Diponkor Mondal

Lifecycle Management Framework for IVDR and EU AI Act Compliant Machine Learning Enabled Medical Device Software

Metropolia University of Applied Sciences
Master of Engineering
Information Technology
Master's Thesis
10 Jun 2025

Preface

This research work was conducted at the Hematoscope Lab, an academic team of physicians, programmers, scanner engineers, and students affiliated with the Hospital District of Helsinki and Uusimaa HUS and HUSLAB. The lab focuses on combining high-resolution automated imaging and deep learning methods to enhance the diagnosis, monitoring, and understanding of hematological diseases. Hematoscope Oy is the spin-off company emerging from this research group.

During the research and writing process, I received invaluable guidance and support from my supervisors, Principal Lecturer Mikael Soini and Mikko Purhonen. Specifically, Mikko's expert insights and constructive feedback significantly enhanced the clarity of my work. I am also truly grateful to Oscar Brück, the principal investigator of the Hematoscope Lab, for his continuous support and inspiration.

Additionally, I would like to extend my heartfelt thanks to my colleagues at Hematoscope Lab and fellow students at Metropolia University of Applied Sciences for their collaboration and meaningful discussions.

As I am still at the early stage of my career in regulatory affairs, writing this thesis has been a great learning experience for me. It helped me understand how artificial intelligence is not only changing MedTech industries but also changing the way we look at quality and regulatory work.

Lastly, and it would be unfair not to mention, a special thanks goes to my wife, Snigdha. Her patience, understanding, and unconditional support during countless late office hours and long nights of work have been truly invaluable.

Diponkor Mondal
Syöpäkeskus HUS, 10 June 2025

Abstract

Author:	Diponkor Mondal
Title:	Lifecycle Management Framework for IVDR and EU AI Act Compliant Machine Learning Enabled Medical Device
Number of Pages:	54 pages + 2 Appendices
Date:	10 Jun 2025
Degree:	Master of Engineering
Degree Programme:	Information Technology
Professional Major:	Medical Technology
Supervisors:	Mikael Soini, Principal Lecturer Mikko Purhonen, MSc. (Tech.)

The application of Machine Learning (ML) in In Vitro *Diagnostic* (IVD) medical software presents significant possibilities for improving diagnostic accuracy, but also introduces additional regulatory hurdles within the European Union (EU) due to the interplay between the In Vitro Diagnostic Regulation (IVDR) and the AI Act. This thesis undertakes a comparative gap analysis of two legislative frameworks to identify and address legal differences and overlaps. The IVDR is governed by scientific validity, safety, and clinical performance, but lacks transparency regarding algorithmic bias and verifiable machine learning components. In contrast, the AI Act brings more structure by providing a comprehensive risk-based approach for high-risk AI systems. However, it lacks detailed clinical validation guidance within IVD boundaries. To mitigate these regulatory hurdles, this thesis proposes an integrated lifecycle management framework that incorporates AI Act requirements into the existing lifecycle framework, complying with the IVDR. The framework supports complete compliance with operability restrictions on the development of ML-based IVD medical software. This study makes a significant academic and practical contribution by laying the groundwork for an open-access regulatory framework that enables startup companies to navigate complex compliance requirements efficiently.

Keywords: IVDR, AI Act, ML in Diagnostics, Lifecycle Management Framework, Regulatory Strategy for IVD Software

---

Contents

List of Abbreviations

1 Introduction 1

2 Research Objectives 3

3 Methods 3

3.1 Research Design: Constructive Research 4

3.2 Regulatory Framework Analysis 5

3.3 Data Collection 8

4 Regulatory and Standards Background 9

4.1 EU Legislation 9

4.1.1 In Vitro Diagnostic Regulation (IVDR) 2017/746 9

4.1.2 Overview of the AI Act and Its Impact on ML-Enabled IVD Software 10

4.2 International Standards for Life Cycle Management 15

4.2.1 ISO 13485 Quality Management System 17

4.2.2 ISO 14971: Risk Management 18

4.2.3 IEC 62304: Software Lifecycle Processes 19

4.3 Regulatory Pathways for ML-Enabled Medical Devices 20

4.4 The Need for a Lifecycle Management Framework 21

5 Results: Gap Analysis and Key Findings 22

5.1 Key Challenges in Regulatory Compliance 23

5.1.1 ML-Specific Guidance Gaps 23

5.1.2 Data Privacy and Security Concerns 24

5.1.3 Transparency and Explainability Requirements 24

5.1.4 Validation and Verification of ML Algorithms 25

5.1.5 Data Bias, Fairness, and Validation Requirements 27

5.2 Comparative Analysis IVDR Vs. AI Act 27

5.3 Gap Analysis: IVDR Vs. AI Act 28

5.4 Bridging the Gap Between the AI Act and IVDR 29

5.4.1 Unified Classification and Risk Management Approach 29

5.4.2 Enhanced Technical Documentation and Traceability 29

5.4.3 Proactive Ethical Compliance and Governance Mechanisms 30
5.4.4 Lifecycle Management Integration 30
5.4.5 Continuous Regulatory Surveillance and Update Mechanism 30
6 Proposed Lifecycle Management Framework 31
6.1 Framework Objectives and Design Principles 31
6.2 Integration of IVDR and AI Act Requirements 32
6.2.1 Mapping IVDR (Annex I & II) into the ML lifecycle 33
6.2.2 Mapping AI Act (Articles 10–15) into the ML lifecycle 34
6.3 Model Development and Validation Process 35
6.3.1 Define Intended Purpose and Risk Classification 36
6.3.2 Data Acquisition and Preparation 36
6.3.3 Model Design and Training 36
6.3.4 Verification and Validation 37
6.3.5 Human Factors Validation and Oversight Testing 37
6.3.6 Technical Documentation and Final Design Review 37
6.3.7 Regulatory Submission and Notified Body Review 38
6.3.8 Deployment and Release 38
6.4 Post-Market Monitoring and Feedback Loop 38
6.5 Adaptive System Lifecycle and Retraining Protocols 40
6.5.1 Locked vs. Learning Approach 40
6.5.2 Triggers for Retraining or Updating the Model 41
6.5.3 Retraining Process 41
6.5.4 Regulatory Submissions 42
6.5.5 Update Strategy: Periodic Rather Than Continuous 42
6.5.6 Version Control and Change Tracking 42
6.5.7 User Communication and Support 43
6.6 Ensuring Transparency and Algorithm Explainability 43
6.7 Addressing Data Privacy and Security Compliance with GDPR 45
7 Discussion and Conclusion 47
References 50
Appendix 1: AI Act Compliance Checklist
Appendix 2: Machine Learning Model Development SOP

List of Abbreviations

AI	Artificial Intelligence
AI Act	Artificial Intelligence Act (EU Regulation 2024/1689)
BSI	British Standards Institution
CAPA	Corrective and Preventive Action
CE	Conformité Européenne (European Conformity)
DPIA	Data Protection Impact Assessment
EUDAMED	European Database on Medical Devices
GDPR	General Data Protection Regulation
GSPR	General Safety and Performance Requirements
IEC	International Electrotechnical Commission
IFU	Instructions for Use
ISO	International Organization for Standardization
IVD	In Vitro Diagnostic
IVDR	In Vitro Diagnostic Regulation (EU 2017/746)
KPI	Key Performance Indicator
MDCG	Medical Device Coordination Group
MDR	Medical Device Regulation (EU 2017/745)
MDSW	Medical Device Software
ML	Machine Learning
NB	Notified Body

PMPF	Post-Market Performance Follow-up
PMSR	Post-Market Surveillance Report
PRRC	Person Responsible for Regulatory Compliance
PSUR	Periodic Safety Update Report
QA	Quality Assurance
QMS	Quality Management System
RA	Regulatory Affairs
RAPS	Regulatory Affairs Professionals Society
SME	Small and Medium-sized Enterprise
SOUP	Software of Unknown Provenance
SOP	Standard Operating Procedure
TD	Technical Documentation
UDI	Unique Device Identifier

1 Introduction

The development and deployment of machine learning (ML) enabled software within the medical device sector represents a transformative advancement in healthcare technology. ML methods enable complex data analysis and predictive modeling. By utilizing these methods on medical data, it is possible to develop ML models and integrate them into medical devices. These devices can be used for different medical tasks, such as personalized diagnostics, early disease detection, and optimized clinical workflows (1). Consequently, medical devices utilizing ML technologies have expanded capabilities far beyond conventional diagnostic and therapeutic methods, promising enhanced precision, speed, and reliability in clinical decision-making processes (1).

However, the integration of ML into medical device software poses unique regulatory and operational challenges due to its dynamic and adaptive nature (2). In contrast to traditional software, ML models can be continuously re-trained with updated datasets, resulting in changes to their performance and behavior over the lifecycle (1). Such adaptations necessitate repeated verification and validation procedures to ensure accuracy, effectiveness, and safety (3). This creates challenges for regulatory frameworks like the EU Medical Device Regulation (MDR) and In Vitro Diagnostic Regulation (IVDR), which were originally designed for static software systems with infrequent updates. ML-based systems require additional activities, such as managing training data and validating evolving model performance. As a result, regulators are beginning to adapt their guidelines to address these dynamic technologies better (4).

In the European Union (EU), the regulatory environment governing IVD medical devices is structured around the In Vitro Diagnostic Regulation 2017/746. It imposes compliance obligations to ensure high standards of patient safety, device reliability, and clinical efficiency. IVDR-regulated devices require extensive clinical performance evaluations, robust risk management processes, clear traceability mechanisms, and continuous post-market surveillance. However, IVDR does not explicitly detail the regulatory treatment of adaptive

ML systems. The recent introduction of the European Artificial Intelligence Act (AI Act) addresses this gap by categorizing most ML-enabled medical devices as high-risk, thus mandating additional AI-specific requirements for manufacturers. The AI Act supplements the existing IVDR by providing guidance on risk management, data governance, transparency, accountability, and comprehensive technical documentation requirements specifically tailored for AI-enabled medical software.

Together, the IVDR and the AI Act create a layered and complex compliance environment for manufacturers developing ML-enabled medical devices. The ability of manufacturers to efficiently manage dual regulatory obligations is crucial for rapid market entry while ensuring device safety (5).

In response to these regulatory challenges, this thesis proposes a lifecycle management framework for ML-enabled medical device software, regulated under the IVDR and the AI Act.

This study adopts the constructive research method to analyze regulatory frameworks and to develop a lifecycle management framework for ML-enabled IVD software for the use of Hematoscope Oy. The structure of the thesis is as follows: Chapter 2 presents the research objectives that guide this work. Chapter 3 explains the methodology, including the constructive research approach and data sources. Chapter 4 offers a comparative analysis of the EU IVDR and the AI Act in the context of lifecycle management. Chapter 5 represents a comparative gap analysis between IVDR and AI Act. Chapter 6 proposes a practical lifecycle management framework tailored to ML-enabled IVD software. Chapter 7 discusses the findings, implications, and limitations. This chapter also concludes the thesis and suggests directions for future research. The appendices include supporting materials, such as the AI Act compliance checklist, a mapping table comparing IVDR and AI Act requirements, and a standard operating procedure (SOP) for developing machine learning models.

2 Research Objectives

This research aims to address the complex regulatory and technical challenges associated with the lifecycle management of ML-enabled IVD software devices. Given the overlapping scopes of the EU IVDR and the AI Act, this study has two objectives. The first objective is to compare the life cycle management approaches of the AI Act and the EU IVDR for class C devices. The aim is to identify how the regulations address the development, deployment, and post-market activities with a focus on ML-related activities. This comparison will help to uncover overlaps, gaps, and complementary elements in both frameworks, providing insights into how ML-enabled medical devices can be developed in compliance with both regulatory regimes.

Based on comparative analysis, the second objective is to develop a practical lifecycle management framework tailored for ML-enabled IVD software. The proposed framework aims to serve as a guideline for developers and regulatory professionals, supporting compliant, safe, and effective development and maintenance of ML-based software throughout its lifecycle.

3 Methods

This chapter presents the methodological approach employed in this study to investigate the regulatory challenges and develop a framework. Given the dual-regulatory landscape involving the IVDR and the EU AI Act, a constructive research method was selected to support both analytical exploration and the development of a practical framework. The chapter outlines the research design, data collection strategies, and regulatory analysis steps that underpin the framework presented in later sections.

3.1 Research Design: Constructive Research

This research employs a constructive research approach suitable for solving practical problems. The constructive research approach follows six distinct stages as follows (6):

1. Identify a practically relevant problem
2. Obtain a general and comprehensive understanding of the topic
3. Construct a solution or a framework
4. Demonstrate that the solution works
5. Evaluate the solution
6. Reflect on the theoretical and practical contribution

This thesis covers stages 1-3. Due to the early-stage development of the software of Hematoscope, stages 4-6, i.e., the demonstration, evaluation, and reflection stages, could not be fully implemented. However, the framework is conceptually adaptable to these follow-up stages and can serve as a basis for subsequent implementation and validation efforts.

To identify the problem, a systematic review of related work and the existing theoretical body of knowledge is required (Figure 1) (6). Together with professionals of Hematoscope and their existing knowledge, experience, and future needs, the problem of dual regulatory complexity of the AI Act and the IVDR was identified.

To obtain the required knowledge for solving the problem, the second stage included an extensive review of relevant regulations, harmonized standards, guidance documents, and scientific literature related to the AI Act and IVDR. These are covered in Chapter 4.

Stage three is a core element of the constructive research approach and the key contribution of this thesis. Here, the problem identification is transformed into a problem solution, utilizing the obtained theoretical framework (7). The identified regulatory challenge is addressed by constructing an integrated

lifecycle framework for ML-enabled IVD medical software. This process is covered in Chapters 5 and 6.

The resulting framework is designed to align with key regulatory expectations while also considering practical constraints observed in a real-world development context. The resulting solution aims to generate both practical utility by addressing real-world needs and epistemic utility by contributing to theoretical understanding. This dual output (Figure 1) leads to both practical relevance and theoretical relevance, which are central objectives of constructive research.

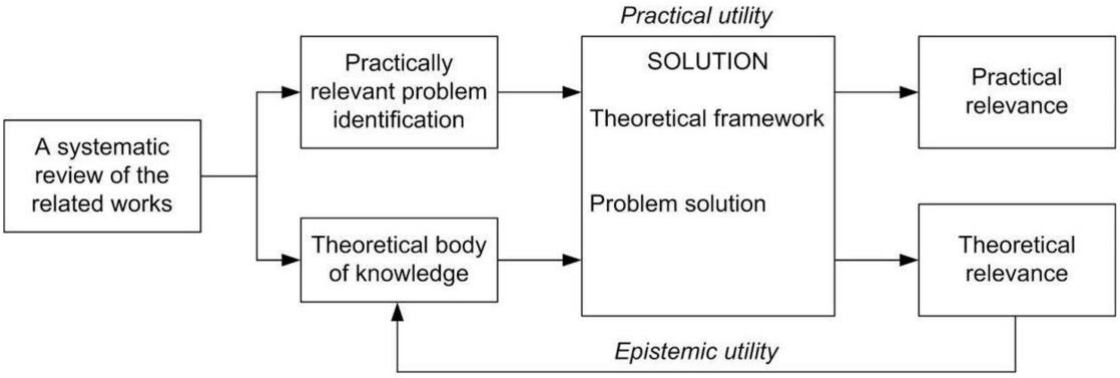

Figure 1. Stages of the Constructive Research Approach(8).

3.2 Regulatory Framework Analysis

The lifecycle management of machine learning-enabled IVD software must align with the European Union's In Vitro Diagnostic Medical Devices Regulation 2017/746. As a directly applicable regulation across EU Member States, the IVDR establishes a comprehensive and binding framework to ensure the safety, performance, and reliability of IVD medical devices, including software.

The regulation emphasizes several critical pillars, such as General Safety and Performance Requirements (GSPRs), risk management, evidence-based performance evaluation, and device classification rules. For IVD software,

compliance involves addressing specific provisions outlined in Annex I of the IVDR and relevant guidance documents issued by the Medical Device Coordination Group (MDCG), such as MDCG 2020-16 (qualification and classification of software) and MDCG 2021-24 (IVD classification examples).

IVDs are subject to classification rules that may place them in higher-risk categories, particularly when they inform critical clinical decisions. Consequently, Notified Bodies are involved in conformity assessments for Class B, C, and D devices. These assessments scrutinize technical documentation, clinical evidence, and risk controls.

The IVDR also indirectly addresses emerging technologies such as machine learning by requiring transparency, performance consistency, and post-market surveillance mechanisms capable of detecting performance drifts or safety concerns. While the regulation does not yet include detailed provisions specific to adaptive ML systems, the existing framework provides a solid foundation for managing risks through risk assessment, verification, validation, and change management.

This analysis focuses on EU IVDR to inform a compliant and robust lifecycle framework tailored to ML-enabled IVD software within the European regulatory landscape.

A structured alignment between regulatory requirements and lifecycle stages reveals how compliance must be maintained across development, deployment, and post-market phases. Table 1 below maps the key regulatory requirements of the IVDR to typical lifecycle stages in the development of ML-enabled IVD software. This alignment helps illustrate how compliance is maintained throughout the development and post-market lifecycle.

