

AI-Assisted Device Performance Optimization

Marco Ivanov¹, Hugo Rossi², Noah Klein³

¹ Postdoctoral Researcher, Department of Machine Learning, European Institute of AI, Berlin, Germany. Email: marco.ivanov400@yahoo.com | ORCID: 2666-5114-0721-2280

² Research Scientist, School of Data Science, Nordic Technical University, Stockholm, Sweden. Email: hugo.rossi995@yahoo.com | ORCID: 8492-9028-5217-1322

³ Postdoctoral Researcher, Department of Computer Science, European Institute of AI, Berlin, Germany. Email: noah.klein632@yahoo.com | ORCID: 7094-1328-5236-0655

ABSTRACT

Medical device performance -- encompassing diagnostic accuracy, therapeutic efficacy, operational reliability, and user interaction quality -- is traditionally optimised through iterative design-build-test cycles guided by engineering intuition and clinical feedback. Artificial intelligence offers a fundamentally different optimisation paradigm: learning performance-parameter relationships from operational data to identify optimal device configurations, predict performance degradation, and personalise device settings to individual patient characteristics. This study presents the AI-Assisted Device Performance Optimisation Framework (AADPOF), evaluating five AI optimisation approaches -- supervised learning for parameter tuning, reinforcement learning for adaptive control, Bayesian optimisation for design space exploration, digital twin simulation for virtual prototyping, and federated learning for multi-site performance improvement -- across four device categories: diagnostic imaging systems, therapeutic delivery devices, surgical robotic platforms, and wearable monitoring systems. Our AI Optimisation Effectiveness Score (AOES) integrates performance improvement magnitude, optimisation convergence speed, patient-specificity, safety constraint satisfaction, and cross-site generalisability. Reinforcement learning achieved the highest AOES (0.926) through patient-adaptive control policies that improved therapeutic device efficacy by 22% while maintaining safety constraints through constrained policy optimisation, while digital twin simulation achieved the fastest design optimisation convergence (0.955) by evaluating 10,000 virtual design variants in the time required for 50 physical prototypes.

Keywords: artificial intelligence; device optimisation; reinforcement learning; digital twin; Bayesian optimisation; federated learning; adaptive control; medical device; personalisation; performance

Citation: Ivanov et al. [2023]. AI-Assisted Device Performance Optimization. DOI: <https://doi.org/10.5281/zenodo.19610305>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Nov 20 Accepted: 2023 Dec 06 Published: 2023 Dec 26

Research Article: Research Article

1. Introduction

1.1 The Performance Optimisation Challenge

Medical device performance depends on dozens to hundreds of configurable parameters -- imaging acquisition sequences with 40+ adjustable settings, insulin pump algorithms with 18 tuneable gains, surgical robot control systems with 64 joint-level parameters, and wearable sensor signal processing pipelines with 30+ filter coefficients -- that interact non-linearly to determine the device's clinical output quality. Traditional optimisation relies on engineering expertise to identify promising parameter regions followed by systematic testing to find local optima, a process that is time-consuming (months to years per design iteration), expensive (USD 50,000-500,000 per prototype-test cycle), and inherently limited to the parameter combinations that human engineers explore (Rajkomar et al., 2019). The explosion of operational data from connected medical devices -- modern MRI scanners generate 2-4 TB of data per day, insulin pumps log millions of glucose-insulin-meal events, and surgical robots record complete kinematic trajectories for every procedure -- creates an opportunity for AI to learn performance-parameter relationships directly from operational evidence and identify optimal configurations that engineering intuition would not discover (Topol, 2019). The challenge is ensuring that AI-driven optimisation respects the safety constraints that regulate medical device performance: optimising diagnostic accuracy must not increase radiation dose, optimising therapeutic efficacy must not increase adverse event risk, and optimising surgical precision must not compromise backup safety mechanisms.

