Exploring the challenges in the harmonization of clinical evaluation of medical device software across EU member states

Anna Polishchuk

2023 Laurea

Laurea University of Applied Sciences

Exploring the challenges in the harmonization of clinical evaluation of medical device software across EU member states

Anna Polishchuk

Global Health and Crisis Management

Master's Thesis

December, 2023

Laurea University of Applied Sciences **Abstract**

Degree Programme in Global Health and Crisis Management

Master of Health Care

Anna Polishchuk

Exploring the challenges in the harmonization of clinical evaluation of medical device software across EU member states

Year 2023 Number of pages 70

The goal of this master's thesis was to explore the challenges of harmonization of medical device software (MDSW) clinical evaluation across EU member states, with a particular focus on proposing practical recommendations for the EU SHAPES pilot project to meet the requirements of the EU Regulations for safety, performance, and clinical benefit. The research questions focused on both exploring the challenges in the harmonization of clinical evaluation of MDSW in the EU and the potential challenges in the EU SHAPES pilot project related to clinical evaluation of the MDSW developed within the project.

The medical device industry is evolving rapidly due to the escalating importance and impact of software designed for medical purposes in enhancing healthcare outcomes. Software plays a pivotal role in diagnosing and treating various medical conditions. Consequently, it becomes imperative for both manufacturers and regulators to possess a clear understanding of when software, functioning as a medical device, must adhere to relevant regulations.

In the dynamic landscape of evolving regulatory frameworks, harmonizing clinical evaluation processes for MDSW in the EU remains challenging. The crucial role of clinical evaluation in ensuring general safety and performance requirements is highlighted by the EU Regulations, such as the Medical Devices Regulation (MDR) and the In Vitro Diagnostic Medical Devices Regulation (IVDR).

This master’s thesis was carried out within the scope of the EU SHAPES pilot project. The EU SHAPES pilot project aimed to establish a comprehensive socio-technical infrastructure encompassing healthcare digital solutions, devices, and support services. The overarching goal is to empower aging individuals by fostering engaged, autonomous, and independent living within the familiar confines of their homes.

To address the research questions, the study engaged both internal and external experts from the EU SHAPES project and the wider European MDSW ecosystem. Qualitative research methods, particularly interviews, were employed to explore non-quantifiable concepts and gather insights based on expert opinions. This master's thesis provides a comprehensive account of the impact of the MDR and IVDR on the clinical evaluation of MDSW, examining practical challenges associated with the EU regulatory framework, national implementation, variations among notified bodies, clinical investigation practices, and the utilization of real-world data (RWD).

The insights derived from the expert interviews contribute to the formulation of recommendations for the clinical evaluation process of MDSW within the SHAPES project with a special focus on thorough qualification and risk class assessments for the software to meet the requirements of MDR/IVDR for general safety and performance, as well as clinical benefit.

Keywords: Medical Device Software, MDSW, Software as Medical Device, SaMD, Clinical evaluation, MDR, EU SHAPES

Contents

Introduction 7
1 Background 8
1.1 Software Used for Medical Purpose 9
1.2 Regulatory Considerations for Medical Device Software 10
1.2.1 EU Regulation Pertaining to Medical Device Software 10
1.2.2 Evaluation Requirements for Medical Device Software 15
1.2.3 Harmonizing Medical AI Regulation: The European Union's Approach 17
1.2.4 Harmonization of Clinical Evaluation: EU and UK Perspectives 17
1.2.5 International Collaborations in Regulation of Medical Device 18
1.3 Enhancing MDSW Regulation through the EU SHAPES Pilot Project 18
1.4 Harmonizing Clinical Evaluation Across EU Member States 20
2 Study Goal, Objectives, and Research Questions 21
3 Methods 21
3.1 Sampling 22
3.2 Data Collection 24
3.3 Data Analysis 25
3.4 PICO Model 26
3.5 Research Integrity and Research Ethics 27
3.6 Research Trustworthiness 27
4 Results 28
4.1 EU-Level Challenges 29
4.1.1 EU Legislation Relevant to MDSW 29
4.1.2 Clinical Evaluation of MDSW 32
4.2 National Level Challenges 35
4.2.1 Notified Bodies 35
4.2.2 National Approaches to Clinical Investigations 35
4.2.3 Classification Differences 36
4.2.4 Use of Real-World Data 37
4.3 Summary 38
5 Discussion 39
5.1 EU-Level Challenges 39
5.1.1 Legislation and Clinical Evaluation 39
5.1.2 National Level Challenges 43
5.2 The SHAPES Pilot Project: Advancing Technical Solutions for Seniors 47

5.2.1 Intended Use of the Software 47
5.2.2 Regulatory Landscape: Software as a Solution, not a Product 48
6 Conclusions 50
7 Recommendations for the EU SHAPES Pilot Project 51
References 53
Figures 60
Tables 60
Appendices 61

Introduction

The harmonization of clinical evaluation processes for medical device software (MDSW) within the EU is steered by EU regulations governing medical devices. The most recent regulations in this domain are the Medical Devices Regulation (MDR) and the In Vitro Diagnostic Medical Devices Regulation (IVDR). MDSW is deemed a medical device under these Regulations when it aligns with their specified definitions. These Regulations extend their implications to AI-based medical device software, encompassing a framework for clinical evaluation to ensure safety, performance, and compliance with essential requirements (Regulation (EU) 2017/745; Regulation (EU) 2017/746).

To achieve harmonization among EU member states a centralized coordination framework known as the Medical Device Coordination Group, MDCG, was established (MDCG Working Groups 2023). The MDCG serves a pivotal role by offering advice and recommendations regarding the interpretation and application of the Regulations. Furthermore, it fosters cooperation, coordination, and the exchange of information among EU member states, facilitating the sharing of exemplary practices. The clinical evaluation process assumes a pivotal role in ascertaining the safety, performance, and clinical benefit of MDSW. This process entails a meticulous and systematic examination of clinical data pertinent to the application of the software (European Commission 2016; Regulation (EU) 2017/745).

The MDCG has issued comprehensive guidance documents aimed at harmonizing clinical evaluation of MDSW. These materials encompass a wide array of topics, including clinical evidence requirements, methodologies for clinical evaluation, post-market clinical follow-up, and issues specific to software (MDCG 2020). They play a crucial role in assisting both manufacturers and notified bodies to assess MDSW compliance across the EU.

The MDR requires clinical evaluation regardless of the risk class of the device to verify its compliance with the general safety and performance requirements (Article 61). The MDR and IVDR introduce stringent clinical evidence and post-market surveillance requirements for medical device software (Regulation (EU) 2017/745; Regulation (EU) 2017/746). The MDR encompasses the need to provide clinical evidence substantiating the safety and performance of the software. If required, clinical investigations must be conducted, and robust post-market monitoring systems must be established to continuously evaluate the software's performance and identify potential risks or issues (MDCG 2020).

The challenges encountered by the healthcare system due to the implementation of digital solutions in recent years necessitate the exploration of practical remedies. One such viable approach to enhance the effectiveness of executing complex strategic tasks is the utilization of pilot projects. The overarching goal of an EU Horizon 2020 SHAPES (Smart and Healthy Ageing through People Engaging in Supportive Systems) pilot project was to establish the first European open ecosystem for supporting and extending the healthy and independent living of older individuals experiencing reduced functionality (Grigoleit 2020).

To foster the implementation of new integrated-care services in Europe, the EU SHAPES pilot project offered comprehensive guidelines, a roadmap, an action plan, and a set of priorities for standardization. It also extended its support to key EU stakeholders in alignment with the SHAPES Ethical Framework (Sarlio-Siintola, Nikula, Aholaakko, Alapuranen & Kartsidis 2019).

This thesis served to explore challenges in harmonization of clinical evaluation of MDSW among EU member states and to formulate practical recommendations for the EU SHAPES pilot project. Since expertise of relevant stakeholders representing the European MDSW ecosystem (Tzintoli 2022) is of a great importance, the study was aimed to actively engage both the EU SHAPES pilot project internal and external experts who shared their experiences and viewpoints. This input contributed to the development of final recommendations for the clinical evaluation process of MDSW within the EU SHAPES pilot project.

In this study, interviews with industry experts delved into the primary consequences of the EU regulatory framework on clinical evaluation harmonization of software qualified as medical device. The nexus between the regulatory framework and the challenges faced by manufacturers of MDSW in ensuring compliance is of paramount significance. These discussions also addressed practical challenges associated with the national implementations, encompassing variations among requirements of notified bodies and competent authorities, clinical investigation practices, and utilization of real-world data (RWD).

1 Background

The background section provides an overview of the new regulatory landscape, qualification and risk classification of MDSW, clinical evaluation, the role of notified bodies, and other pertinent issues. Lastly, it delves into the medical device software developed within the scope of the EU SHAPES pilot project.

1.1 Software Used for Medical Purpose

The EU has a market comprising thousands of medical devices, underscoring the sector's significance not only in delivering essential healthcare services to citizens but also in contributing significantly to the global economy. Medical device solutions encompass a wide range of functions, including diagnosis, prevention, monitoring, prediction, prognosis, treatment, and disease alleviation (Bianchini & Mayer 2022).

The landscape of the medical device industry is evolving rapidly due to the escalating importance and impact of software designed for medical purposes in enhancing healthcare outcomes (Deniz-Garcia et al. 2023). Software plays a pivotal role in diagnosing and treating various medical conditions, spanning both mental health and pathological conditions (Dang, Arora & Rane 2020). The digital health field is witnessing exponential growth, leading to a proliferation of digital health devices, including both MDSW and Artificial Intelligence (AI) - based Medical Device Software (AI-MDSW) within the EU market. Consequently, it becomes imperative for both manufacturers and regulators to possess a clear understanding of when software, functioning as a medical device, must adhere to relevant regulations (Ludvigsen, Nagaraja & Daly 2021).

A crucial criterion for software to qualify as MDSW is that it must possess a medical purpose. MDSW is intended for utilization, either independently or in conjunction with other components, to fulfill the purposes outlined in the definition of a "medical device" as outlined in MDR/IVDR (MDCG 2019). Stand-alone MDSW encompasses software explicitly designed to be used independently as a medical device.

Illustrative examples of stand-alone MDSW include:

- Clinical Decision Support Systems (CDSS): These systems leverage patient data to provide healthcare providers with evidence-based recommendations and alerts, aiding in clinical decision-making (Sutton et al. 2020).
- Medical Imaging Software: Tailored for the interpretation of medical images, such as radiology software employed in the analysis of X-rays, CT scans etc. (Tournier et al. 2019).
- Mobile Health Applications: MDSW facilitates users in managing their health, tracking, and analyzing health-related data, monitoring chronic diseases, and even conducting remote consultations (Levine et al. 2020).

- AI for Disease Diagnosis: Stand-alone MDSW may harness AI algorithms to analyze medical data, including patient records or images, to facilitate illness diagnosis, prediction, and treatment planning (Kumar, Koul, Singla & Ijaz 2022).

AI-based technology has gained prominence in the healthcare sector across various domains, including medical practices, diagnostics, and healthcare system administration (He et al. 2019).

The distinguishing feature between MDSW and AI/ML- MDSW lies in the incorporation of AI and machine learning (ML) capabilities. Machine learning operates on the premise that computer learning systems acquire knowledge from experience and continually enhance their performance over time (FDA 2019; IMDRF 2022). While conventional MDSW primarily focuses on fundamental functions such as device control and data monitoring, AI/ML medical device software advances further. It leverages advanced algorithms and techniques to provide sophisticated analysis, pattern recognition, and decision support (Muehlematter, Daniore & Vokinger 2021).

It is imperative to emphasize that AI/ML medical device software, akin to other medical devices, is subject to rigorous regulatory scrutiny and must adhere to applicable standards and regulations (Gerke, Babic, Evgeniou & Cohen 2020). However, it is worth noting that the existing regulatory framework for MDSW often encounters challenges when dealing with machine learning algorithms that undergo continuous retraining (Kwade 2022). To address important challenges in this field, it is urgently necessary to design guidelines for diagnostic accuracy studies specifically for artificial intelligence (Aggarwal 2021).

1.2 Regulatory Considerations for Medical Device Software

In recent years, the evolving landscape of MDSW within the EU has prompted a heightened focus on regulatory considerations. As technological advancements continue to shape the healthcare industry, the intersection of innovation and regulatory frameworks becomes increasingly intricate. The exploration of regulatory considerations for MDSW within the EU is imperative for stakeholders ranging from manufacturers and healthcare practitioners to policymakers, as they collectively strive to strike a delicate balance between patient safety, technological advancement, and regulatory adherence.

