Available online www.ijpras.com

International Journal of Pharmaceutical Research & Allied Sciences, 2025, 14(4):89-101
https://doi.org/10.51847/JDgGCLTR7N

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Crossing the Pharmaceutical Translation Gap: A Systematic Review of Evidence Maturity, Reproducibility, Manufacturing Readiness, and Implementation Barriers

Erik Van der Meer[1]*, Lieke Jansen[1], Bram de Vries[2], Sophie Claes[3]

[1]Department of Pharmaceutical Translation and Evidence Maturity, Faculty of Pharmacy, Wageningen University, Wageningen, Netherlands.
[2]Department of Reproducibility and Manufacturing Readiness, Faculty of Pharmacy, University of Groningen, Groningen, Netherlands.
[3]Department of Implementation Barriers and Systematic Review, Faculty of Pharmacy, University of Amsterdam, Amsterdam, Netherlands.

***Email:** e.vandermeer@wur.nl*

ABSTRACT

Pharmaceutical innovation frequently generates compelling mechanistic, computational, preclinical, or technological findings that do not progress into reproducible, manufacturable, clinically useful, and sustainably implemented products. This systematic review examined how evidence maturity, reproducibility, external validation, manufacturing feasibility, clinical relevance, and implementation conditions jointly shape pharmaceutical translation. A protocol-driven review design was used, with prespecified questions, eligibility criteria, concept-based searches across bibliographic and article-record sources, DOI and journal-status verification, duplicate and version control, structured extraction, design-specific quality appraisal, and qualitative framework synthesis. Meta-analysis was not planned because the reviewed literature encompassed heterogeneous technologies, experimental systems, comparators, outcomes, and stages of development. The evidence indicates that translation readiness cannot be inferred from technical novelty, internal predictive performance, mechanistic plausibility, or isolated proof-of-concept success. Greater maturity is associated with convergent biological evidence, transparent data provenance, replication, external evaluation, fit-for-purpose human relevance, integrated process development, scalable quality control, clinical utility, and compatibility with organizational and patient-care contexts. However, evidence remains uneven across domains. Computational and preclinical studies frequently lack transportable validation, manufacturing studies often evaluate selected platforms or products, and implementation reports may document adoption without establishing patient benefit or sustainability. Cross-domain comparison is further limited by inconsistent terminology, platform-specific criteria, publication bias, and the absence of universally accepted progression thresholds. The review therefore supports a multidimensional and decision-specific interpretation of translation readiness. Progression should depend on the weakest consequential evidence domain rather than the strongest isolated result, while remaining explicitly conditional on the product, intended use, population, manufacturing environment, and implementation setting.

Key words: *Translational pharmaceutical research, Translation readiness, Reproducibility, Manufacturing feasibility, Clinical relevance, Evidence maturity*

Received: 29 August 2025; **Revised:** 29 November 2025; **Accepted:** 02 December 2025

INTRODUCTION

Pharmaceutical development remains characterized by substantial attrition, although its magnitude varies across therapeutic areas, development phases, and analytical assumptions. Estimates based on clinical-development

databases demonstrate that success rates are sensitive to indication, trial status, and the handling of missing or censored outcomes, while broader analyses attribute failure to interacting limitations in efficacy, safety, pharmacokinetics, development design, commercial strategy, and execution [1, 2]. Translation failure should therefore not be treated as a single late-stage event. It may emerge through a sequence of evidentiary weaknesses that become visible only when a candidate encounters a new biological system, manufacturing environment, clinical population, or implementation context.

Prevailing accounts frequently emphasize scientific novelty or local technical performance while underrepresenting the evidence needed to support progression. Multidimensional target assessment requires attention to biological rationale, human evidence, safety, tractability, tissue exposure, and strategic context rather than target association alone [3]. Similarly, machine-learning methods can contribute to target identification, molecular design, prediction, and clinical-development decisions, but retrospective accuracy or benchmark superiority cannot independently establish pharmaceutical usefulness [4]. Translation readiness is consequently different from technological sophistication: the former concerns whether the available evidence is sufficiently relevant, reproducible, transferable, and decision-ready for the next consequential stage.

Human-relevant experimental systems and precision-medicine programs illustrate the same distinction. Organ-on-chip technologies may improve physiological representation and permit use-specific investigation of pharmacology or toxicity, yet their translational value depends on reproducibility, benchmarking, endpoint validity, interoperability, and evidence that model outputs correspond to the intended human decision [5]. Pharmacogenomics likewise demonstrates that biological association and analytical testing do not automatically produce routine clinical benefit; implementation also depends on variant interpretation, clinical utility, decision support, professional capability, workflow integration, and equitable access [6]. These examples show why evidence maturity, manufacturing readiness, and implementation cannot be evaluated as independent end-stage concerns.

This systematic review evaluates the determinants that distinguish exploratory pharmaceutical promise from credible progression toward external validity, manufacturing feasibility, clinical relevance, and implementation. It uses PRISMA-compatible reporting principles to make the questions, eligibility criteria, information sources, selection logic, extraction framework, appraisal approach, and qualitative synthesis explicit [7]. The review does not propose a universal readiness score, validated predictive model, regulatory pathway, or clinical recommendation. Its contribution is an evidence-bounded cross-domain synthesis that identifies recurrent progression conditions, major uncertainty sources, and the boundaries beyond which technical performance should not be interpreted as pharmaceutical translation.

