



Virtual cells: evolution, applications, and challenges

Yueqi Li^{1,2,#}, Wei Xiong^{3,#}, Kuan Feng^{1,2,#}, Yingying Ge², Jun Wang², Yunjia Ning^{1,2}, Fei Deng^{1,2}, Hualin Wang^{1,2}

Keywords:

Virtual cells, whole-cell models, artificial intelligence, AI Agent, perturbation prediction, digital twin

Citation: Li, Y.; Xiong, W.; Feng, K.; Ge, Y.; Wang, J.; Ning, Y.; Deng, F.; Wang, H. Virtual cells: evolution, applications, and challenges. *AI Agent* 2026, 2, 28. <https://dx.doi.org/10.20517/aiagent.2026.16>

Received: 12 May 2026

First Decision: 11 Aug 2026

Revised: 25 Sep 2026

Accepted: 15 Oct 2026

Published: 27 Oct 2026

Academic Editor:

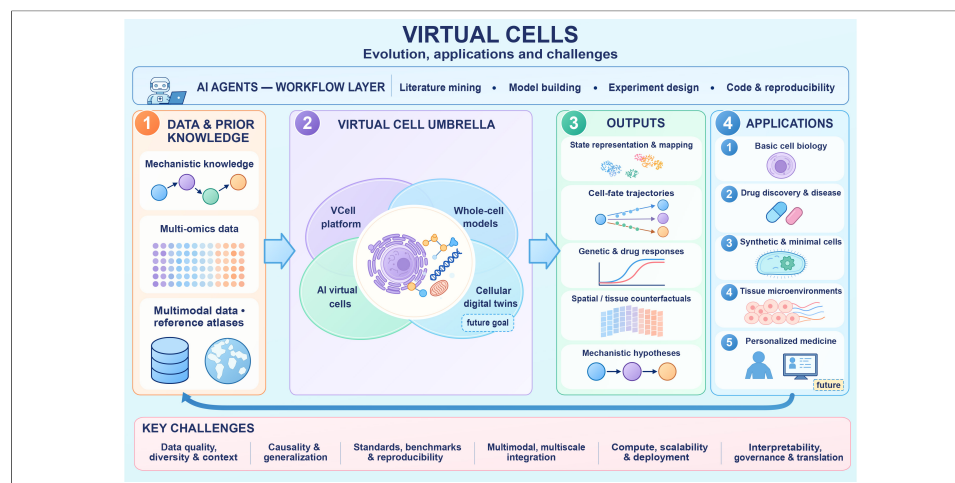
Aloysius Soon

Copy Editor:

Tong Wang

Production Editor:

Tong Wang



Abstract

Predicting cellular behavior is a central goal of systems biology, and virtual cells aim to provide computational representations of living cells that can simulate cellular states, predict responses to genetic or pharmacological perturbations, and support hypothesis testing across biological contexts. Ideally, virtual cells can serve as executable models for exploring cell fate transitions, disease mechanisms, therapeutic responses, and eventually patient-specific digital twins. By integrating multimodal cellular data with artificial intelligence (AI)-driven modeling, virtual cell research is evolving from descriptive simulation toward perturbation-aware modeling with increasing emphasis on evaluating predicted responses to defined interventions. In this review, we trace the development of virtual cells from their conceptual origins to the emergence of AI-driven virtual cells, summarize representative methodological advances and major applications, and discuss the key challenges that continue to constrain progress in the field, including noisy data and incomplete contextual information, limited causal inference and cross-context generalization, fragmented benchmarking and evaluation frameworks, and bottlenecks in computational resources and large-scale deployment. We further highlight the need for open data resources, unified ontologies, reproducible benchmarks, and shared community



¹State Key Laboratory of Virology and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, Hubei, China.

²National Virus Resource Center, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, Hubei, China.

³Department of Orthopedics, Wuhan Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, Hubei, China.

#These authors contributed equally to this work.

Correspondence to: Prof. Yunjia Ning, Prof. Fei Deng, Prof. Hualin Wang, State Key Laboratory of Virology and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, Hubei, China. E-mail: nyj@wh.iov.cn; df@wh.iov.cn; h.wang@wh.iov.cn

infrastructure. Together, this review aims to provide a timely overview of how virtual cells are evolving from mechanistic simulations into AI-driven predictive systems, and how they may ultimately become transformative tools for biomedical research and precision medicine.

INTRODUCTION

A fundamental paradigm shift is underway in the life sciences, with biology rapidly evolving from a science primarily focused on description and classification into one centered on prediction and engineering^[1-3]. The core driving force behind this shift lies in our pursuit of the computability of the most basic unit of living systems^[3,4]. Cellular behavior, from gene expression and signal transduction to metabolic regulation and fate selection, is essentially a complex information processing program operating at the molecular scale^[2,5]. Understanding and simulating this program, constructing a virtual cell, has not only been a goal of systems biology for decades but has also become an inevitable convergence point for the two cutting-edge fields of precision medicine and artificial intelligence (AI)-driven biology^[3,4,6].

Traditional systems biology has used mathematical equations and computational models to analyze intracellular regulatory networks, attempting to understand the emergent nature of life from a mechanistic perspective^[2,7]. Precision medicine urgently needs to predict cellular behavior in specific genetic backgrounds or microenvironments to achieve personalized intervention strategies^[6,8]. In recent years, artificial intelligence technologies, represented by deep learning, have provided a powerful new set of tools capable of automatically learning complex mapping relationships from massive amounts of high-dimensional data^[9,10]. The convergence of these three technologies has made the construction of a virtual cell that can respond to genetic or chemical perturbations and predict cellular state transitions no longer a distant science fiction concept, but rather an operational, assessable, and increasingly competitive frontier in scientific research^[3,11].

Amid profound transformation and rapid growth in the field, the concept of virtual cells warrants re-examination now. Early “Virtual Cell” was primarily a mechanism modeling and simulation environment for cell biology, with its core based on solving differential equations related to reaction networks, diffusion, and spatial processes^[12,13]. However, in the last two or three years, with the rapid growth of single-cell omics^[14], spatial omics^[15], and large-scale perturbation data^[16], the concept of “AI virtual cell” has emerged strongly. It no longer relies on explicitly writing out the equations for every biochemical reaction, but emphasizes learning a generalizable mapping function from cell state to intervention-response on massive, multimodal data^[3,10]. Between 2024 and 2026, a series of milestone events, from the Cell paper defining the field’s agenda^[3] to the global challenge aimed at establishing unified evaluation standards^[17], and the successive releases of large-scale models such as State^[18] and X-Cell^[19], signify that virtual cells have evolved from an academic concept into a systematic engineering project driven by data, models, benchmarks, and industry forces.

Given that the term “virtual cell” has been given multiple meanings in literature spanning over two decades, it is necessary to clearly define it at the outset to avoid conceptual confusion. Here, we use “virtual cell” as an umbrella term for computational models that represent cellular states and simulate or predict biologically meaningful cellular functions, behaviors, or state transitions under defined conditions. We consider three minimum capabilities essential for this designation: (i) an explicit computational representation of a cell or cellular state; (ii) the ability to model at least one biologically meaningful cellular function, behavior, or transition; and (iii) the ability to simulate or predict changes in that cellular behavior under defined conditions, perturbations, or environmental contexts. Experimental validation is not a prerequisite for the terminology itself, but is essential for establishing biological fidelity and practical utility.

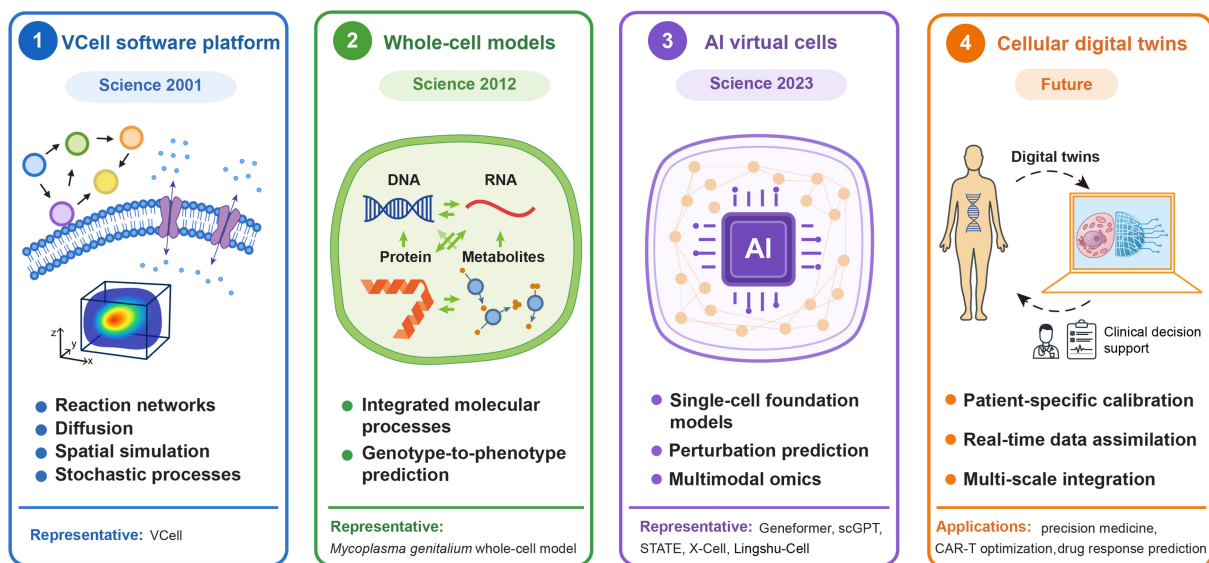


Figure 1. Conceptual landscape of virtual-cell research. scGPT: Single-cell generative pre-trained transformer; CAR-T: chimeric antigen receptor T cell.