Table 1: EU IVDR Requirements with Lifecycle Stages of ML-Enabled IVD Software

Lifecycle Stage	IVDR Requirements & Analysis
1. Planning & Intended Purpose	Article 10 mandates a clear definition of intended purpose.
2. Design & Development	Requires integration of IEC 62304 compliant software lifecycle processes and risk management (ISO 14971). For ML models, this must include design inputs on training data quality and update mechanisms.
3. Risk Management	Continuous process per Annex I, integrating software-specific risks (e.g., algorithmic bias, cybersecurity). Class determination per Annex VIII is especially critical for ML models informing high-risk decisions.
4. Performance Evaluation	Annex XIII requires evidence of analytical, clinical, and scientific validity. ML models may need extensive real-world validation beyond static datasets to meet this obligation.
5. Technical Documentation	Annex II demands structured documentation, including algorithm description, Software development and validation performance metrics, and update traceability. Technical Documentation must clearly describe how the algorithm works, how the software was developed and validated, how its performance is measured, and how updates or changes are managed and tracked. For ML-based software robust change control is essential for maintaining safety and performance.
6. Conformity Assessment	Class B/C/D software requires Notified Body involvement. Algorithms affecting diagnosis and prognosis may enhance classification, triggering stricter conformity processes.

7. Post-Market Surveillance	Articles 78–81 emphasize the importance of real-world monitoring. For ML, this includes detecting performance drift and implementing algorithm updates under controlled, validated procedures.

3.3 Data Collection

A systematic and evidence-based method was employed to support the design of the ML-driven IVD software lifecycle framework. Regulatory compliance, patient safety, and flexibility in adjusting to new technologies were emphasized. Various data were gathered from regulatory authorities, scientific literature, and industry reports to serve as the foundation to conduct the gap analysis and build the framework.

Primary regulatory documents were reviewed, including the EU IVDR 2017/746 and the AI Act 2024/1689. Additional guidance was taken from the MDCG, the International Medical Device Regulators Forum (IMDRF), and the U.S. Food and Drug Administration (FDA) for contextual comparison. Nevertheless, EU-specific regulations were prioritized.

Academic publications were retrieved from databases such as PubMed and Science Direct using search terms including *“machine learning medical devices,” “IVDR compliance for AI,”* and *“adaptive algorithms regulation.”* Peer-reviewed studies published between 2020 and 2025 were selected, with priority given to those addressing lifecycle management and regulatory challenges in healthcare AI.

Industry perspectives were incorporated through the analysis of reports and white papers published by MedTech Europe, the Regulatory Affairs Professionals Society (RAPS), Open Regulatory, and the British Standards Institution (BSI). These materials were used to capture real-world insights and current regulatory expectations.

To validate the practical dimensions of the research, informal consultations were conducted with professionals in machine learning and regulatory affairs. Through the synthesis of regulatory guidelines, academic findings, and industry practices, a foundation was established for the regulatory gap analysis and the lifecycle framework presented in Chapter 6 and Appendices 1.

4 Regulatory and Standards Background

This chapter provides an overview of the regulatory frameworks and international standards that form the foundation for the development, approval, and lifecycle management of IVD software. The focus is on describing the core elements of the EU regulatory environment and harmonized standards relevant to the safe and effective development of software-based medical devices.

4.1 EU Legislation

The European Union is the highest regulatory authority governing the manufacturing and market entry of medical devices within its jurisdiction (9). The EU legislative framework operates at two levels: primary and secondary legislation. Primary legislation consists of EU treaties forming the legal foundation for all regulations, and secondary legislation includes directives and regulations that provide compliance requirements for specific purposes (10). In the medical device field, such requirements include MDR 2017/745 and IVDR 2017/746. Directives require transposition into national laws of member states, allowing flexibility in implementation. In contrast, regulations do not allow modifications and must be followed, as it is stated in all EU member states (11).

4.1.1 In Vitro Diagnostic Regulation (IVDR) 2017/746

The IVDR is the EU's main law for controlling the safety and performance of diagnostic medical devices, including software used to analyze blood, tissue, or

other samples(5). It applies to both traditional laboratory tests and modern digital tools, including ML-enabled software(12).

Under the IVDR, all devices are classified by risk(13). The manufacturer can self-certify low-risk products, which are Class A. However, most ML-based diagnostic software falls into higher-risk categories (Class B, C, or D), which require a review by an independent organization called a Notified Body(14). This process must be completed before the product can be sold and carry the CE mark, which shows that it meets EU safety and quality standards(15).

The regulation outlines detailed requirements that encompass the entire lifecycle of a device, from initial design and development to real-world use. These include:

- Proving that the device works safely and effectively (performance evaluation)
- Managing risks throughout development (risk management)
- Keeping clear technical documentation and records (traceability)
- Monitoring how the product performs after it is on the market (post-market surveillance)

Although the IVDR does not specifically mention artificial intelligence, many of its rules apply to ML-enabled devices, especially those related to software safety, quality management, and performance consistency(16). These requirements form the foundation for CE certification and are meant to protect both patients and users.

4.1.2 Overview of the AI Act and Its Impact on ML-Enabled IVD Software

The EU AI Act (Regulation EU 2024/1689) represents a significant step forward in regulating artificial intelligence systems across Europe. It is the first comprehensive regulatory framework of its kind in the world(17), introducing a risk-based classification approach. The AI Act classifies most ML-enabled IVD software as high-risk due to its impact on critical clinical decisions such as

diagnosis, treatment planning, or prognosis(16). Although not all IVD devices are considered high-risk, many fall under low risk (Class A).

The AI Act defines an AI system as any machine-based solution that can operate with varying levels of autonomy, interpret input data, and generate outputs, such as predictions, recommendations, or decisions, that can influence physical or virtual environments (18). ML-enabled IVD software fits this description, as it provides outputs that clinicians use to guide patient care (19).

To understand how this legislation works, it is necessary to examine its structure. The Act is divided into 13 chapters and several annexes, each covering a specific regulatory aspect, from general principles and risk categories to market oversight and enforcement mechanisms. Figure 2 provides a chapter-wise overview of the AI Act's layout. The image is taken from the Johner-Institute(20).

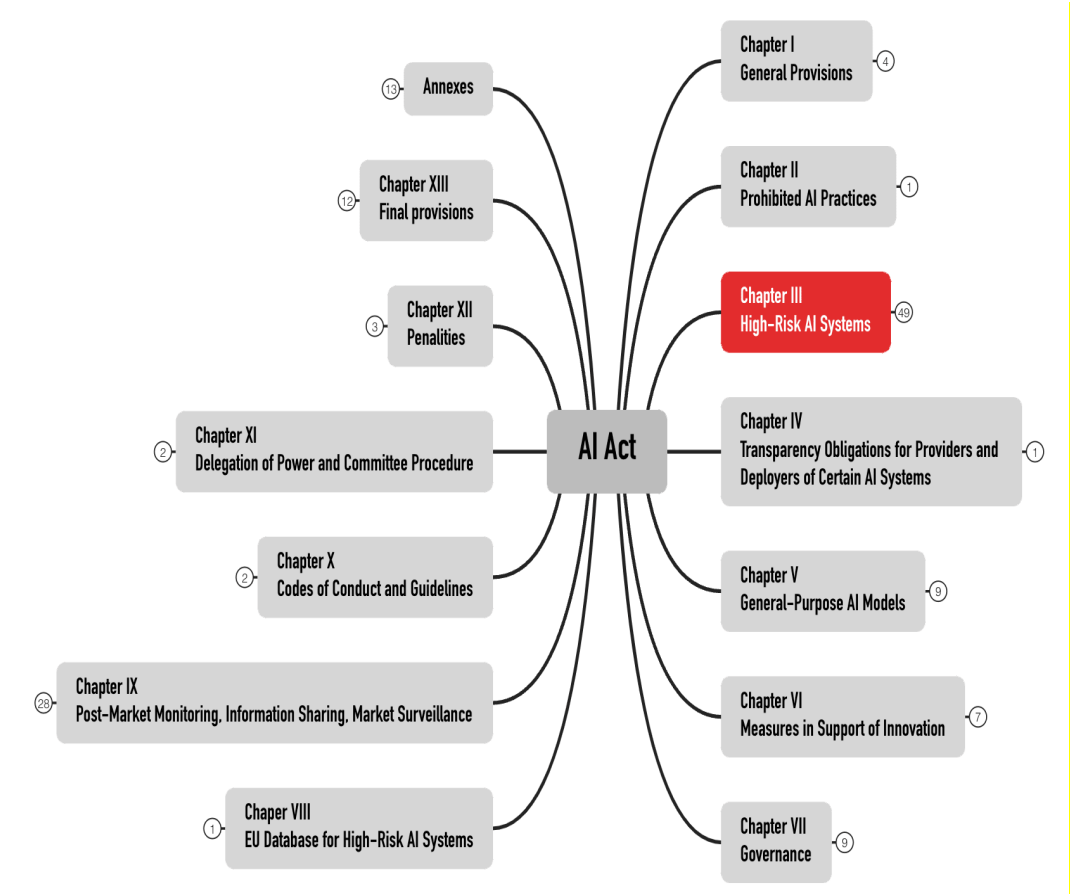

Figure 2. Chapter-wise structure of the EU Artificial Intelligence Act.

Among all sections, Chapter III on High-Risk AI Systems is significant in the context of this thesis and as well as medical device software. It outlines the main compliance requirements for these systems, including risk management processes, data governance, technical documentation, and quality management. Meeting the obligations outlined in Chapter III is crucial to ensuring the safety, reliability, and clinical appropriateness of AI in medical applications.

Chapters VIII and IX introduce obligations for maintaining a centralized EU database for high-risk AI systems and establish post-market monitoring responsibilities. These align closely with the IVDR's expectations for ongoing product oversight(20).

Chapters IV and V shift the focus toward transparency and the ethical deployment of AI. They address general-purpose AI models and emphasize the importance of explainability and human accountability, critical considerations for healthcare, where users must be able to interpret and trust AI outputs.

Chapters X through XIII support the broader infrastructure of the Act, covering procedures like codes of conduct, governance models, enforcement practices, and final provisions. Together, these chapters help create a consistent and harmonized AI regulatory landscape across the EU.

Figure 3 illustrates how Chapter III is structured. The image is taken from the Johner-Institute(20). It consists of five sections and 49 articles, which together define the complete set of compliance responsibilities for developers of high-risk AI systems.

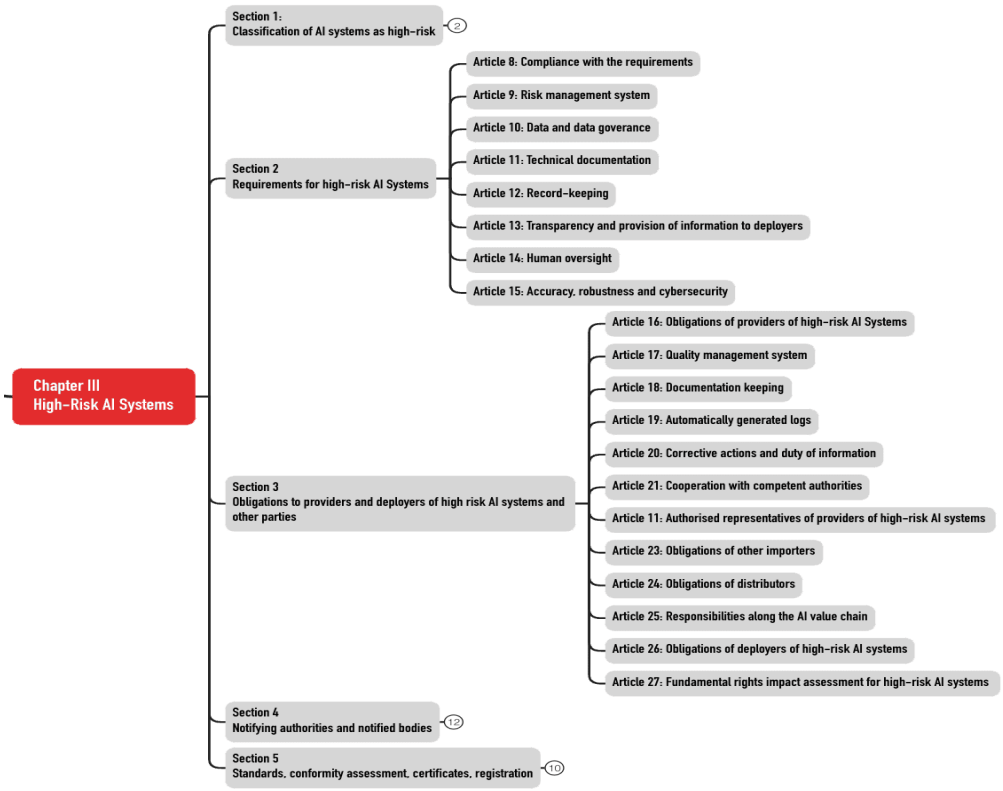

Figure 3. Overview of key compliance requirements under Chapter III of the AI Act.

The AI Act divides AI systems into four risk categories: unacceptable, high, limited, and minimal. Most ML-based IVDs fall into the high-risk category, as stated in Annex II and Article 6, because their outputs can significantly impact health outcomes (21). High-risk systems face the strictest compliance standards, while the other categories carry fewer or no binding requirements (20).

For high-risk systems, manufacturers must meet several key requirements. These include setting up an AI-specific quality management system, preparing detailed technical documentation, establishing continuous monitoring processes, and implementing robust data governance practices (21,22).

Importantly, companies that already have a quality management system based on ISO 13485 for IVDR or MDR compliance do not need to create a completely new system (23). However, they must adapt and extend their existing QMS to incorporate the additional AI-specific requirements outlined in the AI Act. All

these steps are essential to ensure that the software functions accurately, transparently, and ethically (24).

One of the Act's foundational principles is data governance. The law mandates the use of high-quality, representative, and unbiased datasets during development. This is particularly critical in the IVD space, where poor or biased data could lead to incorrect or unfair outcomes. As a result, ML models used in IVD software must be designed to be both interpretable and explainable, allowing healthcare professionals and regulators to understand and trust the process by which decisions are made.

When an ML-enabled IVD software is classified as high-risk under the AI Act, based on its intended purpose and impact on health decisions, it must comply with strict obligations outlined in Chapter III. These include requirements related to risk management, technical documentation, data governance, and post-market monitoring(12).

Notably, the AI Act applies not just to companies based within the EU but also to any provider whose AI systems affect users in the EU(17). This extraterritorial scope ensures that all systems used within the EU must follow the same rules. To help simplify compliance, manufacturers are allowed to incorporate AI Act requirements into their existing ISO 13485 compliant QMS.

Although the Act supports innovation, it also introduces a degree of regulatory complexity, especially for startups and small or medium-sized companies. These businesses may struggle with the combined demands of both the IVDR and AI Act requirements.

The Act officially came into effect on 1 August 2024. Several deadlines have been established, like prohibited AI systems must be withdrawn by February 2025, obligations for general-purpose AI models come into force in August 2025, and full compliance for high-risk systems becomes mandatory by August 2026 (Annex III) and August 2027 (Annex I).

To comply fully, manufacturers of ML-enabled IVD software must ensure they follow AI-specific regulations across the entire product lifecycle, from initial design to development, validation, deployment, and post-market monitoring.

Table 2 below summarizes the core AI Act requirements for ML-based IVD systems, linking each requirement to its legal basis and explaining its application.

Table 2. Key AI Act Requirements Applicable to ML-Enabled IVD Software

Obligation	**Relevant Article(s)**	**Application to IVD Software**
Risk Classification	Art. 6, Annex II	ML-enabled IVDs are often classified as high-risk AI systems it can significantly influence clinical decision.
Risk Management System	Art. 9	Comprehensive risk management processes must be implemented
Data Governance	Art. 10	High-quality, representative, and unbiased datasets are required
Technical Documentation	Art. 11	Detailed documentation demonstrating compliance must be maintained
Transparency and Human Oversight	Arts. 13 & 14	AI systems must be transparent and subject to human oversight
Post-Market Monitoring	Arts. 61 & 62	Continuous monitoring and incident reporting mechanisms are required
Extraterritorial Applicability	Art. 2	Non-EU providers must comply if their AI systems are used within the EU

4.2 International Standards for Life Cycle Management

A suite of harmonized international standards plays a critical role in supporting the lifecycle management of ML-enabled medical device software. These standards help ensure that products are developed, maintained, and monitored

in a way that is safe, effective, and secure. Many of these standards also align with the expectations set by the EU AI Act and the IVDR (21).

At the foundation are ISO 13485, which defines the quality management system requirements for medical devices, and ISO 14971, which focuses on systematic risk management throughout the device lifecycle. These two overarching standards set the structural backbone for software-specific requirements (25).

At the core of software development is IEC 62304, which establishes the life cycle requirements for medical device software, including development, maintenance, risk management, and problem resolution (26). This standard is supported by:

- EN 82304-1, which applies to health software intended to operate independently of hardware medical devices (27).
- IEC 62366-1, which guides usability engineering and user interface design to minimize use-related risks(28).

Cybersecurity is addressed by IEC 81001-5-1, which provides a structured approach for integrating cybersecurity processes across the software development lifecycle, and IEC TR 60601-4-5, which outlines technical requirements to ensure resilience against cybersecurity threats (29).

Figure 4 (generated by ChatGPT) below illustrates the interrelationship between key international standards that support the lifecycle management of medical device software. At the foundation are ISO 13485 and ISO 14971, which provide quality and risk management structures, respectively. These standards support IEC 62304, which governs the software lifecycle, and are complemented by usability and cybersecurity standards relevant to modern medical devices that utilize ML.