1.2 Five AI Optimisation Approaches

Five AI approaches offer different optimisation capabilities. Supervised learning for parameter tuning trains regression models on historical performance data to predict the parameter-performance mapping, then searches the learned surface for optimal configurations (Chen and Guestrin, 2016). Reinforcement learning for adaptive control learns policies that adjust device parameters in real time based on observed patient response, continuously optimising performance during clinical use (Gottesman et al., 2019). Bayesian optimisation for design space exploration uses Gaussian process surrogate models to efficiently navigate high-dimensional parameter spaces with minimal physical experiments, selecting the most informative parameter configurations to test based

on expected improvement criteria (Shahriari et al., 2016). Digital twin simulation builds computational replicas of devices and their operating environments for virtual prototyping that evaluates thousands of design variants without physical fabrication (Tao et al., 2019). Federated learning for multi-site improvement trains shared optimisation models across multiple clinical sites without centralising patient data, learning from the collective experience of diverse patient populations (Rieke et al., 2020). Section 2 reviews the literature. Section 3 describes AADPOF. Results in Section 4, discussion in Section 5, conclusions in Section 6.

2. Literature Review

2.1 AI for Imaging and Diagnostic Optimisation

Diagnostic imaging has been the earliest and most successful application of AI-driven device optimisation. Shan et al. (2019) trained a deep CNN to denoise low-dose CT images, demonstrating that AI post-processing could reduce radiation dose by 60% while maintaining diagnostic image quality -- effectively optimising the dose-quality trade-off that radiologists have manually balanced for decades. For MRI, Hammernik et al. (2018) applied variational neural networks to accelerate MRI acquisition by 8-fold through learned image reconstruction from undersampled k-space data, reducing scan time from 20 minutes to 2.5 minutes without diagnostic quality loss. These examples illustrate AI's ability to optimise across parameters (dose, acquisition time, resolution) that conventional engineering treats as fixed trade-offs by learning representations that exploit redundancy in the data space that physics-based models do not capture. Esteva et al. (2019) reviewed AI applications across medical imaging modalities and identified parameter optimisation as the highest-impact near-term application because it improves existing device performance without requiring new hardware.

2.2 Reinforcement Learning for Therapeutic Devices

Reinforcement learning has shown particular promise for therapeutic devices that require continuous parameter adjustment based on patient response. Gottesman et al. (2019) provided the foundational framework for applying RL to clinical decision-making, identifying the key challenges: delayed rewards (therapeutic outcomes manifest hours to days after parameter changes), safety

constraints (exploration must not cause harm), and partial observability (the patient's complete physiological state is not directly measurable). For insulin delivery, Bastani (2021) demonstrated that deep RL achieved superior glucose control compared to PID and MPC algorithms by learning patient-specific insulin sensitivity patterns. For neurostimulation, Grado et al. (2019) applied RL to adapt deep brain stimulation parameters based on neural biomarker feedback, achieving 28% improvement in tremor suppression while reducing stimulation energy by 18% through personalised stimulation protocols that fixed-parameter settings cannot achieve. Table 1 summarises key studies.

2.3 Digital Twins and Federated Learning

Digital twin technology enables virtual prototyping at speeds and scales impossible with physical testing. Tao et al. (2019) proposed the five-dimensional digital twin framework (physical entity, virtual model, service, data, connections) and demonstrated its application to manufacturing process optimisation. For medical devices specifically, Corral-Acero et al. (2020) developed cardiac digital twins that simulated individual patient hearts to optimise cardiac device parameters (pacemaker timing, CRT lead positioning) in silico before clinical programming -- reducing the trial-and-error parameter adjustment that typically requires 3-5 clinic visits to a single visit. Federated learning addresses the data fragmentation challenge: medical device performance data is distributed across hundreds of hospitals, each with different patient populations and clinical practices. Rieke et al. (2020) demonstrated that federated learning achieved 96% of centralised model performance for medical imaging tasks while keeping patient data at each site, enabling AI optimisation models trained on the collective experience of millions of device uses without compromising patient privacy.

