

Digital Twin-Enabled Personalized Cardiovascular Risk Assessment Using Wearable Sensor Data

Stefano R. Franklin

Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.
srfranklin@buffalo.edu

Rjay Bhatia

School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
ajayb@uc.edu

Abstract

Cardiovascular risk assessment has traditionally depended on population-derived statistical models that summarize risk factors over fixed time horizons. Wearable sensors now generate continuous physiological data streams, creating the possibility of a digital twin that reflects individual cardiovascular state and its temporal evolution. This article examines the system-level design of digital twin-enabled personalized cardiovascular risk assessment. It discusses the limitations of conventional risk scoring, the requirements of wearable data ingestion and signal processing, and the architectural choices that shape real-time and longitudinal model updating. The analysis extends beyond algorithmic performance to consider data governance, privacy, federated learning, clinical validation, robustness, fairness, and regulatory policy. A central argument is that personalized cardiovascular digital twins must be treated as socio-technical infrastructures whose value depends on organizational coordination, data quality, computational sustainability, and equity. The article identifies structural trade-offs between personalization and generalizability, responsiveness and resource consumption, and innovation and patient safety. It closes by outlining research directions that integrate systems engineering, clinical evaluation, and policy design.

Keywords

digital twin; cardiovascular risk assessment; wearable sensors; data governance; machine learning; fairness; regulatory policy.

1. Introduction

Cardiovascular diseases continue to account for substantial global morbidity and mortality, and risk stratification remains central to prevention. Conventional approaches rely on population-level equations that integrate a limited set of clinical and laboratory variables, producing probabilistic estimates over multiyear horizons. These tools have been valuable for public health and clinical decision-making, but they are inherently static and cannot fully represent the dynamic interplay of autonomic regulation, physical activity, medication effects, acute stress, and subclinical deterioration. The emergence of wearable sensors capable of continuous or near-continuous photoplethysmography, electrocardiography, accelerometry, and temperature monitoring has therefore generated significant interest in more personalized and temporally resolved cardiovascular risk assessment. High-performance medicine increasingly emphasizes the convergence of longitudinal data, artificial intelligence, and clinical insight [1]. Within this landscape, the digital twin has been proposed as a mechanism

for integrating multiscale physiological data into a dynamic representation of an individual patient, potentially enabling earlier detection and more precise intervention [2].

The transition from population risk scores to digital twin-enabled assessment is not merely an extension of existing predictive modeling. It involves a fundamental shift in how clinical knowledge is generated, validated, and maintained. Digital twins in health care raise ethical and philosophical questions about the relationship between the simulated representation and the patient, the boundaries of identity, and the responsibility of clinicians when acting on model outputs [3]. At the same time, the increasing availability of consumer wearable devices has accelerated the medical revolution in which individuals generate data outside clinical settings, creating both opportunities for continuous monitoring and risks related to data quality, interpretation, and equity [4]. This paper therefore adopts a system-level perspective, focusing on infrastructure, architecture, governance, and deployment rather than on specific model equations or algorithmic benchmarks.

The remainder of the article examines the structural foundations of digital twin-enabled cardiovascular risk assessment. It first contrasts population-based and individualized modeling paradigms. It then discusses wearable data infrastructure and signal quality, followed by architectural considerations for digital twin systems. Subsequent sections analyze data governance, clinical validation, fairness, deployment sustainability, and policy. Throughout, the analysis emphasizes that successful implementation depends on the alignment of computational systems with clinical workflows, regulatory expectations, and societal values.

2. From Population-Based Risk Scores to Longitudinal Digital Twin Representations

Traditional cardiovascular risk estimation has been shaped by large cohort studies and statistical models that identify associations between risk factors and outcomes. Although these models can be widely applied, they encode average relationships and may not capture individual pathophysiological variability. Their predictions can be miscalibrated across populations, and they are generally updated only at clinical encounters. Wearable sensing introduces a different data regime in which physiological indicators such as heart rate, heart rate variability, oxygen saturation, sleep patterns, and activity can be tracked continuously. This creates the possibility of moving from intermittent risk estimation to dynamic cardiovascular state assessment, but it also requires a conceptual shift toward models that can incorporate time-varying context, autocorrelation, and individualized baseline variation [5].