Table 1. Definitions and distinctions among major concepts related to virtual cells

Concept	What it denotes	Model and data basis	Capability	Evidence boundary	Representative examples
VCell software platform	Software environment for user-specified mechanistic models	Reaction networks; differential-equation, spatial and stochastic solvers	Simulates defined cellular processes under specified conditions	Requires specified mechanisms, parameters and boundary conditions; scope depends on the model	VCell ^[12]
Whole-cell model	Integrative modeling scope for a specified cell or organism	Coupled submodels of annotated molecular processes	Simulates selected cell-level phenotypes, including growth and gene perturbation	Coverage and validation depend on available knowledge; difficult to extend to complex cells	<i>Mycoplasma genitalium</i> model ^[20] , JCVI-syn3A ^[21]
AI virtual cell	Data-driven or hybrid modeling of cell states and, where tested, state changes	Single-cell and spatial profiles; measured intervention responses; learned features	Cell representation and conditional response prediction when trained for that task	Representation alone does not establish response prediction; held-out accuracy does not identify mechanism	scGPT ^[9] , Geneformer ^[10] , STATE ^[18] , X-Cell ^[19]
Cellular digital twin	Individualized, dynamically updated modeling goal that can combine other approaches	Longitudinal multimodal data, calibration and data assimilation (proposed)	Potential for patient-specific prediction and model updating	Continuous synchronization and prospective patient-level validation remain limited	Proposed CAR-T framework ^[22] , organ digital twin concepts ^[23,24]

AI: Artificial intelligence; scGPT: single-cell generative pre-trained transformer; CAR-T: chimeric antigen receptor T cell.

Within this umbrella, VCell, whole-cell models, AI virtual cells, and cellular digital twins describe different dimensions. VCell refers to a specific mechanistic modeling platform; whole-cell modeling describes the scope of biological integration; AI virtual cells describe a data-driven or hybrid modeling paradigm; and cellular digital twins emphasize individualized, dynamically updated representations linked to specific biological systems. These dimensions can overlap and can be combined within a single modeling framework [Figure 1]. To clarify the conceptual boundaries and relationships among these related terms, Table 1 summarizes their definitions, conceptual roles, and relationships to the broader virtual-cell concept.

ADVANCES AND TRENDS IN VIRTUAL CELL RESEARCH

Mechanistic platforms and whole-cell models

Virtual cell research originated from the development of computational cell biology in the late 20th century. In 2001, Loew and Schaff proposed the Virtual Cell (VCell) platform, constructing a general framework for transforming cell biology problems into mathematical models and numerical simulations^[25]. In VCell, users specify molecular species, reactions, cellular compartments, and, for spatial models, cellular geometry; the software translates these biological descriptions into mathematical equations for numerical simulation^[12]. Compartmental models describe reaction kinetics using ordinary differential equations, whereas spatial models use partial differential equations to describe how concentrations change through local reactions and molecular transport^[12]. This coupling makes cellular geometry part of the mechanistic model, allowing researchers to investigate how molecular localization and diffusion influence signaling dynamics^[12]. However, limited by data scarcity and insufficient computing power at the time, model parameters mainly relied on literature inferences, and research focused on local pathways and processes, failing to cover complete cell behavior.

Subsequent developments extended VCell's modeling methods and application scope. Blinov *et al.* introduced compartmental and spatial rule-based modeling, enabling molecular interaction rules to generate reaction networks for deterministic or stochastic simulations across cellular compartments^[26]. Moving-boundary numerical methods developed for VCell were subsequently extended to free-boundary models that explain cell shapes and turning behavior through actomyosin mechanics^[27]. More recently, Lee *et al.* used VCell to investigate how rapidly accelerated fibrosarcoma (RAF) abundance shapes epidermal growth factor receptor-extracellular signal-regulated kinases (EGFR-ERK) signaling^[28]. These studies illustrate VCell's continuing role in connecting explicit biological mechanisms with experimentally testable predictions, while emphasizing the need for experimentally constrained parameters and validation^[28].

With the increasing demand for system integration, whole-cell modeling became a key turning point. In 2012, Karr *et al.* constructed a whole-cell model of *Mycoplasma genitalium* by integrating 28 submodels representing cellular processes, including metabolism, DNA replication, transcription, and translation^[20]. These submodels used process-specific mathematical descriptions and exchanged shared cellular variables, such as metabolite concentrations and molecular counts, at one-second simulation intervals^[20]. This integration allowed changes in one process to affect other processes through their shared cellular state. For example, simulations of single-gene disruptions predicted effects on cellular growth and enabled comparison with experimentally measured gene essentiality, illustrating how molecular perturbations could be connected to whole-cell phenotypes^[20]. However, its applicability was limited to simple cell systems, model construction heavily relied on human knowledge integration, and insufficient parameters and validation data made it difficult to extend to complex eukaryotic cells.

Following the 2012 model, whole-cell modeling expanded to more complex bacterial systems and became a framework for evaluating experimental data consistency. Macklin *et al.* developed an integrated *E. coli* model that cross-evaluated heterogeneous datasets and identified inconsistencies with functional consequences^[29]. Thornburg *et al.* subsequently constructed kinetic and spatial models of the genetically minimal bacterium JCVI-syn3A, linking metabolism, gene expression, and cellular growth^[21]. Their 2026 study extended this approach to a complete spatially resolved cell cycle, incorporating chromosome replication and segregation, membrane growth, and division^[30]. These advances broadened whole-cell modeling from phenotype prediction to data reconciliation and spatial cell-cycle reconstruction. Nevertheless, inconsistent measurements, omitted mechanisms, and the computational cost of spatial simulations remain important challenges^[29,30].

Data foundations and spatial mapping

Breakthroughs at the data level significantly changed the research path. Single-cell transcriptome sequencing technology has achieved large-scale application, enabling precise description of cell states in high-dimensional vector form^[31,32]. The Human Cell Atlas (HCA) project^[33] has systematically built reference cell atlases that provide diverse cellular contexts for model pretraining, annotation, and external evaluation; however, observational data alone cannot determine how cells respond to perturbations. Perturb-seq combines pooled clustered regularly interspaced palindromic repeats (CRISPR)-based genetic perturbations with single-cell RNA sequencing, linking the perturbation assigned to each cell to its transcriptomic readout^[34]. By comparing perturbed cells with controls, researchers can identify affected genes and cellular programs while resolving heterogeneous responses within the population^[16,34]. In the original study, Dixit *et al.* applied this approach to transcription factors regulating the response of dendritic cells to lipopolysaccharide, demonstrating how targeted interventions could reveal the organization of an immune-response circuit^[34]. Such datasets provide experimentally observed intervention-response relationships for training and evaluating virtual-cell predictors. Spatial omics complements these measurements by preserving positional information and enabling analysis of gene expression within tissue architecture^[35]. These advances collectively construct a comprehensive data system encompassing cell state representation, perturbation response, and spatial environment information, providing a crucial foundation for characterizing dynamic cell changes.

AI models for cell representation and perturbation prediction

An early use of deep learning was to place dissociated cells in tissue context. Tangram, published in 2021, aligns single-cell or single-nucleus profiles with spatial measurements to infer where cell populations and unmeasured genes are likely to occur^[36]. This is a prediction about location and spatial expression, rather than a prediction of how cells respond to an intervention.

As single-cell transcriptomic datasets expanded, researchers began training models on large collections of cell profiles to learn relationships among genes and cellular states, then explored how these learned features could support tasks such as cell annotation and disease analysis [Figure 2]. Geneformer, published in 2023, orders genes in each cell by expression relative to a reference corpus and learns to recover masked gene identities from context^[10]. Its learned features were also used to prioritize possible therapeutic targets for cardiomyopathy when disease-specific data were limited^[10]. Single-cell generative pre-trained transformer (scGPT), published in 2024, instead pairs gene identities with discretized expression values and learns cell and gene representations while predicting masked expression values^[9]. Both approaches provide features for tasks such as cell annotation. scGPT was also adapted and evaluated for a perturbation task^[9], but pretraining on observed cell profiles alone cannot establish the accuracy of a model's predictions for new interventions.

