

# The Computational Patient: Building Digital Twins That Actually Matter in Clinical Practice

OCT 10, 2025



Share

---

*Disclaimer: The views and opinions expressed in this essay are solely my own and do not reflect the views, opinions, or positions of my employer or any affiliated organizations*

## Abstract

Digital twins represent one of the most ambitious applications of computational medicine, promising patient-specific simulators capable of predicting therapeutic responses and guiding interventions before they occur in biological reality. This examines the technical foundations, clinical applications, and practical challenges of physiological digital twins in healthcare. Key topics include:

- Current state of organ-level and systems-level physiological modeling
- Integration architectures linking genomic data, cellular models, organ simulations, and real-time sensor streams
- Validation frameworks for establishing model fidelity and clinical utility
- Adaptive calibration methods for maintaining accuracy as patient physiology changes
- Safety and ethical considerations for closed-loop therapeutic recommendations
- Business models and reimbursement pathways for digital twin technologies
- Technical barriers preventing widespread clinical deployment

The analysis argues that while digital twins have demonstrated value in specific domains like cardiac electrophysiology and oncology treatment planning, achieving

the vision of comprehensive patient simulators requires fundamental advances in multiscale modeling, data integration, and validation methodology that remain years away from clinical readiness.

## Table of Contents

1. Introduction: The Seductive Promise of Simulation
2. What Actually Works Today: Digital Twins in Current Clinical Practice
3. The Multiscale Modeling Challenge: From Genes to Physiology
4. Data Integration and the Reality of Clinical Sensors
5. Validation and the Problem of Ground Truth
6. Closing the Loop: Adaptive Calibration When Models Diverge
7. Safety Frameworks for Automated Recommendations
8. The Business Case Nobody Wants to Discuss
9. Conclusion: Realistic Timelines and Achievable Milestones

---

## Introduction: The Seductive Promise of Simulation

Imagine a surgeon preparing for a complex aortic valve replacement who can first perform the procedure virtually on a patient-specific digital replica of the heart, complete with accurate geometry from imaging, individualized tissue properties from elastography, and hemodynamic parameters from catheterization data. The simulation predicts that the planned valve size will create paravalvular leak due to calcium deposits in an unexpected location, prompting a change in approach that prevents complication. Or consider an oncologist selecting chemotherapy regimen who consults a tumor model built from the patient's genomic profile, previous treatment responses, and current metabolic imaging, receiving quantitative predictions about which drug combinations will maximize tumor kill while minimizing toxicity based on

the patient's specific pharmacokinetic and pharmacodynamic parameters. These scenarios represent the aspirational vision of digital twin technology in healthcare: computational models so accurate and personalized that they become indispensable tools for clinical decision-making, allowing physicians to test interventions in silico before exposing patients to risk.

The concept of digital twins originated in manufacturing and aerospace, where virtual replicas of physical systems are used for design optimization, predictive maintenance, and failure analysis. General Electric popularized the term in the context of jet engines, creating computational models that mirror the state of individual engines in service and predict when components will fail before they actually do. The appeal of translating this concept to healthcare is obvious. Human bodies are vastly more complex than jet engines, but they are also systems that can be measured, modeled, and potentially predicted. If we can simulate fluid dynamics in turbines, why not blood flow through coronary arteries? If we can model stress and fatigue in metal structures, why not predict bone fractures or cartilage degradation? The fundamental physics and chemistry underlying biological systems are well understood, and computational power has increased exponentially while costs have plummeted. The pieces seem to be in place for a revolution in personalized medicine driven by patient-specific simulation.

Yet despite decades of research in computational physiology and billions of dollars invested in precision medicine, digital twins remain largely aspirational in clinical practice. A few narrow applications have achieved clinical adoption, but the comprehensive patient simulators imagined in futuristic demos are nowhere near reality. The gap between vision and implementation is not primarily due to insufficient computing power or inadequate physics engines. Rather, it reflects fundamental challenges in biological complexity, data availability, model validation, and the messy reality of integrating computational tools into clinical workflows designed around human decision-making. Understanding these challenges and identifying tractable paths forward requires moving beyond the hype of digital twin marketing materials to examine what actually works, what remains unsolved, and what barriers are truly fundamental versus merely difficult.