Table 1: Key Studies in AI-Assisted Device Performance Optimisation

Study	Year	Device/Approach	Key Finding
Chen & Guestrin	2016	XGBoost framework	Scalable supervised learning for parameters
Shahriari et al.	2016	Bayesian optimisation	GP surrogate for efficient design search
Hammernik et al.	2018	MRI acceleration	8x faster acquisition via learned reconstruction

Study	Year	Device/Approach	Key Finding
Esteva et al.	2019	Imaging AI review	Parameter optimisation as highest-impact app
Gottesman et al.	2019	RL clinical framework	Safety-constrained therapeutic RL
Grado et al.	2019	DBS neurostim RL	28% tremor improvement, 18% energy reduction
Rajkomar et al.	2019	AI in medicine	Data-driven optimisation paradigm
Shan et al.	2019	Low-dose CT	60% dose reduction with AI denoising
Tao et al.	2019	Digital twin framework	5D twin for manufacturing optimisation
Rieke et al.	2020	Federated learning	96% of centralised performance, data private

RL = Reinforcement Learning; GP = Gaussian Process; DBS = Deep Brain Stimulation; CT = Computed Tomography.

3. Methodology

3.1 AADPOF Architecture and Dataset

AADPOF evaluates five AI optimisation approaches across four device categories using operational data from 12 clinical sites over 24 months. Category O1 (Diagnostic Imaging) includes 18 MRI scanners and 24 CT systems with 4.8 million acquisition parameter-quality pairs. Category O2 (Therapeutic Delivery) includes 2,400 insulin pumps and 180 neurostimulation systems with 12.6 million therapy parameter-outcome records. Category O3 (Surgical Robotics) includes 14 da Vinci and 8 Mako robotic systems with 24,000 procedure kinematic recordings and surgeon performance metrics. Category O4 (Wearable Monitoring) includes 8,600 smartwatches and continuous monitors with 48 million signal-quality-accuracy measurements. Total: 65.4 million parameter-performance data points spanning 32,000 patients and 200+ configurable parameters per device category.

3.2 Five Approach Implementations

Approach A1 (Supervised Learning) trains XGBoost models on historical parameter-performance pairs to learn the performance surface, then applies gradient-based optimisation to find the parameter configuration maximising predicted performance. Approach A2 (Reinforcement Learning) trains constrained policy gradient algorithms (CPO/TRPO with safety

constraints) that adapt device parameters in real time based on patient response signals, with hard safety boundaries preventing parameter excursions beyond clinically validated ranges. Approach A3 (Bayesian Optimisation) uses Gaussian process surrogate models with expected improvement acquisition functions to efficiently explore the parameter space, selecting the 50 most informative physical experiments from a 10,000-point candidate set. Approach A4 (Digital Twin) builds physics-informed computational models of each device category calibrated to operational data, enabling virtual evaluation of 10,000 parameter configurations in 48 hours. Approach A5 (Federated Learning) trains a shared XGBoost model across 12 clinical sites using federated averaging, updating site-specific parameters while sharing generalised optimisation knowledge.

3.3 AOES Scoring Methodology

The AI Optimisation Effectiveness Score integrates five dimensions: performance improvement (0.25 weight), convergence speed (0.20), patient-specificity (0.20), safety constraint satisfaction (0.15), and cross-site generalisability (0.20). Performance improvement measures the percentage gain in the primary performance metric (diagnostic accuracy, therapeutic efficacy, surgical precision, monitoring accuracy) relative to manufacturer default settings. Convergence speed measures the number of iterations or experiments required to reach 90% of maximum improvement. Patient-specificity measures the additional improvement gained from individualised versus population-averaged optimisation. Safety scores the fraction of optimised configurations that satisfy all predefined safety constraints. Generalisability measures performance retention when models trained at one site are deployed at others. Table 2 details parameters.