The representation approach then expanded in scale and spatial detail. In 2024, scFoundation used masked expression reconstruction on a large human single-cell collection to learn gene and cell features^[37]. Its gene features were also tested within a separate perturbation-prediction pipeline^[37]; that downstream application should not be confused with the data used for pretraining. Nicheformer subsequently combined dissociated and spatial transcriptomes to represent both a cell and its tissue neighborhood^[38], extending the spatial perspective introduced by Tangram. Universal cell embedding (UCE) pursued a related goal of placing cells from different tissues and species in a shared representation^[39]. These models can describe cell states and spatial context, although their success on those tasks does not by itself demonstrate prediction of biological intervention responses.

A parallel research direction asks how cells change after a defined intervention and uses measured responses to train or test that prediction. Graph-enhanced gene activation and repression simulator (GEARS), published online in 2023, learned from single-cell genetic screens and gene relationships to predict

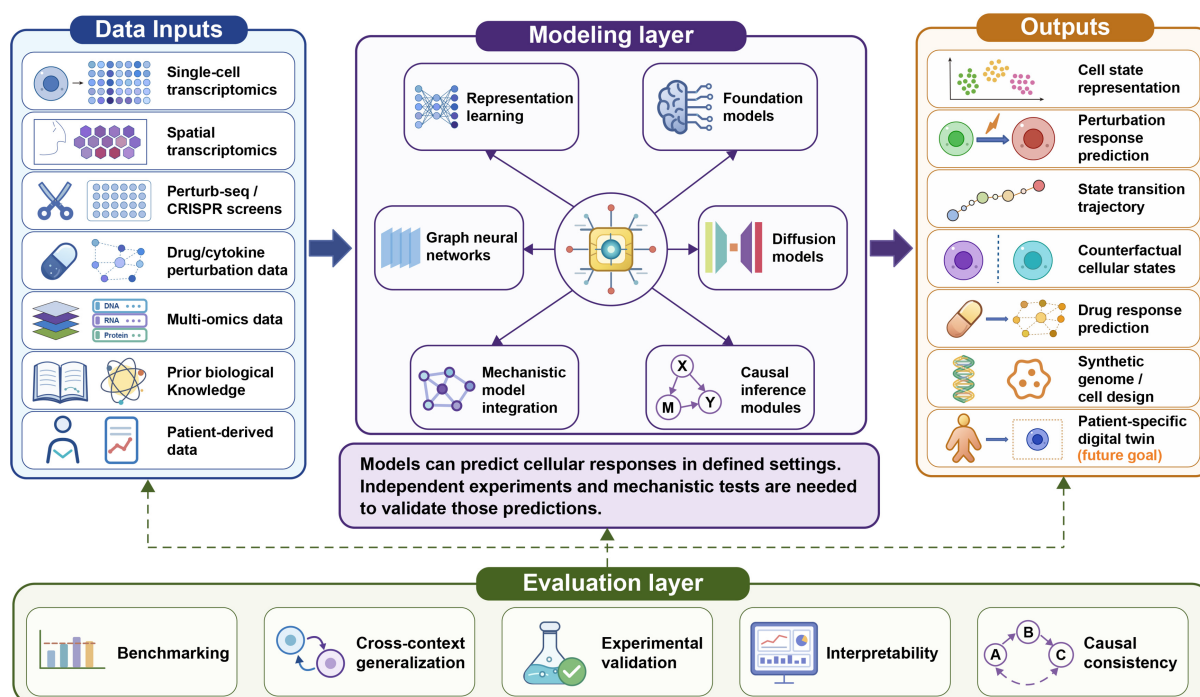


Figure 2. Data-model-prediction and evaluation framework for AI-driven virtual cells. AI: Artificial intelligence.

expression after single or combined gene perturbations, including combinations not seen during training^[40]. The 2024 agenda for AI virtual cells made experimentally testable responses a central goal^[3]. CellFM, reported in 2025, connects the two directions: it was first pretrained on transcriptomes from around 100 million human cells, then its learned gene features were substituted into GEARS and assessed with Perturb-seq response data^[41]. The authors also adapted CellFM using measured perturbations to predict responses to untested gene combinations^[41]. Thus, its perturbation results rely on a response-prediction step, even though its large-scale pretraining was based on cell profiles. The 2025 Virtual Cell Challenge provided shared tasks for comparing response models^[17]. State later reported predictions across several perturbation types and contexts^[18], while X-Cell focused on genetic responses across cell types^[19]. How well these models transfer to new interventions and biological settings remains an empirical question.

Evaluating generalization and causal claims

Model evaluation must keep these two lines distinct. For a cell representation, a zero-shot test asks whether pretrained embeddings support an evaluation task without further task-specific training. Kedzierska *et al.* found that unmodified scGPT and Geneformer embeddings did not consistently outperform simpler methods in cell-type separation and batch integration, and noted overlap between some evaluation and pretraining datasets^[42]. Their result concerns those models and tests; it does not assess every foundation model or gains after fine-tuning. For a response predictor, the withheld unit must be stated: a perturbed gene, a gene combination, a cell type, a donor, or an experimental condition. These splits answer different questions about transfer. A comparison of perturbation-prediction methods also found that deep learning did not consistently outperform simple linear baselines on the examined tasks^[11]. Validation therefore requires clearly defined splits, appropriate baselines, and, where possible, new experimental measurements^[11,17].

Representative results illustrate why performance claims must be interpreted within their specific evaluation settings [Figure 3 and Supplementary Table 1]. In the original GEARS study^[40], GEARS reduced the mean squared error (MSE) for the top 20 differentially expressed genes relative to the next-best baseline by 54.2%

