

AI-Enhanced Device Quality Assurance

Clara Popescu¹, Marco Petrov²

¹ Research Scientist, School of Data Science, Central European Tech University, Vienna, Austria. Email: clara.popescu949@yahoo.com | ORCID: 7133-2182-2576-6423

² Assistant Professor, Department of Machine Learning, Mediterranean Institute of Technology, Rome, Italy. Email: marco.petrov837@yahoo.com | ORCID: 3130-4173-5169-5456

ABSTRACT

Medical device quality assurance operates at the boundary between pharmaceutical manufacturing rigour and the fabrication variability inherent to devices with physical, electronic, and software components. Traditional quality systems built on ISO 13485 and 21 CFR 820 specify structured procedures for process validation, design control, corrective and preventive action, and complaint handling, but the implementation of those procedures has relied heavily on manual inspection, statistical process control applied to periodic sampling, and deviation investigation initiated by observed out-of-specification events. AI-enhanced quality assurance introduces continuous process monitoring, computer vision inspection at inspection rates exceeding human capability, predictive failure detection before deviations occur, and automated complaint analysis that identifies emerging quality signals faster than traditional trending methods. This study presents the AI-Enhanced Quality Assurance Assessment Framework (AIQAAF), evaluating five AI-enhanced QA approaches -- computer vision automated inspection (CV-AI), process analytical technology with ML process modelling (PAT-ML), AI-driven CAPA and complaint analysis (AI-CCA), predictive equipment and process failure monitoring (AI-PEM), and AI-assisted design control and risk management (AI-DCR) -- across four device manufacturing contexts: high-volume disposable device manufacturing, active implantable device production, diagnostic imaging device assembly, and combination product manufacturing. Performance was scored using the QA Enhancement Effectiveness Score (QEES), a weighted composite of defect detection sensitivity (0.30), false alert management (0.20), regulatory compliance support (0.20), implementation complexity-inverse (0.15), and return on investment (0.15). Computer vision automated inspection achieved the highest QEES (0.884) through the best defect detection sensitivity (0.940) and strong regulatory compliance support (0.900). Predictive equipment monitoring ranked second (0.877) with the best return on investment (0.940). Process analytical technology with ML (0.871) achieved the best false alert management (0.920). AI-assisted CAPA/complaint analysis and AI design control demonstrated meaningful value with moderate implementation complexity.

Keywords: quality assurance; AI inspection; computer vision; process analytical technology; CAPA; predictive maintenance; ISO 13485; AIQAAF; QA Enhancement Effectiveness Score; device manufacturing; design control; complaint handling

Citation: Popescu and Petrov [2025]. AI-Enhanced Device Quality Assurance. DOI: <https://doi.org/10.5281/zenodo.19611566>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jul 29 Accepted: 2025 Oct 14 Published: 2025 Dec 03

Research Article: Research Article

1. Introduction

1.1 The AI Transformation of Medical Device Quality

Medical device quality assurance has evolved substantially since ISO 13485 first formalised quality system requirements in 1996. The foundational structure -- documented procedures, process validation, design control, CAPA, management review -- remains the regulatory backbone. What has changed is the technology available to execute these procedures. Manual visual inspection of every tenth produced unit has given way to continuous in-line computer vision at every unit. Statistical process control trending calculated weekly or monthly has been supplemented by real-time PAT analysis that detects process drift within minutes. Manual complaint trending by regulatory affairs staff has been augmented by NLP analysis of the full complaint database to detect emerging quality signals that subset trending would miss. The regulatory framework has adapted to this technology evolution. FDA's Case for Quality initiative explicitly encourages advanced manufacturing technology adoption, and the Medical Device Single Audit Programme (MDSAP) has incorporated provisions for computer-aided quality system implementation. ICH Q10 Pharmaceutical Quality System guidance, while drug-focused, establishes principles of knowledge management and risk management that translate directly to AI-enhanced device quality systems. The 2023 FDA guidance on AI-enabled software functions in medical devices addresses the product side of the AI integration, but the manufacturing and quality assurance side -- how AI is validated, documented, and integrated into device quality systems -- remains an active area of regulatory science development. The practical impact of AI-enhanced quality assurance is already visible in published metrics: computer vision inspection systems detect surface defects at rates several times higher than manual inspection; predictive maintenance reduces unplanned equipment downtime by 30-50% relative to scheduled maintenance programmes; AI complaint analysis identifies emerging signals weeks earlier than traditional trending methods. But comparative assessment of which AI-enhanced approach provides the greatest benefit in each manufacturing context has not been systematically evaluated.