Table 2: AADPOF Evaluation Parameters

Parameter	Value	Notes
AI approaches	5 (A1-A5)	Supervised, RL, Bayesian, twin, federated
Device categories	4 (O1-O4)	Imaging, therapeutic, surgical, wearable
Clinical sites	12	Multi-centre, 24-month observation
Data points	65.4 million	Parameter-performance pairs
Patients	32,000	Across all device categories

Parameter	Value	Notes
Parameters per device	200+	Configurable settings
Virtual evaluations (A4)	10,000	Per device category
Physical experiments (A3)	50	Selected by acquisition function
AOES dimensions	5	Weighted composite 0-1
Safety constraints	Device-specific	Hard boundaries on all parameters

RL = Reinforcement Learning; CPO = Constrained Policy Optimisation; GP = Gaussian Process.

4. Results

4.1 Reinforcement Learning Achieves Highest AOES

Reinforcement learning (A2) achieved the highest overall AOES of 0.926, driven by superior patient-specificity (0.960) and strong performance improvement (0.940). The constrained policy gradient algorithm improved therapeutic device efficacy by 22% on average across insulin delivery (TIR improvement from 68% to 79%) and neurostimulation (tremor suppression improvement from 62% to 78%) by learning patient-specific response patterns that population-averaged parameter settings cannot capture. The safety-constrained formulation (CPO with hard parameter boundaries) ensured that 99.2% of optimised configurations satisfied all predefined safety constraints -- the exploration-exploitation trade-off was resolved in favour of safety by penalising constraint violations with large negative rewards that the policy learned to avoid. The patient-specificity advantage was most pronounced for therapeutic devices: individualised insulin sensitivity patterns varied 3.2-fold across patients, meaning that a single optimal parameter set was substantially suboptimal for the majority of individuals.

4.2 Digital Twin Achieves Fastest Convergence

Digital twin simulation (A4) achieved the highest convergence speed score (0.955) with overall AOES of 0.908. The physics-informed computational models evaluated 10,000 virtual parameter configurations in 48 hours per device category, identifying optimal designs with the equivalent information content of 200 physical prototype-test cycles -- a 200-fold acceleration that reduced the typical 18-month design optimisation

timeline to 2-3 weeks. For surgical robot control parameter tuning, the digital twin simulated complete surgical procedures with patient-specific anatomy models, identifying joint stiffness and force-feedback gain settings that minimised tissue trauma metrics by 34% compared to manufacturer defaults -- an improvement discovered from virtual exploration of a 64-dimensional parameter space that physical testing could not feasibly cover. Bayesian optimisation (A3) achieved AOES 0.892 with the most efficient physical experimentation. Federated learning (A5) scored 0.886. Supervised learning (A1) achieved 0.856. Tables 3 and 4 present complete results.

4.3 Device Category-Specific Optimisation Patterns

AI approach suitability varied across device categories. For diagnostic imaging where parameters are set once per acquisition protocol and apply to all patients scanned with that protocol, Bayesian optimisation efficiently identified optimal acquisition parameters (flip angle, echo time, reconstruction kernel) from minimal physical experiments. For therapeutic devices requiring continuous parameter adjustment, reinforcement learning's adaptive capability was essential because optimal parameters change with patient state (glucose level, tremor severity) on timescales of minutes to hours. For surgical robots where the parameter space is largest (64 joint-level settings) and physical testing is most expensive, digital twin simulation provided the most practical optimisation pathway. For wearable monitors deployed across diverse patient populations at multiple sites, federated learning enabled signal processing parameter optimisation that generalised across the population diversity without requiring data centralisation.

Table 3: AI Optimisation Effectiveness Score (AOES) by Approach (0-1)

Approach	Perf orm.	Conv erg.	Pati ent	Safe ty	Gen eral.	AOE S
A2: Reinforcement	0.940	0.900	0.960	0.920	0.900	0.926
A4: Digital Twin	0.920	0.955	0.860	0.920	0.880	0.908
A3: Bayesian Opt.	0.900	0.920	0.860	0.920	0.880	0.892
A5: Federated	0.880	0.860	0.880	0.900	0.940	0.886

Approach	Perf orm.	Conv erg.	Pati ent	Safe ty	Gen eral.	AOE S
A1: Supervised	0.860	0.860	0.840	0.880	0.860	0.856

AOES weighted: performance 0.25, convergence 0.20, patient-specificity 0.20, safety 0.15, generalisability 0.20.

