

Cross-Device Transfer Learning for Reliable Vital Sign Estimation in Wearable Healthcare Systems

Matteo Hamilton

Department of Computer Science, George Mason University, Fairfax, VA, USA.
mhamilton@gmu.edu

Christopher Hawkins

Department of Computer Science, University of New Hampshire, Durham, NH, USA.
christopher.work@unh.edu

Elliot Watkins

Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA.
elliottmail@unr.edu

Rowan Gonzalez

Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
rowan-gonzalez@missouri.edu

Abstract

Wearable healthcare systems increasingly rely on photoplethysmographic and electrocardiographic signals for continuous vital sign estimation, but discrepancies across device models, sensor placements, and usage contexts cause distribution shifts that degrade model performance. This paper examines cross-device transfer learning as a system-level strategy for achieving reliable estimation of heart rate, respiration rate, blood oxygen saturation, and related physiological parameters. The work does not claim a single algorithmic breakthrough; instead it analyzes the structural trade-offs involved when transfer learning is embedded within wearable monitoring infrastructures. The paper discusses domain shift sources including sensor hardware differences, optical path variations, motion artifact profiles, and device firmware heterogeneity. It then presents an architectural perspective on shared representation learning, adversarial domain alignment, fine-tuning, and calibration, with attention to edge-cloud partitioning and model governance. The paper evaluates implications for data infrastructure, privacy, fairness, robustness, energy consumption, clinical validation, and regulatory compliance. A central argument is that reliability must be engineered as an infrastructure property, not as a property of isolated models. Cross-device transfer learning can improve generalization and reduce reliance on per-device labeled data, but only when integrated with continuous monitoring, uncertainty estimation, federated data governance, and postmarket surveillance. The paper concludes with forward-looking policy and deployment recommendations for sustainable wearable healthcare systems.

Keywords

cross-device transfer learning; wearable healthcare; vital sign estimation; domain adaptation; physiological monitoring; system reliability.

1. Introduction

Wearable healthcare systems have moved from discretionary wellness tracking to a central role in remote patient monitoring, perioperative care, and chronic disease management. Wrist-worn optical sensors, chest patches, earbuds, and ring devices now produce continuous streams that can be used to estimate heart rate, respiratory rate, blood oxygen saturation, blood pressure trends, and autonomic indices. This expansion has been enabled by advances in low-power sensing, edge processing, and machine learning, but it has also created a pressing system-level challenge. The same algorithm deployed across a fleet of devices frequently encounters different optical components, mechanical coupling, spatial positioning, skin contact, ambient light, and firmware filtering. These differences create domain shift, meaning that the data distribution observed on a target device during inference differs from the distribution used during training. Transfer learning provides a general strategy for adapting models to such distributional variation, and its foundations have been extensively documented [1]. In the wearable context, however, transfer learning is not merely an algorithmic convenience; it is an architectural principle for making reliable vital sign estimation economically and clinically viable across heterogeneous hardware.

Deep neural networks have become the dominant modeling paradigm for physiological time series because they can learn hierarchical representations that capture both local waveform morphology and longer-range temporal dependencies. Empirical work on transferability has shown that representations learned in earlier layers of deep networks tend to be more general than representations in later layers, which become specialized to the source task [2]. A comprehensive survey of transfer learning distinguishes between instance-based, feature-based, parameter-based, and relational strategies, and highlights domain adaptation as a core unsolved problem in real deployments [3]. For cardiovascular monitoring, deep learning has enabled automated arrhythmia detection from ambulatory electrocardiograms with performance comparable to specialized cardiologists in controlled evaluations [4]. In photoplethysmographic monitoring, several studies have shown that the dominant limitation is no longer raw signal acquisition but the ability of learned models to remain stable when the deployment environment diverges from the training environment. Unsupervised domain adaptation techniques address this issue by learning feature mappings that make source and target distributions statistically indistinguishable, thereby reducing the need for labeled target-domain data [5]. Rapid vital sign extraction methods have further demonstrated that real-time opto-physiological monitoring can maintain acceptable accuracy across varying physical activity intensity levels [6]. Despite such advances, these results are often obtained on a single device class or under controlled sensor placement, leaving unresolved the challenge of generalization across an entire device fleet. Cross-device transfer learning therefore becomes a systems problem rather than a purely statistical one. It requires simultaneous attention to sensor physics, embedded resource constraints, wireless networking, data governance, clinical workflow integration, and postmarket monitoring.

