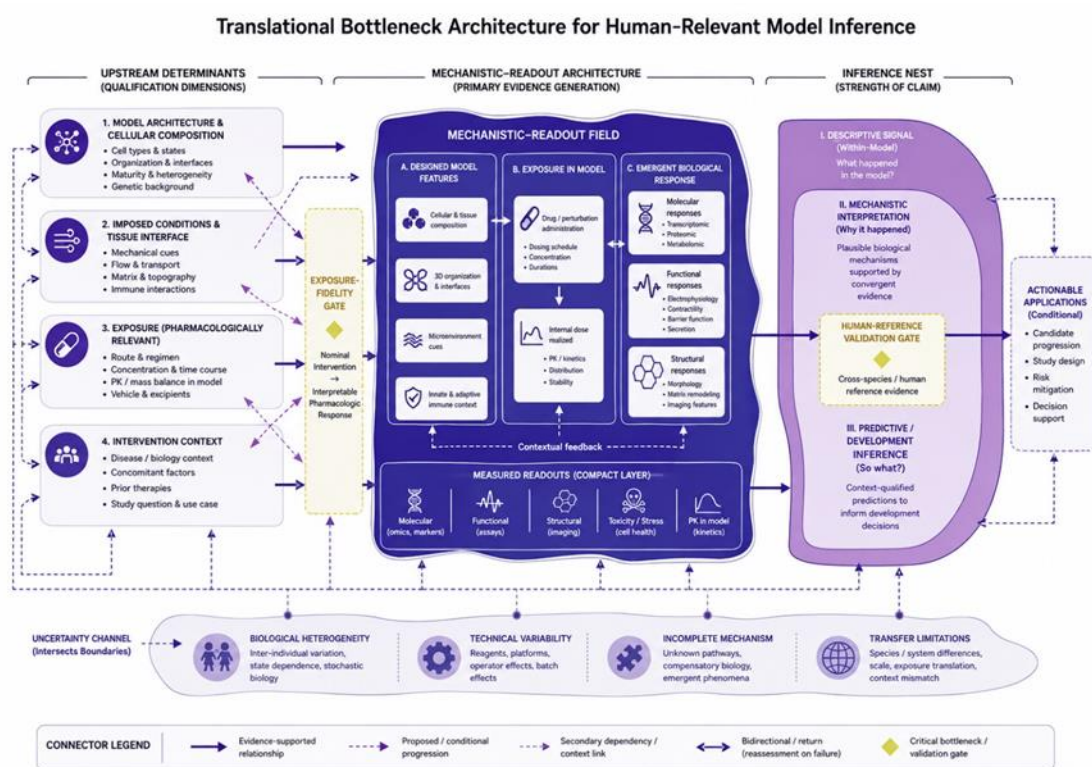


Figure 1. Mechanistic signals become translational evidence only after exposure, interpretation, and human-reference boundaries are crossed

Reproducibility and interlaboratory transfer

A model cannot support a replacement claim if its pharmacological interpretation depends on uncharacterized laboratory-specific conditions. Reproducibility in complex microphysiological systems cannot be reduced to demanding that every measurement vary minimally. Fit-for-purpose quality management instead requires defining which attributes must remain controlled for the intended use, which sources of variability can be monitored, and which differences represent legitimate biological behavior [24]. The acceptable reproducibility target is therefore linked to the decision endpoint, not to a universal coefficient of variation.

Brain organoids illustrate the distinction between biological richness and analytical reliability. Their capacity to preserve human genetics and reproduce aspects of neurodevelopment has generated substantial interest, yet variation in differentiation, maturation, characterization, and analytical practice can constrain interpretation. Reliable analytical platforms and more uniform reporting are consequently prerequisites for comparing findings across laboratories and studies [25]. A human-derived phenotype that cannot be measured consistently remains difficult to qualify pharmacologically.

Importantly, reproducibility is itself multidimensional. In a three-site study using a harmonized cortical-organoid protocol and a single induced pluripotent stem-cell line, cortical progenitor and neuronal cell-type proportions and cortical wall-like organization were reproducible, whereas organoid size and expression of metabolic and cellular-stress genes differed among sites [26]. These results argue against a binary label of “reproducible” or “not reproducible.” A platform can be transferable for one endpoint while remaining insufficiently controlled for another endpoint that is more relevant to a particular drug-development question.