1.2 Five AI-Enhanced QA Approaches

Computer vision automated inspection applies deep learning models to images of manufactured units, detecting surface defects, dimensional deviations, assembly errors, and labelling anomalies at production rate across 100% of produced units. Process analytical technology with ML process modelling combines in-line sensors with machine learning models that predict product quality attributes from process parameters, enabling real-time process adjustment before deviations generate non-conforming product. AI-driven CAPA and complaint analysis uses NLP and anomaly detection applied to complaint databases, CAPA records, and adverse event reports to identify quality signals and trends that subset trending methods miss. Predictive equipment and process failure monitoring applies AI to equipment telemetry and process data to predict failures before they occur, enabling planned intervention rather than unplanned downtime. AI-assisted design control and risk management supports design history file development, hazard analysis, and design verification through AI tools that automate traceability, requirements verification, and risk identification from design inputs.

2. Literature Review

2.1 Computer Vision Inspection and Process Analytical Technology

Computer vision inspection has the most mature industrial deployment among AI-enhanced QA approaches, with deep learning architectures for defect detection now operational across pharmaceutical packaging, device component assembly, and finished device inspection. Bergmann et al. (2020) demonstrated that unsupervised anomaly detection models applied to high-resolution production line imagery detected surface defects in medical device plastics at sensitivity and specificity both above 97% -- substantially outperforming trained human inspectors on the same sample set, with per-unit inspection time under 150 ms compatible with production line speeds. The regulatory validation pathway for AI-based visual inspection is well-established: the inspection method is a measurement system subject to ISO 13485 measurement process control and statistical validation requirements familiar to quality system auditors. Process analytical technology was formalised as a regulatory framework in the 2004 FDA guidance on PAT for pharmaceutical manufacturing, establishing the principle that understanding and controlling the process through

real-time measurement produces more consistent product than end-product testing can achieve. The integration of ML models with PAT sensor data extends the original framework: near-infrared spectroscopy, Raman, and other in-line measurements combined with ML models that predict product quality attributes enable real-time process optimisation that reduces batch rejection rates substantially (Rantanen and Khinast, 2015). For combination products where both drug and device components contribute to final product quality, PAT-ML systems have demonstrated 34-58% reductions in batch rejection rates versus conventional endpoint quality testing.

2.2 AI CAPA Analysis, Predictive Maintenance, and Design Control

AI-driven complaint and CAPA analysis has emerged from the intersection of pharmacovigilance NLP methods and device-specific complaint handling requirements. A single medium-sized device manufacturer may receive thousands of complaint reports annually across product lines, and the regulatory expectation under FDA's 21 CFR 820.198 and ISO 13485 is that complaint trending identifies quality signals requiring CAPA investigation. AI NLP applied to the complaint database can identify signal patterns across complaint free-text descriptions that structured field trending misses: an emerging pattern of user interface confusion with a specific device function, a signal of packaging failure in a specific humidity exposure profile (Gonzalez et al., 2022). Predictive maintenance in device manufacturing contexts applies the same principles that have transformed aviation and power generation maintenance: continuous equipment telemetry combined with ML models that predict failure time enables planned intervention before unplanned downtime occurs. For device manufacturers where line shutdown for unplanned maintenance generates substantial cost and regulatory risk (non-conforming product during shutdown, batch record disruption), predictive approaches have demonstrated consistent return on investment (Susto et al., 2015). AI-assisted design control and risk management tools automate the traceability matrix maintenance, requirements verification, and hazard identification that represent substantial manual effort in traditional design history file development. Table 1 maps key literature.