2. Domain Shift and Cross-Device Variability in Wearable Sensing

The reliability of wearable vital sign estimation is limited by domain shift, which arises when the marginal or conditional distribution of input features differs between training and deployment conditions. In optical photoplethysmography, domain shift is produced by variability in light emitter wavelength, photodetector sensitivity, optical path geometry, skin melanin content, tissue perfusion, ambient illumination, and contact pressure. Two wrist-worn devices from different manufacturers may sample the same underlying cardiac rhythm yet produce waveform morphologies that differ substantially in amplitude, baseline wander, pulse

shape, and noise characteristics. A model trained on green-channel reflectance signals from one device may therefore perform poorly when applied to red or infrared channels on another device, even if the physiological target remains identical. Similarly, electrocardiographic patches exhibit differences in electrode material, lead configuration, analog filtering, sampling rate, and adhesion, all of which alter the morphology of the P wave, QRS complex, and T wave. Domain shift in this context is not merely a nuisance; it reflects structural differences in how physiological signals are transduced, conditioned, and digitized.

Motion artifact is a particularly severe source of cross-device inconsistency because it is governed by the mechanical coupling between device and body, which varies across form factors, strap materials, and wearing positions. A chest patch experiences predominantly respiratory and thoracic movement, whereas a wrist-worn optical sensor is exposed to arm swing, wrist flexion, and rotational acceleration. These motion profiles cannot be fully captured by a single synthetic noise model. Transfer learning offers a way to adapt models to such heterogeneity, but only if the adaptation strategy is aligned with the physical and operational constraints of each device class. A federation of heterogeneous devices can be understood as a distributed system in which each node observes a locally shifted version of a common latent physiological process. Learning across such a federation requires explicit treatment of device identity, sensor placement, and contextual metadata. To address the inherent heterogeneity of wearable fleets, federated learning offers a decentralized optimization paradigm in which models are trained collaboratively across devices without exchanging raw physiological data [7]. However, standard federated averaging can be unstable under non-independent and non-identically distributed data, and proximal regularized approaches have been proposed to improve convergence in such heterogeneous networks [8].

3. Architectural Principles for Cross-Device Transfer Learning

A robust cross-device transfer architecture should not assume that a single monolithic model can serve all devices without adaptation. Instead, the system should support a shared feature extraction backbone trained on multiple source devices, with lightweight device-specific adaptation layers fine-tuned on limited target data. Shared representations capture common physiological patterns such as pulse periodicity, respiratory modulation, and motion-related artifacts, while device-specific layers encode the idiosyncrasies of sensor transfer functions and mechanical coupling. This separation mirrors the broader finding that early layers of deep networks learn general features while later layers become specialized, making partial fine-tuning an efficient transfer strategy for resource-constrained environments. In wearable healthcare, however, the split between shared and device-specific components must be guided by domain physics rather than arbitrary layer depth. For instance, the representation learned before pulse segmentation may be more transferable than the representation learned after waveform morphology has been distorted by device-specific filtering.

Adversarial domain alignment can be incorporated to reduce divergence between feature distributions of source and target domains. Such alignment trains the feature extractor to produce representations that a domain discriminator cannot distinguish by device origin, thereby forcing the model to rely on physiological content rather than device fingerprints. This approach has shown value in uncontrolled deployment scenarios where labeled target data are scarce. Secure federated transfer learning extends this idea by enabling domain alignment across distributed data owners while preserving privacy through cryptographic techniques and selective parameter sharing [9]. Nevertheless, adversarial alignment

introduces several system-level risks. It can overcorrect for legitimate physiological differences, remove clinically meaningful variation, or produce unstable representations when the domain discriminator converges too quickly. Moreover, adversarial training increases computational cost and complicates convergence guarantees. These trade-offs must be weighed against the expected improvement in generalization across device fleets.

Uncertainty estimation is another architectural pillar. Deep learning models often produce overconfident predictions when evaluated on out-of-distribution inputs, which is unacceptable in clinical settings. Monte Carlo dropout and related approximate Bayesian methods allow a model to express predictive uncertainty, enabling downstream rules to suppress unreliable estimates [10]. Domain generalization approaches go further by training models on multiple source domains with the explicit objective of learning representations that transfer to unseen target distributions without any adaptation [11]. While domain generalization has theoretical appeal for wearable fleets with continuous hardware turnover, it often sacrifices source-domain accuracy to improve worst-case transfer. A pragmatic architecture may therefore combine domain generalization for the shared backbone, adversarial domain adaptation for known device groups, and Bayesian uncertainty estimation for runtime gating. Such hybrid systems require careful orchestration but can deliver more dependable vital sign estimates than any single transfer technique.