1.2.1 EU Regulation Pertaining to Medical Device Software

Transition from Directives to Regulations: The regulation of medical devices in the EU underwent a significant transformation, transitioning from a set of Directives, including 93/42 on medical devices (MDD), Directive 98/79 on in vitro diagnostic medical devices, and

Directive 90/385 on active implantable medical devices, to two Regulations - Regulation 2017/745 on medical devices (MDR) and Regulation 2017/746 on in vitro diagnostic medical devices (IVDR). Unlike the Directives, which required incorporation into national legislation, the MDR and IVDR, being Regulations, came into direct force in all EU member states (Regulation (EU) 2017/745; Regulation (EU) 2017/746).

The MDR introduces a risk-based classification system categorizing medical devices into four classes (Class I, IIa, IIb, and III); contingent upon the level of risk they pose to patients. The MDR/IVDR accentuate the necessity for more robust clinical evidence supporting the safety and performance of medical devices, especially for higher-risk devices, accompanied by stricter post-market surveillance requirements encompassing enhanced vigilance reporting, post-market clinical follow-up, and periodic safety update reports. A novel Unique Device Identification (UDI) system is introduced to enhance medical device traceability and post-market monitoring (European Union 2019).

The Regulations recognize medical device software or in vitro diagnostic (IVD) device, contingent upon its intended purpose and the claims made by the manufacturer. In October 2019, the European Commission's MDCG issued guidance on the Qualification and Classification of Software in Regulation (EU) 2017/745 - MDR and Regulation (EU) 2017/746 - IVDR. This guidance defines "software" as a set of operations processing input data and generating output data. Primarily targeted at medical software manufacturers, this guidance delineates the criteria that software must meet to qualify under the MDR (2017). It also offers guidance on the application of classification standards for software under Regulations along with insights into placing products on the market (MDCG 2019).

The MDR/IVDR mandate heightened performance and safety standards for software designed for medical purposes. These requirements parallel those applicable to medical devices and IVD devices, encompassing changes in risk classification, clinical evaluation, risk management protocols, post-market surveillance, and the role of notified bodies. These elements are emphasized as the "key components" of the regulatory framework, aiming to reinforce and harmonize medical device regulations across the EU, thereby ensuring the safety, performance, and effectiveness of software throughout its lifecycle.

Notified bodies: Assessing Quality Management Systems and Technical Documentation

The EU has long championed stringent measures to ensure the safety and regulatory compliance of medical devices (EMA 2021). The establishment of a system of notified bodies entrusted with the task of conformity assessment serves this goal. Notified bodies play a pivotal role in the CE marking process for medical devices by meticulously evaluating these

devices to determine their compliance with the exacting EU regulations and standards (Regulation (EU) 2017/745).

Notified bodies shoulder the responsibility of scrutinizing the manufacturer's Quality Management System (QMS) and technical documentation (Regulation (EU) 2017/745, Annex IX). This arduous assessment process, with the exception of Class I devices, culminates in the issuance of essential Annex certificates, particularly Annex XII under the MDR. Compliance with these stringent prerequisites is an indispensable prerequisite for manufacturers seeking to declare their conformity with EU regulations (Regulation (EU) 2017/745).

However, it is important to consider the concerns raised by scholars (Jarman, Rozenblum & Huang 2020). They have underscored a critical limitation within the current regulatory framework—a limitation stemming from the delegation of conformity assessment to privately held entities, including both manufacturers and notified bodies. This delegation introduces market-driven dynamics into the regulatory process, where companies are granted autonomy in selecting notified bodies for conformity assessment. This competitive landscape may inadvertently shift the focus from regulatory diligence to market competition, prompting concerns about impartiality and diligence in conformity assessment (Jarman et al. 2020).

Jarman et al. (2020) contended that the motivations for national competent authorities to effectively monitor the performance of notified bodies may be insufficient in this context. This highlights an ongoing debate regarding the need for robust regulatory oversight mechanisms to ensure that conformity assessments remain steadfastly rooted in regulatory compliance and patient safety.

Intended medical purpose

In the framework of the MDR/IVDR, the manufacturer is critical in defining the intended use of an MDSW and confirming that it meets the regulatory requirements. This idea makes it clear that having the features described in either definition of a device does not suffice; the manufacturer of the device must also intend for it to be used for one of the medical uses. Software must have a "specific medical purpose" to be considered an IVD or a medical device (Regulation (EU) 2017/745, Article 1).

It is crucial to underscore that mere utilization in a medical setting does not suffice for a device to garner qualification as a medical device. Rather, it is contingent upon the manufacturer's explicit intent for a medical purpose. For instance, several "sports-oriented" devices may monitor organ functionality, but their intended purpose revolves around athletic

rather than medical considerations. Consequently, such devices should not be categorized as medical devices (Ordish, Hannah & Hall 2019).

Moreover, the description of the intended purpose necessitates the inclusion of a statement delineating benefits for the patient. The absence of such a statement precludes the device from being marketed as a medical device, as stipulated in Articles 61-62, Annexes XIV and XV of the MDR.

Qualification of the medical device software under MDR/IVDR

The MDR/IVDR regulations introduce distinct categorizations for various types of software, creating a clear demarcation:

- Software as a Medical Device or In Vitro Diagnostic Device.
- General Software (Excluding Medical or In Vitro Diagnostic Devices).
- Applications for Lifestyle or Wellness Objectives (Excluding Medical or In Vitro Diagnostic Devices).

For manufacturers, comprehending the potential regulatory framework applicable to their software products is helpful. The MDCG 2019-11 "Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 - MDR and Regulation (EU) 2017/746 - IVDR," provides insights into the classification of software used in medical devices and addresses related topics. Furthermore, it distinguishes between stand-alone MDSW, and software integrated into existing medical devices. Stand-alone MDSW refers to software intended for autonomous use, separate from any hardware medical equipment or in vitro diagnostic device. Its purpose is to provide diagnostic, therapeutic, or monitoring functionalities to patients or healthcare providers (MDCG 2019).

The development and utilization of MDSW for medical control and optimization must adhere to the regulatory provisions of the MDR/IVDR. Software qualifies as MDSW if its intended purpose involves processing, analyzing, creating, or modifying medical information. For instance, software that alters data representation for medical purposes falls under this category. Another example is an MDSW smartwatch app designed to send alarm notifications to users and healthcare practitioners upon detecting irregular heartbeats for cardiac arrhythmia detection (MDCG 2019).

Software may qualify as MDSW regardless of its location, whether in the cloud, on a computer, a mobile phone, or as additional functionality within hardware medical devices. Moreover, MDSW may be intended for use by healthcare professionals or laypersons. When

intended for layperson use, manufacturers must adhere to safety and performance requirements outlined in MDR Annex I. 22 and 23.4 (w) or IVDR Annex I. 9.4 and 20.4.2.

Advances in highly accurate sensors available outside of clinical settings have eroded the once-clear boundary between "consumer gadgets" and "medical devices." To address this, the criteria for defining a medical device must strike a balance, being sufficiently rigid to provide manufacturers with predictability regarding their product's classification while retaining flexibility for regulators to regulate unsafe devices. In this regard, it has been recommended that the European Commission update and expand the Handbook on Borderlines and Classification in the Community Regulatory Framework for Medical Devices under the Regulations (Ordish et al. 2019).

Risk classifications

Under the MDR/IVDR software is subject to risk classifications: it is classified based on its intended purpose, associated risks and potential impact on patient safety. The Regulations categorize software into Classes: I, IIa, IIb, and III under MDR (Figure 1) and Classes: A, B, C, D under IVDR. The classification determines the level of scrutiny and conformity assessment procedures required for the device.

The International Medical Device Regulators Forum (IMDRF) provided recommendations for the establishment of the new classification rule (IMDRF/ SaMD/ Working Group 2014). Adjustments in classification have been a major source of concern with respect to medical device software. In contrast to the past, where most software medical devices were Class I, today nearly all will probably fall into Class IIa, IIb or possibly Class III (the highest risk category). Therefore, they will be subject to conformity inspection by a notified body (Keutzer & Simonsson 2020).

Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as Class IIa, except if such decisions have an impact that may cause:

• Death or an irreversible deterioration of a person's state of health, in which case it is in class III;

or

• Serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as Class IIb.

Software intended to monitor physiological processes is classified as Class IIa, except if it is intended for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as Class IIb.

All other software are classified as Class I.

Figure 1: MDR Annex VIII, section 6.3, Rule 11

1.2.2 Evaluation Requirements for Medical Device Software

Within the context of MDR/IVDR compliance, manufacturers assume a pivotal role in the evaluation process, wherein two distinct types of evaluation come into focus: "clinical evaluation" if the software qualifies as a medical device, and "performance evaluation" if the software falls under the category of an IVD. Key legislative documents germane to clinical evaluation encompass MDR, MEDDEV 2.7/1 revision 4 - Clinical evaluation: Guide for manufacturers and notified bodies, as well as MDCG guidance (2020).

Clinical evaluation assumes a pivotal role in assessing the safety, performance, and clinical benefits of medical device software. This process entails a meticulous and methodical analysis of clinical data pertaining to the software's usage. It is defined as "a systematic and planned process to continuously generate, collect, analyze, and assess the clinical data pertaining to a device to verify its safety and performance, including clinical benefits when used as intended by the manufacturer" (Regulation (EU) 2017/745, Article 2). MDR mandates manufacturers to conduct clinical evaluations for all medical devices, including software, unless specific exemptions apply.

The clinical evaluation process for MDSW typically encompasses the following activities:

- Data Collection and Analysis

Manufacturers are tasked with collating and scrutinizing clinical data from diverse sources, spanning clinical investigations, post-market surveillance, scientific literature, and other clinical experiences.

- Safety and Performance Verification

The compiled data must demonstrate the software's adherence to general safety and performance requirements. This evidence must be directly pertinent to the software and its intended purpose.

- Risk-Benefit Analysis

Manufacturers are required to conduct a risk-benefit analysis, juxtaposing the potential risks of the MDSW against its clinical benefits. This analysis aids in ascertaining the appropriateness of the risk-benefit ratio for the software's intended use.

- Continuous Monitoring

It is imperative to continually monitor the software's performance and safety. Manufacturers must establish a post-market surveillance system to amass RWD utilization, encompassing the observation of any side effects or issues encountered.

The "sufficient clinical evidence" standard, necessary for compliance, hinges on three primary factors: the device's intended use, an evaluation of its side effects, and the applicability of the resulting risk-benefit ratio. Manufacturers are generally obliged to elucidate why the provided clinical evidence suffices to meet conformity standards. In certain instances, especially when adequate clinical evidence coverage is unavailable, clinical investigations may be mandated.

Both MEDDEV 2.7/1 rev. 4 and the MDR have significantly impacted the clinical evaluation of MDSW. Notably, MEDDEV 2.7/1 rev. 4 introduced the concept of a clinical evaluation plan (CEP) and emphasized the need for continuous clinical evaluation (European Commission 2016). The MDR focused on risk assessment, particularly benefit-risk analysis, and the requirement to substantiate claimed benefits with evidence. It also highlighted the significance of Post-Market Clinical Follow-up (PMCF).

The culmination of the clinical evaluation process results in the production of a clinical evaluation report (CER). This report relies on critical analysis of relevant scientific literature, an exhaustive assessment of all available clinical investigation outcomes, and a consideration of alternative care or diagnostic options currently available. While creating a CER may require more effort compared to MEDDEV 2.7/1 rev. 3, the entire clinical evaluation process is now more systematic, objective, robust, and comprehensive (Pritchard 2022).

For IVD products, developers must furnish a performance evaluation report as mandated by IVDR. Scientific validity, analytical performance, and clinical performance represent the primary facets of this evaluation, encompassing the connection between an analyte and a clinical or physiological condition, the device's capacity to accurately detect the analyte, and its ability to deliver pertinent results for the intended use.

Interestingly, Ordish et al. (2019) contends that the performance evaluation framework outlined in the IVDR might be better suited for diagnostic machine learning devices compared to the clinical evaluation framework under the MDR. This juxtaposition raises intriguing questions, given that a significant portion of clinical machine learning devices will likely fall under the category of medical devices rather than IVDs.

It is imperative to recognize that the awarding of a CE mark does not absolve the manufacturers of their responsibilities (Pane et al. 2019). In fact, under the MDR/IVDR manufacturers may be obligated to engage in post-market surveillance (PMS), including periodic safety reviews, as integral components of their ongoing surveillance strategy (Pane et al. 2019).