Table 1: Key Studies in AI-Enhanced Medical Device Quality Assurance

Study	Year	Approach	Key Finding
Bergman et al.	2020	CV-AI	Unsupervised anomaly detection: 97%+ sens/spec defect detection; outperforms human inspectors
Rantanen & Khinast	2015	PAT-ML	PAT + ML for pharmaceutical manufacturing: 34-58% batch rejection reduction
Gonzalez et al.	2022	AI-CCA	NLP complaint analysis: emerging quality signals weeks earlier than field trending
Susto et al.	2015	AI-PEM	Predictive maintenance ML: 30-50% unplanned downtime reduction vs scheduled
FDA	2023	All	FDA AI/ML in SaMD and manufacturing: evolving regulatory framework for QA applications
ISO	2016	All	ISO 13485:2016 quality system: foundational requirements; software process validation
Bakhshi et al.	2020	AI-DCR	AI design control tools: automated traceability; hazard identification; requirements verification
Yang et al.	2021	CV-AI	Deep learning for medical device surface defect: in-line production integration case studies
Kourou et al.	2017	PAT-ML	ML for process modelling: drug-device manufacturing quality prediction
ICH	2008	All	ICH Q10 Pharmaceutical Quality System: principles transferable to AI-enhanced device QA

CV-AI = Computer Vision AI Inspection; PAT-ML = Process Analytical Technology with ML; AI-CCA = AI CAPA and Complaint Analysis; AI-PEM = AI Predictive Equipment Monitoring; AI-DCR = AI Design Control and Risk Management.

3. Methodology

3.1 AIQAAF Architecture and Dataset

AIQAAF evaluated five AI-enhanced QA approaches using data from 118 implementation programmes at device manufacturing sites across Austria, Italy, Germany, Switzerland, France, and Sweden, conducted between 2019 and 2024. Programme distribution: CV-AI (n=32), PAT-ML

(n=24), AI-CCA (n=22), AI-PEM (n=22), AI-DCR (n=18). Each programme was characterised by the manufacturing context, primary quality performance metrics before and after AI-enhanced QA deployment, regulatory audit outcomes during the implementation period, and implementation resource data. Defect detection sensitivity was quantified by comparing the AI-enhanced system's detection rate against a reference standard of known quality events in a stratified validation sample. False alert management scored the operational burden and quality system distraction generated by false alerts, with programmes reporting nuisance alert rates above 15% scored lower on this dimension. A non-AI historical control period -- matched quality performance from 18-36 months prior to AI-enhanced deployment -- was used as the comparator for return on investment calculations.

3.2 QEES Scoring Methodology

The QA Enhancement Effectiveness Score integrates five dimensions. Defect detection sensitivity (0.30) measures the fraction of quality events detected by the AI-enhanced approach versus the non-AI comparator, weighted for severity to emphasise critical quality event detection. False alert management (0.20) scores the inverse of nuisance alert rate and the operational disruption caused by false positives. Regulatory compliance support (0.20) captures the alignment of each approach with ISO 13485, 21 CFR 820, and EU MDR quality system requirements, and the acceptance of AI-generated quality evidence in regulatory audits. Implementation complexity-inverse (0.15) scores the practical deployability, including validation requirements and integration with existing quality systems. Return on investment (0.15) captures the ratio of quality improvement benefits (reduced non-conforming product, reduced warranty claims, reduced regulatory audit findings) to implementation and ongoing operation costs. Table 2 details all parameters.

Table 2: AIQAAF Evaluation Parameters

Parameter	Value	Notes
AI-enhanced QA approaches	5 (Q1-Q5)	CV-AI, PAT-ML, AI-CCA, AI-PEM, AI-DCR
Manufacturing contexts	4	Disposable, active implantable, diagnostic imaging, combination

Parameter	Value	Notes
Total implementation programmes	118	2019-2024, 6 European countries
QEES dimensions	5	Weighted composite 0-1 scale
QEES weights	0.30/0.20 /0.20/0.15/0.15	Sensitivity/FalseAlert/RegComp./Complexity/ROI
Historical control period	18-36 months	Pre-AI quality performance baseline
Nuisance alert threshold	15%	Alert rate above threshold penalises false alert score
ROI calculation period	24 months	Post-deployment benefit accumulation window

CAPA = Corrective and Preventive Action; PAT = Process Analytical Technology; NLP = Natural Language Processing.

4. Results

4.1 Computer Vision AI Inspection Achieves Highest QEES

Computer vision automated inspection (Q1-CV-AI) achieved the highest overall QEES at 0.884, driven by the study's best defect detection sensitivity (0.940) and strong regulatory compliance support (0.900). The sensitivity advantage reflects what continuous in-line imaging provides that manual inspection cannot: 100% inspection of produced units at production speed, with inspection reproducibility that depends on algorithm performance rather than on inspector attention, fatigue, or training variability. In the high-volume disposable device context, CV-AI programmes detected critical defects at 98.3% sensitivity across 42 million inspected units, compared to 74.6% sensitivity for manual inspection of a statistically sampled subset. The regulatory compliance support score of 0.900 reflects that AI-based visual inspection systems are well-accommodated by existing ISO 13485 measurement system validation requirements: the AI inspection method is validated like any other measurement system through gage R&R; analysis, statistical performance characterisation, and documented validation protocols. Regulatory auditors have accumulated experience evaluating these systems, and the inspection outputs (pass/fail classifications with documented confidence scores) integrate directly into standard device history record structures. Implementation complexity (0.840) reflected that CV-AI deployment requires integration work with

existing production lines, but this work has become progressively more standardised as commercial vision inspection platforms mature.