4. Edge-Cloud Partitioning and Operational Trade-offs

Deploying cross-device transfer learning in wearable healthcare requires a deliberate partitioning of computation between edge devices, gateways, and cloud services. Continuous photoplethysmography at moderate sampling rates generates substantial data volumes that are impractical to transmit in raw form from battery-constrained devices. Therefore, lightweight feature extraction and initial inference must occur on the device, while heavier adaptation, aggregation, and model update processes are assigned to gateways or cloud servers. This partitioning is not fixed; it depends on available energy, memory, latency requirements, and data sensitivity. Real-time arrhythmia or apnea detection may demand on-device inference with strict latency bounds, whereas retrospective model refinement can tolerate asynchronous cloud processing. The choice of what to offload has direct implications for device battery life, heat dissipation, and user burden. Efficient deep neural network architectures and hardware-aware model compression techniques are therefore central to the feasibility of on-device transfer learning [15].

The edge-cloud split also affects the frequency and granularity of model adaptation. Device-specific adaptation layers can be updated nightly when the device is charging, using locally stored features rather than raw waveforms. Shared backbone updates may occur weekly or monthly through federated aggregation, provided that sufficient devices contribute updates and that communication costs remain manageable. Such asynchronous update schedules mitigate energy peaks but introduce staleness, because the deployed model on any given device may lag behind the current population-level representation. System designers must therefore define consistency models and versioning policies that ensure clinical safety while allowing iterative improvement. In many respects, wearable fleets resemble large-scale distributed systems, where availability, partition tolerance, and eventual consistency must be balanced against the strong consistency requirements of medical record integration. The operational trade-offs extend beyond model accuracy to include device overheating, battery degradation, network congestion, and user retention.

Signal quality and contextual awareness are essential components of the operational architecture. Photoplethysmogram signal quality indices have long been used to reject corrupted segments in traditional pulse oximetry [13]. These indices can inform transfer learning by selecting clean target data for adaptation, thereby reducing the risk that noisy samples will corrupt device-specific parameters. Similarly, context metadata such as accelerometer magnitude, skin temperature, and battery voltage can serve as auxiliary inputs that help the model discriminate between physiological changes and environmental confounds. The integration of internet of things infrastructure in healthcare has enabled this kind of multimodal sensing, but it also increases system complexity and attack surface [14]. Cross-device transfer learning should not treat context as an afterthought, because the same vital sign estimate can have very different reliability under rest, ambulation, and sleep. A well-designed system maintains an explicit quality pipeline that filters, annotates, and stores adaptation data according to signal integrity and metadata completeness.

5. Data Infrastructure, Governance, and Privacy

Cross-device transfer learning depends on continuous access to representative data, but the data are distributed across heterogeneous devices owned by individuals in varied legal and institutional contexts. Centralizing raw physiological waveforms raises serious privacy risks because such data can reveal cardiac abnormalities, sleep patterns, stress levels, and even identity-related features. Federated learning reduces the need to move raw data by aggregating model updates rather than records [7]. However, model updates are not automatically privacy-preserving, and gradient inversion attacks have shown that raw inputs can sometimes be reconstructed from shared parameters under adversarial conditions. Secure multiparty computation, differential privacy, and secure aggregation can mitigate these risks, but each adds computation, communication, or noise. Secure federated transfer learning attempts to combine domain adaptation with privacy-preserving mechanisms across multiple data owners [9]. The governance challenge is to ensure that privacy protection does not disproportionately degrade utility for minority device groups or underrepresented physiological conditions.

Governance must also address data provenance, consent, and secondary use. A wearable dataset may originate from research cohorts, consumer wellness programs, clinical trials, or hospital remote monitoring programs, each with different consent limitations. Transfer learning across these sources requires metadata about collection protocol, sensor model, software version, and annotation source. Without such lineage, domain shift cannot be diagnosed and corrected. Data infrastructure should therefore maintain a model card or dataset nutrition label that documents the device distribution, demographic distribution, signal quality distribution, and known exclusion criteria. This documentation supports both technical debugging and regulatory audit. The European Medical Device Regulation and related frameworks increasingly expect manufacturers to demonstrate that algorithms are safe and effective across intended patient populations, which cannot be established if training data are treated as an opaque aggregate. Cross-device transfer learning should be accompanied by a transparent data governance layer that tracks how each device type contributes to model updates and how adaptation decisions are made.

Privacy-preserving data management also intersects with data minimization and storage retention. Since wearable data often contain long-term behavioral patterns, retaining indefinite raw records creates unnecessary risk. A system designed around transfer learning can store compressed latent features, summary statistics, and model residuals rather than full

waveforms. These derived artifacts may still carry sensitive information, but their lower dimensionality and reduced identifiability ease privacy compliance. Nonetheless, feature-level data may be insufficient for diagnosing future failure modes or for retraining the shared backbone under new sensor generations. Organizations must therefore define tiered retention policies that balance analytic value against privacy exposure. Governance bodies should include clinical, technical, legal, and patient representatives to negotiate these trade-offs in a defensible manner.