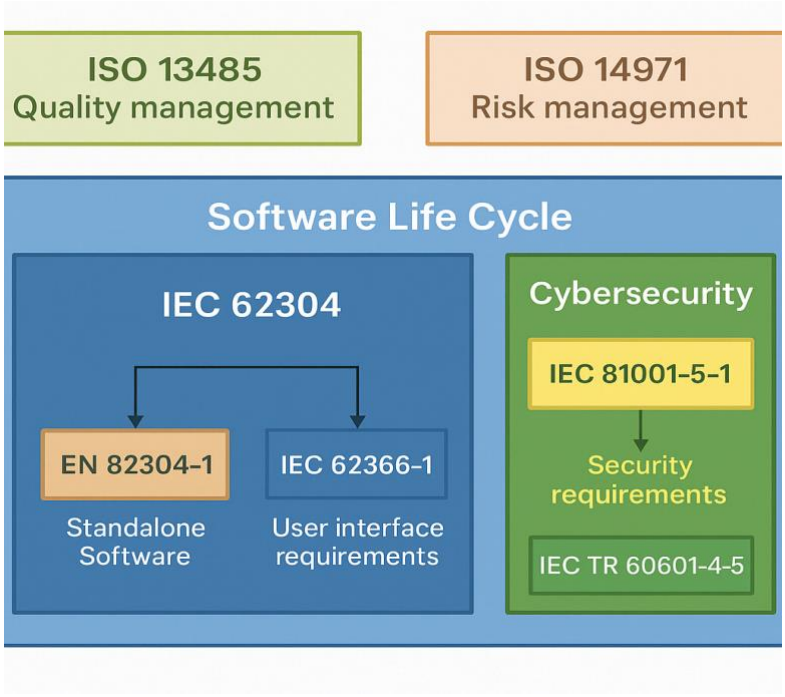

Figure 4. International standards supporting the lifecycle management of medical device software

4.2.1 ISO 13485 Quality Management System

ISO 13485:2016 is the main standard for quality management in the medical device industry. It helps organizations ensure their devices and services meet customer needs and comply with regulatory requirements. The standard is harmonized under both the IVDR and MDR, which means it plays a key role in getting CE certification in Europe (25)

In the context of IVD software, ISO 13485 plays a critical role in structuring internal processes to ensure that software development is controlled, validated, and auditable. It provides a framework for integrating other key standards, such as IEC 62304 (software lifecycle processes), ISO 14971, and IEC 81001-5-1, into the overarching QMS.

While ISO 13485 does not contain software-specific provisions, it acts as the structural backbone that mandates the implementation and documentation of compliant development practices(30). With the introduction of the AI Act, ISO

13485 can also support conformity by ensuring quality processes are in place for machine learning-enabled IVD software, especially in pre-market design controls and post-market surveillance (24).

4.2.2 ISO 14971: Risk Management

ISO 14971 is a globally recognized standard that provides a structured process for managing risk in the medical device industry. It guides manufacturers in identifying, assessing, controlling, and monitoring risks throughout a product's lifecycle (31). The standard applies to hardware and software-based medical devices and supports compliance with EU IVDR (32).

The core principle of ISO 14971 indicates hazard identification, risk assessment, and risk control measures to minimize the potential harm to the patient (31). This standard also provides guidelines for the risk mitigation of the environment. According to this standard, the risk is measured based on severity and probability. It also provides guidelines to mitigate all possible residual risks. The standard follows a life cycle approach, which requires the risk management activities to be incorporated from the initial design phase through to post-market surveillance (33). Even new risks identified after product release can be mitigated according to its continuous improvement process(32).

ISO 14971 is harmonised with IEC 62304, which mandates its application in medical device software development. A comprehensive risk documentation is required for implementing ISO 14971. It also requires hazard analysis, risk control measures, and benefit-risk assessment. Traceability is also important for the implementation of this standard. Traceability between risk management activities and software development processes is essential to demonstrate compliance. Challenges in applying the standard include defining acceptable risk levels, ensuring complete risk identification, and maintaining up to date risk documentation throughout the product life cycle(34).

4.2.3 IEC 62304: Software Lifecycle Processes

IEC 62304 is the main international standard that defines how to manage the software life cycle for medical device software. It provides a structured process for developing, maintaining, and managing risks in software used in medical devices. The standard supports regulatory compliance with both the EU IVDR and the FDA (35).

A key part of the standard is the risk-based classification of software:

- Class A: No risk of injury or damage to health.
- Class B: Possible non-serious injury.
- Class C: Risk of death or serious injury.

The higher the class, the more documentation, testing, and verification are required during development. IEC 62304 defines five major processes that must be followed throughout the software lifecycle. These are shown in Figure 5. The image is taken from sunstonepilot.com(36).

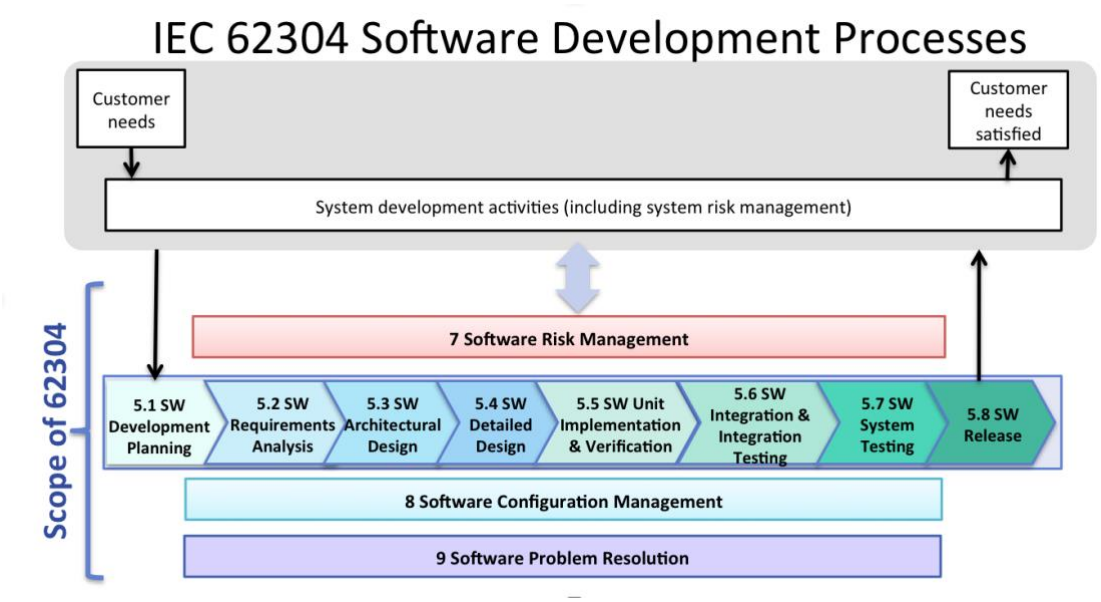

Figure 5. Key processes defined in IEC 62304

IEC 62304 demands extensive documentation, including a software development plan, risk management files, verification and validation reports, and configuration management records. Traceability ensures every software component is linked to risk management and testing.

4.3 Regulatory Pathways for ML-Enabled Medical Devices

Under IVDR, ML-enabled IVD software is categorized as Medical Device Software (MDSW) and is subject to classification rules outlined in Annex VIII, particularly Rules 1.4 and 3.3, which often lead to classification in Class C, especially when the software provides information used for diagnostic or therapeutic decisions with potential impact on patient health.

Key regulatory steps in the pathway include:

1. Risk-Based Classification:
   The risk class of the software determines the route to conformity. Most ML-based diagnostic tools fall under Class C due to their influence on critical clinical decisions. Class B and D classifications are also possible depending on the device's intended use and output risk level.
2. Conformity Assessment:
   For Class C devices, the manufacturer must engage a Notified Body to perform a conformity assessment. This includes an in-depth review of the technical documentation, QMS, and performance evaluation evidence.
3. General Safety and Performance Requirements (GSPRs):
   The device must fulfil the GSPRs listed in Annex I. These include requirements on safety, reliability, risk control, software lifecycle processes, and usability. For ML-enabled systems, this entails ensuring consistent performance, transparency, and a clear description of the intended use, algorithm logic, and limitations.
4. Performance Evaluation:
   Article 56 and Annex XIII mandate a robust performance evaluation, which includes:
   - Scientific Validity

- Analytical Performance
- Clinical Performance

  For ML models, evidence must demonstrate that the algorithm performs reliably across varied datasets and populations. Performance must be monitored throughout the product lifecycle.

5. Technical Documentation:
   Required under Annex II, this includes details of software design, development processes (aligned with IEC 62304), risk management (ISO 14971), and quality processes (ISO 13485). ML-specific documentation should also explain training data selection, validation strategy, and mechanisms for change control or retraining.
6. Post-Market Surveillance (PMS) and Vigilance:
   ML-based IVDs must include a PMS Plan and, for Class C devices, a Periodic Safety Update Report (PSUR). Continuous monitoring of software performance, especially for drift or bias, is essential. Mechanisms must also be in place for corrective actions and incident reporting.

IVDR does not yet include explicit guidelines tailored solely to adaptive ML algorithms. However, existing obligations under Annex I and the increasing guidance from MDCG documents (e.g., MDCG 2021-24, MDCG 2020-16) provide a foundation for regulatory alignment(37). Manufacturers must demonstrate that any learning system remains stable or is updated through a validated and traceable process(38).

Thus, the regulatory pathway for ML-enabled IVD software under IVDR is stringent, risk-based, and performance-driven. It demands early integration of compliance into design and continuous validation throughout the device's lifecycle.

4.4 The Need for a Lifecycle Management Framework

The complexities of machine learning-assisted IVD software are not yet fully covered by existing medical device regulations. Conventional regulatory

frameworks are based on static software, whose operation remains constant after approval(12,16). ML models, by contrast, can shift and adapt over time, requiring continuous reassessment to maintain ongoing safety, accuracy, and clinical appropriateness. Yet, such mechanisms to facilitate changes are not often clear in existing models, and it is uncertain how to validate updates or update documentation. Such deficiencies can hinder innovation and increase compliance risks.

Currently, the EU AI Act imposes strict requirements on algorithm transparency, traceability, and ongoing supervision. These obligations must be met in conjunction with IVDR measures in the areas of clinical performance, risk management, and quality(12,39). The developers are required to comply with both sets of obligations, in addition to ensuring that the products remain both safe and effective throughout their life cycle(40).

To meet these challenges, this thesis suggests an integrated life cycle management approach that harmonizes the essential requirements of the IVDR and the AI Act. These standards are underpinned by harmonized standards, such as ISO 14971 on risk management, IEC 62304 on software life cycle processes, and ISO 13485 on quality systems. By bringing together obligations applicable to AI with medical device rules and quality standards, the proposed framework provides a realistic approach for creating safe, transparent, and compliant ML-based diagnostics.

5 Results: Gap Analysis and Key Findings

This chapter presents the results of the regulatory gap analysis, focusing on the main compliance challenges. First, the limitations of the IVDR and the AI Act are identified. A comparison is then made between them, followed by a proposed strategy to help close these regulatory gaps.

5.1 Key Challenges in Regulatory Compliance

5.1.1 ML-Specific Guidance Gaps

The regulation of ML-enabled IVD software presents ongoing challenges, primarily due to its adaptive nature and the limited availability of tailored guidance within existing EU frameworks. While the IVDR offers a structured approach for traditional medical device software, it lacks specific provisions for managing the evolving behaviour of ML systems, particularly when algorithms are retrained, improved, or modified after market release(14).

In current practice, ML models are not automatically updated in real-time. Instead, updates are typically carried out by human developers through controlled processes. However, it remains unclear when such updates, such as retraining a model with new clinical data, constitute a “significant change” that requires renewed conformity assessment under IVDR Article 110(13).

Although the MDCG 2020-3 guidance outlines structured criteria for assessing software changes(41), it was not designed with adaptive ML systems in mind. The guidance lists factors such as changes to the intended purpose, input data, output performance, or algorithm logic as potential triggers for regulatory reassessment. Yet, it does not address common ML-specific scenarios, such as updates to training datasets, model weights, or learning parameters. This lack of alignment leaves developers uncertain about how to evaluate algorithm changes in practice(3,15).

PMS, required under the IVDR, is intended to monitor the ongoing performance and safety of IVD devices(14). However, there is currently no clear regulatory link between PMS findings and decisions about model updates or retraining in ML systems. For example, if field data shows performance drift, it is unclear whether retraining the model would trigger a new conformity process or be considered part of regular maintenance.

Together, these gaps create uncertainty for manufacturers developing ML-enabled IVD software. Without clear, harmonized guidance on how to assess, document, and act on model changes over time, companies face difficulty balancing regulatory compliance with the iterative nature of machine learning development.

5.1.2 Data Privacy and Security Concerns

ML-enabled medical software relies on large volumes of patient data, raising compliance challenges related to data protection and cybersecurity. The primary issues include:

1. In the EU, GDPR (Regulation 2016/679) imposes strict requirements for automated decision-making and data processing, which affect AI-based diagnostics. The GSPR under IVDR further mandates robust cybersecurity measures (IVDR, Annex I).

2. Continuous validation of AI models using real-world data is essential for performance monitoring. However, GDPR restricts how patient data may be used for algorithmic training, creating barriers to model retraining and improvement(42).

5.1.3 Transparency and Explainability Requirements

The European regulatory landscape places a strong emphasis on the transparency and explainability of ML algorithms used in IVD medical devices. Under the IVDR, manufacturers must ensure that the intended purpose, functioning, and limitations of their devices, including those enabled by ML, are clearly described in the technical documentation and the information provided to users(21). Transparency is critical not only for clinical users to make informed decisions but also for notified bodies and competent authorities to adequately assess device safety and performance.

Similarly, the EU AI Act further reinforces the need for transparency in AI systems, particularly for high-risk applications, which include ML-enabled IVDs(21). The AI Act mandates that high-risk AI systems must be accompanied by appropriate information to enable users to understand and interpret their outputs correctly. This includes clear instructions for use, information about the characteristics, capabilities, and limitations of the AI system, and the logic involved in its functioning, where possible.

Manufacturers must also anticipate scrutiny of these aspects during conformity assessment procedures. Failure to adequately address transparency requirements may hinder market approval or expose manufacturers to post-market challenges, including regulatory sanctions(20). Thus, lifecycle management frameworks for ML-enabled IVDs must embed transparency considerations from the earliest stages of development through post-market surveillance.

5.1.4 Validation and Verification of ML Algorithms

AI-based medical software presents unique validation and verification challenges, primarily due to the black-box nature of machine learning models(43). Regulatory bodies require AI systems to demonstrate reproducibility and explainability, but current frameworks lack standardized evaluation methods (41).

- Reliability and performance assurance: While most ML models used in regulated environments are deterministic, producing the same output for the same input, validating their performance across varied clinical scenarios remains a challenge. Unseen data distributions, data drift, or overfitting may not be detected by traditional testing alone(44).
- Explainability and interpretability requirements: The IVDR emphasizes the need for "explainable AI", but there are no universally accepted criteria to measure explainability in a regulatory context (45).

Table 3 summarizes the key challenges associated with validating and verifying ML-enabled IVD software under the current regulatory framework. As a result, manufacturers face challenges in proving the reliability of AI models in a way that meets current regulatory expectations.

Table 3. Key Validation and Verification Challenges for ML-Enabled IVD Software.

Challenge	Description	Regulatory Relevance
Lack of harmonized V&V standards	No widely accepted methodology exists for validating machine learning models in a regulated medical device context.	IVDR Annex I (GSPR), AI Act Chapter III
Black-box model complexity	The inner workings of complex ML models are often opaque, making it harder to evaluate performance and clinical safety.	AI Act Article 13 (Transparency), IVDR Annex I Section 16
Performance on unseen data	Deterministic models may still underperform when exposed to real-world or out-of-distribution data, posing patient safety risks.	IVDR Post-Market Surveillance (Articles 78–81)
Explainability requirements	Developers must demonstrate interpretability, but regulators do not define clear thresholds or metrics for what is "explainable."	AI Act Article 13, Recital 47; IVDR General Safety Requirements
Validation after updates	It is unclear how often or under what conditions retrained ML models must be revalidated or re-certified.	IVDR Article 110, MDCG 2020-3

5.1.5 Data Bias, Fairness, and Validation Requirements

Regulatory compliance also requires ensuring fairness in AI-driven diagnostic decisions. However, machine learning models can exhibit bias due to factors such as imbalances in training data or underrepresentation of specific patient groups. These issues can lead to unequal performance across demographics and risk undermining clinical trust and safety.

Currently, the IVDR does not provide explicit requirements for identifying or mitigating algorithmic bias. This leaves manufacturers without clear regulatory guidance on how to validate fairness or demonstrate that their models perform consistently across diverse patient populations. While fairness is implied within general safety and performance requirements, there is a lack of harmonized expectations or testing criteria.

As a result, developers must take a proactive approach by integrating subgroup analysis, representative data sampling, and performance audits into the validation process. These measures can help demonstrate compliance with ethical standards and reduce the risk of unintended harm, even in the absence of formal regulatory benchmarks.