In summary, the onus rests with software manufacturers to instill awareness and compliance with the prevailing regulations MDR/IVDR at the earliest stages of development. This proactive approach is indispensable for ensuring requisite traceability across different stages within the software lifecycle (Bianchini, Francesconi, Testa, Tanase & Gemignani 2019).

1.2.3 Harmonizing Medical AI Regulation: The European Union's Approach

In addition to the recent implementation of the Regulations governing medical devices, the EU has embarked on a comprehensive strategy aimed at legislating the burgeoning field of AI. These legislative endeavors address several key concerns arising from the rapid advancements in medical AI. One of the primary instruments of this strategy is the proposed Artificial Intelligence Act (AIA) (European Commission. Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence 2021).

The fundamental objective of the AIA proposal is to ensure the smooth functioning of the EU's internal market by establishing harmonized regulations, particularly in the context of the development, introduction, and utilization of products and services powered by AI technologies. This encompasses AI systems deployed as stand-alone entities, as well as those integrated into broader technological frameworks, thus encompassing a comprehensive approach to AI governance.

1.2.4 Harmonization of Clinical Evaluation: EU and UK Perspectives

Preceding the United Kingdom's departure from the EU, the clinical evaluation processes for software-based medical devices in EU member states were primarily aligned with the standards outlined in the Medical Devices Directive (MDD) (Council E. 1993) and the Active Implantable Medical Devices Directive (AIMDD) (Directive 1990). However, with the advent of the MDR/IVDR efforts towards harmonization have expanded to encompass all EU member states. Simultaneously, the UK has forged its own regulatory framework for medical devices, inclusive of software. The oversight of the UK's regulatory framework falls under the purview of the Medicines and Healthcare products Regulatory Agency (MHRA), which remains in harmony with international regulations (Software and AI as a Medical Device Change Programme 2023). Notably, despite the UK's establishment of an autonomous regulatory

system, ongoing endeavors to unify clinical evaluation standards for software medical devices across both EU member states and the UK persist.

1.2.5 International Collaborations in Regulation of Medical Device

Two pivotal regulatory frameworks in the realm of medical devices are the EU MDR/IVDR and the International Medical Device Regulators Forum (IMDRF). While these are distinct entities, they play indispensable roles in ensuring the safety, efficacy, and quality of medical devices on a global scale.

The EU MDR/IVDR is focused on regulating medical devices and IVD devices within the confines of the European Union. In contrast, the IMDRF operates on a broader, international stage, uniting regulatory authorities from diverse regions with the shared goal of achieving a common understanding and alignment of medical device regulations worldwide. Notably, the Software as a Medical Device (SaMD) IMDRF Working Group (Software as a Medical Device (SaMD) 2017) has been instrumental in shaping the regulatory landscape pertaining to software employed for medical purposes.

The term "Software as a Medical Device", used by IMDRF, encompasses software designed for medical purposes, operating independently of hardware medical devices. The clinical evaluation process for SaMD involves the establishment of a valid clinical association between the software's output and the targeted clinical condition or pathological state. Furthermore, it necessitates the provision of requisite technical and clinical data. Manufacturers of SaMD are also expected to institute ongoing lifecycle processes to continually assess the product's performance in its intended market (IMDRF/SaMD WG/N41 2017). This approach underscores the dynamic nature of regulatory developments in the sphere of software-based medical devices and the imperative of aligning with these evolving standards.

1.3 Enhancing MDSW Regulation through the EU SHAPES Pilot Project

The realm of medical device software regulatory issues stands to gain valuable insights from pilot projects. The healthcare system has confronted formidable challenges in recent years, particularly with the integration of digital solutions. As a response to these challenges, the project-based approach has emerged as an effective methodology, demonstrating its prowess over decades in numerous developed countries. Pilot projects represent a tangible endeavor geared towards enhancing the efficiency of implementing intricate strategic objectives. In this context, the healthcare industry assumes a paramount role, offering fertile ground for the evolution of medical MDSW/AI-MDSW systems.

One noteworthy initiative supported by the EU is the EU Horizon 2020 SHAPES pilot project. The primary aim of SHAPES project was to establish the inaugural European open ecosystem, fostering the widespread adoption of diverse digital solutions. These solutions are geared towards promoting and prolonging the health and independence of elderly individuals facing diminishing functionality and capabilities (Grigoleit 2020).

Central to SHAPES was the development of an interoperable platform that integrates intelligent digital solutions. These solutions encompass the collection and analysis of health, environmental, and lifestyle data among older individuals. Within this ecosystem, connected devices, AI, robotics, big data, and other technologies converge to offer personalized and optimized healthcare delivery. SHAPES operated with the active participation of entities from several EU member states, the UK, and Norway.

To catalyze the robust framework and implementation of integrated-care services across Europe, SHAPES has formulated a set of recommendations. These recommendations provided guidelines, a roadmap, an action plan, and a set of priorities for standardization, actively involving significant EU stakeholders. The recommendations are grounded in empirically supported findings, which underscore the platform's acknowledged value in enhancing autonomy, enabling active living, and fortifying the long-term sustainability of healthcare delivery systems across Europe (Sarlio-Siintola et al. 2019).

Crucially, SHAPES placed a premium on ensuring the sustainable utilization of these services, as well as the prevention and management of health issues commonly encountered by older populations. To achieve these objectives, SHAPES engaged all pertinent stakeholders in the development of digital health services. Through collaboration with technology firms, service advisors, end-users, and beneficiaries, SHAPES endeavored to co-create innovative digital solutions and devices. This collaborative effort encompassed seven pilot teams within the project, covering areas such as medicines control and optimization, psychosocial and cognitive stimulation for enhancing well-being, in-home physical rehabilitation, and cross-border data exchange (Spargo et al. 2021).

Comorbidities, characterized by the presence of multiple chronic illnesses in a single individual, are prevalent among older populations (Atella et al. 2018). Managing medications for individuals with comorbidities is a multifaceted and intricate task that demands continuous monitoring and adjustments to ensure safety and efficacy. MDSW can significantly facilitate this process by offering a range of functionalities, including integration with other health monitoring devices such as wearable sensors or home monitoring systems for data collection and analysis. This data encompasses vital signs, activity levels, glucose levels, and

various other health indicators (Grigoleit 2020). The specific use cases for pilot themes (UC-PT3-00x) are listed in table below (Table 1).

Table 1: List of some use cases of the SHAPES pilot project (Grigoleit 2020)

Use case	"Support Multi-morbid Older Individuals"
UC-PT3-001	In-home decompensation prediction for heart failure patients
UC-PT3-001b	Prediction of stroke by home using blood pressure values
UC-PT3-001c	Advanced telemonitoring of patients with heart failure in home environment
UC-PT3-002	Diabetes self-management, control and prevention
UC-PT3-002b	Monitoring of blood glucose levels to older individuals with diabetes or pre-diabetic, abnormal glucose indications

MDSW serves as a valuable tool for assessing the effectiveness of pharmaceutical regimens, incorporates decision-support algorithms that provide evidence-based recommendations for medication adjustments or alterations based on patient-specific data (Armando, Miglio, de Cosmo & Cena 2023). These systems leverage clinical guidelines, pharmacological databases, and patient records to assist healthcare providers in making informed decisions regarding medication optimization and management. However, it is essential to underscore that MDSW should always be used in conjunction with the oversight of healthcare professionals, and it should never supplant the skill and judgment of medical experts.

1.4 Harmonizing Clinical Evaluation Across EU Member States

The rise of MDSW/AI-MDSW transforms the medical device industry. Harmonizing MDSW/AI-MDSW clinical evaluation is not feasible with EU member states acting alone. Especially, AI-MDSW relies on diverse datasets and can integrate into any product or service within the internal market (IMDRF 2022).

Regulators and notified bodies must acquire expertise to evaluate these digital health devices, including machine learning. Market stakeholders new to the Regulations should understand their obligations under MDR and IVDR. Many MDSW manufacturers are now under MDR, lacking the support traditional medical device manufacturers enjoy (Ordish et al. 2019).

Thus, involving relevant stakeholders representing the European MDSW/AI-MDSW ecosystem is crucial. They contribute to creating practical recommendations for further harmonizing clinical evaluation within the EU.

2 Study Goal, Objectives, and Research Questions

The goal of the study is to explore challenges in the harmonization of clinical evaluation of MDSW across EU member states to meet requirements of the EU Regulations for safety, performance and clinical benefit.

Study Objectives

1. To reveal challenges that are still present in the harmonization of clinical evaluation of MDSW across EU member states.
2. To create recommendations for EU SHAPES project for clinical evaluation of the MDSW.

Research Questions

1. What are the challenges in the harmonization of clinical evaluation of MDSW across EU member states?

2. What are the challenges in the EU SHAPES project related to clinical evaluation of the MDSW developed within the project?

3 Methods

This study focuses on exploring concepts that are not easily quantifiable and are better addressed through qualitative research methods. In research literature, qualitative research has carved out a special niche for itself with the goal of producing information based on the opinions of people (Sandelowski 2004). Qualitative method such as interviews were employed to uncover major themes related to challenges of clinical evaluation of MDSW across EU member states and recommendations to the EU SHAPES pilot project.

Aside from the obvious advantages, it is clear that expert interviews provide researchers with an efficient technique of collecting results, particularly those that are favorable. The reality that the interviewer and interviewee share a common academic education or value system can usually boost the level of inspiration that is required of the expert in engaging in the interview (Bogner, Littig & Menz 2009).

3.1 Sampling

The sampling technique relied on non-probability, heterogeneous, and purposive sampling. The primary objective was to identify major themes, thus heterogeneous sampling was employed to fulfill this objective (not aimed at in-depth information) (Clarke, Braun & Hayfield 2015). Sampling can address different levels (sites, people, events, etc.) depending on your research question and the method you will apply" (Flick 2018).

Eligibility criteria for selecting the participants included holding a senior position in a relevant field, defined as having several years of experience, with a professional background within the EU region, as well as the EU SHAPES pilot project participants involved in the use cases relevant to the focus of this study. These criteria were deemed essential to ensure that interviewees possess substantial knowledge of MDSW clinical evaluation and could provide in-depth responses to study inquiries.

Participants were sourced from various outlets and initially contacted via email. The participant information letter (Appendix 1) and informed consent form (Appendix 2) were included in the invitation email. Additionally, the author participated in person in the 23rd meeting of the International Medical Device Regulators Forum (IMDRF) Management Committee (MC) and Official Observers took place in Brussels, Belgium on 27 and 28 March 2023 ((Brussels, Belgium (hosted by European Commission on behalf of the EU) 2023). In 2023, the meeting was chaired by the EU. Around 300 participants attended in person and a further 200 virtual attendees participated in the first two public days. The author contacted some interview participants in person to agree the schedule for the interviews.

The author actively engaged in a workshop held during the IMDRF meeting, focusing on "The Lifecycle of Medical Devices: Emphasizing Post-Market Activities." This workshop saw active participation from regulatory authorities, industry representatives, and healthcare professionals, who partook in four informative sessions accompanied by panel discussions. The initial two sessions delved into topics like safety notifications, vigilance, and real-world evidence (RWE). Participants actively addressed challenges and presented innovative ideas to enhance existing systems, ultimately ensuring the safety, effectiveness, and performance of medical devices. They explored potential methods for the collection, validation, and

utilization of RWE within the regulatory framework, aiming to enhance post-market surveillance.

The subsequent two sessions, Sessions 3 and 4, revolved around post-market considerations pertaining to software, particularly AI and AI-based MDs. Attendees actively presented and discussed criteria, methodologies, and strategies to monitor the safety and performance of software, including the unique challenges and opportunities associated with gathering or generating data for digital MDs. Additionally, they delved into specific considerations necessary for the post-market oversight of AI-based MDs (International Medical Device Regulators Forum, Brussels 2023).

The author also participated online in the 24th meeting of the International Medical Device Regulators Forum (IMDRF) Management Committee (MC) and Official Observers took place Berlin, Germany 25 and 26 September 2023 (Berlin, Germany (hosted by European Commission on behalf of the EU) 2023). On 26 September 2023, the IMDRF Stakeholder Forum was held where IMDRF MC Members and Official Observers provided regulatory updates. A novel feature introduced based on feedback from the 23rd IMDRF session in Brussels, was the interactive 'flash panel' discussions on two subjects of interest: Unique Device Identification (UDI) as well as Digital Therapeutics.

Altogether 10 informants with a relevant professional background within the EU region and the focus of this study participated in the interviews (Table 2). The positions of the interviewees are generalized to preserve their anonymity.