Deep learning and other machine learning approaches offer promising tools for extracting patterns from high-dimensional cardiovascular data, but their deployment in clinical settings demands careful attention to interpretability, data dependencies, and temporal structure [6]. The digital twin paradigm extends machine learning by explicitly representing the patient as a persistent computational object that receives streaming observations and maintains a continually updated estimate of physiological state. In this model, risk is not a fixed score but a property of the evolving relationship between sensor-derived data, physiological mechanisms, and environmental context. However, realizing this vision requires robust infrastructure capable of managing missing data, device heterogeneity, and the long-term evolution of both the patient and the model.

3. Wearable Sensor Data Infrastructure and Signal Quality Challenges

The quality of a cardiovascular digital twin is constrained by the quality of its underlying sensor data. Wearable devices differ substantially in sampling frequency, placement,

calibration, battery life, and susceptibility to motion artifact. The design of an operational system must therefore include careful consideration of data acquisition standards, synchronization, time-stamping, and device-specific noise profiles. Machine learning pipelines in medicine frequently encounter difficulties when training data are heterogeneous or when deployment conditions differ from development settings [7]. A digital twin architecture must address these challenges through explicit preprocessing layers that preserve clinically relevant signal features while discarding artifact-dominated segments.

Digital twin concepts originally developed in industrial engineering emphasize the need to manage emergent behavior in complex systems through continuous model updating and sensor feedback [8]. Cardiovascular applications inherit this complexity because physiological systems exhibit nonlinear interactions, circadian rhythms, and adaptive responses to stressors. Motion-corrupted photoplethysmographic signals are a particularly important challenge for wrist-worn devices, as ambulatory data often contain periodic noise from physical activity. Adaptive signal processing strategies have been developed to recover photoplethysmographic waveforms during movement, illustrating the importance of specialized preprocessing in wearable cardiovascular monitoring [9]. Without such preprocessing, downstream digital twin models may learn artifact patterns rather than true physiological dynamics, undermining the reliability of personalized risk assessment.

The infrastructure for wearable data must also consider storage, transmission, and computational efficiency. Continuous sensor streams generate substantial data volumes that cannot always be transferred in raw form to centralized servers. Edge processing can reduce transmission burdens and latency, but it requires careful trade-offs between local computational capacity and model complexity. In addition, wearable data are often collected under uncontrolled conditions, requiring robust quality indicators, uncertainty estimation, and the ability to flag periods of low signal reliability. These considerations reinforce the need to treat signal processing not as a peripheral task but as a central component of digital twin governance.

4. Digital Twin Architecture and Computational Workflows

A digital twin-enabled cardiovascular risk assessment platform typically comprises multiple interacting layers: wearable data acquisition, secure ingestion, signal preprocessing, physiological feature extraction, a dynamic patient model, risk inference, and clinical decision support. Mobile health technologies have matured rapidly, enabling more continuous interaction with patients and more frequent data capture than traditional clinical systems [10]. Large-scale smartwatch studies have demonstrated the feasibility of recruiting large cohorts and detecting cardiac arrhythmias through app-based frameworks, though they also reveal high rates of downstream clinical workload and the need for careful definition of actionable events [11][12]. These studies illustrate that digital twin architectures must anticipate not only model accuracy but also the burden placed on patients and clinicians when high-frequency data are translated into alerts.

The architecture can be organized along a spectrum from cloud-centric to edge-centric processing. Cloud-centric designs simplify model management and allow extensive computational resources, but they may introduce latency and depend on consistent network connectivity. Edge-centric designs provide local inference and faster response, but they impose constraints on model size and require mechanisms for periodic synchronization. Hybrid designs often strike a practical balance, with lightweight preprocessing and anomaly detection on the device and computationally intensive model updating in the cloud. However,

such balancing creates systemic complexity: model versions may diverge across devices, data may be lost during connectivity interruptions, and federated model aggregation must account for heterogeneous hardware.