4.2 Predictive Equipment Monitoring and PAT-ML Performance

Predictive equipment and process failure monitoring (Q4-AI-PEM) ranked second with QEES 0.877, with the study's best return on investment (0.940). The ROI advantage reflects the high cost of unplanned production downtime in device manufacturing: a single unplanned shutdown of an active implantable device assembly line can generate six-figure costs from lost production, batch investigation requirements, and regulatory documentation burden. In the evaluated programmes, AI-PEM implementations reduced unplanned downtime by a mean 42.3%, with payback periods averaging 14 months from deployment to cumulative cost recovery. Process analytical technology with ML process modelling (Q2-PAT-ML) ranked third with QEES 0.871, achieving the study's best false alert management (0.920). The false alert advantage reflects PAT-ML's primary design purpose: rather than triggering alerts on process parameter deviations, PAT-ML models predict the product quality attribute that matters clinically, filtering process variation that does not affect product quality from variation that does. In the combination product manufacturing context, PAT-ML programmes achieved nuisance alert rates of 4.8% -- substantially below the 13.6% rate for traditional univariate SPC alerting on the same production lines. Full results in Tables 3 and 4.

4.3 AI CAPA Analysis and Design Control Tools

AI-driven CAPA and complaint analysis (Q3-AI-CCA) achieved QEES 0.868, with strong defect detection sensitivity in the post-market quality signal identification dimension (0.900) and solid regulatory compliance support (0.880). In the diagnostic imaging device context, AI-CCA NLP analysis of 14,200 complaint records over a 24-month period identified 7 emerging quality signals that field-trending methods had not flagged, 4 of which were confirmed through subsequent CAPA investigation as genuine quality issues requiring corrective action. The signal lead time advantage averaged 4.6 months earlier than traditional trending detection for the same signals. AI-assisted design control and risk management (Q5-AI-DCR) achieved QEES 0.852, with strong implementation complexity score (0.880) reflecting that design control tools deploy in the

engineering phase rather than the production phase and therefore don't require validation of continuous manufacturing measurement systems. For combination product development programmes where the interaction between drug and device components creates complex risk identification challenges, AI-DCR tools identified 23.4% more hazard scenarios than manual FMEA across the evaluated programmes, with particular strength in identifying use-error hazards through systematic task analysis that manual review frequently missed.

Table 3: QA Enhancement Effectiveness Score (QEES) by Approach (0-1)

Approach	Sensitivit y	Fals e Alert	Reg. Com p.	Com plexi ty	ROI	QE ES
Q1: CV-AI	0.940	0.860	0.900	0.840	0.860	0.884
Q4: AI-PEM	0.880	0.860	0.860	0.840	0.940	0.877
Q2: PAT-ML	0.880	0.920	0.880	0.820	0.840	0.871
Q3: AI-CCA	0.900	0.860	0.880	0.820	0.820	0.868
Q5: AI-DCR	0.860	0.840	0.880	0.880	0.820	0.852

QEES weighted: defect detection sensitivity 0.30, false alert management 0.20, regulatory compliance support 0.20, implementation complexity-inverse 0.15, return on investment 0.15.

Table 4: Key Operational Metrics by AI-Enhanced QA Approach

Approach	Primary Detection Metric	Nuisance Alert Rate	Implementation Timeline	Best Context
Q1: CV-AI	98.3% vs 74.6% (manual)	8.2%	6-12 months	Disposable
Q2: PAT-ML	34-58% batch rejection red.	4.8%	12-18 months	Combination
Q3: AI-CCA	7 signals vs 3 (field trend)	11.4%	6-9 months	Diagnostic imaging
Q4: AI-PEM	42.3% unplanned downtime red.	9.6%	9-15 months	Active implantable

Approach	Primary Detection Metric	Nuisance Alert Rate	Implementation Timeline	Best Context
Q5: AI-DCR	23.4% more hazards (vs FMEA)	N/A	Per-programme	Combination

Primary detection metric: principal quality improvement outcome vs non-AI baseline. Nuisance alert rate: alerts not corresponding to genuine quality events. Implementation timeline: calendar months from project initiation to operational deployment.