This issue becomes even more consequential when human-relevant systems are used to interpret interpatient variability. Reproducibility analytics incorporating detailed experimental metadata, coefficients of variation, analysis of variance, and intraclass correlation have been proposed to distinguish biological or clinical heterogeneity from experimental variability in microphysiological systems [27]. Without such partitioning, a donor-specific difference can be mistaken for precision pharmacology when it is actually caused by cell handling, medium composition, flow, device behaviour, assay execution, or another technical factor.

Figure 2 is required because transferability depends on the relationship among model reproducibility, measurement reproducibility, biological heterogeneity, and decision utility rather than on a single global measure of assay consistency.

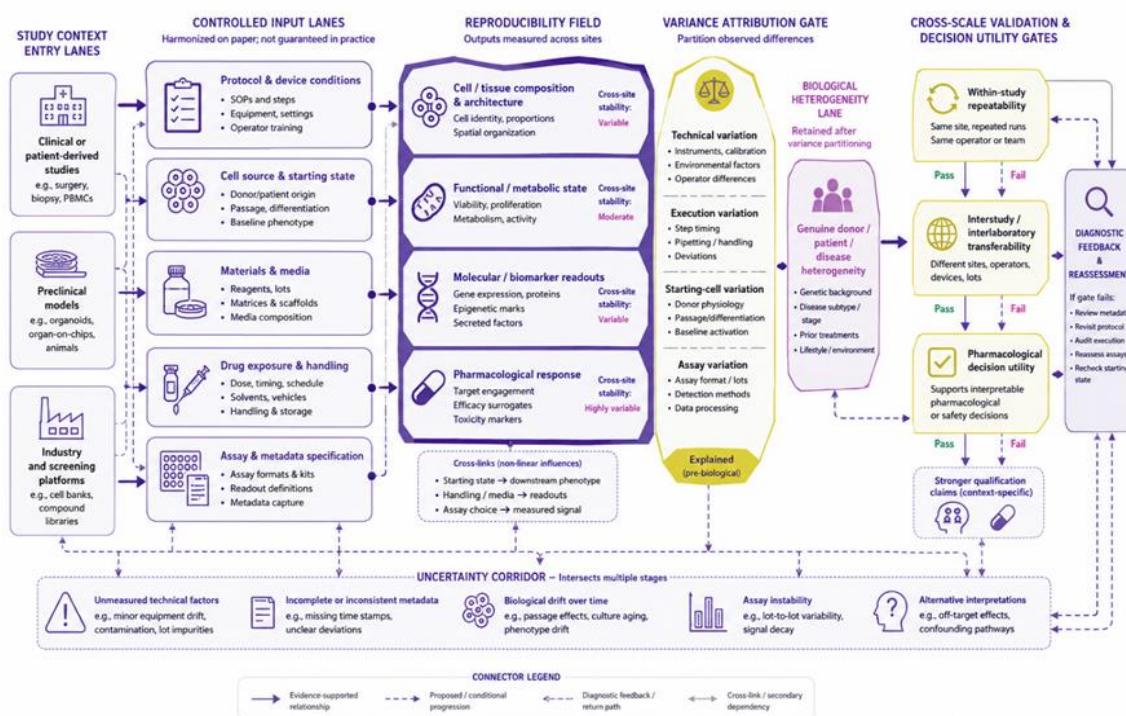


Figure 2. Interlaboratory transfer requires separation of stable biology, technical variance, and decision-relevant heterogeneity

Context-specific qualification

Qualification should attach to a defined pharmacological use rather than to a technology name. The problem is visible in patient-derived organoid drug-sensitivity testing, where expert consensus has emphasized standardized terminology, model characterization, assay procedures, reporting, and clinically interpretable use conditions while recognizing continuing limitations in prediction [28]. A “validated organoid,” therefore, is an incomplete description unless the disease, specimen source, intervention, exposure, endpoint, analytical procedure, comparator, and decision being supported are also specified.

This principle becomes more important as human-relevant models acquire multiple layers of sophistication. A system may have convincing histology but inadequate exposure fidelity; another may reproduce exposure while generating a biomarker whose connection to mechanism is uncertain. Pharmacokinetic translation requires explicit scaling and reference compounds [18], whereas mechanistic interpretation requires quantitative distinction between imposed model features and emergent biological behavior [21]. Reproducibility must likewise be defined for the relevant endpoint rather than assumed from general platform stability [24]. Where patient heterogeneity is itself the intended signal, technical variance must first be separated from genuine biological variance [27].