# What Actually Works Today: Digital Twin in Current Clinical Practice

The most successful clinical applications of digital twin technology are found in cardiac electrophysiology, where personalized models guide ablation procedures for atrial fibrillation and ventricular tachycardia. Companies like Volta Medical and Catheter Precision have developed systems that integrate patient-specific cardiac imaging with computational models of electrical propagation to identify optimal ablation targets. The process begins with high-resolution imaging, typically cardiac MRI or CT, that captures the three-dimensional geometry of the heart chambers. This imaging data is segmented to create a patient-specific mesh representing the cardiac anatomy. The model is then populated with electrical properties, including regions of scar tissue that create reentrant circuits responsible for arrhythmias. During the ablation procedure, the model can be further refined using electrical mapping data collected from catheters inside the heart, creating a computational representation that predicts which tissue regions, if ablated, will terminate the arrhythmia with minimal collateral damage.

These cardiac electrophysiology applications work because they benefit from several favorable circumstances that are not universally present in other clinical domains. First, the underlying physics of electrical propagation in cardiac tissue is relatively well understood and can be simulated with reasonable fidelity using established models like the bidomain equations or monodomain approximations. Second, high-quality imaging of cardiac anatomy is routinely performed as part of clinical care, providing the geometric substrate for patient-specific models. Third, the clinical problem has a clear endpoint that can be verified immediately during the procedure when the arrhythmia either terminates or does not. This tight feedback loop between model prediction and clinical outcome enables rapid validation and refinement. Fourth, the consequences of suboptimal ablation, while serious, are not immediately life-threatening in most cases, creating clinical tolerance for iterative approaches that gradually improve outcomes.

Oncology represents another domain where digital twin concepts have gained cl traction, particularly in radiation therapy planning. Systems like RayStation and Eclipse from Varian create patient-specific models of tumor geometry and surrounding normal tissue based on CT and MRI imaging. These models predict radiation dose distribution throughout the treatment volume, accounting for tis density variations, beam angles, and delivery parameters. The optimization algo search for treatment plans that maximize dose to tumor while respecting toleran limits for organs at risk like spinal cord, heart, or lungs. More sophisticated implementations incorporate biological models that predict tumor control proba and normal tissue complication probability based on the dose-volume histogram generated by different plans. Some experimental systems are beginning to integr functional imaging like PET scans that identify metabolically active tumor regio requiring higher doses and incorporate patient-specific radiosensitivity paramet estimated from previous responses or genomic data.

Beyond these established applications, numerous research groups have develope digital twin systems for other clinical domains that show promise but have not y achieved widespread adoption. Orthopedic surgery has seen development of biomechanical models for surgical planning, particularly in total joint replaceme where patient-specific models can predict implant positioning and stability. Pulmonary medicine has computational fluid dynamics models of airway mechar used for optimizing ventilator settings in acute respiratory distress syndrome. Neurosurgery has brain shift models that predict how tissue deformation during craniotomy affects the accuracy of image-guided navigation. Endocrinology has glucose-insulin dynamics models used in artificial pancreas systems for automat diabetes management. What these applications share is focus on specific, well-d physiological subsystems where the relevant physics is understood, imaging or s provides adequate input data, and clinical decision-making can be meaningfully informed by quantitative predictions.

The common thread in successful digital twin applications is that they solve trac subproblems rather than attempting comprehensive patient simulation. They fo single organ systems operating over time scales where physiological parameters

remain relatively stable. They make predictions about outcomes that can be verified relatively quickly, enabling model validation and refinement. They integrate into existing clinical workflows at points where quantitative analysis adds clear value to surgical planning or treatment optimization. The lesson from these successes is that near-term progress in digital twin technology will come from identifying similar favorable domains rather than pursuing the grand vision of whole-patient simulation that captures everything from molecular biology to organ-level physiology.