Review questions and protocol registration

The review addressed one primary question: among pharmaceutical discovery, preclinical, manufacturing, and implementation studies, which evidence characteristics distinguish exploratory technical promise from credible progression toward externally validated, manufacturable, clinically relevant, and implementable innovation? Secondary questions examined how evidence maturity was defined or demonstrated; which replication and external-validation practices strengthened progression claims; how manufacturing and scale-up readiness were evaluated; which clinical, organizational, economic, and patient-related barriers affected implementation; and which determinants repeatedly explained progression, variation, or failure.

The protocol specified the review questions, evidence domains, eligibility criteria, information sources, concept blocks, duplicate handling, selection logic, extraction fields, design-specific appraisal domains, and structured narrative synthesis before manuscript drafting. It was not prospectively registered, and no registration identifier is reported. Protocol specification was treated as a transparency mechanism rather than as evidence of prospective registry endorsement.

The planned synthesis distinguished descriptive evidence from validation evidence, computational performance from pharmaceutical utility, internal reproducibility from external transportability, laboratory feasibility from manufacturing readiness, adoption from sustained implementation, and implementation activity from patient benefit. Because the eligible literature included heterogeneous designs, systems, technologies, comparators, and outcomes, statistical pooling was neither prespecified nor considered methodologically appropriate. Findings were therefore integrated through qualitative framework synthesis, with conclusions classified narratively as supported, conditional, or emerging rather than assigned numerical certainty scores.

Eligibility criteria and information sources

Eligible sources were peer-reviewed journal articles with verifiable DOIs and confirmed first-quartile status in a relevant subject category. They were required to provide extractable evidence concerning at least one prespecified domain: translational pharmaceutical research, evidence maturity, reproducibility, external validation, patient relevance, manufacturing readiness, scale-up, autonomous experimentation, clinical translation, implementation, sustainability, or progression decision-making. Primary empirical and methodological studies were prioritized for specific claims, while reviews, scientific statements, and evidence-based frameworks were retained when they contributed field-level synthesis or explicit appraisal constructs.

Eligible systems included pharmaceutical discovery programs, molecular targets, predictive models, experimental disease systems, organ-chip platforms, manufacturing processes, therapeutic modalities, clinical implementation programs, and cross-domain progression frameworks. Relevant outcomes included replicability, generalization, human relevance, predictive validity, process integration, scalability, quality control, transferability, clinical utility, adoption, sustainability, patient access, and progression or attrition. A comparator was not mandatory when a study provided methodological or framework evidence, but comparative context was extracted whenever internal and external validation, alternative models, conventional and automated processes, or differing implementation settings were evaluated.

Conference-only papers, preprints without journal publication, books, theses, websites, policy webpages, software documentation, promotional reports, and articles lacking verifiable DOI or journal-quartile information were excluded. Sources were also excluded when they mentioned translation only tangentially, offered unsupported readiness claims, duplicated another publication or correction record, or could not be assigned to a specific review question, evidence category, manuscript claim, table entry, or figure relationship. Information was retrieved through PubMed/MEDLINE and Europe PMC, supplemented by PubMed Central records, publisher pages, DOI metadata, and journal-ranking records. Publisher and DOI records were used to resolve bibliographic discrepancies rather than as substitutes for scientific evidence.

Search strategy and study selection

The search strategy combined pharmaceutical terms with translation, validation, manufacturing, and implementation concepts. The principal logic was:

(pharmaceutical OR drug OR therapeutic) AND (translation OR “translation readiness” OR “clinical translation” OR “evidence maturity” OR “platform readiness” OR attrition) AND (validation OR reproducibility OR manufacturing OR implementation).

Additional searches combined pharmaceutical or preclinical terms with replication, external validation, generalization, robustness, biological heterogeneity, continuous manufacturing, scale-up, process integration, decentralized manufacturing, clinical utility, patient relevance, sustainability, autonomous experimentation, self-driving laboratories, and decision gates. Searches were adapted to the syntax of each information source without changing their conceptual scope.

Study selection followed sequential relevance and verification stages. Titles and abstracts were first assessed against the review domains and unit of inclusion. Potentially relevant articles underwent full-record or full-text examination, followed by verification of publication type, DOI, bibliographic metadata, journal status, direct claim fit, and extractable evidence. Duplicate titles and DOIs, correction notices, preprint–journal duplicates, and alternate versions of the same article were reconciled so that only the definitive eligible journal article was retained.

A second verification pass reassessed each inclusion decision against the prespecified criteria and required every retained source to support a defined manuscript claim, methodological judgment, table cell, figure element, or synthesis category. Articles were not retained solely because they described a promising technology or used the language of translation. This separated initial topic relevance from final evidentiary eligibility and reduced the risk that speculative, adjacent, or weakly matched sources would determine the review conclusions.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for eligibility, extraction, and quality-appraisal framework for crossing the pharmaceutical translation gap.

Table 1. Eligibility, extraction, and quality-appraisal framework for Crossing the Pharmaceutical Translation Gap.