Qualification should also be treated as state-dependent rather than permanently conferred. In metastatic colorectal cancer organoid testing, predictive performance depended partly on treatment and sampling context [7]. Pancreatic organoid studies show substantial pharmacological heterogeneity among patients [13], while interlaboratory cortical-organoid work demonstrates that some model properties can remain consistent even as metabolic or stress-associated outputs change across sites [26]. A model qualified for one decision under one state may consequently require reassessment after a change in treatment line, specimen source, protocol, platform version, exposure regime, or analytical endpoint.

Table 2 is needed because qualification is not a one-time label: its meaning can change as the biological state, exposure state, experimental state, or development question changes over time.

Table 2. Time- and state-dependent conditions that can preserve, narrow, or invalidate a pharmacological qualification claim

Governing condition	Immediate decision consequence Permissible development claim	Transfer boundary	Required next validation Evidence layer Unresolved qualification question
---------------------	---	-------------------	---

Qualification state or window			
Initial model establishment			
Biological identity, source material, protocol, and intended use are defined	Immediate decision consequence		Required next validation
	Establishes what the model is intended to represent		Characterize phenotype, assay behavior, and relevant human reference
	Permissible development claim		Evidence layer
	Descriptive human relevance		Biological/model fidelity
			Unresolved qualification question
Does the represented biology contain the processes needed for the intended drug question?			
Qualification state or window			
Exposure-defined use			
Drug concentration-time conditions and relevant disposition processes are specified	Immediate decision consequence		Required next validation
	Enables pharmacological interpretation of measured response		Benchmark exposure or disposition against human PK or justified scaling
	Permissible development claim		Evidence layer
	Exposure-relevant mechanistic interpretation		PK/exposure
			Unresolved qualification question
Is the model experiencing a clinically meaningful exposure state?			
Qualification state or window			
Mechanistically interpretable use			
Measured response is linked to a defined biological process	Immediate decision consequence		Required next validation
	Supports stronger inference than phenotypic association alone		Demonstrate process relevance and appropriate human physiological reference
	Permissible development claim		Evidence layer
	Mechanistic claim within the tested context		Mechanism/readout
			Unresolved qualification question
Does the readout identify the relevant process rather than merely correlate with perturbation?			
Qualification state or window			
Reproducible operational state			
Critical outputs remain sufficiently stable across runs or sites	Immediate decision consequence		Required next validation
	Allows results to be compared and transferred		Endpoint-specific cross-run or cross-site testing with metadata and variance analysis
	Permissible development claim		Evidence layer
	Transferable assay performance for defined endpoints		Reproducibility
			Unresolved qualification question
Which outputs remain stable enough for the intended decision?			
Qualification state or window			
Clinically benchmarked state			
Model output has been compared with human exposure, response, or toxicity	Immediate decision consequence		Required next validation
	Supports a bounded predictive or decision-support claim		Replicate in relevant external populations or contexts
	Permissible development claim		Evidence layer
	Context-specific predictive utility		Human pharmacology
			Unresolved qualification question
How far can the benchmarked relationship be transported without recalibration?			
Qualification state or window			
Changed biological or operational state			
Patient population, therapy, specimen source, protocol, platform, or analytical method changes	Immediate decision consequence		Required next validation
	Previous qualification may narrow or lapse		Requalification of the affected dimensions
	Permissible development claim		Evidence layer
	Only the portion of the prior claim unaffected by the change		Longitudinal/contextual
			Unresolved qualification question
Which elements of the original qualification remain applicable?			

Benchmarking against human pharmacology

Replacement claims require an external reference that is stronger than visual or molecular resemblance. Clinical pharmacology provides that reference by asking whether model-derived exposures, mechanisms, responses, or toxicities behave in ways that correspond to human observations [2]. For pharmacokinetic applications, this means comparing model outputs with human concentration-time behavior rather than assuming that a nominal in vitro dose represents clinically relevant exposure [18].

Patient-derived tumour organoids provide one form of human benchmark because their ex vivo treatment responses can be compared with matched clinical outcomes. Biliary tract cancer organoids showed high concordance in the small matched clinical subset evaluated [6], while metastatic colorectal cancer organoids demonstrated that predictive performance can be quantified prospectively but remains sensitive to treatment and sampling context [7]. Breast, gastric, gastrointestinal, and rectal-cancer organoid studies similarly show that concordance is achievable when the disease, treatment, specimen, and outcome are clearly defined [9-12]. These studies should not be pooled informally into a universal accuracy estimate; their value lies in demonstrating multiple ways to construct disease-specific external validation.