## **The Multiscale Modeling Challenge: From Genes to Physiology**

The aspiration of truly comprehensive digital twins requires linking models across biological scales, from genomic variants that influence protein function to cellular behaviors that determine tissue properties to organ-level physiology that manifest as clinical phenotypes. This multiscale integration is conceptually straightforward but practically nightmarish. Consider attempting to predict cardiovascular response to a beta blocker in a specific patient. At the genomic level, variants in the *ADRB1* gene encoding beta-1 adrenergic receptors influence receptor density and signaling efficiency. At the protein level, the beta blocker binds to these receptors with an affinity determined by molecular structure and pharmacokinetics influenced by hepatic metabolism genes like *CYP2D6*. At the cellular level, reduced beta-adrenergic signaling decreases cyclic AMP production in cardiomyocytes, affecting calcium handling and contractile force. At the tissue level, reduced myocyte contractility decreases ventricular ejection fraction, while effects on sinoatrial node cells reduce heart rate. At the organ level, these changes manifest as reduced cardiac output, which triggers compensatory mechanisms through baroreceptor reflexes and neurohormonal activation. At the whole-patient level, the net effect depends on baseline hemodynamics, volume status, comorbidities, and concurrent medications.

Building a computational model that accurately captures this cascade requires linking fundamentally different types of models operating at different spatial and temporal scales. Molecular dynamics simulations that predict drug-receptor binding operate on nanosecond time scales and nanometer length scales. Cellular electrophysiology

models like Hodgkin-Huxley type equations operate at millisecond to second time scales and micrometer to millimeter scales. Organ-level finite element models of cardiac mechanics operate at heartbeat time scales and centimeter scales. Whole pharmacokinetic models operate at minutes to hours and represent entire organ compartments. These models use different mathematical formalisms, require different types of input data, and have different computational requirements. Coupling them into a coherent multiscale simulation that maintains numerical stability and biological fidelity is an active area of research with many unsolved problems.

The most common approach to multiscale modeling is hierarchical coupling, where lower-scale models inform parameters used in higher-scale models rather than running simultaneously. For example, molecular simulations might determine binding constants that become fixed parameters in cellular models, which in turn determine tissue properties used in organ models. This approach avoids the computational expense of running all scales simultaneously but loses the ability to capture feedback between scales. In biological reality, changes at the organ level like reduced cardiac output trigger cellular responses through mechanotransduction and neurohormonal signaling that then influence molecular processes. Capturing these bidirectional interactions requires more sophisticated coupling strategies that remain computationally challenging and mathematically complex.

Another approach is multiscale agent-based modeling, where biological entities at different scales are represented as autonomous agents following local rules that give rise to emergent system-level behavior. This approach has been applied successfully to modeling tumor growth, where individual cells are agents with behaviors determined by local chemical gradients, mechanical forces, and genetic state, and macroscopic tumor properties emerge from the collective behavior of millions of cells. Agent-based models are conceptually intuitive and can capture complex interactions, but they are computationally expensive and require careful parameterization to ensure emergent behaviors match biological reality. They also struggle with the validation problem: it becomes difficult to determine whether specific parameter choices that produce realistic system-level behavior do so for the right reasons or through fortuitous cancellation of errors.

The practical reality is that comprehensive multiscale digital twins spanning ger physiology remain research projects rather than clinical tools. The models that have achieved clinical utility typically operate at a single scale or use highly simplified representations of other scales. The cardiac electrophysiology models mentioned earlier might incorporate patient-specific anatomy and tissue properties but use generic cellular models not personalized to the patient's genomics or comorbidities. The oncology models predict radiation dose distribution accurately but use simple biological response models that do not account for tumor heterogeneity, immune system interactions, or microenvironmental factors. These simplifications are necessary to make the problems tractable but also limit model accuracy and generalizability. The gap between what is theoretically desirable and what is practically achievable remains large, and closing it will require not just incremental improvements but fundamental advances in modeling methodology and computational efficiency.

## **Data Integration and the Reality of Clinical Sensors**

Even with perfect models, digital twins are only as good as the data used to parameterize and drive them. The vision of continuously updated patient models responding in real-time to sensor feeds from wearables, implantables, and clinic monitoring systems crashes against the messy reality of healthcare data. Clinical data is fragmented across systems, recorded in inconsistent formats, contaminated with measurement error and artifacts, and often unavailable when and where it is needed. Building digital twins that integrate genomic sequencing, electronic health records, medical imaging, laboratory tests, continuous vital sign monitoring, and patient-reported outcomes requires solving data engineering problems that are at least as difficult as the modeling challenges.