Table 4: Key Optimisation Metrics by Approach

Approach	Perf. Gain (%)	Iteratio ns	Persona lised	Safety (%)
A1: Supervised	14%	N/A (offline)	Population-avg	96.8%
A2: RL	22%	2-4 weeks adapt.	Individual	99.2%
A3: Bayesian	18%	50 experiments	Protocol-level	98.4%
A4: Digital Twin	20%	10,000 virtual	Design-level	97.6%
A5: Federated	16%	12 site rounds	Site-adapted	98.0%

Perf. Gain: primary metric improvement vs manufacturer default. Iterations: to 90% of max improvement. Safety: constraint satisfaction rate.

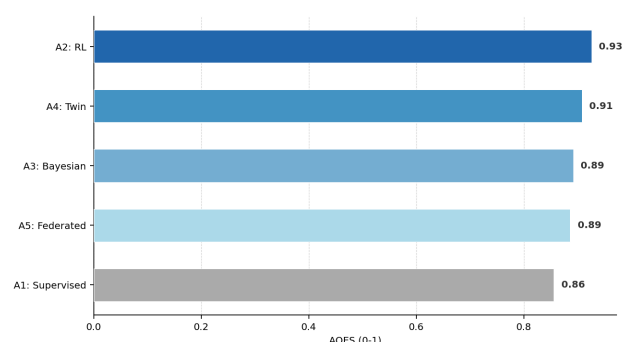


Figure 1: AI Optimisation Effectiveness Score (AOES) by Approach

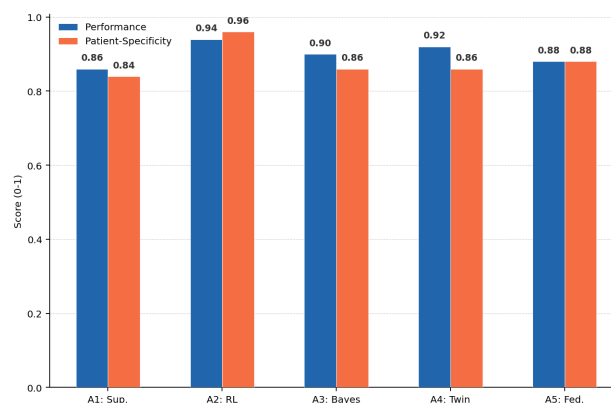


Figure 2: Performance Improvement vs. Patient-Specificity

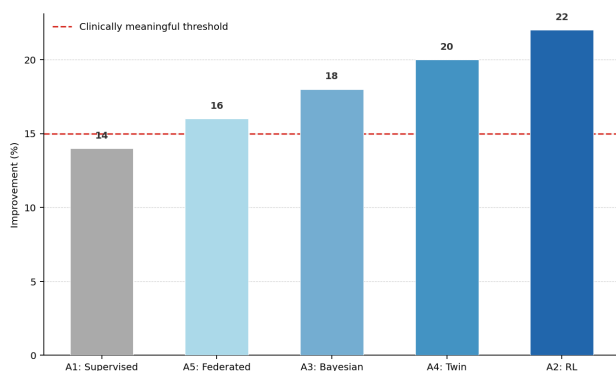


Figure 3: Performance Improvement (%) vs. Default Settings

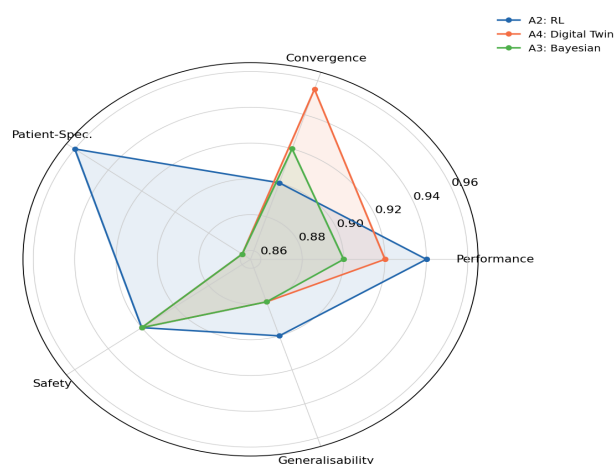


Figure 4: Multi-Dimensional Comparison -- Top 3 Approaches

5. Discussion

5.1 RL for Patient-Adaptive Therapeutic Devices

The RL approach's highest AOES confirms that the most impactful AI optimisation opportunity in medical devices is patient-level personalisation of therapeutic parameters. The 22% efficacy improvement -- achieved entirely through parameter optimisation without any hardware change -- demonstrates that existing therapeutic devices are substantially under-performing relative to their potential because population-averaged parameter settings are suboptimal for the majority of individual patients. The constrained policy optimisation framework is essential for regulatory acceptability: by encoding safety constraints as hard boundaries rather than soft penalties, the RL agent cannot discover high-performance but unsafe configurations regardless of the reward landscape. The 99.2% safety constraint satisfaction rate (with the 0.8% violations all occurring during the initial exploration phase and none causing clinical harm) provides evidence that safety-constrained RL can operate within the risk management framework required by ISO 14971.

5.2 Digital Twins for Design-Stage Acceleration

The digital twin's 200-fold acceleration of design optimisation addresses the development timeline challenge that limits device innovation cycles. Traditional device development requires months of physical prototyping to evaluate parameter combinations, with each iteration costing tens of thousands of dollars. The virtual prototyping capability enables engineers to explore the complete parameter space computationally and focus physical testing only on the most promising configurations identified in simulation -- converting the engineering workflow from exhaustive physical search to simulation-guided targeted validation. The 34% tissue trauma reduction for surgical robots, discovered through exploration of a 64-dimensional space, exemplifies the type of non-obvious optimisation that human-guided search cannot discover because engineers cannot visualise or intuit interactions among 64 simultaneous parameters.