Table 2: Informants (N=10), their ID, position, professional field, and number of words in transcript

ID	Position	Professional field	Number of words in transcript
a1	Regulatory and Academia	Medical Device Software and Regulatory, Academia	9074
p1	EU Policymaker	Medical Device Software and Regulatory	8884
p2	EU Policymaker	Medical Device Software and Regulatory	6850
a2	Regulatory and Academia	Medical Device Software and Regulatory, Academia	9930
m1	MDSW Manufacturer	Health Technology	10526
h1	Scientific and Healthcare Director	Pharmaceutical and Medical Device	7045
s1	Researcher Health Artificial Intelligence Expert	Health Technology	6505
r1	Regulatory	Health Technology Manufacturer and Regulatory	4364
c1	Senior Consultant	Life Science, Health Technology and Regulatory	10127
r2	Research regulatory	Health Technology	3894
Informants (N=10)	Total number of words		77199

3.2 Data Collection

Data collection involved semi-structured interviews conducted via video conferencing over the internet. Given the target audience's likely busy schedules, Microsoft Teams video conferences over the internet were chosen to maximize their time utilization. The interviews were done during April, May, and June 2023.

Semi-structured interviews were chosen due to their flexibility in adapting to the participants' perspectives and building rapport (Appendix 3). This flexibility proves advantageous when dealing with a diverse sample, allowing for adjustments to question order and format based on individual respondent characteristics and professional expertise. Outstanding interviews always benefit from the researchers' openness in adapting their questions to the particular informant and the circumstances of the specific interview (Flick 2018). All interviews were conducted in English.

3.3 Data Analysis

The data analysis for this study adopted a thematic analysis approach as outlined by Clarke et al. (2015), offering the flexibility to employ both deductive and inductive methods. While template analysis, a subtype of thematic analysis involving the initial creation of color codes using a subset of the data, was considered, the chosen approach involves analyzing the entire dataset with a template derived from these initial codes.

The process of thematic analysis, as described by Clarke et al. (2015), encompasses the following key steps:

- Becoming Familiar with the Data: The initial step involves thoroughly reviewing the interview transcripts and making preliminary notes on noteworthy aspects within the dataset.
- Generating Initial Codes: In this stage, the focus is on the creation of initial codes that are firmly rooted in the data itself. Each piece of information in the transcripts is meticulously examined to identify recurring themes, laying the foundation for the smallest thematic units known as codes.
- Identifying Themes: At this juncture, the aim is to identify both sub-themes and overarching themes. Codes are grouped based on their interrelatedness within the data. Codes form the building blocks for sub-themes, which, in turn, contribute to the emergence of broader themes.
- Reviewing Themes: This phase involves revisiting the initial themes to ensure they are well-supported by the data. Themes that lack substantial evidence are either discarded or merged with similar themes. Any overly complex concepts are deconstructed into multiple interrelated themes as necessary.
- Defining and Naming Themes: Each theme and sub-theme is meticulously defined and substantiated with data excerpts from the interview transcripts.

By adhering to this systematic approach, the study aimed to uncover and describe significant themes and sub-themes embedded within the interview data, providing valuable insights into the subject matter at hand. An excerpt of the raw transcript data from Appendix 4 illustrates transcript material and an example of the analysis is included in Appendix 5.

The study's findings have been scrutinized in light of two research questions. Thematic analysis of the data derived from 10 interviews revealed three prominent themes, as presented in Table 4. These themes were subsequently categorized in accordance with the research questions and were distilled from the collective insights provided by all

interviewees. These critical issues were further classified into two distinct groups, encompassing challenges at the EU and national levels, each of which gave rise to several subthemes.

3.4 PICO Model

The PICO model is a logical structure used in evidence-based healthcare and academic research to formulate specific clinical or questions regarding research. PICO is an acronym that means "Population/Problem/Patient," "Intervention", "Context/Comparison", and "Outcome" (Duke University 2023).
According to this framework:

- Population/Problem (P) describes the characteristics of the subject that will be studied.
- Intervention (I) specifies the exposure or intervention under investigation.
- Context (C) aids researchers in determining the efficacy of the intervention by giving within a context.
- Outcome (O) describes the measurable outcomes that the study aims to attain (Duke University 2023).

The PICO framework helps researchers to frame specific questions that drive the planning and implementation of studies, subsequently leading to evidence-based decision-making and expertise growth in a variety of fields of study (Duke University 2023).

Table 3 below describes PICO model composed for the Master study to explore challenges of the harmonization of clinical evaluation of MDSW across EU member states and develop recommendations for the clinical evaluation of EU SHAPES pilot project.

Table 3: PICO model applied to conduct the study within the EU SHAPES pilot project

Problem (P)	The harmonization of clinical evaluation practices among EU member states still poses a challenge for MDSW. The problem revolves around the need to align these practices in accordance with the stringent EU regulations governing safety and performance.
Intervention **(I)**	This study explores the challenges associated with harmonizing the clinical evaluation of MDSW across EU member states. The primary objective is to gain a deep understanding of the hurdles and intricacies involved in this process.
Context **(CO)**	This exploration is conducted within the context of the EU SHAPES project, that is developing MDSW. SHAPES serves as a critical backdrop for this study, offering real-world insights and relevance to the subject matter.
Outcome **(O)**	By providing a set of practical and well-founded recommendations, this thesis aims to contribute valuable insights that can be directly applied within the SHAPES project.

3.5 Research Integrity and Research Ethics

Since the study focused on the expert interviews instead of the collecting of private or potentially sensitive information, the author was not required to get institutionalized authorization in order to carry out the research. Since there was no funding for this study, there was not any bias or detrimental impact on the results and study was carried out without any conflicts of interest.

Potential participants were sent a brief invitation letter via email. This assisted in avoiding participant selection error. In the beginning of each interview participants were given the opportunity to provide informed consent. Subsequently, the recording commenced, with consent reaffirmed for the record. Participants were informed of their right to stop the recording and withdraw from the interview or research at any time, along with the handling of the information they shared. Interviews' transcriptions were done automatically by the Microsoft Teams during the recording process. The transcripts are accessed exclusively by the author of the study to secure the anonymity of the participants.

3.6 Research Trustworthiness

The author included every pertinent finding related to the study topics in order for the discussion to be considered trustworthy. How the researcher uses the facts to back up the

primary points of view and develop a compelling explanation determines how credible the approach to research is (Starks & Trinidad 2007). According to Braun and Clarke (2006), researchers should try to explain each theme's meaning as well as the underlying presumptions and consequences.

Straight quotes from interview participants, according to King (2004), should be included in the final manuscript. To illustrate the themes' widespread application and to help in clarifying specific ideas of interpretation, brief quotes may be used. To give readers a sense of the source material, longer quotation excerpts might be provided. Consequently, the results section includes both longer block quotes and shorter quotes from the transcripts, and each quote was accompanied by a unique identification to show that different contributors were represented throughout the results. Additionally, Table 4 and Appendix 5 contain themes, subthemes, and representative quotes. The thesis's last discussion section covered each of the matters.

4 Results

The current section delves into the study's results, employing a thematic analysis and grounded in the extensive examination of interview transcripts. The study has discerned three themes indicative of challenges inherent at both the EU and national levels, as succinctly presented in Table 4.

Of particular note is the persistent impact of certain challenges on the harmonization of clinical evaluation practices for MDSW across member states of the EU. As this section unfolds, a critical examination of these challenges, both at the EU and national levels, will elucidate the complexities that underscore the landscape of clinical evaluation practices of MDSW.

In the realm of medical devices, the prevailing opinion among some stakeholders was that the regulatory landscape is overly complex. Within the EU, a plethora of legislative documents governs medical device software, with a constant influx of new regulations. Notably, AI software designed for medical applications, while distinct from conventional software, is regulated under the same legal framework.

Table 4: Themes and subthemes derived from thematic analysis

Research question	Theme	Sub-theme
1. Challenges of the harmonization of clinical evaluation of MDSW across EU member states	1. EU-level challenges 2. National level challenges	→ EU legislation relevant to MDSW → Clinical Evaluation of MDSW → Differences in approaches and strategies between notified bodies → National approaches of competent authorities → National approaches to clinical investigations → Legislation on RWD varies across EU countries
2. Recommendations for EU SHAPES project for clinical evaluation of MDSW	3. EU SHAPES project	→ Early planning of regulatory compliance → Engagement with regulatory authorities → Providing participants with a comprehensive understanding of regulatory requirements for clinical evaluation of MDSW → Qualification of Software → Risk classification of MDSW

4.1 EU-Level Challenges

4.1.1 EU Legislation Relevant to MDSW

The majority of experts, in their discussions, frequently referenced MDR/IVDR, as well as the clinical evaluation/ performance evaluation guidance provided by the MDCG (2020). This emphasis was due to the primary focus of the research, which was to explore the challenges related to harmonizing the clinical evaluation of MDSW across EU member states.

Many participants acknowledged the substantial efforts made at both international and EU levels to harmonize clinical evaluation practices for software. Notably, all IMDRF guidance documents are now directly applicable within the EU for software, in addition to the existence of an EU-specific subset of guidance documents.

Several interviewees highlighted both the strengths and limitations of the recent regulatory changes. They often pointed out that medical device software used for medical purposes is

now classified more rigorously and necessitates clinical data for MDSW falling into Class IIa and higher categories. This has the effect of aligning legislation across EU member states.

However, there were concerns expressed by some academics regarding the implementation of the new Regulations for software. As one expert put it,

> A1: What we have in Europe in a harmonized sense is that we have gone from a directive to a regulation, and in a legal sense. Regulation should apply equally, but at the same time, the regulation is not designed for software; it's designed for all medical devices.

In the view of experts, while the EU provides the Regulation, it is still up to individual countries to adopt and define it in their own terms through internal laws. This approach is seen as providing a common basis for the entire EU, enabling a consistent approach in terms of products and solutions.

An expert from software manufacturer highlighted the need for clarity in the interpretation of the MDR across different countries. They emphasized the importance of clarifications to ensure a competitive yet equitable market for digital health in the future.

> M1: Interpretations of the MDR are different from country to country, so clarification is actually what we need. I think that what we have in place is quite good, but sometimes it's difficult to understand what the key objective of the legislator was when they wrote specific elements in the law, and of course, that I will read it in one way and probably another provider of another solution is going to read it in another way. So, clarifications are what we need to do to make sure that we do have a competitive but fair market for digital health in the years to come.

During their discussions, experts frequently touched upon crucial aspects of the new Regulations and guidance, such as the general structure of the Regulations, Definitions, Intended purpose, Risk Classifications, Clinical evaluation, Post-Market Surveillance (PMS), and the use of RWD.

Definition and Intended Purpose

One of the challenges highlighted by these experts pertains to the definition of software as a medical device, which is not always clear. Distinguishing between digital health therapeutics and stand-alone apps, especially when it comes to software as a stand-alone product, can be challenging.

As one expert explained,

> A1: If you look to Article 2.1 of the medical device regulation, it describes a medical device and that definition for the first time introduced things

> like prediction into the definition as a medical purpose. But it's still not easily applicable to technology sometimes. I guess that's the first problem, a problem of definition or regulatory capture.

The intended purpose of MDSW emerged as a recurrent theme during interviews, with experts emphasizing that any software used for medical purposes falls under the category of software as a medical device. Even seemingly simple software, such as a dose calculator, can meet the definition of software as a medical device. Similarly, software that processes patient data to provide predictions or diagnoses also qualifies as a medical device. This has led manufacturers to explore ways to alter the intended purpose to potentially avoid regulatory classification.

One expert noted that when MDSW developers realize that their software falls under medical device regulation, they often seek ways to change their intended purpose to avoid being classified as such:

> A1: When you fall in as a medical device, clearly or not, software developers often get confused or they're trying to figure out a way to avoid device regulation by changing their intended purpose.

Another interviewee pointed out that while there is a way for manufacturers to determine if their software qualifies for medical use, the definition is rather broad:

> C1: But I agree this definition is not specific. There are general definitions, general explanations, general concepts.

Classification

Risk classification was a heavily discussed topic among participants, as it is a critical factor in determining how to meet regulatory requirements for CE marking. Given that software was scarcely mentioned in previous directives, the approach to risk classification has significantly changed between the Directive and the Regulation. In the past, since there were no specific classification rules for software, most software was placed in Class I, making it self-certifiable as Class I devices do not require involvement from a notified body. However, the new Regulation has introduced a substantial change, making the classification of software risk-based. Software developers are now required to consider the potential risks associated with using the device, leading to most products being reclassified as Class IIa or IIb.