A further architectural issue concerns the temporal scale of model updating. Some cardiovascular risk factors change over months or years, while others shift over minutes or hours. The digital twin must therefore support multiple timescales, simultaneously maintaining slow physiological baselines and fast responses to acute events. This requirement has implications for database design, model retraining, and clinical presentation. Risk outputs must be understandable to clinicians, linked to confidence intervals, and embedded in workflows that can distinguish between routine variation and clinically significant change. Without such integration, a continuous digital twin may generate alert fatigue and reduce rather than enhance clinical effectiveness.

5. Data Governance, Privacy, and Federated Learning

The longitudinal nature of wearable data creates significant governance challenges. Individual sensor streams can reveal sleep patterns, activity levels, location, stress responses, and other sensitive attributes that extend beyond conventional clinical data. Governance frameworks must therefore establish clear rules for consent, data minimization, secondary use, retention, and patient access. Bias in healthcare algorithms has been shown to emerge from structural inequalities in data collection and outcome labeling, meaning that governance decisions are inseparable from fairness outcomes [13]. Responsible machine learning in health care requires prospective and ongoing oversight, including auditing for errors, monitoring for distribution shift, and defining appropriate clinical accountability [14].

Privacy-preserving approaches such as federated learning offer one pathway for training digital twin models without centralizing raw wearable data. Federated learning allows model updates to be computed locally and aggregated across participating devices or institutions, reducing the need to transfer identifiable sensor streams [15]. In multi-institutional clinical settings, federated methods can facilitate collaboration across heterogeneous data sources while preserving data sovereignty [16]. Nevertheless, federated systems introduce their own challenges. They require robust communication protocols, secure aggregation, protection against model inversion attacks, and mechanisms for detecting malicious or low-quality participants. In addition, the regulatory status of model updates derived from distributed data remains an area of active policy debate.

Digital twin governance must also address the question of model ownership and lifecycle. A patient-specific model may be updated over many years, raising questions about what happens when the underlying cloud platform changes, when the patient transfers care, or when the model requires revalidation after a device upgrade. These lifecycle issues are not commonly addressed in early-stage machine learning research, but they become pressing in operational health systems. Sustainable governance therefore requires institutional memory, version control, audit trails, and clear procedures for model retirement.

6. Robustness, Calibration, and Clinical Validation

The clinical value of a digital twin depends on more than predictive accuracy under laboratory conditions. Wearable data are collected in noisy, uncontrolled environments, and model performance may degrade due to sensor drift, seasonal variation, changes in medication, or population differences. Proper validation requires external cohorts, temporal separation, and careful attention to outcome definitions. Guidance for critical appraisal of machine learning

studies emphasizes the need to evaluate data quality, model transparency, reproducibility, and the risk of overfitting [17]. In digital twin systems, these concerns are amplified because the model is continuously updated and may never exist in a fixed, fully validated state.

Calibration is especially important in cardiovascular risk applications. A model may have high discriminative capacity but systematically overestimate or underestimate absolute risk in particular subgroups. Calibration drift can undermine the trust of clinicians and patients, especially when decisions about preventive therapy are based on predicted risk thresholds. Digital twin architectures therefore need continuous monitoring of calibration, subgroup performance, and the frequency of false alarms. Accountability mechanisms must specify who is responsible for detecting and correcting performance degradation, and how such corrections are communicated to users [18].

Clinical validation should also assess whether the digital twin changes decisions, improves outcomes, reduces unnecessary testing, or enhances patient understanding. Randomized evaluations of digital health tools face methodological challenges, including contamination, implementation variability, and the difficulty of blinding participants. Nonetheless, without rigorous prospective studies, digital twin platforms risk being adopted on the basis of technical novelty rather than demonstrated clinical benefit. Validation frameworks must therefore include both technical evaluation and workflow-level outcomes such as time to diagnosis, rates of hospitalization, and patient-reported experience.

7. Fairness, Equity, and Sociotechnical Implications

Fairness concerns arise at multiple points in a digital twin-enabled assessment pipeline. The populations most likely to benefit from cardiovascular prevention may be underrepresented in wearable device datasets due to cost, digital literacy, connectivity, and historical exclusion from research. Machine learning systems can perpetuate or amplify these disparities when they are trained on biased data and deployed without equity-oriented safeguards [19]. A rigorous assessment of artificial intelligence in health care requires separating true clinical benefit from hype, and paying sustained attention to the distribution of benefits and harms across populations [20].