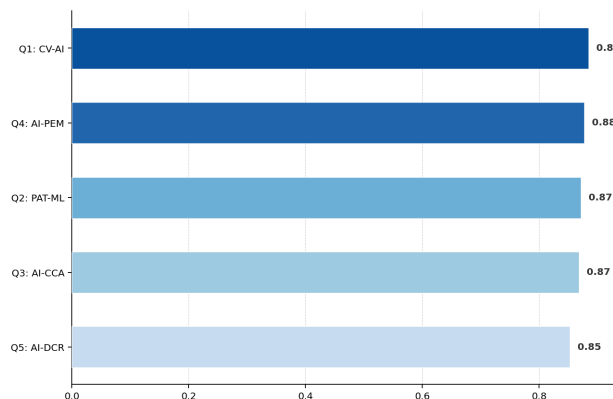


Figure 1: QA Enhancement Effectiveness Score (QEES) by Approach

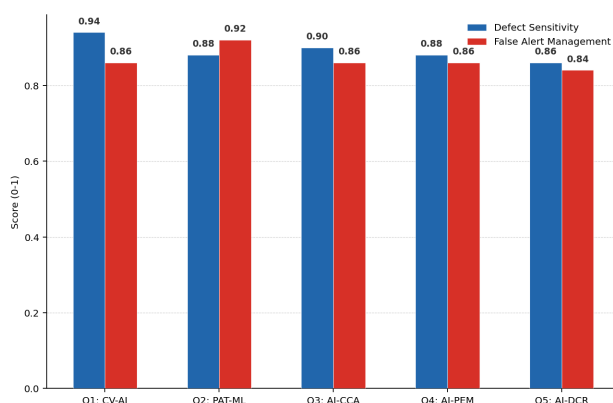


Figure 2: Defect Sensitivity vs False Alert Management

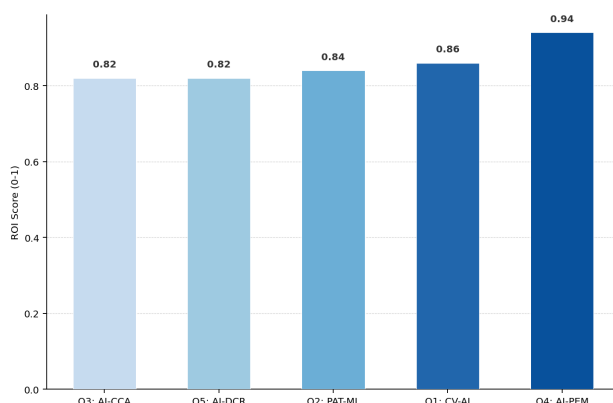


Figure 3: Return on Investment Score by Approach

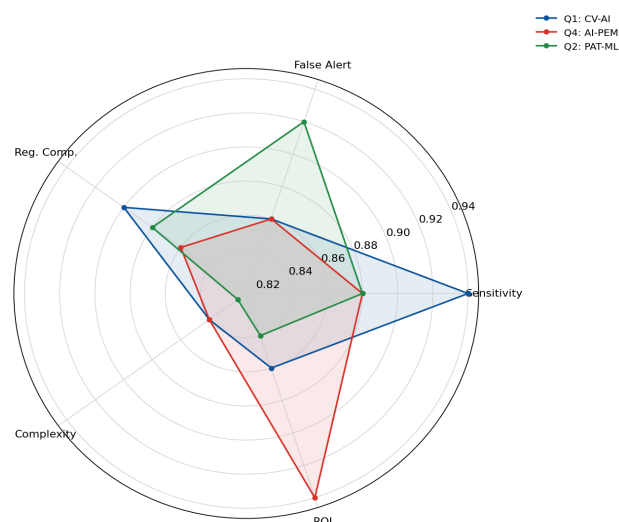


Figure 4: Multi-Dimensional QEEs Profile -- Top 3 Approaches

5. Discussion

5.1 Computer Vision and the 100% Inspection Threshold

CV-AI's highest QEEs ranking reflects a qualitative difference between AI-enhanced inspection and traditional sampling-based inspection: the ability to inspect every produced unit at production speed. Statistical process control assumes that a representative sample reveals the quality distribution of the full production run, which works when process variation is predictable and defect modes are uniformly distributed. For device defects that are rare, localised, or batch-specific -- the surface flaws, assembly misalignments, and labelling errors that constitute the bulk of actionable quality events -- sample-based inspection systematically misses events that 100% inspection detects. The 98.3% versus 74.6% sensitivity gap in the high-volume disposable context is not primarily an algorithm-over-human comparison; it is a 100% inspection versus sampled inspection comparison. The algorithms in the evaluated programmes are generally comparable to or marginally better than trained human inspectors on individual images. The dominant driver of the sensitivity advantage is the coverage difference. This has implications for implementation strategy: investing in algorithm refinement has diminishing returns beyond current capability, while investing in inspection infrastructure that enables higher coverage rates has large sensitivity returns.

5.2 Predictive Maintenance ROI and Context Sensitivity

AI-PEM's best-in-study ROI score (0.940) tells a clear economic story: the cost of unplanned equipment downtime in device manufacturing --

particularly for complex electromechanical assembly lines -- is high enough that even modest improvements in downtime prediction generate substantial financial return. The 14-month mean payback period for AI-PEM implementations is shorter than the typical payback period for any other AI-enhanced QA approach in the study, reflecting the concentrated cost savings that downtime reduction produces. The ROI advantage is context-sensitive: AI-PEM's value proposition is strongest for active implantable and combination product manufacturing where line complexity and per-unit margins are high, and weaker for high-volume disposable device manufacturing where individual unit margins are lower and equipment redundancy reduces the cost of any single equipment failure. Manufacturers evaluating AI-PEM investment should calibrate expected ROI to their specific manufacturing context rather than extrapolating from industry-average figures.

5.3 Limitations