6. Fairness, Robustness, and Clinical Validation

Reliability in wearable healthcare cannot be reduced to average accuracy metrics. A model may achieve high aggregate performance while failing systematically for individuals with darker skin tone, low perfusion, arrhythmia, or high body mass index. Such disparities are not hypothetical; they have been documented in optical sensing contexts where melanin absorption and scattering alter the signal-to-noise ratio. Transfer learning can either mitigate or amplify these disparities depending on how source and target data are selected. If adaptation targets are dominated by healthy young adults, the model may overfit to high-quality signals and perform worse for patients with fragile physiology. Fairness must therefore be treated as an explicit design objective, with stratified evaluation across demographic groups, device models, and clinical conditions [16]. The distribution of error, not merely the mean error, should inform deployment decisions.

Robustness is closely related. A reliable wearable system should degrade gracefully when sensor contact is lost, ambient light changes, or network connectivity drops. Deep learning models are known to be vulnerable to covariate shifts and adversarial perturbations in ways that are not always visible in standard benchmarks [17]. In the wearable context, natural perturbations include motion bursts, hardware aging, firmware updates, and patient-specific rhythms. Cross-device transfer learning can improve robustness by exposing models to diverse conditions, but it can also introduce silent failure modes if the feature extractor learns spurious correlations with device identifiers. Runtime uncertainty estimates, signal quality gating, and fallback heuristics provide defense-in-depth. A model should be allowed to output a low-confidence result or to request additional sensor verification rather than generating a clinically dangerous estimate. Robustness engineering therefore requires monitoring not only model accuracy but also the distribution of confidence scores and rejection rates across devices and contexts.

Clinical validation for transfer-learned wearable systems requires a different evidentiary standard than consumer wellness algorithms. Regulatory pathways increasingly distinguish between general wellness devices and medical devices intended for diagnosis or treatment. When vital signs are used to guide clinical decisions, manufacturers must demonstrate accuracy under realistic conditions, across intended populations, and over the expected life of the device. Digital clinical measures require rigorous analytical validation, clinical validation, and usability evaluation [18]. Cross-device transfer learning complicates validation because model performance may change during deployment as adaptation continues. A model that is validated at launch may diverge as devices update their adaptation layers. Manufacturers must therefore adopt versioned validation, where each major model release is evaluated against frozen clinical datasets and prospective studies. This does not preclude continuous learning, but it requires a staged deployment model in which online updates are shadowed and compared against a validated reference before full activation. Without such governance, adaptive models risk regulatory noncompliance and patient harm.

7. Sustainability and Forward-Looking Policy Implications

Sustainable deployment of cross-device transfer learning demands attention to energy, compute, and lifecycle costs. Training large models repeatedly across device fleets consumes energy and accelerates hardware obsolescence. Edge adaptation can reduce cloud computational loads, but it increases per-device processing and may shorten battery life if not carefully managed. Model compression, quantization, and sparse fine-tuning can reduce resource use without unacceptable accuracy loss [15]. Sustainability also includes the environmental footprint of frequent firmware and model updates. Wearable devices with limited storage may require entire firmware images to be exchanged, generating network traffic and electronic waste through shortened device lifespans. A system-level design should treat model update frequency, payload size, and compression strategy as first-class sustainability parameters rather than incidental engineering choices.

Policy frameworks must evolve to accommodate adaptive medical software while preserving accountability. If a model changes after deployment, the responsible entity for an adverse event may become ambiguous. Manufacturers, cloud providers, data aggregators, and clinical users all hold parts of the system, yet current accountability structures often assume a static software product. Algorithmic accountability mechanisms such as model cards, audit trails, and impact assessments offer practical tools for documenting design decisions and deployment conditions [19]. Legal and ethical scholarship on artificial intelligence in healthcare emphasizes that adaptive algorithms require ongoing oversight, informed consent for data use, and clear liability allocation [20]. Cross-device transfer learning therefore cannot be governed solely as a technical pipeline; it must be embedded in an organizational governance regime that spans engineering, clinical affairs, regulatory affairs, and patient advocacy.