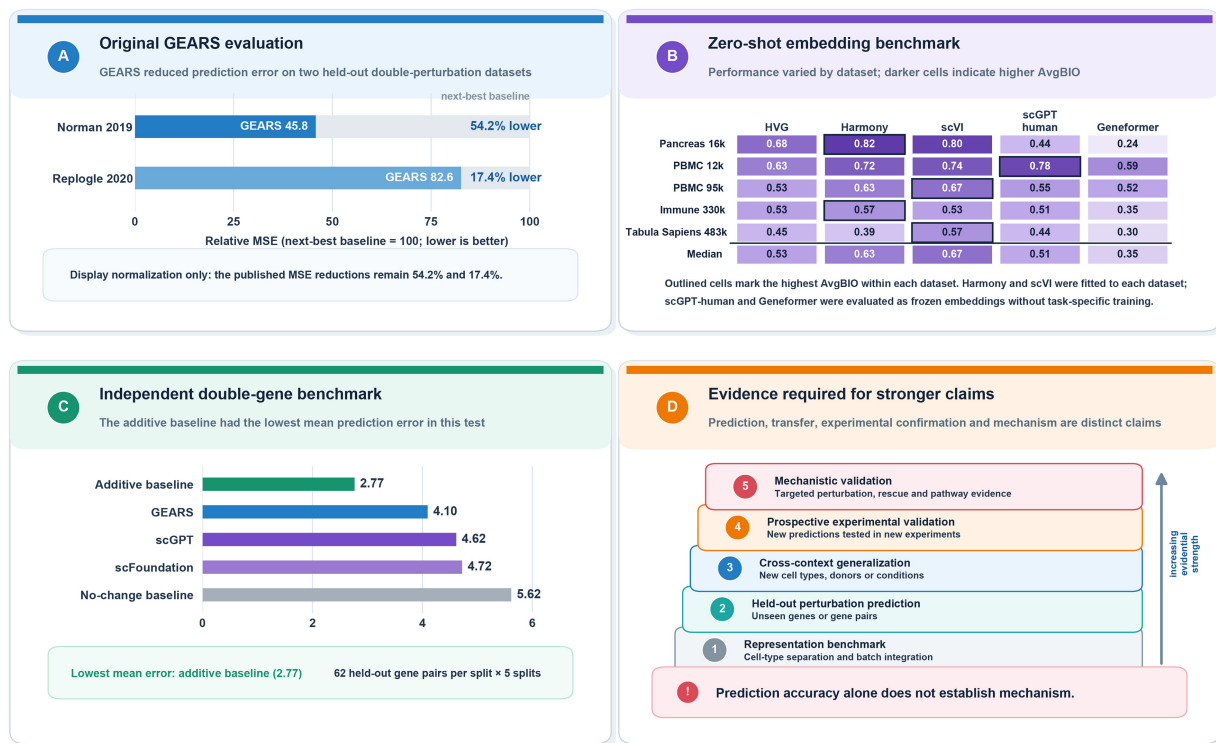


Figure 3. Evaluation-dependent performance of AI virtual-cell models and the evidence required for stronger biological claims. (A) In the original GEARs evaluation, the MSE for the top 20 differentially expressed genes was reduced by 54.2% and 17.4% relative to the next-best baseline on the Norman 2019 and Replogle 2020 K562 double-perturbation datasets, respectively. Baseline MSE was normalized to 100 for visualization. Adapted from reference^[40] under the terms of the CC BY 4.0 License; (B) AvgBio of cell embeddings across five datasets. Darker cells indicate higher scores, outlined cells mark the best-performing method within each dataset, and the bottom row shows the median. Harmony and scVI were fitted to each dataset, whereas scGPT-human and Geneformer were evaluated as frozen embeddings. Adapted from reference^[42] under the terms of the CC BY 4.0 License; (C) Mean L2 prediction errors in an independent double-perturbation benchmark. The additive baseline achieved the lowest mean error, compared with GEARs, scGPT, scFoundation, and the no-change baseline. Adapted from reference^[11] under the terms of the CC BY 4.0 License; (D) Conceptual hierarchy linking increasingly strong biological claims to the corresponding evaluation evidence. AI: Artificial intelligence; GRARS: graph-enhanced gene activation and repression simulator; MSE: mean squared error; AvgBio: average biological conservation; scVI: single-cell variational inference; scGPT: single-cell generative pretrained transformer.

on the Norman 2019 dataset and by 17.4% on the Replogle 2020 dataset, both of which contain K562 double-perturbation data [Figure 3A]. In a separate zero-shot evaluation of cell embeddings, frozen pretrained embeddings from scGPT-human and Geneformer did not consistently outperform highly variable genes (HVG), Harmony, or single-cell variational inference (scVI) across five datasets. This comparison must be interpreted in light of differences in training access, because Harmony and scVI were fitted to each evaluation dataset, whereas scGPT-human and Geneformer were evaluated without task-specific training^[42] [Figure 3B]. In an independent double-perturbation benchmark, the additive baseline achieved the lowest mean L2 prediction error (2.77), compared with GEARs (4.10), scGPT (4.62), scFoundation (4.72), and a no-change baseline (5.62), across 62 held-out gene pairs per split and five splits^[11] [Figure 3C]. Because Figure 3A–C differ in task, training access, data splits, comparators, and metrics, their values should not be directly compared or combined into a single model ranking. Figure 3D summarizes the progressively stronger forms of evaluation required to support claims about representation quality, perturbation prediction, cross-context transfer, prospective experimental validity, and mechanism.

Predicting the effect of an intervention does not by itself explain its mechanism. A controlled perturbation experiment can establish an effect on a measured outcome in the system studied; a model may learn to predict that effect for untested perturbations. Agreement with held-out responses, however, does not identify

Table 2. Representative virtual-cell models and simulation platforms

Model/Platform	Type	Main data/modeling strategy	Main task	Reported evaluation evidence	Evidence limit
VCell ^[12]	Mechanistic platform	Reaction-diffusion and stochastic simulation	Simulate specified cellular pathways	Published compartmental and spatial modeling applications	Requires defined mechanisms and parameters; not a whole-cell model
Mycoplasma genitalium WCM ^[20]	Whole-cell model	Integrated, literature-curated cellular processes	Simulate growth and gene perturbations	Simulated phenotypes compared with measured cellular and knockout data	Model and validation concern a simple bacterium
scGPT ^[9]	Single-cell foundation model	Generative pretraining on single-cell profiles	Representation, annotation and response tasks	Reported annotation/integration and perturbation-task benchmarks	Independent zero-shot embeddings can trail simple baselines
Geneformer ^[10]	Single-cell foundation model	Transfer learning from ranked gene expression	Cell-state and network predictions	Transfer-learning tasks and experimental TEAD4 test; separate zero-shot tests	Zero-shot cell-type/batch performance is dataset-dependent
scFoundation ^[37]	Single-cell foundation model	Pretrained on > 50 million human cell profiles	Expression, annotation and response tasks	Reported task benchmarks, including perturbation prediction with adapted features	Task-specific adaptation; pretrained representation alone is not a response model
Nicheformer ^[38]	Single-cell and spatial foundation model	Joint pretraining on dissociated and spatial transcriptomes	Spatial labels and neighborhood context	Spatial-label transfer and neighborhood-density probing on held-out data	Predicted spatial context is not a perturbation effect
UCE ^[39]	Cell embedding model	Unlabeled single-cell RNA data plus protein-sequence embeddings	Cross-dataset and cross-species cell embeddings	Zero-shot embeddings on new datasets and species; cell-type/batch metrics	Embedding transfer does not validate intervention response
CellFM ^[41]	Single-cell foundation model	Pretrained on 100 million human cells; adapted to perturbation data	Cell annotation and perturbation response	Adamson/Norman held-out perturbations; expression error and correlation	Response results use a trained prediction component; not causal identification
GEARS ^[40]	Perturbation prediction model	Perturb-seq responses plus gene interaction priors	Single- and two-gene response prediction	Held-out single/double gene perturbations; non-additive response tests	Predicted effects do not identify a unique causal mechanism
STATE ^[18]	Perturbation prediction model	Set-based state representation and transition model	Predict expression after perturbation	Held-out perturbation and cellular-context tests with mean baselines	Predictive accuracy varies with test split and metric
X-Cell ^[19]	Diffusion perturbation model	CRISPRi single-cell data with biological priors	Predict genetic perturbation responses	Preprint: held-out contexts including primary T-cell donors	Preprint evidence; transfer across patients or mechanisms untested
Lingshu-Cell ^[43]	Generative transcriptome model	Masked discrete diffusion of single-cell expression	Generate states and predict responses	Preprint: Virtual Cell Challenge H1 and cytokine-response tests	Benchmark prediction; no clinical or mechanistic validation
JCVI-syn3A 4DWCM ^[30]	Spatial whole-cell model	Spatial and kinetic simulation of a minimal cell	Simulate growth and division in 4D	Growth, replication and molecular measures compared with experiments	Validated for one minimal bacterial system
Celcomen ^[44]	Spatial perturbation model	Generative graph model separating intra-/intercellular effects	Generate spatial counterfactuals	Simulations and human tissues; mouse lung tumor knockout spatial data	Causal graph requires assumptions; Visium spots are multicellular

TEAD4: TEA domain transcription factor 4; CRISPRi: CRISPR interference.

the intermediate regulatory interactions that produced them or show that the response will recur in other tissues. Mechanistic claims require targeted follow-up, such as independent perturbations or rescue experiments, while claims of transfer require data from cell types or conditions excluded from training. Table 2 therefore lists the questions addressed by representative models alongside their reported evaluation evidence and limits^[11,40].

AI agents for modeling and experimental design

AI agents can assist with experiment design and with the construction or assessment of virtual-cell models, but an agent coordinating these tasks is distinct from a model that represents or predicts cellular behavior. BioDiscoveryAgent, presented at International Conference on Learning Representations (ICLR) 2025, uses a large language model (LLM) with literature-search and code tools to prioritize genetic perturbations and was evaluated using existing perturbation datasets^[45]. More recent preprint studies illustrate model-building workflows: CellForge coordinates agents to design architectures and generate executable code^[46], while VCHarness uses an AI coding agent to construct and refine perturbation-response models^[47]. The reported evaluations concern experiment prioritization or the performance of the resulting models, not the agent as a cellular model.

Other preprint studies focus on assessing explanations or existing model repositories. VCR-Agent uses retrieved biological knowledge and verifier-based filtering to produce candidate mechanistic explanations^[48], whereas AblateCell reproduces model repositories and performs ablations to assess contributions to benchmark performance^[49]. These computational checks can support research and guide follow-up experiments, but they do not independently confirm a proposed mechanism or cellular response in biological experiments.

Virtual-cell research has therefore expanded in several complementary directions. Mechanistic platforms and whole-cell models specify processes to simulate; single-cell foundation and spatial models characterize measured cellular states; perturbation models forecast responses; and agents help coordinate model development and testing. These approaches may be combined, but their value depends on whether they answer the biological question at hand and hold up under appropriate tests. Table 2 summarizes what representative methods model, what evidence supports their performance, and where validation remains limited. A key next step is to connect these capabilities through experiments that test predicted responses and, where mechanisms are claimed, the proposed explanations for them.