> A1: Some could still possibly stay as Class one, but the classification guidance isn't very clear as to what exactly a certain product will be Class Ia or IIb, that's why we'll need more examples and more precedent.

While there is a classification manual, it contains only a few examples because the regulation includes new rules that were not derived from the previous directive or its accompanying guidance:

> A1: The regulatory capture, the fact that the rules are not specific for clinical evaluation for software, and then the classification rules have changed a lot.

Rule 11 of the Medical Device Regulation's classification scheme specifies that software used for diagnostic or therapeutic purposes, including optimizing therapy, should be classified as at least Class IIa and possibly higher. Consequently, risk management strategies play a pivotal role. Classification hinges on a careful examination of the device's intended purpose, considering how it will be used in light of that purpose.

However, the expert consensus on whether being classified as Class IIa or IIb matters greatly in practice was that it doesn't:

> A2: My opinion is that it doesn't make a very big difference in terms of the challenge of actually getting regulatory approval. So, in either case, de facto, you're going to require a clinical investigation and you're going to require the involvement of a notified body, the quality management system, and all of what that entails, to require a clinical evaluation in detail, to require post-market clinical follow-up, irrespective of whether you're Class IIa or b.

Experts delved into historical aspects related to the former and current features of medical device legislation. The directives, when initially released, were part of the "new approach legislation," which aimed to demonstrate that devices met specific performance and safety standards. According to those rules, all information was treated as commercially confidential:

> A1: Article 20 of the Medical Device Directive was basically applied to prohibit any clinical evidence from ever becoming public.

Experts emphasized the fundamental differences in the development of software products compared to physical devices like stents or implants:

> A1: And I guess the second challenge then, thinking of clinical evidence generation, is that, and the rules that we have in MDR, I guess they are designed with physical products in mind.

4.1.2 Clinical Evaluation of MDSW

Many experts agreed that the clinical evaluation requirements have been significantly harmonized by the EU Commission with the introduction of the new Regulations:

> P1: There might not be any problem with this harmonization of clinical evaluation because that is something that we have harmonized at least by

> regulation for more than 10 years. From this point of view, there's no need for additional, more harmonization.

The new Regulation introduces a systematic approach to clinical evaluation, emphasizing a lifecycle approach that requires considering risks related to the software and where clinical data is needed to support the intended purpose from the initial stages of idea development to prototyping:

> P1: We have not really changed too much on the clinical evaluation. The only thing that we have done is a little bit that it is now a more systematic approach. It is now a life cycle approach.

However, certain challenges persist, particularly from a policymaking perspective, related to the vast number of different products:

> P2: In terms of harmonization, I think we've taken huge steps. What remains a challenge is obviously the ever-growing nature of our sector.

Experts in the field express significant concerns regarding the regulatory landscape for MDSW. They anticipate a proliferation of tens of thousands of medical device apps, many of which may pose clinical risks. However, they find a lack of clear clinical and scientific methodologies in the guidance, primarily because regulations were designed *"with physical devices in mind" (A1)*.

One expert highlights the need for well-defined scientific and clinical methodologies for handling the vast amount of data generated by these apps:

> A1: But if we don't figure out exactly what the scientific and clinical methodologies, we're going to apply to all this data, it could be that, you know, a somewhat avoided opportunity. So hopefully that might generate some more thinking about how to get better clinical and scientific methodology.

Experts from the consulting field emphasize that guidance typically offers directions rather than strict methodologies, which can be both a challenge and an opportunity. Standards play a crucial role in guiding the development process:

> C1: Usually, the standard and the guidelines never say how you need to do things; they usually explain what the requirements are.

One of the critical aspects of assessing a medical device's benefit is clinical evaluation. To determine whether a diagnostic finding from a device is correct, a comparison must be made between those findings and test results from the software. This requires access to clinical evidence or patient data to establish comparisons.

However, obtaining access to clinical evidence is challenging in the EU countries. The stringent Regulations make it difficult to access real-world evidence:

> C1: I think the real-world evidence now here in Spain is very difficult to have access to. In other countries, it could be easier, but in Europe in general, it is not easy.

There is also concern about the lack of a clear approach in the EU for method comparison when it comes to MDSW claiming equivalence. Software is fundamentally different from physical medical devices, making equivalence claims more challenging:

> C1: I guess, for software in particular, because it's so different from many physical medical devices, it's very hard. You know, we have rules like you can claim equivalents for medical devices, but what does it mean to be technically equivalent if you're a stand-alone piece of software?

Furthermore, even though a clinical investigation is not directly required for Class IIa products, experts find it difficult to conceive of how a device could be approved without one. The formal requirement for a clinical evaluation based on clinical data may necessitate a clinical investigation as the only way to obtain acceptable clinical data. This is particularly true for innovative products:

> P1: But if you have difficulties because it's so innovative, then you have no chance to avoid a clinical investigation, which is costly in terms of money, resources, and time.

As a result, many experts believe that, for most MDSW products, clinical investigations will be necessary.

Some experts argue that medical device software is functionally closer to in-vitro diagnostics, making performance evaluation more suitable in many cases:

> P1: For many cases, there's already enough if you can demonstrate that there's a link between the calculation and the results of the software and the condition, it's a little bit similar to what we have developed for the clinical evidence for in-vitro diagnostics.

For software, clinical evaluation differs from that of traditional medical devices. Regulators often require software to produce consistent output for the same input, signifying the absence of technical issues. The next step is to establish clinical validity, demonstrating a connection between the software's parameters and results.

However, consultants working with businesses note that this approach poses challenges because manufacturers often struggle to grasp the significance of scientific, analytical, or clinical performance for MDSW:

> C1: The manufacturer has a problem understanding the performance evaluation for the CE mark because you need to identify scientific validity, analytical evaluation, and clinical evaluation, and it's very confusing for the manufacturer to understand how to obtain analytical evidence for the software.

4.2 National Level Challenges

At the national level, despite efforts to harmonize legislation within the EU, according to the experts' opinion several issues continue to pose challenges to the clinical evaluation of MDSW.

4.2.1 Notified Bodies

Many experts highlight the role of notified bodies as a primary challenge to harmonization. Notified bodies are independent organizations responsible for assessing the technical documentation of medical devices. The new requirements have added significant responsibilities to notified bodies and ensuring that only competent notified bodies are designated is crucial.

> P2: The Member states themselves don't assess the software. We have a system of designation of notified bodies.

However, there is a lack of full harmonization among notified bodies. Differences in approaches and strategies between notified bodies in different member states can create competitive advantages or disadvantages for businesses.

> A2: Challenges and harmonization within Member states in the EU, based on clinical evaluation, may not be perfectly harmonized because the notified bodies are based in individual Member states.

Furthermore, the lack of coordination among notified bodies is a problem. There are no publicly announced rules or coordinated efforts among these bodies.

> A1: The notified bodies don't coordinate together and don't set rules that are announced publicly.

Transparency in the clinical evidence accepted by notified bodies is considered a challenge for the harmonization of clinical evaluation.

> A1: The evidence the notified bodies accept is not published anywhere.

4.2.2 National Approaches to Clinical Investigations

Companies have the option to choose their own notified body in any EU Member State. National competent authorities are responsible for making decisions regarding the risk class of

MDSW. Differences in interpretation between manufacturers and competent authorities can lead to delays in the approval process.

> A2: None of the notified bodies are very much involved in it, but in the end, it's a decision from the competent authority. So you can put it to the national competent authority to actually have a decision on this, but that's time-consuming and that's not happening fast enough.

Sometimes, manufacturers may find their devices falling between different risk classes, and this can result in significant delays when seeking clarification from competent authorities.

> A2: So you ask your competent authority, but it can be very time-consuming to get an answer from your competent authority, and that needs to be improved.

Clinical evaluation is often understood by manufacturers as a type of clinical investigation for software. However, the process of conducting clinical investigations for software products can be challenging and resource intensive. An interviewee highlights the difficulties in aligning requirements and documentation across different countries where clinical investigations are conducted.

> C1: When you go to obtain the approval for a clinical trial, clinical study, you need to obtain approval from your competent authority in your country. And in this case, there are more differences between the information you need to submit depending on the country and the criteria to evaluate this information.

Experts note that despite efforts to standardize procedures for submitting applications for clinical studies, national laws continue to govern the approval process, leading to differences in interpretations.

> A2: I do think there are some challenges in harmonization and the differences in national approaches to clinical investigations are immensely challenging for manufacturers.

4.2.3 Classification Differences

Differences in the classification of medical devices among Member states are a significant challenge. The classification of MDSW is often considered a gray area, and experts find that the classification manual does not provide enough guidance.

> A2: So, they provide a framework, but you're very often still in this “gray area” in terms of class one medical device software, Class IIa or Class IIb, or Class III.

Interpretations of classifications can vary between different regions within countries, adding uncertainty for manufacturers.

> A2: The interpretation of different competent authorities can be different. For example, the interpretation of the German federal states is actually different in terms of what the different classifications of software are.

Overall, harmonizing clinical evaluation procedures and classifications across EU member states remains a complex challenge due to differences in interpretations, regulations, and approaches at the national level.

4.2.4 Use of Real-World Data

The utilization of RWD in the context of MDSW presents both regulatory and practical challenges. Legislation concerning RWD varies across EU countries, leading to diverse interpretations and implementations.

Experts from the SHAPES project were interviewed, as RWD/RWE collection is a crucial component of MDSW within the project.

From a regulatory standpoint, experts point out that there are generally no significant differences concerning RWD within the medical device regulatory framework across EU countries. However, practical issues can arise when healthcare institutions are unwilling or unable to share patient data:

> P1: Sometimes it could happen that some hospitals would say we cannot share those data with you because we do not want and we are not able to delete the names or the birth dates.

Manufacturers face challenges related to data collection, particularly in post-market clinical follow-up plans, where RWD is often essential. Surveys are seen as a potential solution to obtain RWD, but their effectiveness depends on the quality and specificity of the questions:

> C1: I think the most important way to obtain real-world data from the market from the patient and the clinician. And the problem is when the manufacturer develops the survey, sometimes it doesn't include enough questions or specific questions to identify.

Access to clinical information stored in hospitals' records can be challenging due to strict data protection regulations. Additionally, insufficient databases exist for all product types, making it difficult to access the necessary data:

> C1: Other point is to have access to the real-world data in the database of the hospitals? But you need to pay to have access to this database.

These data challenges can pose difficulties in obtaining approval from notified bodies. Often, there isn't enough time to complete all necessary procedures, leaving manufacturers with insufficient data to incorporate into clinical evaluation reports:

> C1: But when the notified body goes to check the post-market clinical follow-up plan, then the notified body finds it's not enough also, and you don't have enough information in your clinical evaluation report. If you don't have enough information or you don't have enough plans to obtain information in your post-market clinical follow-up plan, you don't have enough information about your clinical performance. And this is difficult.

Another aspect emphasized by experts is the potential for different interpretations by authorities in various countries regarding post-market surveillance (PMS) or clinical investigation activities. This underscores the importance of careful planning from the initial stages:

> A2: It could be interpreted in one country that it's a real-world performance data collection exercise, and another country could take the interpretation as a prospective clinical study, and you have to follow the clinical study regulation. So, it definitely could make a difference in the planning.

In Northern Ireland and England, the secondary use of RWD is identified as a challenge due to a lack of secondary use legislation. The SHAPES project's data lake is seen as a valuable resource for both primary and secondary use:

> H1: The challenge for Northern Ireland is we don't have any secondary use legislation, so if you're collecting data for one primary purpose.

These insights shed light on the complex landscape of utilizing RWD in the development and regulation of medical device software, highlighting both regulatory harmonization efforts and practical challenges faced by manufacturers and healthcare institutions.

4.3 Summary

In summary, a nuanced exploration of regulatory considerations for MDSW in the EU revealed distinct focal points at both the EU and national levels. At the EU level, experts concentrated primarily on legislative dimensions and challenges related to the clinical evaluation of MDSW. Conversely, at the national level, the expert assessments highlighted a spectrum of concerns, encompassing variations in requirements among notified bodies for clinical evaluation, diverse national approaches adopted by competent authorities regarding clinical investigations and risk classification, and the utilization of RWD.

The study's cohort comprised a diverse array of interviewees, representing key stakeholders in the medical device software domain, including policymakers, regulators, manufacturers, healthcare specialists, and researchers. This diversity in perspectives significantly influenced the varied viewpoints and approaches adopted by these stakeholders in addressing the identified challenges. Such comprehensive insights into the perspectives of diverse

stakeholders offer a robust foundation for understanding the intricacies of regulatory considerations for medical device software, thereby contributing valuable knowledge to the broader discourse on healthcare technology regulation within the EU.