Evidence domain or	Review question	Eligible evidence type	Key extraction fields	Appraisal or relevance criterion	Synthesis role	Main limitation	Interpretation boundary

review construct							
Translational relevance	Does the study connect a pharmaceutical claim to a consequential development or use decision?	Empirical studies, methodological analyses, reviews, or frameworks with direct pharmaceutical relevance	Intended use; development stage; system; decision context; claimed consequence	Direct relationship between reported evidence and the stated pharmaceutical decision	Establishes whether evidence addresses translation rather than technical novelty alone	Translation terminology may be used without downstream evaluation	Relevance does not establish readiness or successful progression
Target and mechanistic evidence	Does the evidence support biological rationale, tractability, exposure, safety, or patient selection?	Human genetics, target assessment, pharmacology, mechanistic studies	Target rationale; human evidence; safety; tractability; exposure; population	Multidimensional target assessment rather than association alone [3]	Supports early evidence-maturity classification	Associations may be confounded or incomplete	Mechanistic plausibility and association do not establish causation or clinical efficacy
Computational evidence	Does a model generalize beyond its development data and support an intended pharmaceutical use?	Machine-learning, cheminformatics, predictive modelling, decision-support studies	Data provenance; split strategy; comparator; leakage controls; external evaluation; applicability	Performance must be interpreted in relation to data structure, decision context, and external validity [4]	Distinguishes model output from pharmaceutical utility	Benchmark design may exaggerate performance	Internal performance does not establish prospective utility or clinical consequence
Preclinical and human-relevant models	Does the model reproduce relevant biology and predict an intended human outcome?	Animal, cellular, organoid, organ-chip, and translational modelling studies	Model system; endpoint; comparator; biological heterogeneity; replication; human linkage	Fit-for-purpose relevance, reproducibility, and endpoint validity [5]	Evaluates biological and external-validation maturity	Platforms and endpoints are highly heterogeneous	Human resemblance does not establish universal human prediction
Manufacturing and scale-up	Can the product and process be transferred, controlled, scaled, and supplied while maintaining relevant quality?	Process-development, continuous-manufacturing, modality-specific manufacturing, and scale-up studies	Product definition; unit operations; process control; scale; analytics; comparability; supply assumptions	Evidence must address integrated process feasibility rather than isolated laboratory production	Identifies manufacturing constraints that can halt otherwise promising candidates	Many reports concern selected products or platform demonstrations	Laboratory synthesis does not establish commercial or regulator-accepted readiness
Clinical relevance and utility	Does the evidence inform a clinical decision and plausibly improve care for a defined population?	Clinical-development analyses, translational studies, pharmacogenomic programs, utility evaluations	Population; intended decision; comparator; clinical utility; workflow; patient relevance	Biological or analytical validity must be distinguished from clinical usefulness [6]	Connects technical evidence to decision use	Clinical utility may be inferred rather than prospectively measured	Implementation activity does not establish patient benefit

Implementation and sustainability	Can the intervention be adopted, integrated, accessed, and maintained in its intended setting?	Implementation studies, program reports, determinant frameworks, health-technology evaluations	Setting; workforce; infrastructure; informatics; reimbursement; access; sustainability	Multilevel contextual determinants must be assessed rather than adoption alone	Explains variation between technically similar interventions	Context-specific evidence may not transfer across systems	Adoption, scale-up, and sustainability are distinct outcomes
Review transparency and synthesis	Are identification, eligibility, selection, extraction, appraisal, and synthesis sufficiently explicit?	Systematic-review guidance and transparent methodological reports	Questions; sources; search concepts; selection logic; extraction; appraisal; synthesis method	Reproducible reporting and explicit handling of heterogeneity [7]	Supports structured narrative and framework synthesis	Transparent reporting cannot correct weak primary evidence	Review methodology does not validate the underlying technologies
Cross-domain progression	Are evidence, manufacturing, clinical, and implementation conditions considered together?	Multidomain frameworks and studies spanning more than one translation stage	Domain strengths; unresolved uncertainty; dependencies; feedback; progression decision	No consequential domain should be replaced by strength in an unrelated domain	Original proposed synthesis linking domain-specific findings	No universal progression threshold has been validated	The proposed framework is explanatory and decision-organizing, not predictive or regulatory

Data extraction and quality assessment

Structured extraction recorded study design, population or system, technology, comparator, intended decision, evidence domain, validation stage, principal outcome, reproducibility safeguards, manufacturing relevance, implementation relevance, and stated limitations. Extraction emphasized what each source directly demonstrated rather than what its terminology implied.

Quality appraisal was design specific. Experimental studies were examined for comparator adequacy, reporting completeness, biological heterogeneity, replication, and transferability. Computational studies were assessed for data provenance, leakage control, split strategy, comparator selection, applicability, and external evaluation. Manufacturing evidence was appraised for process integration, operating context, analytical control, scale, transferability, and supply assumptions.

Clinical and implementation studies were evaluated for population relevance, workflow integration, contextual determinants, clinical utility, adoption, access, and sustainability. Reviews and frameworks were assessed for transparent evidence selection, distinction between evidence and interpretation, applicability, and acknowledgement of uncertainty.

No common numerical quality score was assigned because such a score would imply comparability among fundamentally different study designs. Instead, limitations were carried forward into the synthesis as domain-specific concerns. This approach identifies threats to interpretation without converting heterogeneous evidence into an unsupported readiness ranking.