REPRESENTATIVE APPLICATION SCENARIOS OF VIRTUAL CELLS

Virtual cells are not confined to a single methodological domain or application field. Instead, their value lies in providing a general computational framework for interpreting, predicting, and engineering cellular behavior across multiple biological and biomedical contexts. As summarized in Figure 4, current and emerging applications of virtual cells can be broadly organized into five interconnected domains: basic cell biology, drug discovery and disease research, synthetic biology and minimal cell design, tissue microenvironment modeling, and personalized medicine with cellular digital twins. These application scenarios differ in their immediate goals, ranging from mechanistic explanation and perturbation prediction to design-oriented engineering and patient-specific decision support, but they share a common reliance on computational representations of cellular states, transitions, and responses.

Explaining cell states and their dynamic transformation

In basic research, virtual cells play a crucial role in providing quantitative explanations of cell states and their transformation processes^[50]. The VCell platform has long supported the construction of a complete lineage from compartmentalized models to three-dimensional reaction-diffusion models, enabling mechanistic numerical simulations of calcium signaling, cell polarization, membrane transport, and signal transduction networks^[50]. Subsequent versions (VCell 7.0, <https://vcell.org/vcell-7-0-with-kinematics-released>) have further introduced reaction-diffusion simulation capabilities under moving boundaries, providing new computational power for explaining the mechanisms of dynamic cell morphological changes such as cell migration and cell shape remodeling^[27]. Simultaneously, AI methods are expanding the scope of explanation from specific pathways to data-driven learning of regulatory relationships. CellProphet can infer dynamic

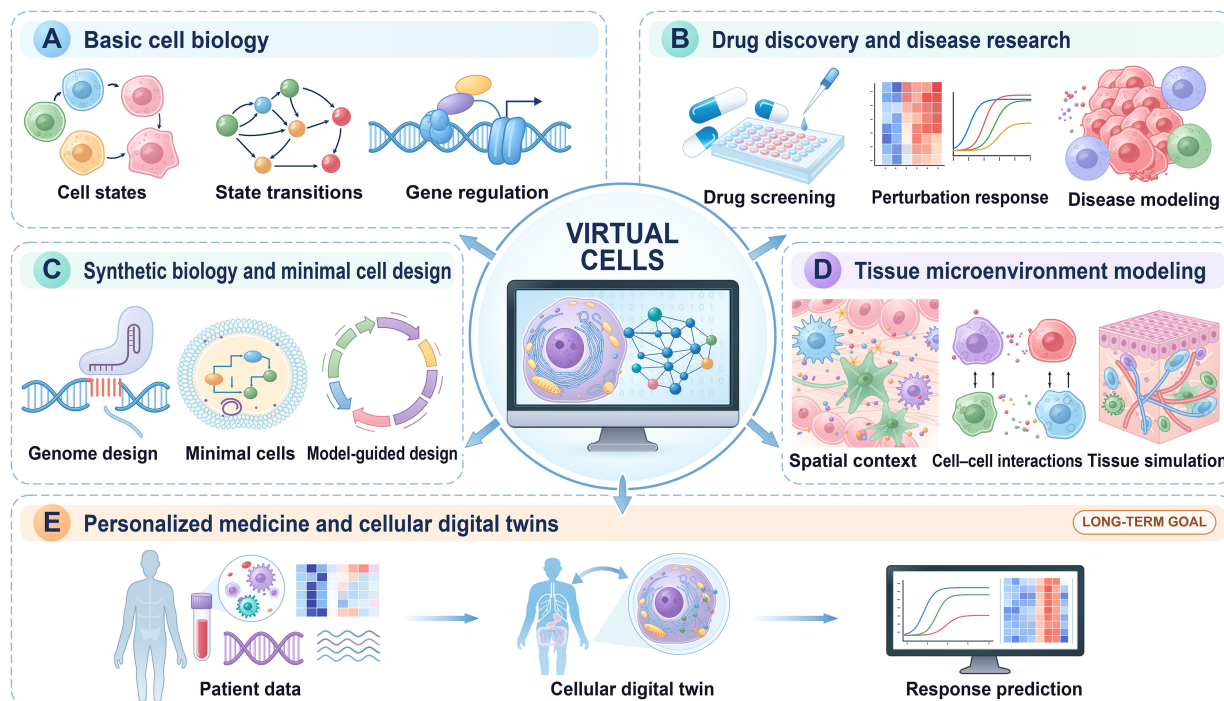


Figure 4. Major research applications and future goals of virtual cells. (A) Basic cell biology: modeling cell states, state transitions, and gene regulation; (B) Drug discovery and disease research: predicting cellular responses to perturbations, prioritizing candidates for experimental screening, and modeling disease-associated cellular states; (C) Synthetic biology and minimal cell design: supporting genome design, minimal-cell modeling, and model-guided engineering; (D) Tissue microenvironment modeling: incorporating spatial context and cell-cell interactions to simulate multicellular behavior; (E) Personalized medicine and cellular digital twins: a long-term goal of integrating patient-specific biological data into continuously updated models to predict cellular responses to candidate treatments.

gene regulatory networks from single-cell differentiation time-series data, thereby explaining how transcription factors regulate cell fate determination over time^[51]. Squidiff can reconstruct cell fate trajectories under developmental and interventional conditions^[52], while VCWorld combines structured knowledge with large language model reasoning to generate interpretable mechanistic hypotheses^[53]. These methods have further advanced the research focus from static state description to the computational explanation of state transition mechanisms^[54].

Drug development and disease research

Beyond helping to interpret biological systems, virtual-cell models can prioritize candidate interventions for experimental testing, extending their potential value in preclinical drug discovery^[55]. In drug development and disease research, a crucial task of virtual cells is predicting cellular responses after interventions such as drugs, cytokines, or gene editing^[55]. Recent large-scale AI-driven virtual cell models have converged on this direction, although they differ in emphasis. Xaira Therapeutics' X-Cell predicts transcriptomic responses to genetic perturbations across cellular contexts, including primary human T cells^[19]. Alibaba DAMO Academy's Lingshu-Cell takes the form of a generative cellular world model capable of simulating transcriptome states and predicting cellular responses to genetic perturbations or cytokine stimulation, with potential implications for drug development, disease mechanisms, and personalized medicine^[43]. XPert offers a more focused example in drug-perturbation prediction. By modeling responses across drug doses and time points, XPert captures transcriptional pharmacodynamic patterns and highlights genes associated with drug response. After pretraining on large-scale perturbation data, XPert improved prediction of treatment-associated transcriptomic changes in held-out breast cancer patients by 15.04% relative to training from scratch; its attention analysis also highlighted genes previously linked to drug resistance^[56]. PRiMeFlow uses flow matching to model heterogeneous single-cell responses to genetic and small-molecule perturbations, improving the recovery of differentially expressed genes in PerturBench, providing a new technical solution for cell state prediction in high-throughput drug screening^[57].

The aforementioned models all point to a trend: virtual cells are gradually evolving from academic demonstrations into actionable tools in preclinical research and drug discovery processes. The advantages of virtual cells are particularly pronounced for highly heterogeneous systems such as tumors, immune systems, and stem cells. The cell states in these systems are extremely diverse, making it difficult for traditional population averaging experiments to fully capture them, while virtual cell models naturally possess the ability to handle heterogeneity.

From an application perspective, AI agents are particularly relevant to virtual screening and perturbation prioritization. Instead of passively scoring a fixed candidate list, an agent can formulate a biological objective, query prior knowledge, select perturbations to test, call a virtual-cell model to estimate responses, analyze failures, and propose the next experimental round. This is consistent with the broader goal of *in silico* phenotypic screening, for which AssayBench has recently been proposed as an assay-level benchmark for evaluating LLMs and agents on CRISPR screen prediction across phenotypic endpoints^[58]. In future drug discovery pipelines, such agents may serve as a control layer that links virtual-cell prediction with wet-lab validation, while human experts remain responsible for biological plausibility, safety constraints, and final experimental decisions.

Synthetic biology and minimal cell design

The ability to design virtual cells has the most direct application value in the field of synthetic biology, namely, the ability to simulate a new cell system in a computer before physically realizing it. In 2022, Thornburg *et al.*'s dynamic simulation of the JCVI-syn3A minimal cell was able to capture interactions among metabolism, genetic information processes, and growth over the cell cycle, providing computational guidance for genome simplification^[21]. The JCVI-syn3A four-dimensional whole-cell model (4DWCM) takes this capability to a new level^[30]. This model accurately reproduces experimental measurements such as 105-minute doubling time, copy number ratio of replication origin to endpoint, mRNA lifetime, and ribosome number, and reveals spatial heterogeneity, random allocation, and metabolic-genetic coupling^[30].

AI methods are also entering the synthetic biology design process. Shcherbakova *et al.* proposed a proof-of-concept framework based on a variational autoencoder, using binary gene presence vectors to model the composition of the bacterial genome, thereby generating candidate reduced *E. coli* gene sets^[59]. In the same year, Gherman *et al.* trained a machine learning agent model to predict cell division with high accuracy by learning from data from the whole-cell model. They eventually designed a simplified *E. coli* cell with about 40% of the genes deleted from the whole-cell model and validated it using the whole-cell model itself^[60].