Genomic data represents the most stable input for digital twins, as germline variants do not change over time and can be measured once to inform lifetime predictions. Whole genome sequencing costs have dropped below one thousand dollars, and clinical exome sequencing is routinely performed for many conditions. However,

interpreting genomic variants to predict physiological parameters remains extremely challenging for most genes. For a small number of pharmacogenes like CYP2D6, CYP2C19, and TPMT, we have established relationships between genotype and metabolism that can inform dosing. For most genes, variant effects on protein function are unknown, and even when effects are known, translating molecular changes to organ-level physiology involves countless unmeasured intermediate variables. Polygenic risk scores that aggregate effects across many variants show modest predictive value for some conditions but are not yet accurate enough to parameterize patient-specific physiological models with confidence.

Medical imaging provides rich anatomical and some functional data but is acquired episodically rather than continuously. A cardiac MRI might be performed once to establish baseline anatomy for a digital twin, but cardiac remodeling occurs over months to years, and the model will gradually become outdated unless updated with new imaging. For acute conditions, this temporal mismatch is less problematic, but for chronic disease management where digital twins could potentially add most value, the static nature of anatomical imaging limits utility. Functional imaging like PET or SPECT provides metabolic or perfusion data that could parameterize physiological models but is expensive, involves radiation exposure, and is performed sparingly. The dream of continuous imaging for model updating remains unrealistic except for specific research settings.

Continuous physiological monitoring from wearables and hospital telemetry provides time-series data that could drive real-time digital twin updates. Consumer devices measure heart rate, activity, and increasingly blood oxygen and heart rhythm. Hospital monitors track vital signs with higher fidelity, including continuous arterial blood pressure, central venous pressure, and cardiac output from invasive monitoring. However, this data is notoriously noisy, contaminated by motion artifacts, sensor malfunction, and physiological irrelevancies like postural changes or eating. Using raw sensor streams to update physiological models requires sophisticated signal processing to extract meaningful physiological variation from noise, a problem that becomes more difficult when models need to run in real-time and make predictions rapidly.

The interoperability problem remains a massive barrier to data integration for digital twins. Even within a single hospital system, data from electronic health records, imaging PACS systems, laboratory information systems, and physiological monitoring platforms reside in separate databases with limited integration. Extracting comprehensive patient data for input to a digital twin requires writing custom integration code for each institution's specific system configuration, a burden that does not scale. National interoperability standards like FHIR help but are implemented inconsistently, and many critical data types lack standardized representations. For digital twins to become practical clinical tools rather than research curiosities, the data integration problem must be solved at the infrastructure level, not repeatedly solved bespoke for each implementation.

Even when data is available and integrated, determining which data is relevant and how to weight different sources when they provide conflicting information requires clinical and statistical sophistication. If a patient's blood pressure measured by a wearable differs from values measured during a clinic visit which were measured by an automated cuff that was later found to be poorly calibrated, which value should be used to parameterize the cardiovascular model? If imaging suggests kidney function is preserved but creatinine is rising, how should the renal component of the digital twin be configured? These data reconciliation problems that human clinicians navigate intuitively through experience and judgment must be codified algorithmically for digital twins, a challenge that requires extensive clinical knowledge engineering.

## **Validation and the Problem of Ground Truth**

The fundamental challenge in validating digital twin models is establishing ground truth against which predictions can be compared. Unlike engineering applications where sensor measurements provide unambiguous system state information, biological systems are only partially observable. We cannot directly measure the myocardial contractility or peripheral vascular resistance that a cardiovascular model predicts. We observe downstream consequences like blood pressure and cardiac output, but these measured variables are influenced by multiple physiological

parameters simultaneously, making it impossible to uniquely validate individual model components. This identifiability problem plagues physiological modeling and makes rigorous validation extremely difficult.

Forward validation, where models trained on one cohort are tested on independent data, provides some assurance of generalizability but does not guarantee physiological accuracy. A model might predict blood pressure accurately by fortuitously balancing errors in cardiac output and vascular resistance, producing the right answer for wrong reasons. When such models are used for counterfactual predictions, like estimating what would happen if a drug were added, the fortuitous error cancellation no longer holds, and predictions become unreliable. Distinguishing models that are truly physiologically accurate from those that merely fit observed data requires validation against perturbations or interventions that were not used in model development.