Evidence-maturity classification

Evidence maturity was defined as the degree to which a claim had progressed from exploratory plausibility toward decision-relevant support. Retrospective accuracy, benchmark performance, or technological novelty represented early evidence unless accompanied by prospective testing, relevant comparators, external evaluation, and connection to an intended pharmaceutical decision [8]. Data provenance, assay context, missingness, chemical similarity, and selection bias further constrained the interpretation of apparently strong computational results [9]. Human evidence could strengthen maturity when it connected a target to clinically observed phenotypes. Genetic support was associated with greater target confidence, while phenotypes linked to target-encoding genes offered safety-relevant information [10, 11]. These associations may improve prioritization but cannot establish efficacy, causation, or complete safety.

Human microphysiological platforms occupied an intermediate and use-specific position. Their relevance depended on reproducibility, physiological fidelity, endpoint validity, benchmarking, and demonstration that outputs informed the intended human decision [12]. Human-derived architecture alone was insufficient evidence of general predictive validity.

The resulting classification distinguished conceptual plausibility, internal technical validation, independent or external evaluation, human relevance, manufacturing feasibility, clinical utility, and implementation readiness. These categories are an original synthesis rather than a validated maturity scale, and progression may remain uneven across them.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for characteristics and article-specific classification of the included evidence for crossing the pharmaceutical translation gap.

Table 2. Characteristics and article-specific classification of the included evidence for Crossing the Pharmaceutical Translation Gap.

Method, platform, or evidence category	Representative application	Reported strength	Validation context	Generalization concern	Contradictory evidence	Evidence gap	Review interpretation
Target assessment	Biological prioritization	Integrates biology, safety, tractability, and human evidence [3]	Portfolio decision	Framework use varies	Strong individual dimensions may not converge	Prospective decision impact	Mature only when dimensions are triangulated
Pharmaceutical machine learning	Prediction and design	Broad pipeline applicability [4]	Mainly retrospective	Leakage, dataset shift, and weak comparators [8, 9]	High benchmarks may not transfer	Prospective external utility	Technical performance is early-stage evidence
Human genetics	Target and safety support	Links genes to development outcomes or phenotypes [10, 11]	Retrospective human data	Association is not causation	Genetic support does not guarantee success	Prospective decision value	Strengthens but does not complete target validation
Organ chips	PK, toxicity, disease modelling	Human-relevant mechanistic representation [5, 12]	Use-specific experiments	Platform and endpoint heterogeneity	Results may be compound specific	Independent multicentre validation	Promising but conditional
Replication evidence	Preclinical reproducibility	Tests transferability of published findings	Repeated or multisite experiments	Protocol incompleteness and model variation	Replication outcomes are heterogeneous	Routine prospective replication	Direct evidence of credibility
Manufacturing studies	Integrated or modality-specific processes	Demonstrate process feasibility	Product- or platform-specific	Scale and transfer limitations	Laboratory integration may not persist commercially	End-to-end robustness evidence	Manufacturing readiness is product dependent
Implementation programs	Clinical adoption and workflow integration	Reveal infrastructure and contextual needs	Health-system programs	Setting dependence	Adoption may occur without patient benefit	Comparative utility and sustainability	Implementation activity is not outcome evidence
Proposed cross-domain classification	Progression decisions	Links evidence, manufacture, clinical use, and context	Original synthesis	Not externally validated	Domains may mature asynchronously	Prospective evaluation	Decision-organizing, not predictive

Reproducibility and external-validation findings

Reproducibility involved both recovery of reported findings and transport across relevant biological conditions. Replication projects exposed difficulties in protocol transfer and effect recovery, while methodological analyses showed that excessive standardization may reduce external relevance when treatment effects depend on biological variation [13, 14].

Transparent reporting was therefore foundational. Adequate description of experimental design, animals, procedures, analysis, and results enables appraisal and attempted replication, although reporting compliance alone cannot guarantee validity [15].

Computational evidence presented parallel concerns. Some ligand-based benchmarks rewarded recognition of familiar chemical structure rather than transferable prediction, making random or permissive splits poor proxies for external generalization [16].

Credible model evaluation consequently required transparent reporting of data, optimization, model development, and evaluation procedures [17]. External validation strengthened confidence only when the external data differed meaningfully from development data and reflected the intended pharmaceutical application. Reproducibility could support translation, but it could not establish manufacturability, clinical utility, or patient benefit alone.

Manufacturing and scale-up readiness

Manufacturing readiness was defined as evidence that a product and process could be integrated, controlled, transferred, scaled, and supplied while maintaining relevant quality. Continuous manufacturing therefore required organizational capability, process understanding, control strategy, and regulatory alignment in addition to connected equipment [18].

An end-to-end tablet platform demonstrated that synthesis, crystallization, filtration, formulation, and tableting could be integrated within a defined process [19]. Its importance was evidentiary rather than universal: feasibility for one product and architecture did not establish commercial robustness across compounds.

Advanced cell and gene therapies introduced different constraints. Decentralized manufacture could shorten distribution pathways but increased demands for comparability, traceability, workforce competence, release testing, logistics, and chain-of-custody control [20].