Quantitative pharmacology offers parallel benchmarks. Linked organ chips have been scaled against human pharmacokinetic behavior for specific compounds [19], and multi-organ systems can relate distribution and tissue response under controlled vascular coupling [20]. In toxicology, blinded evaluation of a Liver-Chip against a defined drug-induced liver injury benchmark demonstrates how a human-relevant platform can acquire an explicit performance claim when the compounds, endpoint, analysis method, and comparator are predetermined [23]. The general lesson is not that every model requires the same validation metric, but that every strong translational claim requires a relevant human reference.

A pharmacological qualification framework

On the basis of the evidence reviewed above, we propose that qualification begin with the development decision, not with the model technology. The intended question should specify the compound or modality, biological context, human population or tissue state, relevant exposure, endpoint, and consequence of error. Human-relevant methodologies are already useful for defined safety questions [1], but the evidentiary requirement changes when their output is intended to alter a clinical-development decision [2].

The proposed framework then evaluates five analytically separate domains: biological fidelity, exposure fidelity, mechanistic-readout validity, reproducibility and transferability, and human-pharmacology benchmarking. Exposure should remain separate because PK translation has distinct scaling requirements [18]. Mechanistic evidence should remain separate because quantitative tissue features and biomarkers do not automatically establish causal relevance [21]. Reproducibility should remain separate because fit-for-purpose quality management and variance attribution determine whether an apparent response can be transported between experiments, sites, or patients [24].

These domains should not be collapsed into a single numerical score. A model with exceptional anatomical fidelity but poorly characterized exposure cannot support an exposure-response claim. Likewise, a model with reproducible morphology cannot support precision pharmacology if between-donor differences are confounded by technical variability [27]. The framework therefore uses a qualitative weakest-relevant-link rule: the maximum permissible claim is constrained by the least-resolved qualification domain that is necessary for that claim.

The strength of the resulting statement should also be graduated. Descriptive fidelity can justify saying that a model reproduces a human feature. Mechanistic qualification can support a claim that a defined process contributes to a response. Predictive qualification additionally requires performance against an external human benchmark, as shown in bounded examples of patient-derived organoid testing and Liver-Chip toxicology [23]. Decision-support or replacement claims require still stronger evidence that performance remains sufficiently reproducible and applicable in the exact development context for which conventional evidence would otherwise be used.

The framework is a proposed synthesis, not an empirically validated scoring system or regulatory standard. Its purpose is to prevent strength in one evidence domain from concealing weakness in another. Cross-site cortical-organoid studies are instructive: stable cellular composition can coexist with variability in size, metabolism, and stress-associated outputs [26]. Qualification must therefore be endpoint-specific and revisable when the model or its intended use changes.

Replacement versus complementary use

Human-relevant models need not satisfy the evidentiary threshold for replacement before they become useful. NAMs can already contribute safety information in specific development settings [1], and organoid and organ-chip platforms can resolve mechanisms, human-specific toxicities, treatment-response heterogeneity, and exposure questions that conventional systems answer poorly. Complementary use is therefore not a lesser scientific status; it can be the most defensible strategy when independent evidence streams address different uncertainties.

Replacement is a stronger claim because it implies that omitting a conventional pathway does not remove information necessary for the intended decision. The clearest candidates are narrow contexts in which the human-relevant model has a specified purpose, reproducible performance, pharmacologically relevant exposure, and an external human benchmark. The blinded Liver-Chip assessment provides an example of such bounded qualification for a defined hepatotoxicity application [23], while fit-for-purpose quality management establishes why reproducibility must remain part of any transfer claim [24].

Where these requirements remain incomplete, triangulation is preferable to premature substitution. A model may reveal clinically relevant biology while lacking human PK translation [18], or it may preserve patient heterogeneity while leaving technical variance insufficiently characterized [27]. In such cases the model can still change compound prioritization, mechanism assessment, biomarker selection, safety interpretation, or clinical-study design without being treated as a complete replacement for existing evidence.

The long-term shift toward human-relevant development should therefore not be framed as a competition between “new” and “conventional” models. The more useful distinction is between evidence streams that are qualified for a particular decision and those that are not. Replacement becomes scientifically defensible when the required pharmacological information is demonstrated to survive the transition from model biology to exposure, mechanism, reproducibility, and human outcome. Until then, complementary use preserves information while the missing qualification evidence is generated.