Prospective clinical trials where model predictions are compared against actual outcomes provide the strongest validation but are expensive and time-consuming; a trial validating a digital twin system for surgical planning might randomize patients to receive surgery planned with or without model guidance and compare complication rates. Such trials require large sample sizes to detect meaningful differences, must carefully control confounding variables, and take years to complete. The slow pace of prospective validation creates a tension between the desire for rigorous evidence and the practical need to deploy promising technologies quickly. This tension is familiar from the broader debate about evidence standards for digital health tools, where some argue that traditional randomized trial evidence should be required before clinical deployment while others contend that observational data and continuous monitoring can provide adequate safety assurance with faster iteration.

Continuous validation after deployment represents a middle path where digital twins are monitored during routine clinical use to detect performance degradation or systematic errors. This requires establishing baseline performance metrics, implementing automated systems to calculate these metrics continuously as new patients are treated, and defining thresholds that trigger alerts when performance deviates from expectations. The challenge is determining appropriate metrics and

thresholds. Should a cardiac surgery planning model be flagged if complications exceed historical averages by five percent, ten percent, or some other threshold? How many cases must be accumulated before conclusions about performance can be made with statistical confidence? How do we account for case mix variations that might make current patients systematically different from historical controls? These questions require careful statistical design and domain expertise to answer appropriately.

The validation problem is compounded when models make predictions about counterfactuals that cannot be directly observed. If a digital twin predicts that a patient would have better outcomes with treatment A than treatment B, and the clinician chooses treatment A based on this prediction, we never observe what would have happened with treatment B, so the prediction cannot be validated. Over large populations, if treatment assignments are randomized or if natural variation in treatment choices provides quasi-experimental variation, counterfactual predictions can be validated in aggregate. But for individual patients, the predictions that are most clinically valuable, those that guide personalized treatment selection, are exactly the ones that are hardest to validate because choosing one treatment precludes observing outcomes from alternatives.

Some researchers have proposed using mechanistic validation, where individual components of multiscale models are validated against experimental data from in vitro systems, animal models, or targeted clinical measurements, with the assumption that a model built from validated components will produce valid system-level predictions. This approach is appealing but rests on the questionable assumption that errors in component models will not interact and amplify when components are coupled. In complex nonlinear systems like human physiology, small errors in multiple components can combine to produce large system-level errors through positive feedback or resonance effects. Mechanistic validation provides confidence but not certainty, and system-level validation remains necessary.

## **Closing the Loop: Adaptive Calibration When Models Diverge**

Even validated digital twin models will inevitably diverge from patient reality over time as physiology changes due to disease progression, treatment effects, aging, and lifestyle factors. Maintaining model accuracy requires adaptive calibration that updates model parameters based on new observations while respecting physiological constraints and avoiding overfitting to noisy data. This is conceptually similar to assimilation problems in weather forecasting, where models are continuously updated based on new observations, but the biological context introduces unique challenges.

The simplest calibration approach is periodic recalibration, where the model is fit to all available data at regular intervals using standard parameter estimation techniques. This works when physiological changes occur slowly relative to calibration frequency and when sufficient data is available to estimate parameters reliably. However, many physiological parameters cannot be directly measured and must be inferred from observable outputs through inverse problems that are often ill-posed, meaning multiple parameter combinations can produce similar outputs. Regularization techniques that constrain parameters to physiologically plausible ranges help but require careful tuning and domain expertise.

Sequential data assimilation methods like Kalman filtering or particle filtering provide more sophisticated approaches that update model state and parameters continuously as new observations arrive. These methods propagate uncertainty through the model, explicitly accounting for measurement noise and modeling errors. Ensemble Kalman filters, which run multiple model instances with perturbed parameters and weight them by likelihood, have been applied successfully in cardiovascular modeling to estimate cardiac output and vascular parameters from noisy pressure waveforms. However, these methods are computationally expensive, require careful specification of observation and model error covariances, and can fail in highly nonlinear systems where the Gaussian assumptions underlying Kalman filtering break down.

Machine learning approaches to model calibration have gained attention recently, particularly using neural networks to learn mapping from observations to model parameters. These methods can handle complex nonlinear relationships and avoid expensive iterative optimization, but they require large training datasets of paired

observations and true parameters, which are difficult to obtain for biological systems where true parameters are often unknown. Hybrid approaches that combine physics-based models with learned calibration functions represent a promising direction that remain largely in the research phase.