Accelerated antibody development illustrated how mature platforms, prior knowledge, parallel activities, and risk-based sequencing may compress timelines [21]. Such acceleration depended on established modality knowledge and could not be generalized to less mature products.

mRNA manufacturing similarly revealed bottlenecks in raw materials, enzymatic production, purification, formulation, analytical control, stability, and coordinated supply [22]. Across modalities, manufacturing readiness was therefore inseparable from product definition and could not be inferred from laboratory production alone.

Table 3 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for comparative evidence synthesis, methodological limitations, and sources of heterogeneity in crossing the pharmaceutical translation gap.

Table 3. Comparative evidence synthesis, methodological limitations, and sources of heterogeneity in Crossing the Pharmaceutical Translation Gap.

Cross-study pattern or configuration	Context or boundary condition	Mechanism or explanatory factor	Observed or reported outcome	Rival explanation	Certainty or rigour concern	Implication	Research priority
Strong internal performance, weak transfer	Computational or preclinical systems	Familiar data, standardized conditions	High local performance	Leakage or narrow applicability	External validation absent	Do not equate accuracy with utility	Prospective heterogeneous testing
Replication varies across settings	Preclinical experiments	Protocol detail and biological variation	Incomplete effect recovery [13]	Context-sensitive biology	Selected studies	Treat reproducibility as use specific	Multisite replication
Integrated manufacture succeeds for selected products	Continuous processing	Connected unit operations and control [19]	Technical feasibility	Product-specific tractability	Limited commercial evidence	Separate demonstration from readiness	Transfer and scale robustness
Platform maturity accelerates development	Established antibodies	Prior knowledge and parallelization [21]	Compressed development activity	Exceptional pandemic conditions	Observational evidence	Acceleration is platform dependent	Comparative platform studies

Modality novelty creates supply bottlenecks	mRNA and advanced therapies	Raw materials, analytics, logistics [20, 22]	Scale-up constraints	Temporary capacity shortage	Rapidly evolving systems	Assess supply early	End-to-end resilience studies
Adoption without demonstrated benefit	Clinical implementation	Workflow and infrastructure	Implementation activity	Selection and contextual effects	Few comparative outcomes	Separate use from benefit	Prospective clinical utility

Figure 1 presents the original conceptual synthesis for systematic-review selection and evidence-classification pathway, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

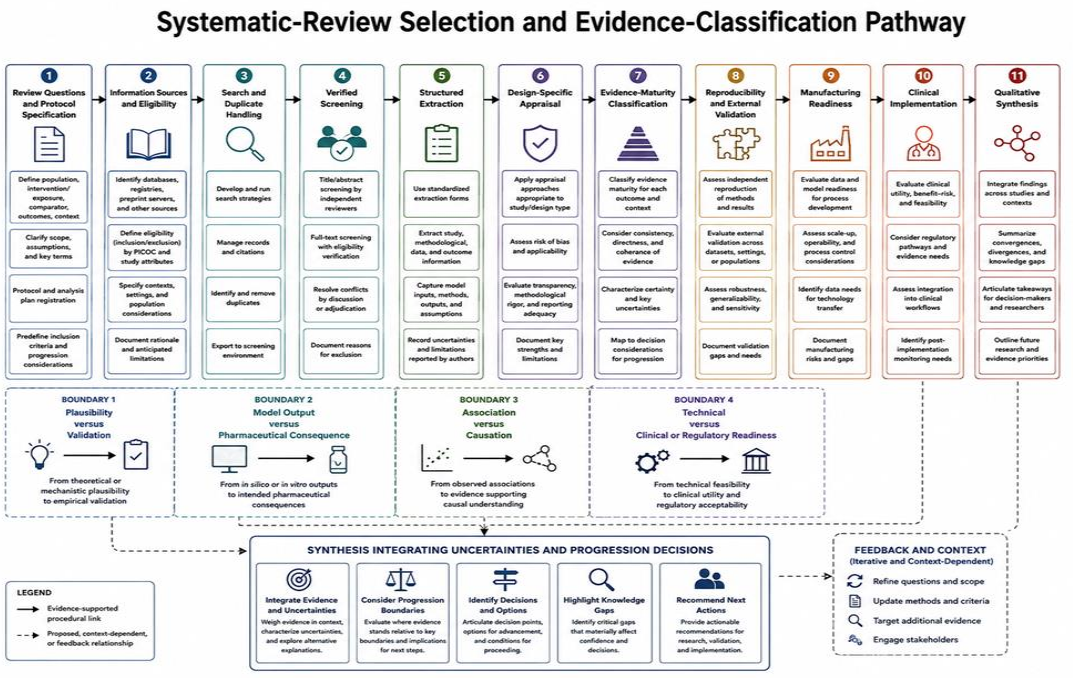

Figure 1. Systematic-Review Selection and Evidence-Classification Pathway

The figure is an original conceptual synthesis developed for "Crossing the Pharmaceutical Translation Gap: A Systematic Review of Evidence Maturity, Reproducibility, Manufacturing Readiness, and Implementation Barriers." Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

Clinical and implementation barriers

Pharmacogenomic programs demonstrated that implementation required coordinated testing, decision support, professional education, workflow design, and cross-institutional alignment [23]. US initiatives additionally identified evidence generation, informatics, stakeholder engagement, reimbursement, and implementation research as recurring needs [24].

Biological evidence itself remained incomplete. Uncertain variant function and imperfect prediction of clinically expressed phenotype limited the movement from genomic association to dependable prescribing decisions [25].