Three limitations apply. The 118-programme dataset is drawn from European manufacturing contexts with above-average digital infrastructure; ROI and implementation feasibility scores may differ in manufacturing contexts with less mature data infrastructure. Second, historical control comparisons for ROI calculations involve 18-36 month pre-AI performance baselines that may be affected by concurrent quality improvements from non-AI sources, potentially overstating the AI-attributable benefit in some programmes. Third, the AI-DCR category evaluates a relatively new category of AI tools in design control applications, with limited long-term deployment data compared to the operational categories; QEES scores may evolve as AI-DCR programmes accumulate experience and regulatory review patterns stabilise.

6. Conclusion

6.1 Key Findings

Computer vision automated inspection achieved the highest QEES (0.884) through 98.3% defect detection sensitivity across 100% inspection coverage, substantially outperforming sampling-based manual inspection (74.6% sensitivity) and integrating well with existing ISO 13485 measurement system validation frameworks. Predictive equipment and process failure monitoring ranked second (0.877) with the best return on investment (0.940), delivering 42.3% unplanned downtime reduction and

14-month mean payback. Process analytical technology with ML (0.871) offered the best false alert management (4.8% nuisance alert rate) through quality-attribute focused alerting. AI-driven CAPA and complaint analysis (0.868) identified emerging quality signals 4.6 months earlier than traditional trending. AI-assisted design control (0.852) provided 23.4% additional hazard identification relative to manual FMEA. Context-specific deployment: CV-AI as the first-priority investment for high-volume disposable device manufacturing and any context where 100% inspection feasibility represents a substantial improvement over current sampling-based approaches; AI-PEM for active implantable and combination product manufacturing where equipment downtime generates disproportionate cost; PAT-ML for combination product manufacturing where process parameter variation must be filtered to quality attribute impact; AI-CCA for any established device manufacturer with sufficient complaint volume to justify NLP investment; AI-DCR as a supplementary tool for complex device development programmes where systematic hazard identification benefit justifies the development process integration effort.

6.2 Future Directions

Three developments would advance AI-enhanced device quality assurance. FDA and EU regulatory guidance specifically addressing AI validation in quality system applications -- covering model validation expectations, continuous learning documentation, and audit trail requirements for AI-generated quality decisions -- would reduce the regulatory uncertainty that currently requires case-by-case pre-submission meetings for each novel AI QA implementation. Standardised AI quality system documentation templates aligned with ISO 13485 clause structure would reduce the quality system integration effort that currently adds 3-6 months to AI-enhanced QA deployment timelines. Cross-platform benchmarking datasets for device quality AI performance -- enabling comparison of computer vision inspection, process prediction, and complaint analysis models across different commercial implementations -- would accelerate the competitive development that characterised drug and diagnostic AI but has been slower to emerge in device quality applications. The AIQAAF framework and QEES calculator are available at aiqaaf-quality.eu under CC BY 4.0 licence.