A critical consideration in adaptive calibration is distinguishing true physiological changes requiring model updates from measurement artifacts or temporary physiological perturbations that should not influence long-term model parameters. For example, if a patient's blood pressure transiently spikes due to pain or anxiety, the cardiovascular model should not be recalibrated to match this spike, as doing so would corrupt baseline parameter estimates. Robust calibration methods must include outlier detection and separate transient from persistent changes, requiring longitudinal analysis and temporal modeling.

The frequency and triggers for recalibration also require careful consideration. Continuous recalibration on every new data point is computationally expensive and can lead to instability if the model constantly chases noisy measurements. Periodic recalibration on fixed schedules may miss important physiological changes between updates. Event-triggered recalibration, where model updates occur only when accumulated prediction errors exceed thresholds or when significant clinical events like hospitalizations or treatment changes occur, provides a middle ground but requires defining appropriate triggers and thresholds for each application.

## **Safety Frameworks for Automated Recommendations**

Perhaps the most challenging aspect of deploying digital twins in clinical practice is ensuring safety when models generate automated treatment recommendations. Unlike decision support systems that provide information for human interpretation, autonomous or semi-autonomous digital twin systems might directly adjust therapy, as in artificial pancreas systems that titrate insulin based on glucose models, or they might generate specific treatment recommendations that clinicians are expected to follow without deep independent analysis. The potential for model errors to directly cause

patient harm creates stringent safety requirements that go far beyond accuracy metrics.

Safety frameworks for model-guided therapy must address several distinct failure modes. First, systematic errors where the model consistently makes incorrect predictions for certain patient subpopulations could lead to widespread harm before being detected. Stratified validation across demographic groups and clinical subpopulations helps identify such biases, but rare subpopulations may have insufficient data for reliable validation, creating unavoidable uncertainty about model performance in edge cases. Second, individual errors where the model makes incorrect predictions for specific patients due to unmeasured confounders, unusual physiology, or data quality issues could cause serious harm even if overall model performance is good. Detecting and mitigating individual errors requires outlier detection, uncertainty quantification, and human oversight mechanisms. Third, distribution shift where the patient population or clinical context changes over time, causing previously validated models to become inaccurate, requires continuous monitoring and adaptation as discussed earlier.

Uncertainty quantification represents a critical safety mechanism, as models should not just make predictions but also provide confidence estimates indicating when predictions are unreliable. Bayesian approaches to modeling naturally produce posterior distributions over predictions that quantify uncertainty due to parameter uncertainty and model structure limitations. Ensemble methods that run multiple model instances with different parameter choices or structural assumptions can estimate prediction variance. Conformal prediction methods provide distributional uncertainty estimates with finite-sample guarantees. Regardless of method, the key requirement is that uncertainty estimates be well-calibrated, meaning stated confidence levels match empirical coverage probabilities. A model that claims ninety-five percent confidence but is actually correct only eighty percent of the time provides false assurance and is dangerous.

Human oversight mechanisms must be carefully designed to be effective without being burdensome. Simply requiring human approval for all model recommendations defeats the efficiency purpose of automation and can lead to alert fatigue where

humans rubber-stamp recommendations without careful review. More sophisticated approaches use risk stratification to direct human attention to high-risk cases while allowing automated implementation in low-risk situations. For example, an artificial pancreas might autonomously deliver small insulin adjustments in response to gradual glucose changes but alert the user before delivering large boluses or when glucose trajectories are rapidly changing. Defining appropriate risk thresholds requires balancing safety against usability and depends on patient preferences, clinical context, and model uncertainty.

Prospective safety trials for model-guided therapy face ethical challenges when the model is hypothesized to improve outcomes. Randomizing patients to receive instead of standard care to validate the model may be ethically problematic if preliminary evidence strongly suggests benefit. Adaptive trial designs that preferentially assign patients to better-performing arms as evidence accumulates can help, but they complicate statistical analysis and may still expose some patients to inferior treatment. Observational studies comparing outcomes before and after model deployment avoid randomization ethics but are confounded by temporal trends and selection effects. There are no easy answers, and safety evaluation methodology for medical AI remains an active area of methodological development.