5.2 Comparative Analysis IVDR Vs. AI Act

The comparative analysis between the IVDR and the EU AI Act elucidates critical intersections and divergences in regulatory expectations. Both regulatory frameworks prioritize patient safety and data integrity but address these issues through distinct regulatory pathways. The IVDR primarily focuses on ensuring the analytical and clinical performance, accuracy, and reliability of diagnostic devices. In contrast, the AI Act emphasizes transparency, accountability, and mitigating ethical risks associated with automated decision-making processes. Specifically, the AI Act introduces rigorous compliance protocols for high-risk AI systems, requiring explicit AI classification, robust technical documentation, and

detailed governance mechanisms that are complementary yet distinct from IVDR mandates.

5.3 Gap Analysis: IVDR Vs. AI Act

Although both the IVDR and the AI Act aim to ensure the safety and performance of digital health technologies, they differ significantly in scope, focus, and operational clarity, especially when applied to machine learning-enabled IVD software.

The IVDR provides a well-established foundation for regulating IVD medical devices, including diagnostic software. It includes detailed requirements for risk management, performance evaluation, technical documentation, and post-market surveillance (14,35). It lacks specific provisions on how to manage dynamic model updates, retraining, or algorithm transparency (15).
In contrast, the AI Act introduces obligations tailored to high-risk AI systems, including requirements for data governance, transparency, human oversight, and continuous risk management. While these provisions address AI-specific concerns, they are often described at a high level and lack the implementation detail needed for medical device developers, particularly in regulated domains like in vitro diagnostics (21).

The key regulatory gap lies in the lack of integration between these two frameworks. The IVDR outlines what must be documented and validated, while the AI Act focuses on how AI systems should be governed and monitored. Yet, neither regulation clearly defines how responsibilities like retraining, explainability, or real world performance monitoring should be handled across both domains (15).

This fragmentation creates challenges for developers, who must interpret overlapping requirements without clear guidance on how to synchronize them. For instance, the AI Act expects continuous oversight and quality control, but IVDR conformity is based on fixed technical documentation and predefined

software functions. There is no harmonized method for managing software updates that may affect both performance and regulatory status(46).
By identifying these disconnects, a structured gap analysis is presented in Appendix 1.

5.4 Bridging the Gap Between the AI Act and IVDR

Bridging the identified regulatory gaps between the AI Act and IVDR requires a structured approach explicitly designed to harmonize the unique demands of AI systems within the existing regulatory framework for medical diagnostic devices. Addressing these gaps is not merely a matter of regulatory compliance; it is essential for fostering innovation, ensuring patient safety, and maintaining public trust in emerging AI technologies used in medical applications. To effectively bridge these gaps, several specific strategies are proposed.

5.4.1 Unified Classification and Risk Management Approach

It’s important to have a strong system for classifying and managing risks that cover both IVDR requirements (like analytical and clinical risks) and the AI Act’s categories of AI-related risk. To do this, AI-specific checklists should be added to existing IVDR risk management processes. This combined approach helps ensure a thorough and consistent way to assess risks, reduces the chance of missing regulatory issues, and makes the compliance process smoother.

5.4.2 Enhanced Technical Documentation and Traceability

A significant gap highlighted between IVDR and the AI Act relates to the documentation and transparency standards. Bridging this requires establishing a harmonized technical documentation protocol that explicitly includes AI Act requirements for transparency and accountability alongside IVDR standards for performance and clinical validation. Creating unified documentation templates that reference both regulations ensures traceability, simplifies regulatory audits

and facilitates smoother market entry processes for AI-enabled medical devices.

5.4.3 Proactive Ethical Compliance and Governance Mechanisms

To address the AI Act's explicit prohibitions on manipulative, exploitative, or discriminatory AI practices, which are currently unaddressed by IVDR, comprehensive governance mechanisms must be introduced. These mechanisms involve regular ethical risk assessments, algorithmic transparency audits, and detailed bias evaluations integrated into existing medical device lifecycle management processes. Explicit documentation and reporting of these assessments will fulfil AI Act-specific ethical compliance while complementing IVDR's focus on patient safety and effectiveness.

5.4.4 Lifecycle Management Integration

An integrated lifecycle management approach is essential to bridge regulatory differences effectively. This involves explicitly embedding AI Act-specific elements such as robustness evaluations, system transparency, and non-discriminatory practices within the existing IVDR-oriented lifecycle management frameworks guided by IEC 62304 and ISO 13485 standards. Such integration will facilitate regulatory coherence, reduce redundant compliance activities, and enhance operational efficiencies.

5.4.5 Continuous Regulatory Surveillance and Update Mechanism

Finally, establishing a continuous surveillance and update mechanism to monitor changes in regulatory expectations, standards, and best practices under both IVDR and the AI Act is critical. A dedicated monitoring protocol will proactively identify and respond to evolving regulatory requirements, thereby maintaining compliance and reducing risk throughout the product lifecycle.

Through these structured and integrated strategies, the identified regulatory gaps between the AI Act and IVDR can be systematically addressed, providing

clarity, ensuring compliance, and enabling the innovative potential of AI-driven medical diagnostic solutions.

6 Proposed Lifecycle Management Framework

This chapter presents a structured lifecycle management framework for ML-enabled IVD software, developed to address regulatory requirements under the IVDR and the AI Act. The goal is to support manufacturers in staying compliant with regulations while also keeping up with the changing technical and clinical needs.

6.1 Framework Objectives and Design Principles

The proposed lifecycle management framework is built to meet two core objectives:

- Ensure safety, performance, and regulatory compliance throughout the lifecycle of ML-based diagnostic tools.
- Enable continuous learning and improvement of models in response to real-world data, without compromising compliance or patient safety.

To achieve these aims, the framework is grounded in five design principles:

1. **Regulatory-by-Design:** Compliance with the IVDR and AI Act is integrated from the earliest stages of development. Requirements from ISO 13485, IEC 62304, and ISO 14971 are embedded throughout the process, not added retrospectively. For example, data collection must meet the AI Act's data governance standards (Article 10), while model reliability and state-of-the-art practices follow IVDR expectations. This ensures that all design outputs are audit-ready and aligned with GSPRs.
2. **Risk Management and Human-Centric Design:** Following ISO 14971 and AI Act Article 9, risk management is continuous and proactive. The framework emphasizes the identification of ML-specific risks such as data

bias, model drift, or misclassification. In line with AI Act Article 14, human-centric design is prioritized. AI must assist "not replace" clinical decision-making. Features like interpretable outputs and override options ensure that human oversight remains in control of the diagnostic process.

3. **Lifecycle Traceability and Iterative Improvement:** Traceability is enforced at every step from requirements through to verification. Each model version is linked to specific training data and test outcomes. Drawing from CRISP-DM and the V-model, the framework supports controlled iteration: post-market data can be used for model improvement, but only through defined change control and revalidation under IEC 62304. This enables adaptive learning while upholding medical device safety standards.
4. **Transparency and Accountability:** Transparency is both a regulatory and ethical imperative. In line with AI Act Article 13, users and regulators must be able to understand how the model works and how decisions are made. Development choices, such as data sources, model architecture, and performance thresholds, are documented with clear justifications. Responsibilities are clearly defined, including the role of the PRRC, as required by IVDR Article 15.
5. **Integration with Quality Management System (QMS):** The framework is embedded within the company's ISO 13485 QMS. It complements existing design and development processes, rather than operating in isolation. Standard procedures such as design reviews, verification planning, and change control are extended to include AI-specific aspects like data bias review and ML model validation. This alignment ensures that teams across functions, regulatory, software, and data science, work within a unified, compliant process.

6.2 Integration of IVDR and AI Act Requirements

Developing an ML-enabled medical device necessitates navigating two regulatory regimes: the IVDR, which governs the device's safety, performance, and technical documentation, and the EU AI Act, which imposes specific

obligations on the AI system within the device. In this framework, critical requirements from both regulations are mapped into the lifecycle phases, ensuring that compliance is built into the development process from the beginning.

6.2.1 Mapping IVDR (Annex I & II) into the ML lifecycle

The IVDR provides structured requirements for software-based medical devices, including ML components. The framework aligns the following elements with lifecycle activities.

1. **Intended Purpose and Classification**: During early design, the ML component's intended use is clearly defined. The device is classified under IVDR, and a formal intended purpose statement is created. This guides data selection and performance target definition in compliance with IVDR Annex II.
2. **General Safety and Performance Requirements (GSPRs)**: Annex I specifies GSPRs. For software, Sections 16.1 and 16.2 are key: they require repeatability, reliability, and state-of-the-art performance. These are implemented through version control, documented workflows (aligned with IEC 62304), independent dataset verification, and clinical validation. GSPRs are translated into concrete design inputs and verification criteria.
3. **Technical Documentation**: Annex II outlines documentation requirements, including device description, risk controls, testing evidence, and performance results. Under this framework, documentation is generated continuously throughout development, eliminating the need for retroactive compilation and ensuring readiness for conformity assessment.
4. **Conformity Assessment and Change Management**: Formal review gates (e.g., Design Release Review) confirm completion of technical documentation and performance claims before CE declaration. IVDR's PMS and vigilance requirements govern post-market activities. Software

update protocols follow MDCG 2020-3, ensuring that changes are evaluated for their significance and potential impact on compliance.

6.2.2 Mapping AI Act (Articles 10–15) into the ML lifecycle

The AI Act introduces specific requirements for high-risk AI systems, which include most ML-enabled IVD software. In this framework, these requirements are woven into the development process to ensure they are addressed proactively, not added on afterward.

1. **Data Governance (Article 10)**
   The framework emphasizes good data practices from the start. Training and testing datasets are reviewed for quality, representativeness, and fairness. Any known limitations, such as demographic imbalances or missing data, are documented. These steps help avoid unintended bias and ensure the AI model learns from relevant and trustworthy data.
2. **Technical Documentation (Article 11)**
   AI-specific documentation, such as how the algorithm was trained, what assumptions were made, and what risks were identified, is compiled as part of the main technical file. This documentation supports transparency and provides the basis for future audits or conformity assessments.
3. **Record-Keeping (Article 12)**
   Logs are incorporated into the system to trace input-output behavior, training sessions, and other important actions. Such logs enable traceability throughout the life of the product and provide insight into how the system acts in various contexts, especially during post-market monitoring.
4. **Transparency to Users (Article 13)**
   The system’s purpose, capabilities, and limitations are explained clearly in the user instructions. Wherever possible, basic details about how the model works are included, for example, what kind of data it uses or how it makes predictions. This helps end users (like clinicians) make informed decisions when relying on the AI.

5. **Human Oversight (Article 14)**
   Human supervision is built into the design. Features such as warnings, override buttons, and interpretable outputs allow users to step in if something looks odd. The system is tested to make sure users understand the outputs and know how to respond appropriately.
6. **Accuracy, Robustness, and Cybersecurity (Article 15)**
   Before release, the model is tested under a variety of conditions to check how stable and accurate it is. This includes scenarios with noisy or borderline data. Cybersecurity risks are also assessed, especially if the software connects to a network or handles sensitive patient information.

Where differences exist between the IVDR and AI Act, for instance, between the AI Act's encouragement of continuous learning and IVDR's requirement for stability, the framework resolves the conflict by treating each updated model as a new device version subject to complete validation and review. Similarly, concerns about data access under the AI Act are addressed by maintaining structured archives and privacy-compliant traceability logs, which can be made available for inspection when required.

In this integrated approach, compliance with both the IVDR and the AI Act is embedded into a unified development process. Each lifecycle phase addresses dual regulatory requirements simultaneously. This reduces duplication, streamlines conformity assessment preparation, and ensures that AI-enabled diagnostic tools are safe and compliant with legal requirements.

6.3 Model Development and Validation Process

To make the framework more concrete, this section provides a walkthrough of the main stages of the ML lifecycle, highlighting how each aligns with design control and regulatory checkpoints.

6.3.1 Define Intended Purpose and Risk Classification

The ML model's intended clinical or diagnostic function is first defined, and its regulatory classification is determined. For this use case (Hematoscope), the device has been classified as Class C under the IVDR, necessitating Notified Body involvement. Simultaneously, the software qualifies as a high-risk AI system under the AI Act due to its IVD designation. Establishing these classifications at the outset ensures that the appropriate regulatory context is applied from the beginning. This stage produces an Intended Use statement and a regulatory plan, including notification to a Notified Body and allocation of resources for compliance.

6.3.2 Data Acquisition and Preparation

This phase involves collecting and preparing training, validation, and test datasets. Data governance requirements (AI Act Article 10), privacy laws (e.g., GDPR), and IVDR Annex I provisions are addressed by ensuring that data is ethically sourced, anonymised where necessary, and representative of the target population. Documentation, such as data quality assessments and dataset descriptions, is generated to support technical documentation. Where required, a Data Protection Impact Assessment (DPIA) is conducted. By formally establishing this phase, early risk mitigation is enabled through structured data validation and bias analysis.

6.3.3 Model Design and Training

The ML model is designed (including algorithm and architecture selection) and trained using the designated dataset. All design inputs, including regulatory and user requirements, are tracked throughout the development process. Justifications for algorithm selection are linked to state-of-the-art standards, as required by the IVDR and AI Act. Version control and secure coding practices, as outlined in IEC 62304, are followed. Once trained, initial performance metrics

are recorded and reviewed in an internal checkpoint before advancing to verification.

6.3.4 Verification and Validation

Validation activities ensure the model meets clinical expectations, including robustness, reliability, and cybersecurity (e.g., through penetration testing or adversarial robustness checks). Domain experts review performance to ensure clinical adequacy. If performance thresholds are unmet, the model is returned to prior phases for refinement. A formal Design Verification Review is conducted to confirm that the design outputs match the design inputs, thereby fulfilling the AI Act Article 15 requirements.

6.3.5 Human Factors Validation and Oversight Testing

Following verification, a dedicated human factors validation is conducted, as per AI Act Article 14 and IEC 62366. Representative users interact with the system in simulated environments to assess ease of use, interpretability, and the potential for use errors. Misunderstandings or reliance issues are documented and addressed through design refinements or updates to instructions. A Usability Validation Report and an updated IFU are generated as key outputs. This phase ensures that the product not only performs technically but is safe and effective for real-world use.

6.3.6 Technical Documentation and Final Design Review

All development outputs are consolidated into the technical documentation, as per IVDR Annex II and AI Act Annex IV. Risk management summaries are updated, and a draft Declaration of Conformity is prepared. A final design review meeting is held, including senior management and the PRRC, to assess full readiness. Documentation completeness, QMS conformance, and support structures are evaluated. Approval at this stage permits the product to proceed toward regulatory submission.

6.3.7 Regulatory Submission and Notified Body Review

Following internal approval, the technical documentation is submitted to the Notified Body for conformity assessment, as required for Class C IVDs. The framework accommodates the time and effort needed for this review. Any deficiencies identified by the Notified Body are addressed through additional testing or documentation. If required under the AI Act, the system may also be registered or certified through the appropriate AI compliance channels. Regulatory approvals must be secured before market release.

6.3.8 Deployment and Release

Upon receiving regulatory clearance, the model is deployed to production and released to users. Training materials and support plans are distributed alongside the product. The release marks the transition to the operational monitoring phase. A PMS plan, prepared earlier, now becomes active. Quality systems transition to the production and post-production phases, triggering monitoring and issue-reporting mechanisms.

6.4 Post-Market Monitoring and Feedback Loop

After deployment, the framework continues through a robust PMS process, combined with a structured feedback loop. These mechanisms ensure the ongoing safety, effectiveness, and compliance in real-world use.

Unlike traditional surveillance that relies mostly on user-reported complaints, this approach includes active performance monitoring tailored to AI systems. Key performance indicators (KPIs), such as false negatives, user overrides, or discrepancies with confirmatory tests, are tracked. Where permitted, telemetry data is collected to monitor how the system performs across different settings and populations. Local installations may send de-identified summaries; cloud-based systems can report directly. These metrics help detect performance drift, input distribution shifts, and emerging edge cases.

Alongside quantitative data, the framework gathers qualitative user feedback. This is done through surveys, embedded feedback forms, or clinical working groups. Even if issues do not qualify as reportable incidents, they may reveal usability concerns, like unclear AI explanations or alert fatigue. These inputs are reviewed regularly and may lead to updates in user guidance, interface design, or training materials without altering the model itself.

All activities are conducted within a formal PMS system, as required by the IVDR. A written PMS plan defines what data will be collected, how often it will be analyzed, and who is responsible. A Post-Market Performance Follow-Up (PMPF) plan is also developed. This may involve running the model on new clinical data periodically to confirm it still meets performance expectations. If serious incidents occur, such as a harmful misdiagnosis, investigations are launched, and reports are submitted within the required regulatory timelines (typically 15 days). Depending on the findings, corrective actions may include model retraining, software patches, or updated user training.

Class C devices require PSURs, typically submitted annually. These summarize safety trends, model performance, benefit-risk assessments, and any corrective measures taken. Real-world data collected through active monitoring form the evidence base for these reports.

A distinctive feature of this framework is its formal link between post-market findings and model updates**.** A multidisciplinary AI Lifecycle Committee, usually including data scientists, clinicians, and regulatory staff, reviews PMS data at set intervals. If performance issues or guideline changes are detected, they may recommend a model update. Each proposed change is evaluated to determine whether it qualifies as a significant change under MDCG 2020-3, and if so, it triggers a new validation cycle.

The feedback loop also extends to how the AI is used in practice. Follow-up studies and user interviews may reveal over-reliance or underuse of the system. For instance, if users are ignoring AI suggestions, further training may be

needed. If users rely too heavily on incorrect outputs, design refinements or clearer alerts may be implemented.