Cellular behavior prediction in tissue environments

The combination of spatial omics and AI methods is expanding the application of virtual cells from single-cell level prediction to perturbation response modeling in tissue microenvironments. Celcomen is one of the important advances in this direction^[44]. This model utilizes a mathematical causal framework and generative graph neural networks to decouple intracellular and intercellular gene regulatory programs from spatial transcriptome data, generating perturbed counterfactual spatial transcriptome atlases, providing researchers with samples that are difficult to obtain experimentally^[44]. Another important tool is BioInformatics WalkThrough (BIWT), a software that can directly initialize agent-based spatial models from single-cell and spatial molecular data, transforming static omics snapshots into evolvable multicellular dynamics simulations^[61]. Spatial context is represented differently in these two approaches. Celcomen constructs a graph of neighboring cells or spots from measured spatial coordinates and models gene interactions within

cells and across their neighborhoods. Its simulated perturbations yield testable spatial transcriptomic predictions: predicted responses to *Tgfbr2* and *Jak2* knockouts were compared with measured knockout lesions in mouse lung tumors^[44]. BIWT instead preserves observed cell locations and identities when initializing an agent-based tissue model. Its reported results show how alternative spatial initializations affect simulated tumor-immune dynamics and clustering; independent validation of prospective intervention outcomes was not reported^[61].

This application direction corresponds to a significant upgrade in the concept of virtual cells, moving from virtual cells to virtual cells in tissue. Future clinical applications, such as predicting the response of immune cells to immunotherapy in the tumor microenvironment and predicting the mobilization and differentiation pathways of stem cells after tissue damage, will inevitably require models to simultaneously understand both the intrinsic cellular state and the extrinsic cellular environment. As Celcomen demonstrated, models capable of generating counterfactual spatial transcriptome maps enable researchers to explore intervention scenarios that are not directly observable in real-world experiments due to ethical or technological limitations.

Personalized medicine and cell-level digital twins

Among all application scenarios, personalized medicine represents the most ambitious and long-term goal of virtual cells. Its ideal vision is to build a continuously updated computational model for each patient, integrating biological data such as genomics, transcriptomics, and proteomics to predict the patient-specific cellular responses to candidate drugs, thereby guiding optimal treatment strategies.

Bhardwaj *et al.* define virtual cells as digital cell twins, considering them a core paradigm for integrating multimodal biological data and achieving predictive and mechanistic understanding of cell function^[62]. Vallée further clarifies that the next generation of digital twins must adopt a multi-scale framework, covering the entire chain of data from molecules to cells, tissues, organs, and even clinical behavior and the environment^[23]. Lyu *et al.*'s review of organ digital twins systematically reviews the methodological framework for constructing digital twins of human organs, particularly emphasizing the crucial role of physical information AI in enhancing model fidelity, scalability, and personalization capabilities^[24].

In specific disease areas, conceptual exploration of digital twins has already begun in chimeric antigen receptor T cell (CAR-T) cell therapy. Related reviews define digital twins as a transformative approach that leverages high-dimensional patient data to provide virtual representations of the treatment-tumor trajectory and outline its development path in optimizing CAR-T therapy^[22]. Noetik's OCTO-virtual cell platform (<http://www.noetik.ai/aacr-2025>), showcased at American Association for Cancer Research (AACR) 2025, represents a more practical multimodal discovery engine, and was presented alongside Perturb-map, a separate high-throughput *in vivo* platform for therapeutic target identification and immune phenotype modeling.

However, it is crucial to recognize that true patient-level cell digital twins are still a significant distance away. Current virtual-cell studies draw on cell lines, model organisms, primary human cells, and patient-derived data, but integrating these sources into continuously updated models for individual patients remains challenging. While X-Cell has achieved breakthroughs in zero-shot generalization - generalizing to human primary T cells from multiple donors - this is still at the cell type level rather than the individual patient level. XPert's evaluation on patient transcriptomic data provides an initial example of transfer from preclinical perturbation datasets, while prospective testing would be needed to assess its use in treatment decisions. To truly realize personalized medical applications of digital cell twins, multiple engineering and scientific challenges need to be overcome, including data assimilation, real-time updates, cross-scale coupling, and clinical validation. Bhardwaj *et al.* also pointed out that computational scalability, parameter inference, and ethical considerations remain key bottlenecks in the development of digital cell twins^[62].

CONCLUSION AND OUTLOOK

In this review, we comprehensively traced the conceptual evolution of virtual cells from early mechanistic modeling frameworks to contemporary AI-driven paradigms, highlighting the transition from descriptive simulation toward predictive and perturbation-aware modeling. We summarized key technological advances, including the integration of single-cell and spatial omics, the emergence of large-scale foundation models and AI agents, and the parallel development of mechanistic and data-driven approaches. Together, these advances have significantly expanded the capability of virtual cells to model cellular states, predict responses to genetic and chemical perturbations, and support applications across basic research, drug discovery, synthetic biology, and personalized medicine.

Despite these encouraging developments, several fundamental challenges continue to limit the realization of fully functional virtual cells. Current models remain constrained by noisy and incomplete data, limited capacity for causal inference, and insufficient generalization across diverse biological contexts. In addition, the lack of standardized benchmarking frameworks and unified evaluation criteria hampers systematic comparison and validation of different modeling approaches. From an engineering perspective, the scalability of models, computational resource requirements, and difficulties in integrating multimodal and cross-scale data further restrict large-scale deployment and real-world applications. These challenges are interdependent and span data generation, model design, evaluation, computation, interpretability, and clinical translation. [Table 3](#) summarizes the major bottlenecks facing virtual cell research, their impact on model development and application, and potential directions for future improvement.

Looking forward, the advancement of virtual cells will require coordinated progress across multiple dimensions. First, improving data quality and diversity through large-scale, high-resolution, and perturbation-rich datasets will be essential for enhancing model robustness and generalizability. Second, methodological innovation is needed to bridge the gap between correlation-based learning and true causal inference, potentially through the integration of mechanistic knowledge with data-driven models. Third, the establishment of standardized benchmarks, open data ecosystems, and shared community infrastructure will play a critical role in accelerating reproducibility and fair evaluation. Finally, advances in computing infrastructure and model efficiency will be necessary to enable scalable, real-time applications.

Beyond these technical and infrastructural challenges, the rise of AI agents introduces new requirements for governance, provenance, and validation. Because agents can autonomously retrieve evidence, generate code, select experiments, and revise model designs, their outputs must be accompanied by transparent execution logs, versioned datasets, reproducible environments, and explicit records of human oversight. Future agentic virtual-cell systems should therefore combine tool-grounded retrieval, verifier agents, repository-level reproduction, prospective wet-lab validation, and benchmark tasks that measure not only prediction accuracy but also reasoning reliability and experimental usefulness.

In the long term, virtual cells are expected to evolve toward fully realized cellular digital twins that can continuously integrate multimodal data, adapt to individual-specific contexts, and support predictive and actionable decision-making in biomedical research and clinical practice. Achieving this vision will require not only technological breakthroughs but also interdisciplinary collaboration across biology, computer science, and medicine. Rather than serving merely as computational replicas of biological cells, virtual cells may ultimately become experimental partners that reshape how hypotheses are generated, tested, and translated. As the field moves from simulation to prediction and from prediction to intervention, virtual cells have the potential to redefine the future of biological discovery and precision medicine.