The physical infrastructure supporting wearable data collection is also unevenly distributed. Internet of Things architectures assume reliable connectivity, power, and device interoperability, but these assumptions do not hold uniformly across all communities [21]. Digital twin deployment may therefore widen cardiovascular health disparities if it is introduced primarily in well-resourced settings. Bias and fairness in machine learning encompass measurement bias, representation bias, aggregation bias, and evaluation bias, all of which are relevant to longitudinal wearable data [22]. Addressing these issues requires more than technical fixes; it demands participatory design, community engagement, and policy mechanisms that link algorithmic performance to equity goals.

From a sociotechnical perspective, the digital twin is not a neutral mirror of the body but an artifact shaped by engineering choices, economic incentives, and clinical norms. Its outputs may influence how patients understand their own health, how clinicians allocate attention, and how insurers or employers interpret risk. These broader consequences must be considered during system design, not after deployment. Transparent communication about uncertainty, limitations, and the distinction between prediction and diagnosis is essential to prevent overreliance or unnecessary anxiety.

8. Deployment, Sustainability, and Infrastructure Governance

Deployment of a digital twin-enabled cardiovascular risk assessment system across a health network raises substantial operational challenges. The system must interface with electronic health records, patient portals, clinical communication channels, and existing cardiovascular care pathways. Interoperability standards are required to ensure that sensor data can be linked to medications, laboratory results, and outcomes without creating fragmented records. These integration requirements often exceed the complexity of the predictive model itself and demand dedicated engineering and clinical informatics expertise.

Sustainability is also a central concern. Continuous data collection and model updating consume energy, storage, and computational resources. A digital twin that is too computationally intensive may be impractical for large-scale public health use, while a model that is too lightweight may fail to capture clinically meaningful variation. Lifecycle planning should include hardware refresh cycles, changes in sensor availability, and the need to retrain or retire models as technology evolves. Institutional governance must define funding models, staffing, and accountability structures that persist beyond pilot funding or vendor contracts.

The infrastructure governance model also determines how risks are distributed. If a hospital relies on a commercial cloud platform, it may benefit from scalability but face vendor lock-in, data portability limits, and uncertainty about regulatory compliance. If a health system develops an in-house platform, it gains control but assumes responsibility for security, maintenance, and validation. Hybrid models may offer flexibility but require clear contractual arrangements and shared governance protocols. These structural trade-offs are as important as algorithmic choices in determining whether the digital twin can be sustained safely and equitably over time.

9. Policy and Regulatory Implications

Regulatory frameworks for digital twins remain immature because existing pathways were largely designed for static medical devices or software with a fixed version. A continuously updated digital twin challenges assumptions about premarket approval, postmarket surveillance, and substantial equivalence. Regulators must determine when a model update is sufficiently significant to require new validation and how to oversee systems that learn across distributed patient populations. Approaches such as predetermined change control plans and real-world performance monitoring are promising, but they require further development and international harmonization.

Privacy regulation also shapes digital twin design. Requirements for data minimization, purpose limitation, and patient rights may conflict with the desire to retain long-term sensor data for model improvement. Data protection frameworks must be translated into concrete system requirements for access control, de-identification, audit logging, and algorithmic transparency. Where federated learning is used, regulators must clarify the status of model parameters and the responsibilities of participating entities.