Over time, post-market data also builds a body of real-world evidence**,** which can support future regulatory submissions. For example, when expanding indications or avoiding additional clinical trials. This evidence helps demonstrate ongoing control over the system’s performance and risk.

The loop is only complete when feedback is acted upon, changes are implemented, and outcomes are monitored again. This cycle aligns with ISO 13485’s continuous improvement philosophy and ensures that each iteration of the AI system is safer, smarter, and more usable than the last.

6.5 Adaptive System Lifecycle and Retraining Protocols

Machine learning introduces new challenges to traditional medical device lifecycles. Unlike fixed software, ML models can improve when retrained on new data. This framework supports such updates, but only through a controlled, traceable, and regulation-aligned process.

6.5.1 Locked vs. Learning Approach

There is a "locked model" approach being employed, which means the deployed model does not alter unless it is been updated according to a formal process. It guarantees consistent performance, makes it simple to manage changes, and aids in regulatory compliance. While the deployed model stays fixed, internal learning continues, and new data is regularly collected and analyzed to guide future updates when needed. When significant improvements are identified, a new development cycle is initiated, allowing for enhancements to be introduced through planned updates while preserving regulatory compliance.

6.5.2 Triggers for Retraining or Updating the Model

The framework defines clear conditions under which a model update is initiated:

1. **Performance Degradation:** If post-market monitoring reveals a decline in performance (e.g., reduced sensitivity), the cause is investigated. Where retraining is appropriate, a new development cycle is initiated.

2. **Availability of New Data:** When relevant new datasets are acquired that could enhance model performance, retraining may be considered. Updates are performed in defined intervals rather than continuously. It will minimize disruption and regulatory complexity.

3. **Expanded Use Cases:** If the model is intended for a broader application (e.g., a new clinical indication or setting), it is treated as a new project. Existing model components may be reused, but a full regulatory review is performed.

4. **Regulatory or Safety-Driven Changes:** If updates are mandated by authorities or triggered by safety concerns, prompt action is taken. The traceability of previous training and validation enables efficient, controlled retraining.

6.5.3 Retraining Process

When updates are justified, the following retraining process is followed:

1. New data are collected and, where appropriate, combined with existing datasets to prevent loss of prior knowledge.

2. System requirements and risk analysis are updated to reflect new inputs or intended uses.

3. The model is retrained either from scratch or by adapting the previous version, and all iterations are version-controlled.

4. Performance metrics are reviewed and compared against previous versions. If any performance degradation is identified, further tuning or revision is initiated.

6.5.4 Regulatory Submissions

For each model update, its regulatory impact is assessed. Regulatory Affairs Specialist determines whether authorities must be notified or whether updated documentation must be submitted. Under IVDR, substantial changes may necessitate renewed Notified Body assessment. The required documentation is updated accordingly.

In cases where the AI Act applies, the conformity assessment may need revision, including re-registration of the system. Change records are maintained within the document control system to ensure traceability and audit readiness.

6.5.5 Update Strategy: Periodic Rather Than Continuous

To reconcile freedom of adaptation with regulatory compliance, updates are being planned at intervals instead of being delivered continuously. It minimizes disruption to users and prevents unnecessary regulatory submissions. Each update is systematically validated to meet safety and performance requirements. In critical instances, such as safety-related problems, updates can be expedited through the Corrective and Preventive Action process.

6.5.6 Version Control and Change Tracking

All model versions and associated training data are archived. A comprehensive Model History File is maintained, documenting datasets used, performance achieved, and differences between model iterations. This documentation

supports both internal review and external audits and satisfies AI Act traceability requirements.

6.5.7 User Communication and Support

When the updated version is released, users are notified through official communication means. The IFU or safety notices are updated as needed. User support is made available to inform users about any changes, particularly when the behaviour of the model has been changed. For instance, when the updated version enhances detection levels, user instructions are provided accordingly.

6.6 Ensuring Transparency and Algorithm Explainability

Transparency and explainability are essential components of trustworthy and effective AI in healthcare. This proposal integrates these principles across all stages of development and deployment, addressing three key domains: regulatory transparency, user-facing transparency, and internal technical transparency. The goal is to ensure that the ML model does not function as a "black box" but as a clinically usable, auditable, and understandable system.

Regulatory transparency is ensured through comprehensive technical documentation. The technical file includes detailed descriptions of the model’s architecture, training approach, performance metrics, limitations, and rationale for design decisions. For instance, the choice of model architecture is explicitly justified based on clinical relevance, interpretability, or supporting literature. The demographic characteristics of the training dataset and any strategies used to mitigate algorithmic bias are also documented. This level of detail supports compliance with the AI Act and the IVDR and serves as a record for internal teams, external auditors, and regulators. When questions arise, such as how the model accounts for a specific patient factor, answers are traceable in the documentation.

User-facing transparency is achieved through features embedded in the software interface, enabling clinicians to interpret the AI's outputs confidently and safely. These features include:

Confidence Scores: The output is accompanied by a confidence value whenever possible. Rather than simply presenting a binary result, the software may display "Positive (97% confidence)," helping users assess result certainty and determine when additional scrutiny is necessary.

Explanation of Contributing Factors: When supported by the model, the interface may display which features influenced the decision. Techniques such as SHAP (Shapley Additive Explanations) or feature importance analysis are used to highlight relevant inputs. It will help clinicians understand how the model reached a conclusion.

User Guidance and Training Materials: The IFU includes detailed interpretation guidance. For example, it may explain that a "Positive" result means a specific pattern was detected with high confidence and that confirmatory testing is recommended in ambiguous cases. During onboarding, clinicians are introduced to the model’s principles, such as its reliance on statistical correlations rather than medical reasoning and what types of input or conditions may affect its output. This instruction improves user understanding and builds trust.

Visual and Interactive Explanations: In models using image or graphical data, visual overlays such as heatmaps or highlight indicators may be used to show which regions or features contributed most to a result. For tabular data, graphical plots or highlighted lab values can serve a similar function. These tools enhance user familiarity with the model’s decision patterns over time.

The framework recognizes that more complex models, such as deep neural networks, often yield better performance but can be difficult to interpret. During the design review process, developers assess whether a simpler, more interpretable model, such as logistic regression, might be sufficient. If a complex

model is used, external tools are integrated to generate approximate explanations that can be reviewed during audits or presented to users.

Uncertainty is handled explicitly within the framework. When the model is unsure, the system displays a warning, such as “Low confidence – retest recommended,” to prevent blind reliance. These alerts are included during model design and user interface development to promote responsible decision-making. When errors or unexpected behavior occur, the system supports root cause analysis through comprehensive logging. Each AI output is logged along with input context and model confidence. This makes it possible to trace back decisions and determine whether any user misunderstanding or system weakness was involved. In cases where the issue could affect patient safety, Field Safety Notices may be issued, and corrective measures are initiated. The process is transparent and designed to build user trust.

The framework also addresses ethical transparency. In jurisdictions where it is required, patients are informed that ML was involved in generating a diagnostic suggestion or treatment recommendation. Clinicians are encouraged to explain this clearly, such as stating: “A computer system trained on past cases helped analyze your results and suggested this outcome based on similar patterns.” This will ensure that the AI system supports safe and informed use in clinical settings and helps to reinforce the essential role of human judgment in AI-assisted healthcare.

6.7 Addressing Data Privacy and Security Compliance with GDPR

Data privacy and security are fundamental requirements in designing ML systems in the health sector. The framework embeds these values throughout the entire life cycle of the ML-based device to meet the European Union's GDPR. This is not merely compliance with the law, it's about gaining public confidence and living up to the expectation of ethics in the clinical setting.

Privacy by design in accordance with Article 25 of the GDPR is implemented from the very beginning of the development process. It is initiated at the data

acquisition level and extends to model training, deployment, and post-market surveillance. Special attention is given in each step of data handling to make sure that personal data is gathered, processed, and stored lawfully, in a fair manner, and transparently.

When data is collected or prepared to train and validate the models, personal data is pseudonymized or anonymized wherever possible. This might mean eliminating explicit identifiers such as name or patient number and substituting them with random codes. Where data such as sex or age is needed to support the operation of the models, it is only provided in the minimum amount, according to the data minimization principle of the GDPR(42).

Any use of personal data is first reviewed and documented before data processing is commenced. Retrospective clinical data is typically allowed in many instances under informed consent or legitimate provisions of secondary use. Even where data are properly anonymised and it is not possible to identify individuals, GDPR is not applicable, but the regime remains committed to rigorous standards of data ethics(42). In any project dealing with personal data, it is necessary to seek the advice of the organization’s Data Protection Officer (DPO) to confirm the data strategy and compliance with privacy legislation. For large-scale or new health data processing projects, a DPIA is performed. This risk assessment determines the likelihood of re-identification, assesses the proportionality of use, and suggests suitable measures. Encryption, access controls, logging, and secure storage are often put in place. The DPIA is undertaken early in the development process, before training any models, and is used to inform both technical and organizational approaches.

The design also honors data subjects' rights under the GDPR. Users are entitled to access their data, to edit errors, or to request the removal of their data. Models already built on past data cannot be retroactively altered, but safeguards are implemented to enable future sets to exclude data from training in the case of such a request. This forward-focused solution is in line with the new idea of "machine unlearning," which attempts to eliminate certain data

points from subsequent iterations of models without compromising performance(47).

This practice will ensure that sensitive health data is processed lawfully and ethically, with safeguards in place to protect individual rights and maintain transparency. Through early planning, clear documentation, and continuous risk assessment, the framework demonstrates how regulatory obligations, and innovation can be successfully balanced.

7 Discussion and Conclusion

The findings of this study highlight several important considerations for the regulatory management of ML-enabled IVD software within the European Union. The comparative analysis between the IVDR and the AI Act revealed both overlaps and critical gaps in their scope and implementation. While the IVDR focuses on clinical safety, scientific validity, and post-market surveillance, the AI Act addresses emerging risks specific to artificial intelligence, including algorithmic bias, lack of transparency, and inadequate human oversight. These frameworks operate with different emphases, one grounded in traditional medical device regulation and the other in AI risk governance.

The gap analysis confirmed that many of the ML-specific requirements are not adequately reflected in the current IVDR structures. For instance, the AI Act's mandatory classification of high-risk systems and its prohibition of manipulative or exploitative practices have no equivalents under IVDR(12,38). Similarly, the AI Act introduces novel obligations such as automated logging, bias mitigation, and interpretability that are not captured in the GSPRs of the IVDR(43). This fragmentation contributes to uncertainty for developers and regulatory professionals, particularly in the absence of harmonized guidance that would clarify how overlapping requirements should be interpreted or operationalized in practice.

This uncertainty may lead to delays in product development or over-engineering of compliance strategies, ultimately hindering innovation in ML-driven

diagnostics. Developers of ML-enabled IVDs may find themselves navigating parallel expectations, such as those for software safety under IVDR and transparency under the AI Act, without a coherent framework that bridges both. As the regulatory environment continues to evolve, the lack of alignment between the two regimes may increase the risk of misinterpretation or non-conformance, especially in smaller organizations with limited regulatory expertise.

These concerns were echoed at the 2025 RAPS Euro Convergence in Brussels, which I attended. During the conference, regulators and industry leaders emphasized the growing complexity of aligning AI Act requirements with existing device regulations. A key takeaway, as reported by Regulatory Focus, was the need for harmonized implementation practices, clearer conformity pathways, and proactive guidance to support innovation without sacrificing safety or transparency(39). These discussions reinforced the importance of developing integrated frameworks like the one proposed in this thesis; frameworks that can operationalize dual compliance in practice and ease the burden on early-stage MedTech developers.

Considering these challenges, the proposed lifecycle management framework offers a practical solution to the current regulatory disconnect. By embedding AI Act-specific controls within an IVDR-compliant development process, the framework supports a regulatory-by-design approach. It enables systematic documentation, validation, and oversight by aligning AI Act requirements, such as algorithmic transparency and data governance, with the technical, risk management, and quality system expectations outlined in ISO 13485, ISO 14971, and IEC 62304. This alignment not only supports compliance but also reduces redundancies and regulatory ambiguity by harmonizing overlapping demands across both legislations.

Nonetheless, the proposed framework is not without limitations. The scope of this thesis is confined to the EU regulatory space, excluding international standards such as the FDA’s Good Machine Learning Practice or emerging frameworks in Asia-Pacific jurisdictions. Furthermore, the framework remains

theoretical. It has not yet been validated through implementation or public consultation. Its effectiveness may be influenced by future revisions to the AI Act and its delegated acts or implementing guidance. As such, this proposal should be considered a foundation for continued refinement, rather than a fully mature compliance solution.

To advance this work, future studies should test the framework in real-world development settings and gather feedback from regulatory bodies and industry stakeholders. This would help determine its practical feasibility and allow for refinements based on empirical data. In addition, expanding the scope to explore global convergence with other regulatory systems and adapting the framework to address continuous learning models would improve its utility and resilience in a rapidly evolving field.

By situating the findings of this study within the broader context of dual regulatory obligations, this discussion underscores the importance of structured, harmonized, and practical approaches to AI compliance in IVD software. The lifecycle framework presented here is one such effort, aiming to bridge regulatory gaps and reduce implementation friction in the development of safe and trustworthy ML-enabled diagnostics.

In conclusion, this thesis contributes to the emerging field of AI regulation in healthcare by providing a structured, compliance-oriented approach to managing the lifecycle of ML-enabled IVD software. Startups and small companies in the early stages of development can use this framework as a practical guide to proactively embed regulatory compliance into the design and deployment of their products. While further refinement and validation are necessary, the proposed framework provides a solid foundation for aligning dual regulatory pathways and supporting the safe and responsible deployment of AI in medical diagnostics.

References

1. Johnson KB, Wei W, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021 Jan 12;14(1):86–93.
2. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019 Mar;19.
3. Onitiu D, Wachter S, Mittelstadt B. How AI challenges the medical device regulation: patient safety, benefits, and intended uses. J Law Biosci. 2024 Apr 9;
4. Hanna MG, Pantanowitz L, Dash R, Harrison JH, Deebajah M, Pantanowitz J, et al. Future of Artificial Intelligence—Machine Learning Trends in Pathology and Medicine. Modern Pathology. 2025 Apr;38(4):100705.
5. Kahles A, Goldschmid H, Volckmar AL, Ploeger C, Kazdal D, Penzel R, et al. Structure and content of the EU-IVDR: Current status and implications for pathology. Vol. 44, Pathologie. Springer Medizin; 2023. p. 73–85.
6. Kasanen E, Lukka K, Siitonen A. https://research.aalto.fi/en/publications/the-constructive-approach-in-management-accounting-research. 1993. The Constructive Approach in Management Accounting Research.
7. Hudspith S. Beyond utility: A framework for designing user experience. In: Proceedings of the Human Factors and Ergonomics Society. Human Factors and Ergonomics Society, Inc.; 1997. p. 447–50.
8. Dagiene V, Gudoniene D, Burbaite R. Semantic web technologies for e-learning: Models and implementation. Informatica (Netherlands). 2015 Sep;26:221–40.

9. Aurélie Mahalatchimy. Regulating Medical Devices in the European Union. The Oxford Handbook of Comparative Health Law Edited by David Orentlicher and Tamara Hervey, Oxford University Press,. 2020;
10. Niemiec E. Will the EU Medical Device Regulation help to improve the safety and performance of medical AI devices? Digit Health. 2022;8.
11. Medical Devices – Guidance on the Implementation of Regulations (EU) 2017/745 and (EU) 2017/746. 2022.
12. Santra S, Kukreja P, Saxena K, Gandhi S, Singh O V. Navigating regulatory and policy challenges for AI enabled combination devices. Vol. 6, Frontiers in Medical Technology. Frontiers Media SA; 2024.
13. Dombrink I, Lubbers BR, Simulescu L, Doeswijk R, Tkachenko O, Dequeker E, et al. Critical Implications of IVDR for Innovation in Diagnostics: Input From the BioMed Alliance Diagnostics Task Force. Hemasphere. 2022 Jun;6(6):e724.
14. Van Deutekom HWM, Haitjema S. Recommendations for IVDR compliant in-house software development in clinical practice: A how-to paper with three use cases. Clin Chem Lab Med. 2022 Jun;60:982–8.
15. Aboy M, Minssen T, Vayena E. Navigating the EU AI Act: implications for regulated digital medical products. NPJ Digit Med. 2024 Sep 6;7(1):237.
16. Azzouzi M El. https://easymedicaldevice.com/ai-medical-devices/. All AI Medical Devices are High-Risk?
17. Navigating the EU Artificial Intelligence Act: Essential Insights for Providers and Deployers. https://page.bsigroup.com/l/73472/2024-09-03/2c8ccqh/73472/1725406104Zhqq8V5C/BSI_EU_AI_Act_Infographic.pdf?utm_source=pardot&utm_medium=email&utm_campaign=gl-rs-cross-lg-nss-ot-mpd-mp-euaiactwhitepaperpromotion-0924.