CONCLUSION

Human relevance is an important property of a translational model, but it is not itself proof of predictive validity. Organoids, tissue models, and organ-on-chip systems can reproduce clinically meaningful responses, human-specific mechanisms, quantitative pharmacokinetics, toxicity, and interpatient variation. They can also exhibit disease specificity, exposure dependence, technical variability, endpoint-specific reproducibility, and incomplete transfer across laboratories or clinical contexts.

The proposed pharmacological qualification approach separates biological fidelity, exposure fidelity, mechanistic-readout validity, reproducibility, and human benchmarking so that each development claim is limited by the evidence it actually requires. This avoids both premature replacement and unnecessarily conservative rejection of useful human-relevant systems. The practical objective is not to prove that one model class is universally superior, but to determine which model is sufficiently qualified for which decision, under which conditions, and with what remaining uncertainty.

Human-relevant models should therefore replace conventional translation pathways selectively rather than categorically. Where the relevant exposure, mechanism, transferability, and human benchmark are established, replacement may be scientifically justified. Where one or more critical domains remain unresolved, complementary use is more defensible. Pharmacological qualification provides the boundary between those two states.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Beilmann M, Adkins K, Boonen HCM, Hewitt P, Hu W, Mader R, et al. Application of new approach methodologies for nonclinical safety assessment of drug candidates. *Nat Rev Drug Discov.* 2025;24(9):705-25. doi:10.1038/s41573-025-01182-9
2. Cao Y, Polacheck W. New approach methodologies: What clinical pharmacologists should prepare for. *Clin Pharmacol Ther.* 2025;118(6):1269-72. doi:10.1002/cpt.70046
3. Loewa A, Feng JJ, Hedtrich S. Human disease models in drug development. *Nat Rev Bioeng.* 2023;1(8):545-59. doi:10.1038/s44222-023-00063-3
4. Vandana JJ, Manrique C, Lacko LA, Chen S. Human pluripotent-stem-cell-derived organoids for drug discovery and evaluation. *Cell Stem Cell.* 2023;30(5):571-91. doi:10.1016/j.stem.2023.04.011
5. Zhou L, Huang J, Li C, Gu Q, Li G, Li ZA, et al. Organoids and organs-on-chips: Recent advances, applications in drug development, and regulatory challenges. *Med.* 2025;6(4):100667. doi:10.1016/j.medj.2025.100667
6. Ren X, Huang M, Weng W, Xie Y, Wu Y, Zhu S, et al. Personalized drug screening in patient-derived organoids of biliary tract cancer and its clinical application. *Cell Rep Med.* 2023;4(11):101277. doi:10.1016/j.xcrm.2023.101277
7. Tan T, Mouradov D, Lee M, Gard G, Hirokawa Y, Li S, et al. Unified framework for patient-derived, tumor-organoid-based predictive testing of standard-of-care therapies in metastatic colorectal cancer. *Cell Rep Med.* 2023;4(12):101335. doi:10.1016/j.xcrm.2023.101335
8. Harter MF, Recalain T, Gerard R, Avignon B, Bollen Y, Esposito C, et al. Analysis of off-tumour toxicities of T-cell-engaging bispecific antibodies via donor-matched intestinal organoids and tumouroids. *Nat Biomed Eng.* 2024;8(4):345-60. doi:10.1038/s41551-023-01156-5
9. Wu H, Wang W, Zhang Y, Chen Y, Shan C, Li J, et al. Establishment of patient-derived organoids for guiding personalized therapies in breast cancer patients. *Int J Cancer.* 2024;155(2):324-38. doi:10.1002/ijc.34931
10. Zhao Y, Li S, Zhu L, Huang M, Xie Y, Song X, et al. Personalized drug screening using patient-derived organoid and its clinical relevance in gastric cancer. *Cell Rep Med.* 2024;5(7):101627. doi:10.1016/j.xcrm.2024.101627
11. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernández-Mateos J, Khan K, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science.* 2018;359(6378):920-6. doi:10.1126/science.aao2774
12. Ganesh K, Wu C, O'Rourke KP, Szeglin BC, Zheng Y, Sauv  CEG, et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med.* 2019;25(10):1607-14. doi:10.1038/s41591-019-0584-2
13. Tiri  H, Belleau P, Engle DD, Plenker D, Desch nes A, Somerville TDD, et al. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov.* 2018;8(9):1112-29. doi:10.1158/2159-8290.CD-18-0349
14. Low LA, Mummery C, Berridge BR, Austin CP, Tagle DA. Organs-on-chips: Into the next decade. *Nat Rev Drug Discov.* 2021;20(5):345-61. doi:10.1038/s41573-020-0079-3
15. Ingber DE. Human organs-on-chips for disease modelling, drug development and personalized medicine. *Nat Rev Genet.* 2022;23(8):467-91. doi:10.1038/s41576-022-00466-9
16.  zkan A, LoGrande NT, Feitor JF, Goyal G, Ingber DE. Intestinal organ chips for disease modelling and personalized medicine. *Nat Rev Gastroenterol Hepatol.* 2024;21(11):751-73. doi:10.1038/s41575-024-00968-3
17. Alver CG, Drabbe E, Ishahak M, Agarwal A. Roadblocks confronting widespread dissemination and deployment of organs on chips. *Nat Commun.* 2024;15(1):5118. doi:10.1038/s41467-024-48864-3
18. Keuper-Navis M, Walles M, Poller B, Myszczy syn A, van der Made TK, Donkers J, et al. The application of organ-on-chip models for the prediction of human pharmacokinetic profiles during drug development. *Pharmacol Res.* 2023;195:106853. doi:10.1016/j.phrs.2023.106853
19. Herland A, Maoz BM, Das D, Somayaji MR, Prantil-Baun R, Novak R, et al. Quantitative prediction of human pharmacokinetic responses to drugs via fluidically coupled vascularized organ chips. *Nat Biomed Eng.* 2020;4(4):421-36. doi:10.1038/s41551-019-0498-9