5 Discussion

This study aimed to delve into the challenges associated with the harmonization of clinical evaluation of MDSW across member states of the EU. Two primary research questions framed the study: the first question aimed to identify the challenges to clinical evaluation harmonization among EU member states, while the second focused on proposing evidence-based recommendations for the EU SHAPES pilot project for clinical evaluation of MDSW. This discussion synthesized the findings in response to these questions, providing insights into the current state of the harmonization efforts and offering potential solutions to the identified challenges.

5.1 EU-Level Challenges

Given the research's core objective of investigating harmonization challenges in the clinical evaluation of MDSW across EU member states, the interviewees primarily referred to the MDR/IVDR (2017), along with the associated MDCG's guidance (MDCG 2019; MDCG 2020), in their responses.

5.1.1 Legislation and Clinical Evaluation

Participants in the study collectively acknowledged the significant progress achieved at both international and EU levels concerning the harmonization of clinical evaluation for software. The adoption and adaptation of IMDRF (2017) guidance within the EU have laid a strong foundation for harmonization. However, this harmony is met with substantial complexity due to the sheer diversity of medical device types, including innovative digital health solutions, numbering approximately 500,000 distinct categories (Melvin & Torre 2019). This multitude of device types presents a formidable challenge to regulators, auditors, and consultants alike, as they grapple with applying generally applicable horizontal guidance to such a vast array of products.

Moreover, the evolving regulatory landscape in the EU, particularly with respect to the digital market, adds another layer of complexity. While existing Regulations encompass both physical medical device and stand-alone software designed for medical purposes, the suitability of these Regulations for software-only scenarios has been questioned (Gilbert et al. 2021;

Granlund et al. 2022). The current regulatory framework, originally crafted for traditional medical devices, must now adapt to the unique characteristics of digital health solutions, demanding thorough scrutiny and modification.

One fundamental challenge that emerges from this complexity is the correct qualification of MDSW. While the MDR (2017) introduced amendments to the definition of a medical device, including the incorporation of predictive elements for the first time, the application of this definition to technology remains convoluted (Ordish et al. 2019). Interviewees pointed out that the definition of a medical device remains somewhat general, leaving room for interpretation and misclassification. Manufacturers, in response, sometimes attempt to alter the intended purpose of their products to navigate these complexities.

This regulatory complexity has implications for industry stakeholders who grapple with navigating the intricacies of compliance. Interviewees expressed frustration, citing confusion over regulatory requirements and a lack of clarity regarding regulatory expectations. This uncertainty creates challenges in determining the appropriate pathway for securing the CE mark.

Technological advancements, such as mobile apps and wearables, fall under MDR if they are intended for medical purposes, make health benefit claims, or pose potential risks to patients. Consequently, many health technology companies opt for cautious marketing strategies, avoiding explicit medical claims and positioning their devices as 'health and wellness' products rather than tools for disease management (Singhal & Cowie 2021).

The next significant challenge in the harmonization of MDSW clinical evaluation across the EU revolves around the risk classification landscape. This transformation is particularly evident when transitioning from the MDD to the MDR. Under the new regulatory framework, most devices will be classified as at least Class IIa, necessitating the involvement of notified bodies in the compliance process. According to the recently introduced classification rule 11 of the MDR Annex VIII, software intended to provide information used in diagnostic or therapeutic decision-making processes must be classified as at least Class IIa. If the software's usage is likely to significantly deteriorate a person's health, it must be classified as Class IIb. Consequently, stakeholders widely anticipate that most standalone medical software will be categorized as Class IIa or higher, aligning with the EU's definition of a medical device (Regulation (EU) 2017/745, Article 2).

This shift marks a substantial change for the sector, introducing a risk-based approach to software classification. However, this new classification paradigm introduces a "gray" area, and differences in classification approaches may arise across different EU member states

(Ravizza et al. 2021). This lack of harmonization can further complicate the regulatory landscape for MDSW.

As previously discussed in the Results section, manufacturers may encounter devices that straddle the border between Class I and Class IIa. If a device falls into Class I, it can be submitted for registration directly with a competent authority, bypassing the need for a notified body. However, for Class II devices erroneously registered as Class I, manufacturers can only engage with competent authorities, and obtaining responses from competent authorities can be time-consuming.

According to Keutzer and Simonsson (2020) analyzing classification Rule 11 in isolation indicates that software used in critical applications like drug dosage calculations, diagnostics, or therapy planning could fall into Class III due to the potential life-threatening consequences of errors. If the risk of death is highly unlikely, it could fall into Class IIb, defined as devices where a mistake can lead to a significant deterioration of an individual's health (Keutzer & Simonsson 2020).

It's important to note that AI/ML devices are subject to the same regulatory legislation as other MDSW. However, the MDR (2017) doesn't directly address AI/ML-based devices, creating uncertainty among industry stakeholders. The critical issue revolves around the application of classification Rule 22 to AI/ML-based systems, potentially elevating the risk classification of these technologies to Class III. This classification should be clear to manufacturers from the outset of product development to facilitate a smoother market introduction for AI/ML-based systems (Granlund et al. 2022).

The European Commission's proposed Artificial Intelligence Act (AIA) (2021) represents the initial legislative effort to standardize laws for AI systems, including those used in medical devices. However, a misalignment between AIA classifications of AI systems and the MDR has been identified. MDSW producers must navigate these classifications, considering both AIA and MDR requirements and commitments. Terminology-related issues also pose challenges, with terms like "non-high-risk" potentially leading to the erroneous perception that certain AI systems are not high-risk (Zapata, Patil, Ward, Loughran & McCaffery 2023).

Considering these changes, the up classification of risk classes for MDSW entails significantly more work for software manufacturers and extended development periods (Scholtes, Behrend, Buedenbender, Volker & Keywan 2018). With the new risk classification requirements, most manufacturers will need to undergo assessment by notified bodies, evaluating all technical documentation, and performing clinical evaluations for their MDSW products. Equivalence claims to other CE-marked devices are increasingly viewed as

impractical under the new legislative requirements. MEDDEV 2.7/1 rev. 4 has made claiming equivalence to other commercially available medical devices more challenging, necessitating access to the technical documentation of the original market product to demonstrate equivalence.

In cases where clinical equivalence is asserted, a signed agreement with the original product's manufacturer is required, rendering this option less feasible from a commercial standpoint (Bayrak & Safak Yilmaz 2022). Instead, clinical equivalence has become nearly unattainable for MDSW products. Consequently, MEDDEV 2.7/1 rev. 4 has underscored the importance of clinical investigations in demonstrating benefit-risk analysis and general safety and performance requirements compliance for medical software, aligning with the rules of MDR 2017/745.

Clinical evaluation for software diverges somewhat from the approach used for traditional medical devices. An interviewee emphasized that software must first demonstrate the ability to consistently produce the same output for the same input, indicating the absence of technical issues. Subsequently, clinical validity must be established, demonstrating a clear correlation between the software's parameters and its outcomes. However, the interviewed expert advising manufacturers on regulatory matters acknowledges that comprehending the significance of scientific, analytical, or clinical performance for MDSW can pose additional challenges for MDSW manufactures.

Several interviewees pointed out that MDCG guidance on clinical evaluation lacks explanations regarding the scientific or clinical methodologies that industry stakeholders might expect. This absence of detailed guidance complicates the harmonization of clinical evaluation practices, leaving manufacturers uncertain about the best approaches. Consequently, MEDDEV 2.7/1 rev. 4 continues to be widely utilized (Pritchard 2022). Additionally, it appears that methodological weaknesses are very widespread throughout research on medical AI devices (Niemiec 2022).

While regulatory requirements, along with guidelines and harmonized standards, form the foundation of medical device development, it falls on manufacturers to optimize each stage of the development process. The use of harmonized standards is a powerful tool for demonstrating compliance with regulatory requirements, although it remains optional for manufacturers (Granlund et al. 2020). For instance, in the case of MDSW, manufacturers must implement an ISO 13485-compliant quality management system based on necessary conformity assessment procedures. However, some ISO 13485 requirements are not applicable to software-only devices. In the absence of official guidance, procedures followed by notified

bodies may differ, potentially impacting manufacturers' equitable opportunities (Granlund et al. 2020).

Given the variation in clinical data weight, particularly for medical devices, using a method tailored to the device, such as the Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence, is essential. The highest level of evidence according to OCEBM is a systematic review of randomized trials (OCEBM Levels of Evidence Working Group 2011).

The National Institute for Health and Care Excellence (NICE) has also provided guidance on clinical evidence requirements for digital health technologies, varying based on the type and purpose of the intervention (Unsworth et al. 2021). Therapeutic interventions have stricter requirements (Ravizza et al. 2021). However, clinical evidence from systematic reviews, primarily drawn from randomized controlled trials and non-randomized controlled cohort studies, may only be available for disorders with a substantial body of published literature (Pritchard 2017).

In essence, the challenges at the EU level stem from the evolving nature of medical software and its interaction with existing regulations. The multiplicity of device types, intricate classification rules, and a lack of clarity in certain regulatory definitions all contribute to the complexity surrounding clinical evaluation harmonization.

5.1.2 National Level Challenges

Beyond the EU-level challenges, interviewees highlighted several challenges pertaining to individual member states. Notified bodies, competent authorities, clinical investigations, and RWD utilization were discussed within the context of national-level challenges.

Notified bodies

The transformation brought about by the MDR significantly impacts the role of notified bodies in the certification process for MDSW. However, both notified bodies and expert panels are expected to face a shortage of individuals possessing the necessary skills to conduct assessments and audits, which could lead to substantial delays and increased costs for manufacturers (Munro 2017).

Notified bodies must establish and implement internal rules and processes for assessing clinical evaluation reports and related data to ensure they have adequate personnel, particularly those with clinical expertise, to handle clinical evaluation examinations (European Commission, MEDDEV 2.7/1 revision 4, 2016).

However, there's a lack of uniformity among notified bodies. Their approach to applying standards to MDSW, especially when artificial intelligence is involved, remains uncertain. Additionally, the role played by health authorities in regulatory procedures varies from country to country. Such discrepancies in the notified body system and their varying interpretations pose challenges to the harmonization of clinical evaluation processes (Ravizza et al. 2021).

According to a few interviewees the system of notified bodies is not fully harmonized, and any lack of convergence between them becomes relevant in a competitive industry where the CE mark applies throughout the EU. Some member states may adopt a more flexible strategy, while others may have a different approach. Many businesses still prefer to work with notified bodies in their own country, which can be either a relative competitive advantage or a disadvantage depending on the circumstances.

Under the MDR (2017), clinical evaluation requires "sufficient" clinical evidence. Notably RWD needs to be collected for more than ten years for devices placed on the market based on equivalence to meet initial MDR requirements. However, ongoing clinical evidence must be based on the manufacturer's own clinical data (Giefing-Kröll & Laumen 2022).

Existing RWD platforms highlight significant obstacles to using RWD for backing both safety and effectiveness evaluations. These issues include errors in data collecting, a lack of the process of randomization poor data collection efficiency, and a generalization of outcomes (Valla 2023). To meet these "sufficient" requirements, a mixture of data sources is recommended by the British Standards Institution (BSI) United Kingdom (UK). Each source has its advantages and shortcomings. Proactive survey gathering offers knowledge from real-world use, while registry data ensures a variety of data from different locations, healthcare providers, and patients, both compliant and non-compliant (Giefing-Kröll & Laumen 2022).

Competent Authorities and Clinical Investigations

Clinical studies for medical devices are often referred to as "Clinical Investigations" rather than "Clinical Studies" or "Clinical Trials" (Döerr, Khalili & D’souza 2023). However, there are several challenges associated with clinical investigations for MDSW at the national level.

It is important to note that, unless explicitly exempted by specific regulations, the MDR mandates manufacturers to conduct clinical investigations for all medical devices, including software. Section 2.4, Chapter I of Annex XV, Clinical Investigation, requires clinical investigations to be in line with the clinical evaluation plan. Section 1.5, Chapter II of Annex XV, requires that the clinical investigation application to include, among other information, "details and/or reference to clinical evaluation plan." Consequently, clinical investigations need to adhere to the data provided in the clinical development plan, which is an element of the clinical evaluation plan.

In order to confirm that AI-MDSW systems are safe and effective, an exhaustive, prospective clinical evaluation will be vital. This evaluation will use clinically applicable performance indicators, which go in addition to assessing technical correctness to consider how AI-MDSW influences clinical results, overall level of healthcare etc. (Kelly, Karthikesalingam, Suleyman, Corrado & King 2019).