18. khan B, Fatima H, Qureshi A, Kumar S, Hanan A, Hussain J, et al. Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector. Vol. 1, Biomedical Materials and Devices. Springer Nature; 2023. p. 731–8.
19. Pavlidis G. Unlocking the black box: analysing the EU artificial intelligence act's framework for explainability in AI. Law Innov Technol. 2024 Jan 2;16(1):293–308.
20. Johner ProfDrC. https://blog.johner-institute.com/iec-62304-medical-software/ai-act-eu-ai-regulation/. What the AI Act means for medical device and IVD manufacturers.
21. EUR-Lex. EU AI Act, Regulation (EU) 2024/1689.
22. Mitch. Software in Medical Devices, by MD101 Consulting. https://blog.cm-dm.com/post/2025/05/16/BSI-White-paper%3A-The-EU-AI-Act-meets-the-MDR.
23. The Impact of the EU AI Act on the Development and Use of Medical Devices. [cited 2025 Apr 17]; Available from: https://www.hunton.com/insights/legal/the-impact-of-the-eu-ai-act-on-the-development-and-use-of-medical-devices
24. Shantanu. https://roninlegalconsulting.com/policy-brief-harmonised-standards-for-the-eu-ai-act/. 2024. Policy Brief: Harmonised Standards for the EU AI Act - Ronin Legal.
25. International Organization for Standardization. ISO 13485:2016 Medical devices — Quality management systems — Requirements for regulatory purposes. https://www.iso.org/standard/59752.html.
26. International Organization for Standardization. IEC 62304:2006 Medical device software — Software life cycle processes. https://www.iso.org/standard/38421.html.
27. International Organization for Standardization. IEC 82304-1:2016 Health software Part 1: General requirements for product safety. https://www.iso.org/standard/59543.html.
28. International Organization for Standardization. IEC 62366-1:2015 Part 1: Application of usability engineering to medical devices. https://www.iso.org/standard/63179.html.

29. International Organization for Standardization. IEC 81001-5-1:2021 Health software and health IT systems safety, effectiveness and security Part 1: Principles and concepts. https://www.iso.org/standard/76097.html.
30. Speer J. https://www.greenlight.guru/blog/document-control. 2020. Document Control for Medical Device Companies: The Ultimate Guide.
31. Ginsberg, MSc R, Dahlke, MSc M. Safety risk management of software. https://www.raps.org/news-and-articles/news-articles/2022/3/safety-risk-management-of-software. 2022;
32. ISO 14971:2019. https://www.iso.org/standard/72704.html.
33. Admin. https://2amagazine.com/key-strategies-for-medical-device-registration-in-taiwans-medical-sector/. Key Strategies for Medical Device Registration in Taiwan's Medical Sector - 2A Magazine.
34. ISO International Organization for Standardization. ISO 14971:2019 – Medical devices — Application of risk management to medical devices. . 2019.
35. Fink M, Akra B. Comparison of the international regulations for medical devices–USA versus Europe. Injury. 2023 Oct;54.
36. https://sunstonepilot.com/2018/09/fda-software-guidances-and-the-iec-62304-software-standard/iec-62304-software-development-process/ [Internet]. IEC 62304 Software Development Process - Sunstone Pilot, Inc.
37. Update - MDCG 2020-16 Rev.2 - Guidance on Classification Rules for in vitro Diagnostic Medical Devices under Regulation (EU) 2017/746 - February 2023. https://health.ec.europa.eu/latest-updates/update-mdcg-2020-16-rev2-guidance-classification-rules-vitro-diagnostic-medical-devices-under-2023-02-10_en.
38. MDCG 2021-24 - Guidance on classification of medical devices. https://health.ec.europa.eu/latest-updates/mdcg-2021-24-guidance-classification-medical-devices-2021-10-04_en.
39. https://www.raps.org/News-and-Articles/News-Articles/2025/5/Euro-Convergence-Experts-discuss-AI-Act-s-future-

a?utm_campaign=regulatory_focus&utm_content=332967604&utm_medium=social&utm_source=linkedin&hss_channel=lcp-306225 [Internet]. Euro Convergence: Experts discuss AI Act's future and impact on medtech.

40. Busch F, Kather JN, Johner C, Moser M, Truhn D, Adams LC, et al. Navigating the European Union Artificial Intelligence Act for Healthcare. Vol. 7, npj Digital Medicine. Nature Research; 2024.
41. Vatsal Chhaya Ms, Kapil Khambholja P. The SaMD regulatory landscape in the US and Europe. 2021.
42. Minssen T, Rajam N, Bogers M. Clinical trial data transparency and GDPR compliance: Implications for data sharing and open innovation. Sci Public Policy. 2021 Apr 24;47(5):616–26.
43. Lewis D, Lasek-Markey M, Golpayegani D, Pandit HJ. Mapping the Regulatory Learning Space for the EU AI Act. 2025 Feb;
44. FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) -. 2019.
45. Medical Device Coordination Group. Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR . 2019.
46. Wu K, Wu E, Rodolfa K, Ho DE, Zou J. Regulating AI Adaptation: An Analysis of AI Medical Device Updates. 2024.
47. Guo C, Goldstein T, Hannun A, van der Maaten L. Certified data removal from machine learning models. In: 37th International Conference on Machine Learning, ICML 2020. International Machine Learning Society (IMLS); 2020. p. 3790–800.

Appendix 1: AI Act Compliance Checklist

This appendix presents the AI Act compliance checklist that maps with relevant AI act Articles to the corresponding phases of the ML-enabled IVD software lifecycle framework. It highlights the regulatory alignment with IVDR requirements, IEC 62304, and ISO 13485, and identifies key regulatory gaps or overlaps.

Requirement	Description	Guidance	IVDR Mapping	IEC 62304	ISO 13485	Gaps/Overlaps
AI Act Article 6-Classification	High-risk AI classification must be confirmed.	In device classification file, cite both IVDR and AI Act criteria. No extra process needed, just cross-reference.	IVDR Article 5, Annex VIII: Classify IVD devices based on intended use and risk.	Not directly applicable; does not address device (Classification based on risk class) classification.	7.1	Gap: IEC 62304 and ISO 13485 do not explicitly address AI/high-risk classification.
Article 5 -Prohibited practices	Do not use AI systems that manipulate	Add a section in risk management and design	Annex I (GSPRs), Article 5	N/A	7.2.1, 7.2.2	Gap: AI-specific prohibitions not fully covered

	people, exploit vulnerable groups, or enable social scoring. Avoid real-time biometric surveillance in public unless strictly authorized by law.	documentation explicitly stating compliance with AI Act Article 5. Reference this in IVDR GSPR checklist.				
Article 8- Compliance/integration	Allows integration of AI Act and IVDR compliance activities	Use a single set of technical documentation and QMS records for both IVDR and AI Act. Clearly indicate in your index/table of contents where AI	Article 10, Annex II, IX	4.1, 5.1, 5.2, 5.3	4.1, 4.2.1	Overlap: Integration permitted

		Act requirements are addressed.				
Article 9 – Risk Management	Requires a risk management system for AI	Expand risk management file to include AI-specific risks (e.g., bias, data drift, algorithmic errors). Document AI-specific risk controls.	Annex I, Section 16.2: Risk mitigation for software.	7.1	7.1, 7.2, 7.3, 7.5.6	Overlap: All require risk management. Include ML bias analysis
Article 10 – Data Governance	Datasets must be relevant, representative, and error-free to minimize bias.Sets requirements for data quality,	Validate dataset diversity (e.g., age, gender, ethnicity); document checks in Data Management Plan, per SOP-009 (Section 4.2.1).	Annex II, Section 6.1: Dataset validation for performance claims.	5.2.1	7.3.3	Overlap: All require data quality. Ensure dataset diversity checks in SOP-011(ML development).

	representativeness, and bias					
Article 11 -Technical docs	Requires technical documentation for AI	Ensure your SOPs and templates are updated with AI-specific references.	Annex II	5.1.1, 5.1.2, 5.1.3	4.2.3, 4.2.4, 7.3.2	Overlap: All require technical docs
Article 12 – Logging and Traceability	High-risk AI systems must log inputs, outputs, and key decisions to support audit and PMS.	Add a data management section to technical documentation, describing dataset quality, bias checks, and traceability. Reference this in IVDR and software lifecycle docs.Ensure logs	Annex I (GSPRs 17.2), Annex III	5.1.1, 6.2 & 7	7.5.6, 8.2.1, 8.2.4	Gap: None of the standards explicitly require logging of AI-specific events (input, output, model decisions).

		are readable, secure, and linked to performance monitoring.				
Article 13 – Transparency IFU	Requires clear instructions and transparency for users	Document ML model logic in Technical File. Add AI-specific explanations (e.g., how the model works, limitations) to IFU and user manuals.	Annex II, Section 6.2: Document software functionality/design.	5.1.1, 5.1.2	7.5.1, 7.5.3.2.1	Overlap: IFU required by all
AI Act Article 14 – Human Oversight	Requires human oversight to minimize risks	Document human oversight procedures in risk management and IFU. Describe how users can override	Annex I (GSPRs 22.2, 23.4)	5.1.1, 5.1.2, 5.7	7.2.1, 7.2.2, 7.3.3	Overlap: All support oversight. Document override protocols in User Manual

		or monitor AI outputs.				
AI Act Article 15 – Accuracy, Robustness, and Cybersecurity	Sets requirements for accuracy, robustness, and cybersecurity	Expand software verification/validation and cybersecurity documentation to include AI-specific robustness, accuracy metrics, and feedback loop controls. Reference these in risk management and PMS plans.	Annex I (GSPRs 9, 17, 18), Annex II	5.1.1, 5	7.3.6, 7.5.6, 7.6, 8.2.1	Gap: AI Act emphasizes PCCP; IVDR/IEC less specific. Define PCCP in SOP-011, per SOP-003 (4.4).
Article 16- Provider obligations	Ensure your AI system meets all EU AI Act obligations before placing it on the	In QMS, add a checklist for AI Act-specific obligations (e.g., registration,	Article 10, 13, 16, 24, 26, 48, 50	5.1.1, 5.1.2, 9	4.2.1, 4.2.3, 7.2.1, 8.3, 8.5	Overlap: Most obligations already required

	market or putting it into service. This includes proper design, documentation, conformity assessment, and CE marking.	accessibility). Reference this in QMS procedures.				
AI Act Article 17- Quality Management System	A Quality Management System (QMS) is mandatory for high-risk AI systems	Describes how QMS covers the design, development, testing, and lifecycle management of AI systems.	Article 10, Annex IX (QMS)	5.1.1, 5.1.2, 6	4.1, 4.2, 8.1, 8.2, 8.5	Overlap: QMS required by all
Article 18 - Documentation retention	Specifies retention period for documentation	Update SOP ocument Management to explicitly include a	Article 10(8), Annex II	5.1. 1, 5.1.2	4.2.4	Overlap: Retention required by all

		10-year retention period for documentation related to high-risk AI systems under the AI Act.				
AI Act Article 43 – Notified Body Engagement	Manufacturers of high-risk AI systems must carry out a conformity assessment before placing the system on the market.	Use IVDR conformity assessment route but ensure notified body reviews AI Act requirements as part of the same process.	Article 48, Annex IX/X/XI	5.1.1, 5.1.2, 5.5	8.2.1, 8.2.2, 8.2.3	Gap: IEC 62304 and ISO 13485 support documentation but do not cover NB-specific conformity assessment workflows.
Article 61 -Informed consent for real-world testing	Requires informed consent for real-world testing	Update clinical/performance study procedures to include AI-specific	Article 58, Annex XIV	N/A	7.2.3, 7.5.6	Gap: More explicit consent for AI

		consent language and documentation.				
Article 62 - SME/start-up support	Provides support and simplifications for SMEs	Set up real-time monitoring in PMS Plan; update PSUR with data drift findings, per SOP-010 (Annex III).	Recital 13, Article 24		4.1.1, 4.2.1	Gap: Additional support for SMEs
Article 73 -Serious incident reporting	Requires reporting of serious incidents involving AI	Integrate AI-specific incident reporting into existing IVDR vigilance and ISO 13485 CAPA system. Reference AI Act in PMS/vigilance SOPs	Article 82, Annex III	5.5	8.2.3, 8.3, 8.5	Overlap: Incident reporting required by all

Appendix 2: Machine Learning Model Development SOP

This appendix contains the standard operating procedure (SOP) for the machine learning model development lifecycle at Hematoscope.

SOP-011

Machine Learning Model Development for IVD Software under EU AI Act & IVDR

Author / Date:	
Reviewer / Date:	
Approver / Date:	

Table of Contents

1 Purpose 3

2 Scope and Applicability 3

3 Roles and Responsibilities 4

4 Documentation and Records 6

5 Machine Learning Model Development Lifecycle Process 7

5.1 Data Acquisition and Preparation 8

5.2 Model Design and Training 9

5.3 Verification and Performance Evaluation 11

5.4 Technical Documentation and Design Review 14

5.5 Deployment and Release 15

5.6 Post-Market Monitoring and Feedback Loop 18

5.7 Continuous Improvement and Model Update Protocol 20

6 References 27

Version History 28

1. Purpose

This Standard Operating Procedure defines the structured process for planning, developing, verifying, deploying, and maintaining machine learning models within Hematoscope's in-vitro diagnostic (IVD) software products.

The SOP ensures that ML components are developed in compliance with:

- The EU In Vitro Diagnostic Regulation (IVDR, 2017/746) includes Annex I and Annex II requirements.
- The EU Artificial Intelligence Act, specifically obligations for high-risk AI systems (Articles 9–15)
- Relevant harmonized standards, including ISO 13485, IEC 62304, and ISO 14971.

The document supports the regulatory-grade development of ML-driven functionality, incorporating proper design controls, risk management, traceability, and validation. It ensures that ML models meet clinical performance expectations and can be confidently integrated into Hematoscope's IVD medical devices.

2. Scope and Applicability

This SOP applies to the development of all machine learning (ML) models that are intended to be incorporated into Hematoscope's IVD software products.

It is particularly relevant for software classified as Class C devices under the IVDR, where ML plays a critical role in clinical decision support, result interpretation, or diagnostic output.

The SOP complements the following company procedures:

- **SOP-009: Product Design and Development** – This SOP extends SOP-009 with ML-specific activities during the design, implementation, and verification phases.

- **SOP-003: Product Lifecycle Process** – This SOP aligns with SOP-003 for activities such as change control, post-market monitoring, and risk-based updates to the ML model.

The SOP ensures that ML components fulfill:

- IVDR General Safety and Performance Requirements (Annex I) and Technical Documentation obligations (Annex II),
- AI Act obligations for high-risk systems, including data quality (Article 10), technical documentation (Article 11), logging (Article 12), transparency (Article 13), human oversight (Article 14), and robustness (Article 15).

The processes defined herein are mandatory for any ML model integrated into Hematoscope's regulated medical software. Exceptions or deviations require documented justification and prior approval from the Quality Assurance (QA) and Regulatory Affairs (RA) team.

3. Roles and Responsibilities

The following roles are responsible for executing, supporting, and ensuring compliance with this SOP throughout the machine learning ML model development lifecycle:

- **Product Owner / Lead System Designer**
    - Defines the intended purpose and regulatory classification of the ML-enabled IVD device.
    - Approves key documents such as the Data Management Plan, Model Design Specification, and Design Review outputs.
    - Ensures ML-related development aligns with the product's clinical and regulatory strategy.
- **Data Scientist / ML Engineer**
    - Prepares and curates datasets for model training, validation, and testing.
    - Validation set refers to dataset that is used to "guide" the model in the right direction during model training.

- Testing Dataset is the dataset that the model does not see during the training.
- Designs and trains ML models; documents architecture, training parameters, and model performance.
- Ensures data quality and bias control in compliance with Article 10 of the AI Act.
- Maintains traceability and reproducibility of model versions.

- **Risk Manager**

  - Applies Hematoscope's risk management process (SOP-004) to ML-specific risks.
  - Identifies, analyzes, and mitigates hazards introduced by the ML component.
  - Ensures risk controls are verified, and residual risks are documented according to ISO 14971 and IVDR Annex I.

- **Quality Assurance (QA) & Regulatory Affairs (RA)**

  - Ensures design controls (e.g., requirements, verification, validation) are followed and documented.
  - Oversees compliance with IVDR Annex II and AI Act Article 11 (Technical Documentation).
  - Conducts formal Design Reviews and manages conformity assessment activities.
  - Verifies that controlled documentation is stored in the Design History File (DHF) and Medical Device File (MDF).

- **Software Engineers (Integration & Deployment)**

  - Integrate validated ML models into the production software system.
  - Manage model configuration, version control, and deployment infrastructure (e.g., cloud or on-premises).

  - Implement cybersecurity and logging mechanisms (AI Act Articles 12 & 15).
  - Support release testing and documentation generation.

- **Clinical and Scientific Affairs**
  - Ensure model validation aligns with intended clinical use and diagnostic performance expectations.
  - Contribute to the Performance Evaluation Report and support post-market follow-up for ML functionality.

- **All personnel involved in ML-related activities must:**
  - Complete training on this SOP before working on ML development tasks.
  - Document and report any deviations from this SOP. QA/RA must approve deviations.
  - Maintain compliance with Hematoscope's QMS and applicable regulations.

4. Documentation and Records

ML model development produces a set of records and work products that become part of the design and technical documentation of the device. Key outputs and documents include:

- **Data Requirements and Preparation Documentation**: A description of the data sets used for model training, validation, and testing. This includes data source provenance, inclusion/exclusion criteria, data preprocessing steps, dataset statistics, and evidence of data quality (e.g., label validation results). Any Data Management Plan or dataset specification is documented and stored (as part of design inputs or an annex to the requirements). This documentation addresses AI Act Article 10 by demonstrating that

training/validation data are relevant, representative, and as complete/error-free as possible.