Regulatory oversight for digital twin systems that provide treatment recommendations varies by jurisdiction and clinical application. The FDA considers such systems as medical devices requiring premarket review, with classification depending on risk level. Systems that provide treatment recommendations without human review are typically Class III devices requiring premarket approval with clinical trials. Systems that provide recommendations subject to human review may qualify for less stringent Class II clearance. However, existing device regulations designed for static products and fit awkwardly with continuously learning models that evolve after deployment. The FDA has proposed adaptive frameworks for software-as-a-medical-device that would allow modifications without full recertification if developers demonstrate adequate quality systems and monitoring, but these frameworks are still being refined and have not yet been applied to autonomous therapeutic systems.

# The Business Case Nobody Wants to Discuss

Despite technical progress and clinical promise, digital twin technology faces significant commercial challenges that receive insufficient attention in the research literature and conference presentations. Building comprehensive digital twin systems requires enormous investment in modeling, validation, data integration, and clinical workflow engineering. The companies pursuing this vision face long development timelines, uncertain regulatory pathways, unclear reimbursement, and difficult questions about value capture. Understanding these commercial realities is essential for entrepreneurs and investors evaluating opportunities in this space and for researchers whose work will only achieve impact if it finds sustainable paths to deployment.

The cost structure of digital twin development is heavily front-loaded. Developing high-fidelity physiological models requires teams of domain experts, modelers, and software engineers working for years to build, validate, and refine the models. Integrating these models with clinical data streams requires extensive engineering to handle multiple data formats and systems. Validating models requires clinical partnerships, data access, and often prospective studies that are expensive and time-consuming. Once these upfront costs are incurred, marginal costs per patient are relatively low, as computational costs have declined substantially and models can be deployed at scale. This cost structure favors large markets where development costs can be amortized over many patients but makes niche applications economically challenging.

Reimbursement for digital twin technology remains unclear in most applications. Payers are accustomed to reimbursing for procedures, drugs, and devices but lack established pathways for paying for computational models or simulation services. Some digital twin applications can be bundled with existing reimbursed services. Cardiac electrophysiology mapping systems are reimbursed as part of the ablation procedure CPT code. Radiation therapy planning systems are embedded in the radiation oncology payment bundle. But standalone digital twin services that pro-

treatment planning or optimization separate from procedures face reimbursement challenges. Creating new CPT codes requires demonstrating value to the AMA (AMA Editorial Panel), a slow process that requires evidence of clinical adoption and outcomes improvement. Value-based contracts that tie payment to outcomes rather than services provide alternative reimbursement paths but require payers willing to enter novel contracting arrangements and sophisticated outcomes measurement infrastructure.

The value proposition for digital twins must be compelling enough to justify the cost, which requires demonstrating either improved outcomes or reduced costs compared to standard care. Outcome improvements must be substantial enough to matter clinically and statistically significant enough to convince skeptical payers and clinicians. Cost reductions are often easier to demonstrate but may face resistance from stakeholders whose revenue depends on current high-cost care patterns. A digital twin system that reduces surgical complications sounds like clear value creation, but hospitals and surgeons who derive revenue from treating complications may not enthusiastically adopt technology that reduces their income unless reimbursement models change to reward value rather than volume.

Market access and distribution channels present additional challenges. Healthcare delivery organizations are the ultimate customers for digital twin technology, but sales cycles are long, purchasing decisions are complex and involve multiple stakeholders, and implementation requires significant integration work. Academic medical centers are early adopters willing to experiment with novel technology but represent a small fraction of total care delivery. Community hospitals and physician practices where most care occurs are more conservative and have limited technical capacity for implementing sophisticated computational tools. Building distribution channels that reach beyond academic early adopters requires either building direct sales forces, which are expensive and slow to scale, or partnering with established medical device or health IT vendors that have existing relationships but may lack expertise in computational modeling.

Intellectual property strategy for digital twins is complex because the technology combines physiological models based on published science, software implementa-

that may be patentable but face enablement challenges, and training data that is proprietary but is not directly protectable. Patents on mathematical models or algorithms are difficult to enforce and face patent eligibility challenges under recent USPTO guidance. Trade secret protection for model implementations requires limiting access in ways that may conflict with clinical transparency requirements. Some companies have pursued data network effects strategies where proprietary clinical datasets used for model training and validation create competitive moats; data access agreements and privacy regulations constrain data accumulation strategies.