One of the primary challenges is that manufacturers are now expected to conduct more pre-market clinical studies (Fraser et al. 2020; Melvin 2022), despite clinical investigations may not be well-suited for MDSW. Another significant challenge is the absence of a single EU standard for clinical investigations of MDSW. The MDR outlines strict regulations for clinical investigations, but it does not provide specific details on the characteristics of such investigations. The level of clinical evidence generated by developers can vary widely, contributing to the lack of harmonization (Ravizza et al. 2021).

Moreover, some interviewees shared that there are notable variations in the requirements of competent authorities for the approval of clinical investigations. These differences further disrupt harmonization efforts and create different conditions for manufacturers seeking to enter the EU market. For example, in Finland the Finnish Medicines Agency (Fimea) makes the decision whether to approve clinical trials on Class IIa and IIb devices. (Laki lääkinnällisistä laitteista 719/2021 - Säädökset alkuperäisinä - FINLEX ® s.a.)

With the MDR making equivalence claims more challenging, the number of clinical investigations and PMCFs is expected to increase significantly. Harmonization is crucial in this context, especially for multicenter investigations that involve data collection from various hospitals and countries to provide more accurate data for medical software's intended use (Kamusheva et al. 2022; Ravizza et al. 2021).

Clinical trials for medical devices are distinct from drug trials. The MDR mandates that manufacturers of high-risk devices prepare comprehensive overviews of their evidence. Full

transparency is essential for making accurate assessments when using innovative medical. Technologies (Deep, Rana & Sharma 2019).

The European Database for Medical Devices (EUDAMED) aims to enhance transparency by providing a platform for improved data coordination and transparency for medical devices marketed in the EU (European Union 2019). However, some interviewees believe that there will continue to be difficulties with harmonization and national differences in the approval of clinical investigations. While standardized procedures for submitting study applications may exist, national laws still govern whether those studies can be approved and “international trials often fall under the remit of multiple regulations” (Negrouk, Lacombe & Meunier 2018) This underscores the challenge of how each country interprets criteria for conducting clinical investigations (Lalova et al. 2020). Therefore, there is a need for more detailed regulations in the EU regarding the clinical investigation of MDSW to ensure uniform standards of efficacy and safety for software (Ravizza et al. 2021).

Secondary use of health data

The concept of secondary use of medical data has gained popularity in recent years. The use of health data acquired for main medical needs in secondary applications such as research, epidemiological studies, and quality improvement efforts is referred to as secondary usage.

Secondary data are defined as data that is used for a purpose other than the one for which it was obtained. The coordinated and optimized use of secondary data within data networks has considerable potential for application in health-related studies. These data can be collected through various resources such (wearables, mobile phone apps, electronic health records etc.). For example, the authors outline relevant sources and methodologies for secondary data processing (Näher et al. 2023).

Finland has built a strong regulatory framework regulating the secondary use of health records. The Act on the Secondary Use of Health and Social Data (Laki sosiaali- ja terveystietojen toissijaisesta … 552/2019 - Säädökset alkuperäisinä - FINLEX ® s.a.) is an essential part of legislation governing data sharing. This legislation allows individuals to select not to data sharing while highlighting the necessity of informed consent. In many ways, Finland's strategy to the secondary use of health data is remarkable illustrating the possibility of ethical data sharing while protecting patient confidentiality and rights. The legal, ethical, and technological foundations that have been established have created an optimal setting for utilizing health data for the advancement of healthcare and scientific research. As other countries attempt to use health data for secondary purposes, Finland's system offers a useful case study.

In summary, national-level challenges related to variations in competent authorities and notified bodies requirements, the lack of standardized approaches in the EU for clinical investigations and RWD utilization, as well as transparency issues pose obstacles to the harmonization of clinical evaluation processes for MDSW across EU member states. Addressing these challenges will be essential for promoting a more cohesive regulatory environment for MDSW within the EU.

5.2 The SHAPES Pilot Project: Advancing Technical Solutions for Seniors

The EU SHAPES pilot project represented a groundbreaking endeavor aimed at addressing the challenges posed by aging populations. This initiative sought to establish a comprehensive socio-technical infrastructure encompassing health and care digital solutions, devices, and support services. The overarching goal was to empower aging individuals by fostering engaged, autonomous, and independent living within the familiar confines of their homes (Seidel et al. 2022).

5.2.1 Intended Use of the Software

At the forefront of this endeavor, the EU SHAPES pilot project strives to harness advanced technologies to enable enhanced health management and caregiving for seniors. One of the project's technical contributors, EDGENEERING (referred to as EDGE), is actively developing the eCare system. This system constitutes a pioneering remote patient monitoring platform designed to serve healthcare and social care providers. “eCare (EDGE) Remote monitoring platform which collects and displays wellbeing and health data gathered manually or automatically (using connected devices like blood pressure monitor and weight scale) in the home environment” (Spargo & Goodfellow 2023).

The eCare system seamlessly integrates with a diverse array of devices, including IoT sensors, medical and wellness devices. Various software components, ranging from MDSW to AI/ML (artificial intelligence/machine learning) algorithms, are possible to integrate into the system's front-end to address different use cases and personalized interventions. Crucially, validated and calibrated medical devices such as blood pressure monitors and pulse oximeters find their place within the homes of citizens (Spargo & Goodfellow 2023).

The eCare system operates as a conduit for timely patient data, affording caregivers the ability to make informed decisions promptly and prevent unnecessary hospital consults. By transmitting crucial health-related information to healthcare professionals, the system enables a comprehensive assessment of the patient's well-being and medical condition over time. It facilitates not only the collection of data from a variety of sensors embedded in the

devices but also leverages direct patient input, through questionnaires or simple feedback forms (Spargo & Goodfellow 2023).

Significantly, the system does not intrinsically provide decision support to healthcare professionals. Instead, it empowers them to promptly assess the patient's condition through real-time data, prompting immediate actions such as contacting the patient, adjusting medication, or scheduling consultations. This interaction is facilitated through the system's interface, ensuring that both caregivers and patients are kept well-informed (Spargo & Goodfellow 2023).

5.2.2 Regulatory Landscape: Software as a Solution, not a Product

The software, designed to deliver health and wellness information for prevention, treatment, and disease monitoring, falls under the qualification of a medical device according to the MDR (2017). However, it was emphasized by some interviewees that this software is still in its developmental stage, and it has been crafted specifically for the needs of the SHAPES project. This distinction makes it more of a dynamic solution in collaboration with the client rather than a finished product, and it reflects the adaptability required to meet specific client needs, infrastructure, and services. Importantly, this approach also allows for innovation without the immediate burden of certification, which can be both time-consuming and costly.

Notably, it was discussed by interviewees that the eCare system software, is perceived as an internal development designed for Portugal's National Hospital in-house use. The proof-of-concept stage of this project necessitates scientific validation, distinct from the validation processes typically employed for medical devices. The Portugal National Hospital actively participates in developing the software for broader market application, assessing it in line with its specific objectives and interests (Spargo & Goodfellow 2023).

The regulatory landscape of the MDR (2017) also recognizes the necessity for flexibility to accommodate unique circumstances. The MDR permits the availability of special devices in specific situations, provided certain conditions are met. This includes devices manufactured and used exclusively within the same health institution (Article 5.5), and devices authorized by competent authorities in the interest of patient safety or public health (Article 59). Such devices are not subject to formal clinical evaluation before they are in use, nor are they required to bear the CE marking. A healthcare organization that manufactures devices under Article 5.5 must publicly declare that the devices meet general safety and performance requirements. In such cases, the health organization must furnish documents detailing the

software's design and performance, backed by clinical experience, and be ready to take corrective actions if necessary (Beckers, Kwade & Zanca 2021).

One notable aspect of the software's journey through the regulatory landscape is the classification of its risk level. A MDSW manufacturer from the SHAPES project believes that their software should fall into risk Class I. However, this assumption is met with skepticism by the interviewed regulators, consultants, and academics, who suggest that it may, in fact, belong to at least Class IIa, given the significant changes in risk classifications brought about by the MDR.

Considering the complexities of risk assessment, it becomes evident that even seemingly straightforward software applications, such as those monitoring vital physiological parameters like heart rate and blood pressure and transmitting data for heart disease management, could be subject to higher classifications (MDCG 2019). In some instances, they might even be categorized as Class IIb rather than Class IIa devices; this shift in classification might be further amplified as predictive and prognostic functions are integrated, broadening the scope of medical software subjected to regulatory review.

In the interviewee’s opinion, manufacturers may find themselves dealing with devices straddling the border between Class I and Class IIa. If a device is classified as Class I, it can be submitted for immediate registration with a competent authority without the involvement of a notified body. However, when a Class II device is mistakenly registered as Class I, manufacturers are restricted to working solely with competent authority, potentially leading to delays and errors in approvals. Resolving these classification uncertainties and streamlining interactions with the appropriate regulatory authorities are ongoing challenges in the regulatory landscape.

The landscape of medical device software has evolved significantly in recent years, necessitating a comprehensive evaluation of its clinical benefit. It is no longer sufficient to merely demonstrate that the software functions; there must also be robust evidence of its statistically significant efficacy in delivering the intended clinical benefit (Ravizza et al. 2021). However, there exists a nuanced perspective within the regulatory framework. An interviewed policymaker argues that, under various conditions of intended use, demonstrating the accuracy of the eCare system software may suffice. This view highlights the importance of tailoring evaluations to specific use cases and contexts. Moreover, the software may require the establishment of long-term studies, extending beyond the initial CE marking, to continuously assess the benefits it offers to patients during the PMCF period.

An additional consideration is the classification of some software as IVD devices. According to the guidance provided by the MDCG (2020), if the software heavily relies on data obtained exclusively from in vitro diagnostic medical devices, it qualifies as an IVD MDSW, falling under the jurisdiction of Regulation (EU) 2017/746. This qualification hinges on whether the software's intended purpose is significantly driven by data sources originating from these diagnostic devices.

The adaptability and customization inherent in the EU SHAPES pilot project allow for the development of software solutions tailored to specific clients, infrastructures, and services. While the regulatory pathway for medical device software can be complex, it is critical to ensure compliance, patient safety, and the delivery of effective healthcare solutions. The eCare system software serves as an example of how developers are considering regulatory compliance, risk assessment, and the evolving nature of software functionalities (Spargo & Goodfellow 2023). The collaboration between developers, healthcare institutions, and regulatory authorities is pivotal in navigating this dynamic landscape and ensuring that innovative technologies like those within the EU SHAPES project can fulfill their potential in supporting healthy and engaged aging.

Additionally, the EU SHAPES pilot project was a notable endeavor aiming to create an ecosystem that facilitates compliance with new legislative requirements, especially for small and medium-sized companies. This collaborative approach recognizes the evolving landscape of MDSW, with many software products falling into the minimum risk Class IIa, thereby necessitating extensive clinical evaluation with clinical investigations, particularly for innovative solutions.

6 Conclusions

Despite the recent big achievements of the harmonization of the requirements for general safety, performance, and clinical benefit in the field of medical devices, clinical evaluation processes for medical device software (MDSW) in the EU is a complex task. It is not fully harmonized due to the different reasons, including the dynamic nature of medical software, the huge number and diverse range of MDSW types as well as its complicated interaction with existing regulatory frameworks. Intricate qualification and risk classification rules, ambiguities within regulatory definitions, lack of the technology specific regulation and methodology for clinical evaluation contribute to the intricacies surrounding the harmonization of clinical evaluation across the EU member states.

Nationally, challenges such as disparities in competent authorities and notified bodies requirements and transparency, lack of standardized approaches for clinical investigations, as well as national differences in RWD utilization concerns exacerbate the complexity of harmonizing MDSW across EU member states.

The EU SHAPES pilot project has emerged as a significant initiative, creating an ecosystem that among other things aimed to facilitate compliance with new legislative requirements, especially for small and medium-sized companies.

The harmonization of clinical evaluation of MDSW in the EU requires ongoing collaboration, transparency, and a deep understanding of the evolving regulatory landscape. By addressing identified challenges, embracing collaborative endeavors like the EU SHAPES pilot project, and continuously adapting to the evolving nature of MDSW, the EU can lay the foundation for a more streamlined and effective approach to ensuring the safety and performance of medical device software, benefiting both industry stakeholders and the broader healthcare ecosystem.