- **Model Development Artifacts:** The source code (e.g., training scripts, model architecture definitions) and trained model files (e.g., learned parameters) are maintained in the configuration management system (Git repository, with appropriate version control tags). A Model Training Log or Report is generated, capturing the training configuration (hyperparameters, data used, date, responsible engineer) and results (training curves, final model performance on training and internal validation sets). This ensures reproducibility and traceability of the model version according to Article 12's record-keeping obligation.
- **Design and Risk Documentation Updates:** The introduction of an ML model is reflected in the system design specification and architecture description (as part of design outputs). For example, the software architecture document (see SOP-009) will include the ML module, its interfaces, and any Software of Unknown Provenance (e.g. ML libraries like TensorFlow/PyTorch), which are handled per SOP-009 Section 4.3 on SOUP. The Risk Management File is updated with ML-specific risk analysis and controls (bias mitigation, failure mode effects if the model misclassifies, etc.), referencing the risk control measures implemented (e.g., algorithm performance requirements, human oversight steps).
- **Verification and Validation Records:** All testing results related to the ML model are documented. This includes Design Verification test cases that link to ML requirements (e.g., verifying the model meets specified accuracy on a validation dataset) and their results, as well as Design Validation or performance evaluation outcomes demonstrating the model's clinical performance on independent test data or clinical samples. These results are summarized in a Verification Report and/or Performance Evaluation Report, which becomes part of the technical file (IVDR Annex II, 6.1). The verification evidence also supports AI Act Article 15 by quantifying the model's accuracy and robustness, which will be declared in the Instructions for Use.

- **Technical File and Change Records:** The Technical Documentation (Annex II) is maintained to include the ML model information: algorithm description, training data summary, performance metrics, the intended purpose of the ML (which must align with the device's intended use), and any limitations. A traceability matrix is updated to ensure that each ML requirement is linked to design implementation and verification outcomes (fulfilling design control and AI Act documentation mandates). When the model is finalized, a Design Release Review (DRR) (as per SOP-009) is conducted, and minutes are recorded, confirming that the ML development outputs meet all input requirements and that the device (including the ML component) is ready for release. Suppose the ML model or its training process is changed post-release. In that case, change requests are handled via the change management process (SOP-003, sec. 4.4.5), and all changes are logged (with justification and impact assessment on safety/performance). Version history is maintained to document each released model version and its validation in line with configuration management procedures (e.g., semantic versioning of models and software).

All documentation is filed in the Medical Device File (MDF) for the product or within the controlled repositories (e.g., requirements and design docs in SharePoint QMS, code in GitHub). The outputs listed above correspond to, or are included in, the documents enumerated in SOP-003 Table 3.1 (e.g., user requirements, design specifications, risk management file, verification/validation plans and reports, etc.). Records must be retained and updated throughout the product lifecycle to support regulatory submissions, audits, and post-market assessments.

5. Machine Learning Model Development Lifecycle Process

The ML model development process is divided into phases (5.1–5.7) that correspond to stages in the software lifecycle, each with specific objectives, regulatory checkpoints, and outputs. Table 1 provides a summary of these phases. This process is aligned with the overall product development and quality management

processes (SOP-009 and SOP-003), ensuring that traditional software development controls and additional AI-specific requirements are satisfied in tandem.

5.1 Data Acquisition and Preparation Activities

In this initial phase, the project team identifies and obtains the data required to develop and evaluate the machine learning (ML) model. Sources may include retrospective clinical datasets (e.g., laboratory results, annotated images), published datasets, or prospectively collected real-world data. The data must align with the intended purpose of the IVD device and adequately reflect the target user population.

A formal data collection protocol is established to define data sources, inclusion/exclusion criteria, and handling procedures for sensitive data. For any dataset involving personal health information, data must be collected and processed by applicable data protection laws (e.g., GDPR, HIPAA). Patient consent or data anonymization steps are performed and documented before use.

After collection, data undergoes a structured preprocessing workflow. This includes cleaning (e.g., removal of errors and duplicates), normalization, and partitioning into distinct training, validation, and test sets. To enhance generalizability and avoid overrepresentation of specific classes, methods such as stratified sampling and dataset balancing are applied. Data augmentation may be employed when additional synthetic data is needed, especially to strengthen underrepresented groups or clinical edge cases.

Exploratory data analysis (EDA) is performed to understand the distributions, assess coverage of clinically relevant subgroups, and detect anomalies or gaps. If significant gaps are found, mitigation strategies such as resampling, sourcing supplemental data, or refining inclusion criteria are implemented. At the end of this phase, the curated datasets are stored with documented versioning and traceability.

Quality and Compliance

In compliance with Article 10 of the AI Act, datasets used for ML development must be relevant, representative, complete, and free of significant errors. The development team ensures this by documenting the source, structure, statistical properties, and limitations of each dataset. Data quality control includes subgroup analysis to detect and mitigate bias, checks for completeness, and verification of label accuracy.

The Data Management Plan includes all preprocessing steps, dataset metadata, quality control metrics, and justification for dataset suitability based on the intended use. Any dataset used for training or testing is version-controlled and linked to a corresponding model version.

All outputs from this phase, datasets, metadata, and the Data Management Plan, are stored in the controlled design environment (e.g., SharePoint, GitHub) and linked to the product's Design History File (DHF) and Medical Device File (MDF).

5.2 Model Design and Training

Activities

In this phase, the development team defines and implements the ML model architecture based on product requirements, performance targets, and the characteristics of the training data. The choice of algorithm and architecture is guided by technical feasibility, explainability needs, regulatory constraints, and the intended clinical context of use.

Commonly used architectures include neural networks, ensemble models, or classical statistical learners, depending on the complexity of the task. Feature engineering or transformation is performed when necessary to align input data with model design assumptions.

Once the architecture is finalized, the model is implemented using SOUP components, such as TensorFlow or PyTorch. These tools are managed under Hematoscope's configuration management and tool validation procedures, in

accordance with IEC 62304 requirements for SOUP. All code, configurations, and scripts used in training are version-controlled (e.g., Git), ensuring full traceability.

Model training is conducted in a controlled computing environment. The training process includes the definition of hyperparameters, random seeds, optimization functions, and convergence criteria. Intermediate results are monitored to detect issues such as overfitting, divergence, or anomalous behavior. Cross-validation, learning curve analysis, and early stopping techniques are used to ensure robust model performance and generalization.

If multiple model candidates are trained, their performance is compared using predefined evaluation metrics. The best-performing candidate is selected for verification based on test set results. Training artifacts such as logs, performance plots, and interim weights are stored as part of the model training record.

Monitoring and risk control

The model design and training phase follows the software development lifecycle principles outlined in IVDR Annex I,16.2, and complies with Articles 9, 13, and 14 of the AI Act, focusing on risk management, transparency, and human oversight.

All design decisions, including algorithm selection, interpretability strategies, and feature selection, are recorded in the Model Design Specification. Configuration management is strictly enforced to ensure traceability, reproducibility, and rollback capability throughout the model development cycle.

To support clinical safety and user trust, the design incorporates features that enhance transparency and oversight, such as output confidence scores, intermediate visualizations, and threshold-triggered alerts. These controls are documented in system-level design inputs and linked to human factors and usability requirements.

The Risk Management File is updated continuously during this phase. Potential hazards, such as subgroup-specific performance limitations or unintended model behavior, are identified and addressed. Risk control measures may include software-

level fail-safes, interface-based warnings, or constraints in the intended use. Residual risks are reviewed and accepted in alignment with ISO 14971.

Quality and Compliance

The model design and training phase adheres to software lifecycle principles as required by IVDR Annex I 16.2 and the AI Act's requirements for risk management (Article 9), human oversight (Article 14), and transparency.

All design decisions, including algorithm selection, model interpretability strategies, and feature set justification, are documented in the Model Design Specification. The use of configuration management ensures reproducibility and enables rollback or audit if needed.

Controls to support human oversight and clinical usability are considered at the design stage. This may include designing the model to output confidence scores, generate intermediate explanations, or integrate threshold-based alerting mechanisms. These features are captured in system-level design inputs and linked to human factors requirements.

The Risk Management File is updated throughout the model training phase. Potential hazards introduced by model behavior (e.g., systematic underperformance on a subgroup) are identified and addressed through risk control measures. These may include interface warnings, model fallback logic, or limitations in the instructions for use.

The final output of this phase includes the trained model with a unique version identifier, updated training documentation (including dataset version used), and an updated Design Specification. These are stored in the Design History File (DHF) and linked to corresponding verification planning documents.

Outputs

The outputs of this phase include a fully trained and traceable machine learning model that is ready for formal verification and performance evaluation. The model is assigned a unique version identifier and stored within a controlled repository under configuration management.

Supporting documentation is generated to ensure full traceability and regulatory compliance. This includes the finalized Model Design Specification, which details the selected algorithm, model architecture, and performance expectations. Accompanying this are the training configuration files, hyperparameters, training logs, and a summary of the datasets used, each linked to specific model iterations through version control.

These outputs demonstrate that the model has been developed in a reproducible and controlled environment. They serve as the foundation for subsequent verification and validation phases and are archived in the Design History File (DHF) as part of the product's technical documentation. Alignment with IVDR Annex II and AI Act documentation requirements is maintained throughout.

5.3 Verification and Performance Evaluation

Activities

In this phase, the finalized model from 5.2 is subjected to rigorous verification testing and performance evaluation. The goal is to demonstrate that the model meets all specified requirements and performance criteria and that it generalizes well to independent data. The team designs test cases that cover:

- **Functional correctness:** Does the model produce the expected output format and integrate correctly with the rest of the software (e.g. input/output interfaces)?
- **Analytical performance:** Using a reserved testing dataset (not seen in training), the model's clinical performance metrics are measured. For a diagnostic ML model, this includes sensitivity, specificity, positive predictive

value, negative predictive value, etc., as applicable. These results are compared against acceptance criteria derived from clinical requirements or statistical targets set during design.

- **Robustness tests:** The model is tested under various conditions – e.g., slightly noisy or perturbed inputs, edge cases, boundary values – to observe behavior. The aim is to ensure the model's performance is stable and robust (AI Act Art. 15 emphasizes robustness and cybersecurity). For example, if the model is for image analysis, tests may include images of lower quality to see if performance degrades gracefully. If the model uses randomness (some do, in inference), multiple runs are tested to ensure consistency (repeatability).
- **Security and fault tolerance:** Although this is primarily handled at the software system level, the ML model is assessed for any vulnerability e.g. susceptibility to obviously adversarial inputs that could realistically occur. Also, verify that errors from the model are handled properly by the system (for instance, if the model fails to return a result, the system should handle that error). Cybersecurity considerations (as per AI Act Art. 15) for the ML component might include ensuring the model file is protected from tampering and that model outputs cannot be easily manipulated by malicious inputs. These aspects might be tested in collaboration with software security experts.
- **Before approval**: Updated models are benchmarked against the currently deployed version. Acceptance criteria are based on performance improvement or at least non-inferiority. Any model failing to meet these benchmarks is rejected or revised, as part of internal release control.

All verification activities follow the Verification Plan established earlier (SOP-009 and SOP-003 reference having a verification plan). Each requirement (functional or performance) that involves the ML algorithm is verified with at least one test. Traceability is maintained so that its clear which test result corresponds to which requirement (typically in a Requirements Traceability Matrix).

In parallel, for regulatory compliance, a formal Performance Evaluation for IVD is carried out. For an ML-based IVD, this means compiling evidence that the device meets its performance claims and is effective for its intended purpose. The

verification testing on independent data contributes to this evidence (analytical performance). If available, results from clinical validation studies (e.g., comparing the ML's output to a reference method on clinical samples) are included or referenced. IVDR requires demonstration of clinical performance or a justified performance evaluation instead of clinical data. For high-risk diagnostics, often a prospective clinical performance study would be needed unless a well-curated archive of clinical specimens is used. The Performance Evaluation Report will aggregate the analytical performance (from lab/bench tests) and any clinical performance data.

Outcomes and Decision

Verification results are compared against pre-defined acceptance criteria. If the model meets all criteria, it is considered verified. A Verification Report is created to summarize all test results. The report highlights the model's accuracy and other key performance metrics.

In line with AI Act Article 15, these metrics must be included in the Instructions for Use. The team ensures that the reported figures are accurately documented for this purpose.

If any test fails to meet the criteria, corrective action is required. The team evaluates whether to collect new data, revise the model design, or retrain the model returning to earlier phases (5.1 or 5.2) if necessary. All changes are tracked through the formal design change control process.

Minor issues may be corrected with small adjustments, such as threshold tuning. More significant problems might need a redesign or a new training cycle. The model continues through this verify-correct loop until it meets requirements or is withdrawn. Each change is followed by regression testing to confirm that no new issues were introduced.

Once verification is successful, the Risk Management File is updated with any residual risks identified during testing. For example, if the model performs slightly worse in a specific subgroup but is still clinically acceptable, the risk is documented

with a justification. Related warnings or limitations are included in the user information to comply with IVDR.

Finally, an internal Design Verification Review is held. QA and Regulatory Affairs confirm that the verification evidence is complete and acceptable. Once approved, the project moves to the documentation and release preparation phase.

5.4 Technical Documentation and Design Review Activities

After successful verification, the development team compiles all design and development outputs into the Technical Documentation required under IVDR Annex II. This includes detailed descriptions of the ML model's intended purpose, algorithm structure, training data, performance metrics, and any associated risks and mitigations. The documentation must demonstrate how the model conforms to design inputs and meets applicable regulatory and safety requirements.

The Technical Documentation also integrates AI-specific information in anticipation of future compliance under the EU AI Act. This may include descriptions of model logic, data quality controls, transparency features, human oversight provisions, and justification for design choices.

A formal Design Review is conducted to evaluate the completeness and adequacy of the ML development outputs. Depending on the stage, this may take the form of a Design Output Review (DOR) or Design Release Review (DRR). The review board, typically including representatives from R&D, QA, RA, and Clinical Affairs, assesses whether the model meets all requirements, whether verification evidence is sufficient, and whether the risk management documentation is complete.

The reproducibility of the model development process is confirmed during the review. This involves verifying that all model artifacts, code, data, model weights, and documentation are under configuration control and traceable. Additionally, the target deployment environment (e.g. cloud infrastructure or on-premises setup) is reviewed to ensure readiness and alignment with design transfer procedures.

Any issues identified during the review are recorded, assigned for resolution, and closed before final approval.

Outcomes

The outcome of the Design Review is a documented decision on whether the model and system are approved for release. The decision is recorded in the meeting minutes and may include one of the following statuses:

- Approved
- Approved with contingencies
- Not approved

The minutes clearly state the review's conclusions, any actions required, and assigned responsibilities for follow-up.

If the design is approved, it is declared a release candidate. From this point forward, the design is considered frozen; any further changes must be made through formal change control processes under SOP-003. Informal adjustments or iterations are no longer permitted.

Upon sign-off, the product (including the ML component) is considered ready for release preparation. For high-risk devices, Regulatory Affairs will use the finalized documentation to compile the Technical File for submission to a Notified Body. For Class C IVDs, an NB audit of the Technical Documentation is mandatory. The ML model documentation, including verification and risk evidence, will be included as part of the submission package.

The successful Design Review provides internal assurance that the ML model meets quality, safety, and performance requirements. Simultaneously, preparations for deployment and market release may proceed under controlled conditions, as defined in the next phase.

5.4 Deployment and Release

Activities

After formal design approval, the machine learning (ML) model is transferred from the development environment to the production environment as part of the release process. Deployment may occur via a cloud infrastructure (e.g., Azure) or as part of an installable on-premises software package. This phase follows the design transfer and configuration management procedures outlined in SOP-009.

Deployment includes configuring the production environment to match the validated development environment. This involves installing the correct library versions, ensuring hardware compatibility (e.g., GPU availability), and validating the environment using Infrastructure-as-Code tools if applicable. The deployed model must be identical to the version that passed verification. To ensure this, checksums or digital signatures may be used to confirm model file integrity.

A set of final release verification tests, often called installation or deployment verification, is conducted after deployment. These tests confirm that the model functions as expected in the production setting. They also verify that system behavior has not changed due to environmental differences. If applicable, uninstallation and update procedures are also tested to validate the robustness of deployment mechanisms.

Concurrently, the user-facing documentation is finalized. The Instructions for Use (IFU), quick start guides, and other labeling materials are updated to comply with both IVDR and AI Act requirements. The documentation must include:

- A clear statement of the ML model’s intended purpose and how it fits into the clinical workflow.
- Performance metrics (e.g., 95% sensitivity, 90% specificity), validated in prior phases.
- Limitations of the model, including any known performance gaps or conditions where outputs should be interpreted with caution.

- Warnings and precautions to ensure safe use and prevent overreliance on automated results.
- A statement identifying the model as an AI/ML-based component and instructions for human oversight, in line with anticipated transparency obligations under the AI Act.

The model-enabled software is assigned a version number using semantic versioning, with changes in the ML model triggering at least a minor version increment. A Unique Device Identifier (UDI) is generated for the software in accordance with IVDR Article 24 and Annex III. The Declaration of Conformity is then updated to reflect this release, formally stating that the software version (including the ML component) conforms with applicable requirements.