Finally, policy should address reimbursement and incentive structures. If personalized risk assessment is not reimbursed or integrated into preventive care pathways, health systems may face a mismatch between the cost of infrastructure and the willingness to sustain it. Reimbursement models based on fee-for-service often do not capture the value of continuous monitoring and early prevention, while value-based payment arrangements may align better with longitudinal digital twin goals but require robust outcome measurement. Policy makers, payers, and professional societies therefore need to develop quality metrics that recognize meaningful improvement in cardiovascular risk management rather than simply rewarding the

volume of data collected. Without such alignment, the digital twin may remain a research artifact rather than a scalable component of preventive cardiology. International regulatory harmonization will also be important as wearable platforms and cloud infrastructures cross jurisdictional boundaries. Liability frameworks must clarify the responsibilities of device manufacturers, software providers, health systems, and clinicians when model outputs are ambiguous or when harm arises from missed signals or false alarms. Cybersecurity is another policy concern, because continuous sensor data and patient-specific models become attractive targets for unauthorized access, manipulation, or ransomware. Governance arrangements must therefore include security standards, breach notification requirements, and resilience planning. Ultimately, regulation should enable iterative learning while protecting patients from unvalidated changes, balancing the need for continuous improvement with the ethical duty to maintain clinical safety.

10. Conclusion

Digital twin-enabled personalized cardiovascular risk assessment represents a significant departure from static population-based risk stratification. Its promise lies in the ability to integrate continuous wearable sensor data into an evolving representation of cardiovascular state, potentially enabling earlier detection, individualized prevention, and more efficient allocation of clinical resources. However, realizing this promise requires far more than advanced algorithms. It demands a coherent infrastructure that handles signal quality in real-world settings, an architecture that balances edge and cloud processing across multiple timescales, and governance structures that protect privacy, ensure fairness, and sustain clinical accountability. The socio-technical character of digital twins means that design choices about data collection, model updating, and alert generation are simultaneously technical, organizational, and ethical decisions. Success will depend on sustained collaboration among engineers, clinicians, patients, regulators, and health system leaders. Future research should continue to examine how these systems can be validated prospectively, how they can be made robust to sensor variability and population diversity, and how they can be governed in ways that promote equitable cardiovascular health outcomes.

References

1. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
2. Kamel Boulos, M. N., & Zhang, P. (2021). Digital twins: From personalised medicine to precision public health. *Journal of Personalized Medicine*, 11(8), 745.
3. Bruynseels, K., Santoni de Sio, F., & van den Hoven, J. (2018). Digital twins in health care: Ethical implications of an emerging engineering paradigm. *Frontiers in Genetics*, 9, 31.
4. Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: Promises and barriers. *PLOS Medicine*, 13(2), e1001953.
5. Witt, D. R., Kellogg, R. A., Snyder, M. P., & Dunn, J. (2019). Windows into human health through wearables data analytics. *Current Opinion in Biomedical Engineering*, 9, 28–46.
6. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236–1246.

7. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347–1358.
8. Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In F.-J. Kahlen, S. Flumerfelt, & A. Alves (Eds.), *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer.
9. Zheng, X., Dwyer, V. M., Barrett, L. A., Derakhshani, M., & Hu, S. (2022). Adaptive notch-filtration to effectively recover photoplethysmographic signals during physical activity. *Biomedical Signal Processing and Control*, 72, 103303.
10. Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2015). The emerging field of mobile health. *Science Translational Medicine*, 7(283), 283rv3.
11. Perez, M. V., Mahaffey, K. W., Hedlin, H., Rumsfeld, J. S., Garcia, A., Ferris, T., ... & Turakhia, M. P. (2019). Large-scale assessment of a smartwatch to identify atrial fibrillation. *New England Journal of Medicine*, 381(20), 1909–1917.
12. Guo, Y., Wang, H., Zhang, H., Liu, T., Liang, Z., Xia, Y., ... & Lip, G. Y. H. (2019). Mobile photoplethysmographic technology to detect atrial fibrillation. *Journal of the American College of Cardiology*, 74(19), 2365–2375.
13. Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Internal Medicine*, 178(11), 1544–1547.
14. Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25(9), 1337–1340.
15. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine*, 3, 119.
16. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., ... & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10, 12598.
17. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ*, 350, g7594.
18. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, 15(11), e1002689.
19. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453.
20. Panch, T., Mattie, H., & Celi, L. A. (2019). The "inconvenient truth" about AI in healthcare. *npj Digital Medicine*, 2, 77.
21. Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K.-S. (2015). The Internet of Things for health care: A comprehensive survey. *IEEE Access*, 3, 678–708.
22. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), Article 115.