Implementation was also shaped by interactions among the technology, intended condition, value proposition, adopters, organization, wider system, and adaptation over time [26]. The updated Consolidated Framework for Implementation Research similarly organized determinants across intervention, setting, individual, and process domains [27].

These findings distinguished installation from implementation and implementation from benefit. Clinical uptake, decision-support availability, or testing activity did not establish improved outcomes, equitable access,

affordability, or sustainability. Patient relevance therefore required evidence extending beyond technical deployment.

Cross-domain synthesis and determinants of progression

Automation offered a possible mechanism for connecting hypothesis generation, synthesis, testing, and iterative learning, but each component retained separate validation requirements [28]. Autonomous experimentation was therefore not synonymous with autonomous pharmaceutical translation.

Robotic chemistry demonstrated that machine learning could guide experimental exploration of reaction space [29]. AI planning was also linked to automated flow synthesis for selected organic compounds [30]. These studies established component integration, not downstream therapeutic efficacy or manufacturing readiness.

Human microphysiological systems provided a different form of integration. Coupled organ chips generated quantitative pharmacokinetic predictions for selected drugs [31], while a Liver-Chip reproduced selected human and species-specific toxicities [32]. Their value remained dependent on use, compound, endpoint, and independent validation.

Across domains, progression was most credible when evidence triangulation, external validation, human relevance, manufacturability, clinical utility, and implementation fit aligned around the same intended decision. Strength in one domain could not compensate automatically for a consequential weakness in another.

The reviewed evidence therefore supported a "weakest consequential domain" interpretation. A candidate may be scientifically persuasive yet unmanufacturable, scalable yet clinically irrelevant, or clinically promising yet impossible to implement equitably. Progression decisions should identify which uncertainty could reverse the next decision and require evidence directed at that uncertainty.

Table 4 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for evidence certainty, unresolved gaps, and research priorities for crossing the pharmaceutical translation gap.

Table 4. Evidence certainty, unresolved gaps, and research priorities for Crossing the Pharmaceutical Translation Gap.

Evidence-maturity domain	Current evidence position	Methodological weakness	Translation barrier	Needed validation	Reporting priority	Decision implication	No-overclaiming boundary
Target evidence	Human and mechanistic evidence may strengthen confidence	Association and selection bias	Weak causal linkage	Independent functional and clinical triangulation	Provenance and conflicting evidence	Progress only for defined use	Genetic support does not guarantee success
Computational prediction	Strong retrospective capability	Leakage and dataset shift	Poor transportability	Prospective external evaluation	Splits, comparators, calibration	Restrict claims to applicability domain	Model output is not pharmaceutical utility
Preclinical models	Useful mechanistic evidence	Limited heterogeneity and replication	Weak human transfer	Multisite, fit-for-purpose validation	Protocol and endpoint detail	Match model to decision	Human resemblance is not clinical equivalence
Manufacturing	Feasibility shown for selected platforms	Product-specific evidence	Scale, control, supply, transfer	End-to-end robustness studies	Operating scale and quality attributes	Assess before late progression	Laboratory production is not readiness
Clinical utility	Established in selected contexts	Limited comparative outcomes	Uncertain decision benefit	Prospective utility evaluation	Population and workflow context	Require decision-relevant outcomes	Adoption is not patient benefit
Implementation	Multilevel determinants are recognized	Context dependence	Infrastructure, access, sustainability	Comparative and longitudinal evaluation	Adaptation and equity	Plan implementation early	Activity is not sustainability
Autonomous platforms	Component integration demonstrated	Narrow tasks and environments	Unvalidated end-to-end transfer	Closed-loop prospective evaluation	Human oversight and	Treat as enabling infrastructure	Automation is not autonomous translation

					failure modes		
Proposed cross-domain readiness	Convergent concept	No validated universal threshold	Asynchronous domain maturity	Prospective decision-impact studies	Residual uncertainty by domain	Gate on consequential weakness	Not a regulatory or predictive score

Evidence gaps and translational recommendations

Reproducibility improvement requires coordinated changes in study design, transparency, reporting, incentives, and evaluation rather than isolated procedural fixes [33]. Translation programs should therefore prespecify the evidence required to challenge a claim, not only evidence expected to confirm it.

Multidimensional progression frameworks may improve decision discipline by linking target confidence, tissue exposure, safety, patient selection, and value [34]. However, company-level experience cannot establish that one framework will improve outcomes across organizations or modalities.

Development, feasibility, evaluation, and implementation should be treated as iterative and context dependent rather than as a simple linear sequence [35]. Evidence generated during manufacturing or implementation may require redesign, renewed validation, or reconsideration of the intended population.

Field-specific recommendations emphasize relevant models, reproducibility, patient stratification, multidisciplinary continuity, and connected preclinical and clinical decisions [36]. Replication evidence likewise supports treating transparent and reproducible methods as translational infrastructure rather than optional documentation [37].