This phase also includes internal release authorization. A formal production release approval is issued, documenting that the product has passed all checks and is cleared for market distribution. If continuous integration/continuous deployment (CI/CD) tools are used, the release is executed by QA through validated pipelines. The software package is archived under configuration control, and the release version is tagged in the code repository.

Release and regulatory steps

With everything in place, the company finalizes the Declaration of Conformity (DoC), asserting that the device (with the ML model version X.Y) meets the IVDR requirements. The software is assigned a Unique Device Identifier (UDI) if not already, and the version number is fixed (semantic versioning is applied such that any ML model change would increment at least the minor or major version). The product is then registered with the national competent authority as required (in Finland, Fimea). Registration is essentially a notification, not an approval, but it’s a legal step confirming that the device can be placed on the market.

Finally, the software is released, meaning it is made available for distribution or deployment to end users. Internally, this requires a formal production release approval (often documented by a signed record stating that all checks are passed

and QA authorizes the release). If using continuous integration/continuous deployment (CI/CD) pipelines, QA might press the “release” button at this point. The released package is archived (for configuration management), and the release is tagged in the repository (matching the version in the labeling).

Outputs

Key outputs of this phase include:

- Released Software Package (or deployment in the cloud) containing the ML model, which is now the actual medical device delivered to customers.
- Release Notes documenting new features (if any), the ML model version, and any known issues or limitations. It may also indicate compliance info (like “Developed under ISO 13485 QMS, conforms to IVDR”).
- Installation/Deployment Verification Report confirming the software build and environment match what was validated, and summarizing results of final checks.
- Instructions for Use and labeling finalized, which now incorporate AI Act transparency requirements (though AI Act is not yet in effect, we proactively include those details to demonstrate commitment to transparency and to ease future compliance)
- Design Release Review record (if the formal release was gated by a final review, its report).
- Device Master Record update (in manufacturing terms, ensure all components, including the model binary, are captured in the device master file for reproducibility).

At this point, the product with the ML model is on the market. The focus now shifts to monitoring its performance in the field and maintaining it.

5.6 Post-Market Monitoring and Feedback Loop

Activities

Once the ML-enabled device is in use, Hematoscope undertakes active post-market surveillance (PMS) to ensure the product continues to be safe and effective. For an ML model, post-market monitoring has the added dimension of checking that the model's performance does not degrade over time or in new use contexts. The team implements a feedback loop whereby real-world data and user feedback are collected and fed back into the lifecycle:

- **Data collection from the field:** Depending on the device setup, this might include automatic collection of de-identified usage data (with appropriate user consent and data protection). For example, the device might log each prediction the ML model makes along with the eventual confirmed outcome (ground truth) if obtainable. These logs provide a growing dataset of real-world cases. The AI Act Article 12 specifically requires high-risk AI systems to be designed with automatic logging. In practice, our software logs model inputs, outputs, and key decision steps (while respecting privacy). These logs are stored securely for analysis.
- **User feedback and complaints:** Standard channels such as customer support (Jira Service Desk or equivalent) gather any user-reported issues. If a user notices an incorrect or implausible ML output, that might be reported as a complaint. The Quality system classifies and investigates these per SOP-003 (feedback and complaint handling). Serious incidents (e.g., an erroneous result that led to a wrong diagnosis and patient harm) would trigger vigilance reporting to regulators. All such incidents are also evaluated to see if they indicate a systematic model issue.
- **Performance tracking:** At defined intervals (maybe quarterly or bi-annually), the team re-evaluates the model's performance on new data accumulated from the field. If ground truth data can be obtained (e.g. for cases where follow-up lab results are available), the model's sensitivity/specificity on those new cases is calculated and compared to the expected performance. Statistical trending is applied (per PMS plan) to detect any statistically significant drop in performance or any emerging bias. For instance, it might be

observed that the model is underperforming in a subpopulation that was not well-represented in the original training data; this would be a trigger for action.

All these activities are documented in PMS reports. For Class C IVDs, a Periodic Safety Update Report (PSUR) is required to be updated at least annually, summarizing PMS findings, including any trend in false results, etc. The ML model's performance monitoring findings would be part of that PSUR. If post-market performance follow-up (PMPF) was deemed necessary (likely yes for an ML diagnostic to confirm it works as intended in the real world), then a PMPF plan would be executed, possibly collecting prospective data or doing studies in the background. The results go into a PMPF evaluation report, which feeds into possibly improving the model or confirming its stability.

Importantly, the risk management process continues in the post-market phase. Any new hazards identified from real-world use (e.g. a scenario where the model fails) lead to an updated risk assessment and possibly new risk controls. For example, if we learn that certain interfering substances in samples cause the model to misclassify, we add a warning in IFU or implement a software update to detect such cases.

The human oversight aspect in real use is also monitored. We ensure, through user feedback that users are indeed interpreting the AI output correctly and not over-relying on it blindly (addressing automation bias as discussed in AI Act Art. 14). If we find users misunderstand the AI output, we might enhance training or the UI to improve interpretability (e.g. provide a more explainable output or confidence level).

The feedback loop is not just for safety surveillance but also for improvement. Real-world data can highlight where the model could be improved. We might find, for example, a new pattern in data that the model didn't see before and struggles with. This insight will inform the retraining or update decisions in the next phase. We maintain a data repository of new cases that could be used in future model versions. However, any such improvement must be weighed against regulatory requirements e.g., we cannot simply deploy a self-updating model without validation. So, the feedback informs a controlled update process.

Outputs

The outputs of this phase are primarily documentation and records:

- Regular PMS reports (internal) and PSURs (for regulators) capturing the findings related to the ML performance and any incidents.
- Customer feedback records and CAPA (Corrective and Preventive Action) records if issues with the ML are discovered and need correction.
- Updates to the risk management file and possibly the labeling (if new warnings or contraindications need to be added due to post-market findings, this is done via a labeling update and communicated as an advisory notice if needed).
- Over time, a curated collection of new validated data is gathered, which can be used to enhance the model. This dataset is documented and stored with version control (as “real-world data 2025-Q3” for instance), ready to feed into the next development cycle if an update is pursued.

The ongoing monitoring ensures that the ML model remains under watch throughout its life on the market, and that we remain compliant with both IVDR’s PMS requirements and the AI Act’s expected provisions for continuous monitoring and logging. If at any point the model’s performance is found to be unacceptable or a serious risk emerges, the company has procedures (per SOP-003) to initiate a field safety corrective action (e.g. temporarily disable the ML function or recall the software) to protect patients. Thankfully, with vigilant monitoring, we aim to detect any issues early and address them proactively.

5.7 Continuous Improvement and Model Update Protocol

Activities

This phase is a loop-back point to the development lifecycle whenever an update to the ML model is needed. Continuous improvement of the model might be driven by several factors:

- declining performance (detected via PMS),

- availability of significant new training data (e.g., new patient cohorts, or expanded use to new conditions),
- or new algorithmic techniques that could enhance performance.

In a regulated environment, any change to the ML model must be carefully managed. Hematoscope employs a Change management process (SOP-003 section 4.4.5) to assess and implement model updates:

- First, a Change Request is submitted (for example, “Update model to version 2.0 with additional training data from region X to improve sensitivity”). This request includes the reason for change, description of proposed changes, and preliminary risk assessment of the change.
- A cross-functional team, including regulatory, quality assurance, technical, and clinical representatives, assesses whether a proposed change qualifies as significant under IVDR and MDCG guidance. Changes that affect the intended use, safety, or performance of the device typically require notification to the Notified Body (NB) or a new conformity assessment. While retraining a model on new data does not inherently constitute a significant change, it may become one if it alters the model’s intended purpose or negatively impacts performance. Therefore, the team evaluates each case individually and prepares for a Technical File update or NB involvement only when necessary. **(**Note: Under the AI Act, any substantial modification to a high-risk AI system will require a reassessment of compliance, even if performance is maintained.)
- If the change is approved internally, the development team initiates a new development cycle, essentially starting again at phase 5.1 with the new data or revised design. The SOP-011 process is repeated: new data curation (which might largely involve adding the new data to the old and ensuring quality), re-training the model (phase 5.2), and so on. It’s not starting from zero, since a lot of the infrastructure and even documentation can be updated rather than created afresh, but all steps (verification, etc.) are performed for the new model version. This ensures the updated model undergoes the same rigor as the initial one, maintaining compliance.

Human oversight and continuity

During improvement, we ensure that any new model iteration still upholds the human interpretability and control measures. For example, if we were to switch to a more complex model architecture, we would consider whether that affected the ability to explain results to users (transparency). We would not sacrifice the human oversight features for performance. Also, training of end-users might be needed if the update changes the user interaction or output interpretation. These considerations are part of the change implementation plan.

Regulatory submission for updates

If the updated model changes the intended use or a key performance claim, a fresh conformity assessment may be required. Even if not, we must update the Declaration of Conformity and inform our Notified Body of the change (most likely scenario for a significant software change). We prepare an updated Technical Documentation bundle highlighting what changed in the model (often, a “change report” or summary of new validation). The AI Act (draft) would also require re-registration of the AI system if there’s a substantial modification. We track these regulatory triggers carefully, following guidance in place at the time of the update.

Outputs

The outputs of an improvement cycle include a new version of the ML model and software (with a new version number), along with all revised documentation:

- Revised Technical File sections (e.g., new performance data, updated algorithm description, new risk mitigations, if any).
- Model Update Validation Report: a targeted document summarizing what was changed and evidence that the updated model is as good as or better than the previous, without introducing new risks.
- Change control records linking the initial request, approval, and verification of implementation.
- Updated Instructions for Use if anything for the user changes (for example, if the model’s indications for use broaden, or performance claims are improved).

- Communication to users about the new version, if needed (especially if it's an update they need to install).

Once the updated model passes all checks, it effectively goes through deployment (phase 5.5) and into monitoring (5.6) again. The lifecycle thus continues.

Importantly, Hematoscope maintains a mindset of continuous learning but controlled deployment. Unlike non-regulated software, where "move fast and break things" might apply, here every ML model update is treated with formal diligence. This SOP ensures that improvements reach patients safely and effectively, preserving trust in the device and complying with regulatory obligations. Throughout this continuous improvement, transparency is key: we document changes, we inform stakeholders (users, regulators) of updates and their rationale, and we always ensure that a human is in the loop either in development or usage to oversee the AI.

By following this, Hematoscope can confidently integrate machine learning technology into its IVD devices while maintaining the high standards of quality, safety, and efficacy mandated by IVDR and emerging AI regulations. The SOP aligns innovation with compliance, enabling us to leverage data-driven improvements to patient care responsibly.

To support implementation, Table 1 provides a structured summary of the ML lifecycle phases, making it easier for employees to follow and apply the SOP in practice.

Table 1. Summary of ML Lifecycle Phases (Phases 5.1–5.7)

Phase	Key Objectives	Regulatory Touchpoints	Expected Outputs
5.1 Data Acquisition & Preparation	Build and curate high-quality datasets for model development; ensure data is	AI Act Art. 10 (data governance: data must be "relevant, representative, free of errors and complete" IVDR Annex I (device performance must be	Curated training, validation, and test datasets; Data documentation (data sources, preprocessing, quality checks, bias analysis);

	relevant to intended use, representative of diverse cases, and properly annotated. Identify and mitigate biases or data gaps early.	supported by appropriate data). GDPR for any personal data used (ensure consent/anonymization).	Data Management Plan approval.
5.2 Model Design & Training	Define model architecture and parameters based on requirements; implement and train the ML model on training data. Monitor training process to avoid overfitting and ensure reproducibility . Incorporate any risk controls in	IVDR Annex I §16.2 (software development lifecycle principles applied to ML AI Act Art. 9 (risk management during development) & Art. 14 (design for human oversight – e.g. include interface features enabling Use of validated tools (IEC 62304) and consideration of standards (e.g. IEC/TR 24029 for ML).	Model design specification (algorithm choice and rationale); Source code in version control; Training logs and scripts; Initial trained model (prototype) and training evaluation results.

	design (e.g. built-in checks).		
5.3 Verification & Performance Evaluation	Evaluate the trained model against predefined acceptance criteria. Verify that the model meets design requirements (accuracy, specificity, etc.) using validation and test datasets. Perform error analysis, stress cases, and ensure robustness across expected operating conditions.	IVDR Annex I (GSPR on performance and safety: e.g. demonstrate diagnostic sensitivity/specificity AI Act Art. 15 (ensure accuracy, robustness, cybersecurity – test model for edge cases and resilience If applicable, IEC 62366 (usability) – confirm users can interpret outputs.	Verification test cases and results (traceable to requirements); Model performance metrics on independent test set (meeting or exceeding thresholds); Updated Risk Assessment (residual risks documented, new hazards mitigated); Verification Report and/or Performance Evaluation Report drafted.
5.4 Technical Documentation & Review	Compile all design and development outputs into the technical	IVDR Annex II (Technical Documentation: document device description, design,	Updated Technical File section for ML (incl. model description, dataset summary, validation

	file. Conduct formal design review of the ML component (e.g. part of Design Review meetings) to ensure completeness and readiness for regulatory submission or release. Resolve any open issues prior to approval.	and evidence of conformity including ML algorithm details); AI Act Art. 11 (prepare detailed technical documentation before placing on market Internal QMS design review requirements (per SOP-009) apply – e.g. Design Output Review ensuring all inputs are addressed	results, algorithm change log); Design Review minutes and approval decision (e.g. Design Release Review signed-off); Go/no-go decision for deployment (or iteration if criteria not met).
5.5 Deployment & Release	Deploy the validated model within the product's production environment and perform final release checks. Ensure configuration management	IVDR Annex I §16.4 (specify minimum IT environment and security requirements for software IVDR Article 24 & Annex III (UDI and device identification for software versions); AI Act Art. 13 (provide clear instructions for use detailing AI system	Deployed model in production (e.g. cloud service or packaged software) with confirmation of environment setup; Installation/Deployment Verification Report (ensuring no regression in performance in production setting);

	for reproducibility. Generate user-facing documentation (IFU) that includes necessary information about the ML functionality. Officially release the software (version) for use.	capabilities, performance, and limitations. Ensure CE marking and Declaration of Conformity are in place before release.	Updated Instructions for Use and labeling that include ML-related information (e.g. intended purpose, performance metrics, warnings about algorithm limits. Software version number and Unique Device Identifier (for the ML-enabled software) assigned and documented; Declaration of Conformity updated to cover the software version with ML.
5.6 Post-Market Monitoring & Feedback	Continuously monitor the model's performance and safety in real-world use. Collect user feedback, incident reports, and periodic performance	IVDR Post-Market Surveillance (Article 78 and Annex III): execute PMS plan to gather and analyze performance data and user feedback; IVDR Annex XIII (for class C, perform Post-Market Performance Follow-up to confirm model continues to meet clinical performance).	Post-Market Surveillance Reports and/or PMPF reports including ML performance observations; Field data (e.g. aggregated model output vs outcomes) and complaint records (e.g. in Jira Service Management); Routine model

	data. Detect data drift or performance degradation over time. Implement a feedback loop to feed new relevant data or findings into model improvement.	AI Act Art. 12 (ensure automatic logging of events for traceability – system logs inputs, outputs, and outcomes for audit. AI Act Art. 15 (ongoing accuracy and robustness – monitor that the AI continues to perform consistently). Also comply with vigilance requirements (serious incidents reported via MIR).	evaluation reports (e.g. re-testing model on new data at defined intervals); Updates to Risk Management File and possibly software risk classification if operating environment changes.
5.7 Continuous Improvement & Model Updates	Manage changes to the ML model in a controlled manner. Decide when model retraining or updates are needed (e.g. significant new data, performance drop, emerging biases). For any update,	IVDR Article 27(3) and MDCG guidance on “significant changes” to software: determine if a model update (e.g. retraining with new data) is a significant change requiring NB notification or new conformity assessment. AI Act Art. 14 (establish human oversight procedures so that updates do not compromise the ability for humans to validate or override AI	Change Request documentation (justification for model update, impact analysis on safety/performance); Approved change control records (or new development plan if major change); New version of model and software (with new version identifier); Regression test and verification results for the updated model; Addendum to

	initiate a new development cycle under design control. Maintain transparency about changes and obtain necessary regulatory approvals for significant changes. Ensure human oversight remains effective during and after updates.	decisions. AI Act Art. 15 & 11: re-assess accuracy/robustness and update technical documentation whenever the model is modified.	Technical Documentation (describing the modification and updated performance evidence); Notification to regulators if required (for significant changes or trend reporting).

References

Internal Documents

- **SOP-004** – Risk Management
- **SOP-005** – Feedback Management and Customer Complaints
- **SOP-006** – Change Management Process
- **SOP-007** – Problem Management Process
- **SOP-008** – Vigilance
- **R3007** – Reporting Requirements and Reporting Register

Applicable Standards and Regulations

- **IVDR 2017/746** – In Vitro Diagnostic Medical Devices Regulation
- **EU AI Act 2024/1689** – Regulation on Artificial Intelligence
- **ISO 13485:2016** – Medical Devices – Quality Management Systems
- **ISO 14971** – Application of Risk Management to Medical Devices
- **IEC 62304** – Software Life Cycle Processes for Medical Device Software
- **IEC 62366** – Application of Usability Engineering to Medical Devices

Version History

Date	Version	Change	Impact of the changes on other guidelines	Education / Chapter / Information needs (+ target group)	Responsible person
	1	The first version.	N/A	Training and reading of the document for employees working on development and quality management activities.	