Table 3. Key challenges and priorities for virtual-cell research

Challenge	Current evidence gap	Consequence	Priority for progress
Incomplete and heterogeneous data ^[3,31,32]	Batch effects and measurement noise ^[31,32] ; incomplete multimodal and spatial measurements ^[3,31] ; perturbation datasets cover specific experimental systems ^[16,34]	Technical variation can obscure biological signals ^[31,32] ; missing measurements limit the representation of cellular states ^[3,31]	Standardize metadata, preprocessing and quality control ^[31] ; integrate multimodal and spatial data ^[3,31] ; expand perturbation datasets ^[3,16,34]
Limits of causal interpretation ^[3,44]	Accurate response prediction alone does not establish a mechanism ^[3,44] ; causal identifiability depends on stated model assumptions ^[44]	Mechanistic interpretation remains conditional on identifiability assumptions ^[44] and experimental support ^[3,10]	State causal assumptions and assess their applicability ^[44] ; validate candidate mechanisms through targeted perturbations ^[10,34] ; use biologically explicit mechanistic models as complementary frameworks ^[20,29]
Cross-context generalization ^[11,18,42]	Performance depends on held-out perturbations ^[11] and cellular contexts ^[18] ; zero-shot foundation models do not consistently outperform simpler methods ^[42]	Success in one evaluation setting does not establish transfer to other perturbations or contexts ^[11,18,42]	Use prespecified perturbation and context holdouts ^[11,18] ; report zero-shot and fine-tuned performance separately ^[42] ; quantify predictive uncertainty ^[3]
Fragmented benchmarks ^[11,17,42]	Evaluation tasks, data splits and metrics differ across studies ^[11,17,42] ; benchmark conclusions depend on the baselines included ^[11,42]	Model rankings and apparent gains can change with the evaluation setting ^[11,42]	Report held-out units and data splits ^[11,18] ; audit pretraining and test-data overlap ^[42] ; include simple baselines ^[11,42] and task-specific metrics ^[11,17,18]
Computational demands ^[11,23,41]	Large foundation models require substantial training resources ^[41] ; coupled multiscale simulations can be computationally expensive ^[23]	Runtime and GPU-memory requirements affect practical model comparison ^[11] ; simulation costs can constrain iterative calibration and real-time updating ^[23]	Report hardware, runtime and memory use ^[11] ; release reusable code ^[41] ; improve computational tractability and infrastructure for multiscale models ^[23]
Interpretability ^[3,10,44]	Learned representations can suggest regulatory candidates ^[10] , but biological interpretation requires experimental validation ^[3,10] and explicit modeling assumptions ^[44]	Model-derived regulatory hypotheses require biological testing before they support mechanistic claims ^[3,10]	Combine model-derived hypotheses with targeted experiments ^[10] ; disclose causal assumptions ^[44] ; use biologically explicit mechanistic models as complementary frameworks ^[20,29]
Clinical validation ^[22,23,55]	Transfer from preclinical drug screens to patient transcriptomes has been evaluated ^[56] , while large-scale prospective validation of fully multiscale patient models remains limited ^[23]	Preclinical predictive performance alone does not establish patient-level clinical utility ^[23,55] ; CAR-T digital twins remain an emerging translational approach ^[22]	Calibrate and update models using longitudinal patient data ^[22,23] ; test generalizability on independent prospective datasets ^[55] ; evaluate clinical claims in prospective studies ^[23]

GPU: Graphics processing unit; CAR-T: chimeric antigen receptor T cell.

DECLARATIONS

Authors' contributions

Made substantial contributions to the conception and design of the survey and led the writing of the original draft: Li, Y.

Performed literature investigation and contributed to the writing of the original draft: Xiong, W.; Feng, K.

Participated in literature investigation, provided critical resources, and assisted in writing review and editing: Ge, Y.; Wang, J.

Supervised the entire project, administered the process, reviewed and edited the manuscript: Wang, H.; Deng, F.; Ning, Y.

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

During the preparation of this manuscript, the AI tool ChatGPT (OpenAI; version GPT-5.6 Sol, released 2026-07-09) was used to assist with English-language editing and to explore preliminary layouts for [Figures 1-4](#) and the graphical abstract. The final Figures were redrawn by the authors in Adobe Illustrator. The tool

did not influence the study design, data collection, analysis, interpretation, or the scientific content of the work. All authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Financial support and sponsorship

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China (No. T2600524), the Postdoctoral Fellowship Program of CPSF (No. GZC20260806), and the National Science and Technology Major Project for Prevention and Control of Emerging and Major Infectious Diseases (No. 2025ZD01901800).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

Supplementary Materials

[Supplementary Materials](#)

REFERENCES

1. Cameron, D. E.; Bashor, C. J.; Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **2014**, *12*, 381-90. [DOI PubMed](#)
2. Kitano, H. Systems biology: a brief overview. *Science* **2002**, *295*, 1662-4. [DOI PubMed](#)
3. Bunne, C.; Roohani, Y.; Rosen, Y.; et al. How to build the virtual cell with artificial intelligence: priorities and opportunities. *Cell* **2024**, *187*, 7045-63. [DOI PubMed PMC](#)
4. Tomita, M. Whole-cell simulation: a grand challenge of the 21st century. *Trends. Biotechnol.* **2001**, *19*, 205-10. [DOI PubMed](#)
5. Bray, D. Protein molecules as computational elements in living cells. *Nature* **1995**, *376*, 307-12. [DOI PubMed](#)
6. Collins, F. S.; Varmus, H. A new initiative on precision medicine. *N. Engl. J. Med.* **2015**, *372*, 793-5. [DOI PubMed PMC](#)
7. Alon, U. Biological networks: the tinkerer as an engineer. *Science* **2003**, *301*, 1866-7. [DOI PubMed](#)
8. Ashley, E. A. Towards precision medicine. *Nat. Rev. Genet.* **2016**, *17*, 507-22. [DOI PubMed](#)
9. Cui, H.; Wang, C.; Maan, H.; et al. scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat. Methods.* **2024**, *21*, 1470-80. [DOI PubMed](#)
10. Theodoris, C. V.; Xiao, L.; Chopra, A.; et al. Transfer learning enables predictions in network biology. *Nature* **2023**, *618*, 616-24. [DOI PubMed PMC](#)
11. Ahlmann-Eltz, C.; Huber, W.; Anders, S. Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. *Nat. Methods.* **2025**, *22*, 1657-61. [DOI PubMed PMC](#)
12. Resasco, D. C.; Gao, F.; Morgan, F.; Novak, I. L.; Schaff, J. C.; Slepchenko, B. M. Virtual cell: computational tools for modeling in cell biology. *Wiley. Interdiscip. Rev. Syst. Biol. Med.* **2012**, *4*, 129-40. [DOI PubMed PMC](#)
13. Slepchenko, B. M.; Schaff, J. C.; Macara, I.; Loew, L. M. Quantitative cell biology with the Virtual Cell. *Trends. Cell. Biol.* **2003**, *13*, 570-6. [DOI PubMed](#)
14. Abdulla, S.; Aevertmann, B.; Assis, P.; et al. ; CZI Cell Science Program. CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *Nucleic. Acids. Res.* **2025**, *53*, D886-900. [DOI PubMed PMC](#)
15. Yao, Z.; van Velthoven, C. T. J.; Kunst, M.; et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **2023**, *624*, 317-32. [DOI PubMed PMC](#)
16. Replogle, J. M.; Saunders, R. A.; Pogson, A. N.; et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell* **2022**, *185*, 2559-75.e28. [DOI PubMed PMC](#)
17. Roohani, Y. H.; Hua, T. J.; Tung, P. Y.; et al. Virtual cell challenge: toward a Turing test for the virtual cell. *Cell* **2025**, *188*, 3370-4. [DOI PubMed](#)
18. Adduri, A. K.; Gautam, D.; Bevilacqua, B.; et al. Predicting cellular responses to perturbation across diverse contexts with State. *Cell* **2026**, *189*, 5914-31.e20. [DOI PubMed](#)