20. Ronaldson-Bouchard K, Teles D, Yeager K, Tavakol DN, Zhao Y, Chramiec A, et al. A multi-organ chip with matured tissue niches linked by vascular flow. *Nat Biomed Eng.* 2022;6(4):351-71. doi:10.1038/s41551-022-00882-6
21. Nahon DM, Moerkens R, Aydogmus H, Lendemeijer B, Martínez-Silgado A, Stein JM, et al. Standardizing designed and emergent quantitative features in microphysiological systems. *Nat Biomed Eng.* 2024;8(8):941-62. doi:10.1038/s41551-024-01236-0
22. Jang KJ, Otieno MA, Ronxhi J, Lim HK, Ewart L, Kodella KR, et al. Reproducing human and cross-species drug toxicities using a Liver-Chip. *Sci Transl Med.* 2019;11(517). doi:10.1126/scitranslmed.aax5516
23. Ewart L, Apostolou A, Briggs SA, Carman CV, Chaff JT, Heng AR, et al. Performance assessment and economic analysis of a human Liver-Chip for predictive toxicology. *Commun Med (Lond).* 2022;2(1):154. doi:10.1038/s43856-022-00209-1
24. Pamies D, Ekert J, Zurich MG, Frey O, Werner S, Piergiovanni M, et al. Recommendations on fit-for-purpose criteria to establish quality management for microphysiological systems and for monitoring their reproducibility. *Stem Cell Reports.* 2024;19(5):604-17. doi:10.1016/j.stemcr.2024.03.009
25. Sandoval SO, Cappuccio G, Kruth K, Osenberg S, Khalil SM, Méndez-Albelo NM, et al. Rigor and reproducibility in human brain organoid research: Where we are and where we need to go. *Stem Cell Reports.* 2024;19(6):796-816. doi:10.1016/j.stemcr.2024.04.008
26. Glass MR, Waxman EA, Yamashita S, Lafferty M, Beltran AA, Farah T, et al. Cross-site reproducibility of human cortical organoids reveals consistent cell type composition and architecture. *Stem Cell Reports.* 2024;19(9):1351-67. doi:10.1016/j.stemcr.2024.07.008
27. Miedel MT, Varmazyad M, Xia M, Brooks MM, Gavlock DC, Reese C, et al. Validation of microphysiological systems for interpreting patient heterogeneity requires robust reproducibility analytics and experimental metadata. *Cell Rep Methods.* 2025;5(4):101028. doi:10.1016/j.crmeth.2025.101028
28. Xiang D, He A, Zhou R, Wang Y, Xiao X, Gong T, et al. Building consensus on the application of organoid-based drug sensitivity testing in cancer precision medicine and drug development. *Theranostics.* 2024;14(8):3300-16. doi:10.7150/thno.96027