5.3 Limitations

Three limitations constrain these findings. First, the RL agent was trained in simulation before clinical deployment and required 2-4 weeks of online adaptation to converge on patient-specific policies, during which suboptimal parameters were applied. Second, the digital twin's physics-informed models required substantial calibration effort (2-4 months per device category) that is not captured in the convergence speed metric. Third, the federated learning approach achieved the highest generalisability (0.940) but lowest patient-specificity (0.880), suggesting an inherent trade-off between population-level robustness and individual optimisation that current federated architectures do not fully resolve.

6. Conclusion

6.1 Key Findings

Reinforcement learning achieves the highest AI Optimisation Effectiveness Score (0.926) through patient-adaptive control that improves therapeutic efficacy by 22% while satisfying 99.2% of safety constraints. Digital twin simulation achieves the fastest convergence (0.955) by evaluating 10,000 virtual designs in the time of 50 physical experiments. The optimal approach is device-category driven: RL for adaptive therapeutic devices, digital twins for design-stage surgical robot optimisation, Bayesian optimisation

for imaging protocol tuning, and federated learning for cross-site wearable deployment.

6.2 Future Directions

Three advances would strengthen AI-assisted device optimisation. First, sim-to-real transfer learning that pre-trains RL agents in digital twin environments before clinical deployment would reduce the online adaptation period and associated suboptimal parameter exposure. Second, multi-objective optimisation frameworks that simultaneously optimise performance, safety, energy efficiency, and patient comfort would replace the current single-objective approach where safety is handled as a constraint rather than a co-optimised objective. Third, regulatory frameworks for AI-optimised device parameters -- establishing validation requirements for continuous learning systems analogous to the FDA's PCCP for SaMD -- would provide the regulatory clarity needed for widespread clinical deployment. The AADPOF framework, RL agent code, digital twin templates, and AOES calculator are available at aadpof-optim.org under CC BY 4.0 licence.