Figure 2 presents the original conceptual synthesis for translation-gap map across evidence, manufacturing, and implementation domains, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

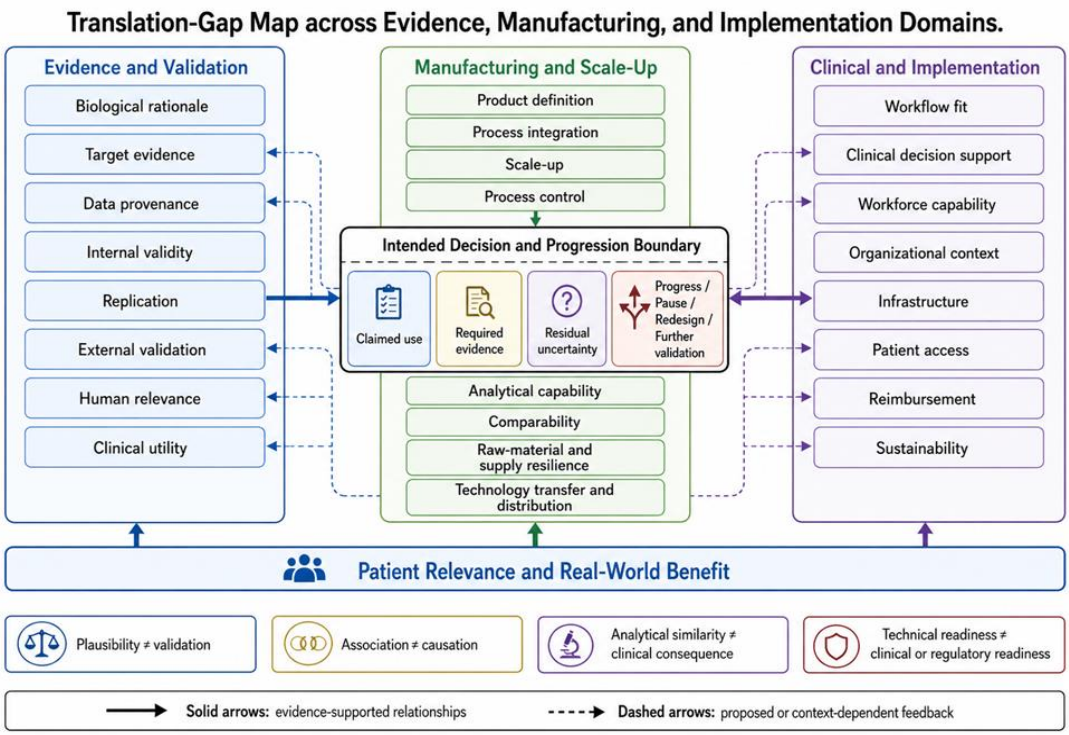

Figure 2. Translation-Gap Map across Evidence, Manufacturing, and Implementation Domains

The figure is an original conceptual synthesis developed for "Crossing the Pharmaceutical Translation Gap: A Systematic Review of Evidence Maturity, Reproducibility, Manufacturing Readiness, and Implementation Barriers." Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

Review limitations

The evidence base was heterogeneous in design, terminology, system, modality, comparator, and outcome. This precluded defensible meta-analysis and limited direct comparison across translation domains.

Searches relied on accessible bibliographic, article-record, publisher, DOI, and journal-ranking sources. Restricted database access, language and indexing effects, and publication bias may have reduced retrieval of relevant negative or unpublished evidence.

Quality appraisal was domain specific rather than numerical. This preserved design relevance but prevented calculation of a common certainty score. Replication evidence was itself affected by protocol availability, feasibility, and context [13].

Cross-domain synthesis necessarily involved interpretation. Implementation frameworks developed in broader health-technology settings may not transfer completely to every pharmaceutical modality [26]. The review identifies recurring determinants but cannot estimate their independent causal effects or provide a universal probability of progression.

CONCLUSION

Crossing the pharmaceutical translation gap requires more than scientific novelty, technical performance, or isolated proof of concept. The reviewed evidence supports a multidimensional interpretation in which biological credibility, reproducibility, external validation, human relevance, manufacturing feasibility, clinical utility, and implementation conditions must be aligned with an intended decision. Progression should remain conditional on the weakest consequential domain and on the specific product, population, process, and setting. The resulting framework is an evidence-bounded synthesis for organizing uncertainty and identifying the next required validation; it is not a validated readiness score, clinical recommendation, regulatory pathway, or deployment model.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. Biostatistics. 2019;20(2):273-86. doi:10.1093/biostatistics/kxx069
2. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. 2022;12(7):3049-62. doi:10.1016/j.apsb.2022.02.002
3. Emmerich CH, Martinez Gamboa L, Hofmann MCJ, Bonin-Andresen M, Arbach O, Schendel P, et al. Improving target assessment in biomedical research: The GOT-IT recommendations. Nat Rev Drug Discov. 2021;20(1):64-81. doi:10.1038/s41573-020-0087-3
4. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-77. doi:10.1038/s41573-019-0024-5
5. 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
6. Roden DM, McLeod HL, Relling MV, Williams MS, Mensah GA, Peterson JF, et al. Pharmacogenomics. Lancet. 2019;394(10197):521-32. doi:10.1016/S0140-6736(19)31276-0
7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372. doi:10.1136/bmj.n71
8. Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: What is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. Drug Discov Today. 2021;26(2):511-24. doi:10.1016/j.drudis.2020.12.009

9. Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: What is realistic, what are illusions? Part 2: A discussion of chemical and biological data. Drug Discov Today. 2021;26(4):1040-52. doi:10.1016/j.drudis.2020.11.037
10. Nguyen PA, Born DA, Deaton AM, Nioi P, Ward LD. Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. Nat Commun. 2019;10(1):1579. doi:10.1038/s41467-019-09407-3
11. Hingorani AD, Kuan V, Finan C, Kruger FA, Gaulton A, Chopade S, et al. Improving the odds of drug development success through human genomics: Modelling study. Sci Rep. 2019;9(1):18911. doi:10.1038/s41598-019-54849-w
12. 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
13. Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, et al. Investigating the replicability of preclinical cancer biology. Elife. 2021;10. doi:10.7554/eLife.71601
14. Voelkl B, Altman NS, Forsman A, Forstmeier W, Gurevitch J, Jaric I, et al. Reproducibility of animal research in light of biological variation. Nat Rev Neurosci. 2020;21(7):384-93. doi:10.1038/s41583-020-0313-3
15. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. PLoS Biol. 2020;18(7). doi:10.1371/journal.pbio.3000410
16. Wallach I, Heifets A. Most ligand-based classification benchmarks reward memorization rather than generalization. J Chem Inf Model. 2018;58(5):916-32. doi:10.1021/acs.jcim.7b00403
17. Walsh I, Fishman D, Garcia-Gasulla D, Titma T, Pollastri G, Harrow J, et al. DOME: Recommendations for supervised machine learning validation in biology. Nat Methods. 2021;18(10):1122-7. doi:10.1038/s41592-021-01205-4
18. Badman C, Cooney CL, Florence AJ, Konstantinov K, Krumme M, Mascia S, et al. Why we need continuous pharmaceutical manufacturing and how to make it happen. J Pharm Sci. 2019;108(11):3521-3. doi:10.1016/j.xphs.2019.07.016
19. Domokos A, Nagy B, Gyürkés M, Farkas A, Tacsi K, Pataki H, et al. End-to-end continuous manufacturing of conventional compressed tablets: From flow synthesis to tableting through integrated crystallization and filtration. Int J Pharm. 2020;581:119297. doi:10.1016/j.ijpharm.2020.119297
20. Harrison RP, Ruck S, Rafiq QA, Medcalf N. Decentralised manufacturing of cell and gene therapy products: Learning from other healthcare sectors. Biotechnol Adv. 2018;36(2):345-57. doi:10.1016/j.biotechadv.2017.12.013
21. Kelley B. Developing therapeutic monoclonal antibodies at pandemic pace. Nat Biotechnol. 2020;38(5):540-5. doi:10.1038/s41587-020-0512-5
22. Rosa SS, Prazeres DMF, Azevedo AM, Marques MPC. mRNA vaccines manufacturing: Challenges and bottlenecks. Vaccine. 2021;39(16):2190-200. doi:10.1016/j.vaccine.2021.03.038
23. van der Wouden CH, Cambon-Thomsen A, Cecchin E, Cheung KC, Dávila-Fajardo CL, Deneer VHM, et al. Implementing pharmacogenomics in Europe: Design and implementation strategy of the Ubiquitous Pharmacogenomics Consortium. Clin Pharmacol Ther. 2017;101(3):341-58. doi:10.1002/cpt.602
24. Volpi S, Bult CJ, Chisholm RL, Deverka PA, Ginsburg GS, Jacob HJ, et al. Research directions in the clinical implementation of pharmacogenomics: An overview of US programs and projects. Clin Pharmacol Ther. 2018;103(5):778-86. doi:10.1002/cpt.1048
25. Lauschke VM, Ingelman-Sundberg M. Emerging strategies to bridge the gap between pharmacogenomic research and its clinical implementation. NPJ Genom Med. 2020;5:9. doi:10.1038/s41525-020-0119-2
26. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to scale-up, spread, and sustainability of health and care technologies. J Med Internet Res. 2017;19(11). doi:10.2196/jmir.8775
27. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci. 2022;17(1):75. doi:10.1186/s13012-022-01245-0
28. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17(2):97-113. doi:10.1038/nrd.2017.232
29. Granda JM, Donina L, Dragone V, Long DL, Cronin L. Controlling an organic synthesis robot with machine learning to search for new reactivity. Nature. 2018;559(7714):377-81. doi:10.1038/s41586-018-0307-8

30. Coley CW, Thomas DA 3rd, Lummiss JAM, Jaworski JN, Breen CP, Schultz V, et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. Science. 2019;365(6453). doi:10.1126/science.aax1566
31. 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
32. 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
33. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. Nat Hum Behav. 2017;1(1):0021. doi:10.1038/s41562-016-0021
34. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. Nat Rev Drug Discov. 2018;17(3):167-81. doi:10.1038/nrd.2017.244
35. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. BMJ. 2021;374. doi:10.1136/bmj.n2061
36. Tocchetti CG, González A, Backs J, Pollesello P, Rainer PP, Schiattarella GG, et al. How to facilitate seamless translation from basic concepts to new heart failure drugs: A scientific statement of the Heart Failure Association of the ESC. Eur J Heart Fail. 2025;27(8):1379-92. doi:10.1002/ejhf.3709
37. Errington TM. Building reproducible bridges to cross the "valley of death". J Clin Invest. 2024;134(1). doi:10.1172/JCI177383