The most successful commercial models for digital twins to date have been those that integrate tightly with existing workflows and revenue streams rather than attempt to create entirely new categories. Cardiac mapping systems generate revenue through capital equipment sales and disposable catheter sales rather than charging per simulation. Radiation therapy planning systems are sold as modules within comprehensive oncology IT platforms. Digital twin startups that have struggled are often those attempting to create standalone services without clear integration points into existing care delivery and payment models. The lesson for entrepreneurs is that business model innovation is as important as technical innovation, and commercial viability depends on navigating the complex political economy of healthcare delivery.

## **Conclusion: Realistic Timelines and Achievable Milestones**

Digital twins in healthcare occupy a peculiar position in the technology hype cycle. The concept has moved beyond the peak of inflated expectations, where breathless futurism dominates, but has not yet reached the plateau of productivity where the technology delivers consistent value at scale. We are somewhere in the trough of disillusionment, where initial enthusiasm has been tempered by reality but genuine progress continues in specific domains. Understanding where we are and what realistic progress looks like over the next decade requires separating science fiction from achievable engineering.

Over the next five years, we should expect incremental expansion of digital twin applications in domains where the technology already demonstrates value. Cardiac electrophysiology mapping will become more sophisticated, incorporating better cellular models, more detailed anatomical reconstruction, and tighter integration with robotic ablation systems. Oncology treatment planning will expand beyond radiation therapy to include surgical planning and systemic therapy selection, though the models will remain relatively simple and focused on specific cancer types where substantial clinical data exists. Orthopedic surgical planning will achieve broader adoption as imaging costs decline and biomechanical models improve. Diabetes management using glucose-insulin models will become more autonomous as continuous glucose monitors and insulin pumps become more capable and integrated.

What we should not expect in the next five years is comprehensive whole-patient digital twins that integrate multiscale models spanning genomics to physiology for complex multi-organ disease management. The technical barriers to multiscale integration, data availability limitations, validation challenges, and safety concerns are too substantial to be solved quickly. Research demonstrations will continue to impress, but clinical deployment of truly comprehensive digital twins remains a decade away at minimum. Claims from companies or research groups suggesting otherwise should be viewed with skepticism and evaluated critically for what is actually being delivered versus what is being promised.

The path forward requires focusing on tractable subproblems that deliver clear clinical value, building validation evidence rigorously through prospective studies, developing business models that align with healthcare payment realities, and investing in unsexy but essential data integration infrastructure. It requires honest acknowledgment of what we do not know about human physiology and limitations of current modeling approaches. It requires collaboration between modelers, clinicians, data scientists, software engineers, and regulatory experts rather than siloed work within disciplines. Most importantly, it requires patience and realism about time, as the transformation of medicine through digital twins will be gradual and uneven rather than sudden and universal.

For health tech entrepreneurs and investors, the opportunity in digital twins is real but requires careful evaluation of specific applications and business models. The technical feasibility must be rigorously assessed rather than assumed based on compelling demos. The clinical value proposition must be validated with evidence rather than anecdote. The commercial path to market must be clearly articulated with realistic assumptions about reimbursement, adoption timelines, and competitive dynamics. Companies that meet these criteria may find substantial opportunities in narrow domains where digital twin technology is mature enough for clinical deployment. Those chasing the broader vision of comprehensive patient simulation should understand they are making long-term bets on fundamental research advances that may or may not materialize on relevant timelines.

The promise of digital twins, of having computational replicas of patients that enable testing interventions before implementing them in biological reality, remains compelling and worth pursuing. But realizing this promise requires moving beyond hype to engage seriously with hard technical problems, clinical realities, regulatory requirements, and economic constraints. The companies and researchers who succeed will be those who balance ambition with realism, who celebrate incremental progress while maintaining focus on ultimate goals, and who build sustainable businesses and research programs capable of weathering the long journey from concept to clinical impact.



2 Likes • 1 Restack

[← Previous](#)

[Next](#)

## Discussion about this post

**Comments**

Restacks



Write a comment...

© 2026 Thoughts on Healthcare · [Privacy](#) · [Terms](#) · [Collection notice](#)  
[Substack](#) is the home for great culture