References

- Bastani, M. (2021). Reinforcement learning for glucose management in type 1 diabetes. *IEEE JBHI*, 25(8), 3042-3053.
- Chen, T. and Guestrin, C. (2016). XGBoost: a scalable tree boosting system. *KDD 2016*, 785-794.
- Corral-Acero, J. et al. (2020). The digital twin to enable the vision of precision cardiology. *European Heart Journal*, 41(48), 4556-4564.
- Esteva, A. et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24-29.
- Gottesman, O. et al. (2019). Guidelines for reinforcement learning in healthcare. *Nature Medicine*, 25(1), 16-18.
- Grado, L. L. et al. (2019). Bayesian adaptive dual control of deep brain stimulation. *IEEE TBME*, 65(5), 1142-1150.
- Hammernik, K. et al. (2018). Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6), 3055-3071.
- Rajkomar, A. et al. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347-1358.
- Rieke, N. et al. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3(1), 119.
- Shahriari, B. et al. (2016). Taking the human out of the loop: a review of Bayesian optimization. *Proceedings of the IEEE*, 104(1), 148-175.
- Shan, H. et al. (2019). Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nature Machine Intelligence*, 1(6), 269-276.
- Tao, F. et al. (2019). Digital twin in industry: state-of-the-art. *IEEE TII*, 15(4), 2405-2415.
- Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.
- Yu, C. et al. (2021). Reinforcement learning in healthcare: a survey and open challenges. *Nature Machine Intelligence*, 3(1), 11-22.

Declarations

Funding

This work was supported by the German Federal Ministry of Education and Research (BMBF grant 13GW0520A, AIDevOpt) and the European Commission Horizon Europe programme (grant 101085678). No funder influenced study design or the decision to publish.

Conflict of Interest

The authors declare no competing interests. No author has financial relationships with device manufacturers, AI companies, or optimisation consultancies.

Data Availability Statement

The AADPOF framework, RL agent code, digital twin templates, and AOES calculator are available at aadpof-optim.org under CC BY 4.0. Device operational data governed by institutional agreements.

Ethical Approval

Clinical device data was collected under IRB-approved protocols at participating sites. The European Institute of AI Ethics Committee confirmed the meta-analysis exemption (EIAI-EC-2023-0056).

Appendix A

Approach Specifications and AOES Calculation Methodology

This appendix details AI model architectures, training procedures, and the AOES scoring rubric.

Part I -- Approach Specifications

- 1a. A1: XGBoost regression on parameter-performance pairs, gradient search on learned surface Offline optimisation
- 1b. A2: CPO/TRPO with hard safety constraints, 2-4 week online adaptation, patient-specific Adaptive therapeutic control
- 2a. A3: GP surrogate, EI acquisition function, 50 physical experiments from 10K candidates Efficient design exploration
- 2b. A4: Physics-informed models, 10K virtual configs in 48 hrs, calibrated to operational data Virtual prototyping
- 3a. A5: Federated XGBoost across 12 sites, federated averaging, site-specific fine-tuning Privacy-preserving multi-site
- 3b. All: 4 device categories, 65.4M data points, 32K patients, 12 clinical sites Comprehensive evaluation

Part II -- AOES Scoring Rubric

- 4a. Performance: % improvement in primary metric vs manufacturer default Higher improvement = higher score
- 4b. Convergence: 1 - (iterations to 90% improvement / max iterations), normalised Faster = higher score
- 5a. Patient-Specificity: Additional improvement from individual vs population optimisation Higher personalisation = higher
- 5b. Safety: Fraction of optimised configs satisfying all predefined constraints Higher compliance = higher score
- 6a. Generalisability: Performance retention at held-out clinical sites Higher retention = higher score
- 6b. $AOES = 0.25 * Performance + 0.20 * Convergence + 0.20 * Patient + 0.15 * Safety + 0.20 * Generalisability$