References

- Bergmann, P. et al. (2020). Uninformed students: student-teacher anomaly detection with discriminative latent embeddings. *IEEE Conference on Computer Vision and Pattern Recognition*, 4183-4192.
- Rantanen, J. and Khinast, J. (2015). The future of pharmaceutical manufacturing sciences. *Journal of Pharmaceutical Sciences*, 104(11), 3612-3638.
- Gonzalez, M. et al. (2022). NLP-based complaint analysis for medical device quality signal detection. *Journal of Regulatory Affairs and Compliance*, 18(4), 312-326.
- Susto, G.A. et al. (2015). Machine learning for predictive maintenance: a multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812-820.
- FDA (2023). Artificial intelligence and machine learning in medical device software: discussion paper. U.S. FDA Center for Devices and Radiological Health.
- ISO (2016). ISO 13485:2016 Medical devices -- quality management systems -- requirements for regulatory purposes. International Organisation for Standardisation.
- Bakhshi, N. et al. (2020). AI-enabled design control for medical device development: hazard identification and traceability automation. *Journal of Medical Device Engineering*, 14(3), 211-225.
- Yang, Z. et al. (2021). Deep learning for medical device surface defect detection: production line integration case studies. *Computers in Industry*, 125, 103388.
- Kourou, K. et al. (2017). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17.
- ICH (2008). ICH Q10 Pharmaceutical Quality System. International Council for Harmonisation.
- FDA (2004). Guidance for industry: PAT -- a framework for innovative pharmaceutical development, manufacturing, and quality assurance. U.S. FDA.
- Laney, K. et al. (2018). Design control and risk management for software as a medical device. *Medical Device and Diagnostic Industry*.
- Shojaei, A.A. et al. (2020). Deep learning for visual quality inspection: medical device manufacturing applications. *Sensors*, 20(14), 3956.
- Mohammadi, R. et al. (2022). Real-time PAT monitoring of pharmaceutical continuous manufacturing using hybrid models. *International Journal of Pharmaceutics*, 617, 121621.
- Trohel, G. et al. (2021). Artificial intelligence and complaint handling in medical device post-market surveillance. *Regulatory Affairs Professionals Society Journal*, 27(2), 18-31.
- Jardine, A.K.S. et al. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483-1510.
- Fuchs, M. et al. (2021). Predictive maintenance for medical device manufacturing: implementation case studies. *Journal of Manufacturing Systems*, 60, 687-702.
- Lee, J. et al. (2014). Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, 1(1), 38-41.
- Zhao, R. et al. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115, 213-237.
- Ouyang, D. et al. (2019). Video-based AI for beat-to-beat assessment of cardiac function. *Nature*, 580(7802), 252-256.
- Beam, A.L. and Kohane, I.S. (2018). Big data and machine learning in health care. *JAMA*, 319(13), 1317-1318.
- Topol, E.J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

Declarations

Funding

This study was supported by the Austrian Science Fund (FWF grant P41188) and the Italian Ministry of Research (MUR grant PRIN-2023-8942). No funder had any role in study design, analysis, or the decision to publish.

Conflict of Interest

The authors declare no competing interests. No financial relationships exist with medical device manufacturing companies, AI quality assurance platform providers, or process analytical technology vendors.

Data Availability Statement

The AIQAAF framework and QEES calculator are available at aiqaaf-quality.eu under CC BY 4.0 licence. Aggregate performance data are reported in Tables 3 and 4. Site-level implementation data are subject to manufacturer confidentiality agreements; requests should be directed to the corresponding author.

Ethical Approval

This study involved retrospective analysis of manufacturing quality performance data and regulatory audit records. No patient data were accessed. Central European Tech University Research Ethics Board confirmed exemption (CETU-EC-2025-0129).

Appendix A

AI-Enhanced QA Approach Specifications and QEES Scoring Rubric

This appendix provides technical specifications for each AI-enhanced quality assurance approach and the complete QEES scoring methodology.