19. Wang, C.; Karimzadeh, M.; Ravindra, N. G.; et al. X-Cell: scaling causal perturbation prediction across diverse cellular contexts via diffusion language models. *bioRxiv* **2026**, Epub ahead of print. [DOI](#)
20. Karr, J. R.; Sanghvi, J. C.; Macklin, D. N.; et al. A whole-cell computational model predicts phenotype from genotype. *Cell* **2012**, *150*, 389-401. [DOI PubMed PMC](#)
21. Thornburg, Z. R.; Bianchi, D. M.; Brier, T. A.; et al. Fundamental behaviors emerge from simulations of a living minimal cell. *Cell* **2022**, *185*, 345-60.e28. [DOI PubMed PMC](#)
22. Aghamiri, S. S.; Amin, R. The potential use of digital twin technology for advancing CAR-T cell therapy. *Curr. Issues. Mol. Biol.* **2025**, *47*, 321. [DOI PubMed PMC](#)
23. Vallée, A. Multi-scale digital twins for personalized medicine. *Front. Digit. Health.* **2026**, *8*, 1753906. [DOI PubMed PMC](#)
24. Lyu, Y.; Li, Z.; Tran, V.; et al. Building digital twins of different human organs for personalized healthcare. *arXiv* **2026**, arXiv:2601.11318. Available online: <https://doi.org/10.48550/arXiv.2601.11318> (accessed 20 October, 2026).
25. Loew, L. M.; Schaff, J. C. The Virtual Cell: a software environment for computational cell biology. *Trends. Biotechnol.* **2001**, *19*, 401-6. [DOI PubMed](#)
26. Blinov, M. L.; Schaff, J. C.; Vasilescu, D.; Moraru, I. I.; Bloom, J. E.; Loew, L. M. Compartmental and spatial rule-based modeling with virtual cell. *Biophys. J.* **2017**, *113*, 1365-72. [DOI PubMed PMC](#)
27. Nickaen, M.; Novak, I. L.; Pulford, S.; et al. A free-boundary model of a motile cell explains turning behavior. *PLoS. Comput. Biol.* **2017**, *13*, e1005862. [DOI PubMed PMC](#)
28. Lee, S. H.; Myers, P. J.; Mendrzycki, M. C.; et al. Systems-level consequences of low RAF abundance for EGFR-ERK signaling. *Biophys. J.* **2026**, *125*, 881-900. [DOI PubMed PMC](#)
29. Macklin, D. N.; Ahn-Horst, T. A.; Choi, H.; et al. Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation. *Science* **2020**, *369*, eaav3751. [DOI PubMed PMC](#)
30. Thornburg, Z. R.; Maytin, A.; Kwon, J.; et al. Bringing the genetically minimal cell to life on a computer in 4D. *Cell* **2026**, *189*, 2582-97.e27. [DOI PubMed](#)
31. Heumos, L.; Schaar, A. C.; Lance, C.; et al. ; Single-cell Best Practices Consortium. Best practices for single-cell analysis across modalities. *Nat. Rev. Genet.* **2023**, *24*, 550-72. [DOI PubMed PMC](#)
32. Lähnemann, D.; Köster, J.; Szczurek, E.; et al. Eleven grand challenges in single-cell data science. *Genome. Biol.* **2020**, *21*, 31. [DOI PubMed PMC](#)
33. Regev, A.; Teichmann, S. A.; Lander, E. S.; et al. ; Human Cell Atlas Meeting Participants. The human cell atlas. *Elife* **2017**, *6*, e27041. [DOI PubMed PMC](#)
34. Dixit, A.; Parnas, O.; Li, B.; et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **2016**, *167*, 1853-66.e17. [DOI PubMed PMC](#)
35. Ståhl, P. L.; Salmén, F.; Vickovic, S.; et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **2016**, *353*, 78-82. [DOI PubMed](#)
36. Biancalani, T.; Scalia, G.; Buffoni, L.; et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods.* **2021**, *18*, 1352-62. [DOI PubMed PMC](#)
37. Hao, M.; Gong, J.; Zeng, X.; et al. Large-scale foundation model on single-cell transcriptomics. *Nat. Methods.* **2024**, *21*, 1481-91. [DOI PubMed](#)
38. Tejada-Lapuerta, A.; Schaar, A. C.; Gutgesell, R.; et al. Nicheformer: a foundation model for single-cell and spatial omics. *Nat. Methods.* **2025**, *22*, 2525-38. [DOI PubMed PMC](#)
39. Rosen, Y.; Roohani, Y.; Agrawal, A.; Samotorčan, L.; Quake, S. R.; Leskovec, J.; Tabula Sapiens Consortium. Universal cell embedding provides a foundation model for cell biology. *Nature* **2026**, *656*, 183-91. [DOI PubMed PMC](#)
40. Roohani, Y.; Huang, K.; Leskovec, J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat. Biotechnol.* **2024**, *42*, 927-35. [DOI PubMed PMC](#)
41. Zeng, Y.; Xie, J.; Shangguan, N.; et al. CellFM: a large-scale foundation model pre-trained on transcriptomics of 100 million human cells. *Nat. Commun.* **2025**, *16*, 4679. [DOI PubMed PMC](#)
42. Kedzierska, K. Z.; Crawford, L.; Amini, A. P.; Lu, A. X. Zero-shot evaluation reveals limitations of single-cell foundation models. *Genome. Biol.* **2025**, *26*, 101. [DOI PubMed PMC](#)
43. Zhang, H.; Yuan, G. H.; Yuan, C.; et al. Lingshu-Cell: a generative cellular world model for transcriptome modeling toward virtual cells. *arXiv* **2026**, arXiv:2603.25240. Available online: <https://doi.org/10.48550/arXiv.2603.25240> (accessed 20 October, 2026).
44. Megas, S.; Chen, D. G.; Polanski, K.; et al. Celcomen: spatial causal disentanglement for single-cell and tissue perturbation modeling. *Nat. Commun.* **2026**, *17*, 4126. [DOI PubMed PMC](#)

45. Roohani, Y.; Lee, A.; Huang, Q.; et al. BioDiscoveryAgent: an AI agent for designing genetic perturbation experiments. In *Proceedings of the International Conference on Learning Representations*, 2025; pp. 26417-66. https://proceedings.iclr.cc/paper_files/paper/2025/hash/4252dc94531833029000f85dc5fac792-Abstract-Conference.html. (accessed 2026-10-20).
46. Tang, X.; Yu, Z.; Chen, J.; et al. CellForge: agentic design of virtual cell models. *arXiv* **2026**, arXiv:2508.02276. Available online: <http://arxiv.org/abs/2508.02276v1> (accessed 20 October, 2026).
47. Cheng, X.; Li, P.; Guo, H.; et al. Harnessing AI to build virtual cells. *bioRxiv* **2026**, Epub ahead of print. DOI
48. Jang, Y.; Zhu, L.; Fawkes, J.; Denton, A. K.; Beaini, D.; Noutahi, E. Towards autonomous mechanistic reasoning in Virtual Cells. *arXiv* **2026**, arXiv:2604.11661. Available online: <https://doi.org/10.48550/arXiv.2604.11661> (accessed 20 October, 2026).
49. Xia, X.; Yao, C.; Tsoi, M.; et al. AblateCell: a reproduce-then-ablate agent for virtual cell repositories. *arXiv* **2026**, arXiv:2604.19606. Available online: <https://doi.org/10.48550/arXiv.2604.19606> (accessed 20 October, 2026).
50. Slepchenko, B. M.; Loew, L. M. Use of virtual cell in studies of cellular dynamics. *Int. Rev. Cell. Mol. Biol.* **2010**, *283*, 1-56. DOI PubMed PMC
51. Peng, R.; Gao, Y.; Chen, Y.; Wang, J. CellProphet: dissecting virtual cell differentiation through AI-powered dynamic gene regulatory network inference. *bioRxiv* **2025**, Epub ahead of print. DOI
52. He, S.; Zhu, Y.; Tavakoli, D. N.; et al. Squidiff: predicting cellular development and responses to perturbations using a diffusion model. *Nat. Methods.* **2026**, *23*, 65-77. DOI PubMed PMC
53. Wei, Z.; Ma, R.; Wang, Z.; Li, Z.; Song, S.; Zheng, S. VCWorld: a biological world model for virtual cell simulation. *arXiv* **2026**, arXiv:2512.00306. Available online: <https://doi.org/10.48550/arXiv.2512.00306> (accessed 20 October, 2026).
54. Rafelski, S. M.; Theriot, J. A. Establishing a conceptual framework for holistic cell states and state transitions. *Cell* **2024**, *187*, 2633-51. DOI PubMed
55. Ma, C.; Zhang, H.; Rao, Y.; et al. AI-driven virtual cell models in preclinical research: technical pathways, validation mechanisms, and clinical translation potential. *NPJ. Digit. Med.* **2025**, *9*, 25. DOI PubMed PMC
56. Guo, Y.; Zhang, H.; Hu, H.; et al. Modelling drug-induced cellular perturbation responses with a biologically informed dual-branch transformer. *Nat. Mach. Intell.* **2026**, *8*, 96-112. DOI
57. Yan, Z.; Wu, Y.; Ji, M. X.; et al. PRiMeFlow: capturing complex expression heterogeneity in perturbation response modelling. *arXiv* **2026**, arXiv:2604.13986. Available online: <https://doi.org/10.48550/arXiv.2604.13986> (accessed 20 October, 2026).
58. De Brouwer, E.; Edwards, C.; Wu, A.; et al. AssayBench: an assay-level virtual cell benchmark for LLMs and agents. *arXiv* **2026**, arXiv:2605.10876. Available online: <https://doi.org/10.48550/arXiv.2605.10876> (accessed 20 October, 2026).
59. Shcherbakova, A.; Sharma, K.; Chen, Y.; et al. Designing minimal *E. coli* genomes using variational autoencoders. *bioRxiv* **2025**, Epub ahead of print. DOI
60. Gherman, I. M.; Sharma, K.; Rees-Garbutt, J.; et al. Accelerated design of Escherichia coli reduced genomes using a whole-cell model and machine learning. *Cell. Syst.* **2025**, *16*, 101392. DOI PubMed
61. Bergman, D. R.; Johnson, J. A. I.; Naji, M.; et al. BIWT: a bioinformatics walkthrough for embedding spatial multiomics in agent-based models for virtual cells. *Bioinformatics* **2025**, *41*, btaf571. DOI PubMed PMC
62. Bhardwaj, S.; Kumar, G.; Yang, H.; et al. Virtual cells: from conceptual frameworks to biomedical applications. *arXiv* **2025**, arXiv:2509.18220. Available online: <https://doi.org/10.48550/arXiv.2509.18220> (accessed 20 October, 2026).

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.