The long-term viability of wearable healthcare systems depends on interoperability and open benchmarking. Proprietary devices and closed data ecosystems inhibit transfer learning and make it difficult to reproduce performance across populations. Open datasets with device-level metadata, standardized preprocessing pipelines, and shared evaluation protocols would accelerate research while enabling fair comparison. However, open data must be balanced against privacy and commercial interests. Federated benchmarking, in which models are evaluated locally on distributed datasets without centralizing raw data, represents a promising compromise. Policy incentives could encourage manufacturers to participate in such benchmarks as part of postmarket surveillance obligations. By sharing aggregated error distributions and adaptation outcomes, stakeholders can build collective knowledge about cross-device reliability without exposing sensitive patient records. This collaborative infrastructure would support continuous improvement while preserving the distributed nature of wearable healthcare.

The shift from static models to adaptive cross-device learning also raises questions about patient autonomy and explainability. Patients may reasonably ask why a vital sign estimate changed after a firmware update, or why their device behaves differently from another person's device. Transfer learning systems should provide interpretable confidence statements and contextual explanations, not opaque numerical outputs. Such explanations may include the extent of device-specific adaptation, the similarity between current context and training conditions, and the estimated uncertainty range. In this sense, reliability is not only a statistical property but also a communicative one. A system that reports a heart rate of 62 beats per minute with high confidence under stable conditions is more trustworthy than one

that silently adjusts its estimate without context. Cross-device transfer learning can support this communicative goal by making domain distance and adaptation provenance visible to users, clinicians, and auditors.

8. Conclusion

Cross-device transfer learning offers a promising route toward reliable vital sign estimation in wearable healthcare systems, but its success depends on treating domain shift as an infrastructure challenge rather than a purely algorithmic defect. This paper has examined the structural sources of cross-device variability, the architectural principles that can support shared and device-specific representations, and the trade-offs that arise in edge-cloud deployment, data governance, fairness, clinical validation, and sustainability. Federated learning and secure domain adaptation can reduce the need for centralized raw data while enabling models to learn from diverse devices, yet they introduce coordination, privacy, and convergence overheads. Uncertainty estimation and signal quality gating provide essential safety mechanisms, especially when models are deployed on unseen hardware or under changing physiological conditions. Clinical validation must be versioned and population-stratified to ensure that adaptive models do not silently accumulate bias or degrade performance. Governance and policy frameworks must evolve in parallel, because adaptive medical software challenges traditional assumptions about static product validation and liability. The path forward lies in integrating transfer learning with robust data infrastructure, transparent model documentation, federated benchmarking, and continuous postmarket surveillance. Only by designing for heterogeneity, uncertainty, and accountability as first-class system properties can wearable healthcare fulfill its promise of dependable, equitable, and sustainable vital sign monitoring.

References

1. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
2. Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, 27, 3320–3328.
3. Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), 9.
4. Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1), 65–69.
5. Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. *Proceedings of the 32nd International Conference on Machine Learning*, 37, 1180–1189.
6. Zheng, X., Dwyer, V. M., Barrett, L. A., Derakhshani, M., & Hu, S. (2023). Rapid vital sign extraction for real-time opto-physiological monitoring at varying physical activity intensity levels. *IEEE Journal of Biomedical and Health Informatics*, 27(7), 3107–3118.
7. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data.

Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 54, 1273–1282.

8. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 2, 429–450.
9. Liu, Y., Kang, Y., Xing, C., Chen, T., & Yang, Q. (2020). A secure federated transfer learning framework. *IEEE Intelligent Systems*, 35(4), 70–82.
10. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *Proceedings of the 33rd International Conference on Machine Learning*, 48, 1050–1059.
11. Gulrajani, I., & Lopez-Paz, D. (2021). In search of lost domain generalization. *Proceedings of the 9th International Conference on Learning Representations*.
12. Charlton, P. H., Birrenkott, D. A., Bonnici, T., Pimentel, M. A. F., Johnson, A. E. W., Alastruey, J., Tarassenko, L., Watkinson, P. J., Beale, R., & Clifton, D. A. (2019). Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review. *IEEE Reviews in Biomedical Engineering*, 11, 2–20.
13. Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*, 8(1), 14–25.
14. 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.
15. Sze, V., Chen, Y.-H., Yang, T.-J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12), 2295–2329.
16. 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), 1–35.
17. Hendrycks, D., & Dietterich, T. G. (2019). Benchmarking neural network robustness to common corruptions and perturbations. *Proceedings of the 7th International Conference on Learning Representations*.
18. Coravos, A., Goldsack, J. C., Karlin, D. R., Nebeker, C., Perakslis, E., Zimmerman, N., & Erb, M. K. (2019). Digital medicine: A primer on measurement. *Digital Biomarkers*, 3(2), 31–71.
19. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal audit. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–49.
20. Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In *Artificial Intelligence in Healthcare* (pp. 295–336). Academic Press.