7 Recommendations for the EU SHAPES Pilot Project

Based on the information provided by the experts regarding the regulatory landscape and the medical device software within the future EU projects clinically evaluating medical software it is recommended:

1. To recognize the complexity of regulatory processes, particularly market approval, in international projects involving multiple organizations and early planning across diverse entities.
2. To proactively engage with regulatory authorities, such as national competent authorities and notified bodies, early in the development process, and seek guidance and clarifications on regulatory requirements to avoid compliance issues later.
3. To develop a clear and standardized approach to regulatory compliance within international projects by collaboration in defining and implementing the co-created standards.
4. To provide participants with a comprehensive understanding of regulatory requirements and appropriate regulatory approach, including requirements for clinical evaluation of MDSW.
5. To conduct developers thorough qualification and risk assessments for the software to meet requirements of MDR/IVDR for safety and performance, considering factors such as intended use.

6. To prepare developers for ongoing regulatory compliance efforts, including periodic assessments and adaptations to changing regulations by implementing compliance as a continuous process, even after the initial CE marking.

These recommendations aim to assist developers within the EU SHAPES project and similar initiatives in navigating the dynamic regulatory landscape for medical device software while ensuring compliance, patient safety, and the delivery of effective healthcare solutions. Collaboration, documentation, and proactive engagement with regulatory authorities are key strategies for success in this evolving field.

References

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D.S.W., Karthikesalingam, A., King, D., Ashrafian, H. & Darzi, A. 2021. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. npj Digital Medicine, 4, 1. https://doi.org/10.1038/s41746-021-00438-z

Armando, LG., Miglio, G., de Cosmo, P., Cena, C. Clinical decision support systems to improve drug prescription and therapy optimisation in clinical practice: a scoping review. BMJ Health Care Inform. 2023; 30 (1): e100683. [Accessed 23 September 2023]. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10163516/

Atella, V., Piano Mortari, A., Kopinska, J., Belotti, F., Lapi, F., Cricelli, C. & Fontana, L. 2018. Trends in age-related disease burden and healthcare utilization. Aging Cell, 18, 1. https://doi.org/10.1111/acel.12861

Bayrak, T. & Safak Yilmaz, E. 2022. What Will Be the Economic Impact of the New Medical Device Regulation? An Interrupted Time-Series Analysis of Foreign Trade Data. Value in Health Regional Issues, 29, 1–7. https://doi.org/10.1016/j.vhri.2021.07.010

Beckers, R., Kwade, Z. & Zanca, F. 2021. The EU medical device regulation: Implications for artificial intelligence-based medical device software in medical physics. Physica Medica, 83, 1–8. https://doi.org/10.1016/j.ejmp.2021.02.011

Berlin, Germany (hosted by European Commission on behalf of the EU) 2023. [Accessed 23 October 2023]. URL: https://www.imdrf.org/meetings/berlin-germany-hosted-european-commission-behalf-eu.

Bianchini, E. & Mayer, C.C. 2022. Medical Device Regulation: Should We Care About It? Artery Research, 28, 2, 55–60. https://doi.org/10.1007/s44200-022-00014-0

Bianchini, E., Francesconi, M., Testa, M., Tanase, M. & Gemignani, V. 2019. Unique device identification and traceability for medical software: A major challenge for manufacturers in an ever-evolving marketplace. Journal of Biomedical Informatics, 93, 103150. https://doi.org/10.1016/j.jbi.2019.103150

Bogner, A., Littig, B. & Menz, W. 2009. Interviewing Experts. Springer.

Brussels, Belgium (hosted by European Commission on behalf of the EU) 2023. [Accessed 23 October 2023]. URL: https://www.imdrf.org/meetings/brussels-belgium-hosted-european-commission-behalf-eu

Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 2, 77–101. https://doi.org/10.1191/1478088706qp063oa

Gerke, S., Babic, B., Evgeniou, T. & Cohen, I.G. 2020. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. npj Digital Medicine, 3, 1. https://doi.org/10.1038/s41746-020-0262-2

Clarke, V., Braun, V., & Hayfield, N. 2015. Thematic analysis. Qualitative psychology: A practical guide to research methods, 3, 222-248.

Council, E. 1993. Council directive 93/42/EEC concerning medical devices. Official Journal of The European Communities, Luxembourg. [Accessed 23 July 2023]. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31993L0042

Dang, A., Arora, D. & Rane, P. 2020. Role of digital therapeutics and the changing future of healthcare. Journal of Family Medicine and Primary Care, 9, 5, 2207. https://doi.org/10.4103/jfmpc.jfmpc_105_20.

Deniz-Garcia, A., Fabelo, H., Rodriguez-Almeida, A.J., Zamora-Zamorano, G., Castro-Fernandez, M., Alberiche Ruano, M. del P., Solvoll, T., Granja, C., Schopf, T.R., Callico, G.M., Soguero-Ruiz, C. & Wägner, A.M. 2023. Quality, Usability, and Effectiveness of mHealth Apps and the Role of Artificial Intelligence: Current Scenario and Challenges. Journal of Medical Internet Research, 25, e44030. https://doi.org/10.2196/44030

DIRECTIVE, H. A. T. 1990. Council Directive 90/385/EEC of 20 June 1990 on the approximation of the laws of the Member States relating to active implantable medical devices. Official Journal L, 189 (20/07), 0017-0036. [Accessed 13 April 2023]. URL:https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:31990L0385

Duke University. Medical Center Library & Archieves. 2023. Evidence-Based Practice: PICO. [Accessed 17 October 2023]. URL: https://guides.mclibrary.duke.edu/ebm/pico

Döerr, B., Khalili, S. & D'souza, J. 2023. Clinical investigations for medical devices. Medical Writing, 32, 1, 58–64. https://doi.org/10.56012/tpkn1389

EMA 2021. Medical Device Regulation comes into application - European Medicines Agency. [Accessed 18 July 2023]. URL: https://www.ema.europa.eu/en/news/medical-device-regulation-comes-application

European Commission. MEDDEV 2.7/1 revision 4. 2016. Clinical evaluation: A guide for manufacturers and notified bodies under Directives 93/42/EEC and 90/385/EEC.

European Union. 2019. Unique Device Identification (UDI) System under the EU Medical Device Regulations 2017/745 and 2017/746. [Accessed 13 September 2023]. URL: https://ec.europa.eu/docsroom/documents/36664

European Commission. 2021. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. [Accessed 16 February 2023]. URL: https://eurlex.europa.eu/legal-content/EN/TXT/?uri=COM:2021:0206

Flick, U. 2018. Designing Qualitative Research. SAGE Publications.

Food and Drug Administration. 2019. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD). [Accessed 29 September 2023]. URL: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

Fraser, A.G., Byrne, R.A., Kautzner, J., Butchart, E.G., Szymański, P., Leggeri, I., de Boer, R.A., Caiani, E.G., Van de Werf, F., Vardas, P.E. & Badimon, L. 2020. Implementing the new European Regulations on medical devices—clinical responsibilities for evidence-based

practice: a report from the Regulatory Affairs Committee of the European Society of Cardiology. European Heart Journal, 41, 27, 2589–2596. https://doi.org/10.1093/eurheartj/ehaa382

Giefing-Kröll, C., & Laumen, G. 2022. How the EU Medical Device Regulation is affecting the medical device landscape. An interview with Suzanne Halliday, the Regulatory Head of BSI, Medical Devices Notified body. Medical Writing, 31, 62-64.

Gilbert, S., Fenech, M., Hirsch, M., Upadhyay, S., Biasiucci, A. & Starlinger, J. 2021. Algorithm Change Protocols in the Regulation of Adaptive Machine Learning-Based Medical Devices. Journal of Medical Internet Research, 23, 10, e30545. https://doi.org/10.2196/30545

Granlund, T., Stirbu, V. & Mikkonen, T. 2022. Medical Software Needs Calm Compliance. IEEE Software, 39, 1, 19–28. https://doi.org/10.1109/ms.2021.3117292

Grigoleit, S. 2020. Deliverable D6. 1-SHAPES Pan-European Pilot Campaign Plan. [Accessed 20 March 2023]. URL: https://shapes2020.eu/wp-content/uploads/2020/11/D6.1-SHAPES-Pan-European-Pilot-Campaign-Plan.pdf

He, J., Baxter, S.L., Xu, J., Xu, J., Zhou, X. & Zhang, K. 2019. The practical implementation of artificial intelligence technologies in medicine. Nature Medicine, 25, 1, 30–36. https://doi.org/10.1038/s41591-018-0307-0

IMDRF/ SaMD/ Working Group. 2014. "Software as a medical device": possible framework for risk categorization and corresponding considerations. In International Medical Device Regulators Forum. [Accessed 29 November 2023]. URL: "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations (imdrf.org)

IMDRF/SaMD WG/N41. 2017. Software as a Medical Device (SaMD): Clinical Evaluation. [Accessed 29 November 2023]. URL: https://www.imdrf.org/documents/software-medical-device-samd-clinical-evaluation

IMDRF. 2022. Machine Learning-enabled Medical Devices: Key Terms and Definitions. [Accessed 13 March 2023]. URL: https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions

International Medical Device Regulators Forum. 2023. OUTCOME STATEMENT 23rd Management Committee Meeting of the International Medical Device Regulators Forum 27-31 March 2023 Brussels, Belgium. [Accessed 05 May 2023]. URL: EU, Brussels_Belgium_Outcome Statement_IMDRF 23rd session.pdf

Jarman, H., Rozenblum, S. & Huang, T.J. 2020. Neither protective nor harmonized: the crossborder regulation of medical devices in the EU. Health Economics, Policy and Law, 16, 1, 51–63. https://doi.org/10.1017/s1744133120000158

Kamusheva, M., Németh, B., Zemplényi, A., Kaló, Z., Elvidge, J., Dimitrova, M., Pontén, J., Tachkov, K. & Mitkova, Z. 2022. Using real-world evidence in healthcare from Western to Central and Eastern Europe: a review of existing barriers. Journal of Comparative Effectiveness Research, 11, 12, 905–913. https://doi.org/10.2217/cer-2022-0065

Kelly, C.J., Karthikesalingam, A., Suleyman, M., Corrado, G. & King, D. 2019. Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine, 17, 1. https://doi.org/10.1186/s12916-019-1426-2

Keutzer, L. & Simonsson, U.S. 2020. Medical Device Apps: An Introduction to Regulatory Affairs for Developers. JMIR mHealth and uHealth, 8, 6, e17567. https://doi.org/10.2196/17567

King, N. 2004. Using templates in the thematic analysis of text. In C. Cassell & G. Symon (Eds.), Essential guide to qualitative methods in organizational research (257–270). London, UK: Sage.

Kumar, Y., Koul, A., Singla, R. & Ijaz, M.F. 2022. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. Journal of Ambient Intelligence and Humanized Computing, 14, 7, 8459–8486. https://doi.org/10.1007/s12652-021-03612-z

Kwade, Z. 2022. Clinical evaluation of software. RF Q, 2(1), 23-34.

Laki lääkinnällisistä laitteista 719/2021 - Säädökset alkuperäisinä - FINLEX ® s.a. [Accessed 13 September 2023]. URL: https://finlex.fi/fi/laki/alkup/2021/20210719#Pidm46494958395040

Laki sosiaali- ja terveystietojen toissijaisesta… 552/2019 - Säädökset alkuperäisinä - FINLEX ® s.a. [Accessed 13 September 2023]. URL: https://www.finlex.fi/fi/laki/alkup/2019/20190552.L

Lalova, T., Padeanu, C., Negrouk, A., Lacombe, D., Geissler, J., Klingmann, I. & Huys, I. 2020. Cross-Border Access to Clinical Trials in the EU: Exploratory Study on Needs and Reality. Frontiers in Medicine, 7. https://doi.org/10.3389/fmed.2020.585722

Levine, D.M., Co, Z., Newmark, L.P., Groisser, A.R., Holmgren, A.J., Haas, J.S. & Bates, D.W. 2020. Design and testing of a mobile health application rating tool. npj Digital Medicine, 3, 1. https://doi.org/10.1038/s41746-020-0268-9

Ludvigsen, K., Nagaraja, S. & Daly, A. 2021. When Is Software a Medical Device? Understanding and Determining the “Intention” and Requirements for Software as a Medical Device in European Union Law. European Journal of Risk Regulation, 13, 1,78–93. https://doi.org/10.1017/err.2021.45

Deep, M. A., Rana, A.C. & Sharma, P.C. 2019. Regulation and Clinical Investigation of Medical Device in the European Union. Applied Clinical Research, Clinical Trials and Regulatory Affairs, 6, 3, 163–181. https://doi.org/10.2174/2213476x06666190821095407

Medical Device Coordination Group (MDCG). 2019. Guidance on qualification and classification of software in regulation (EU) 2017/745–MDR and regulation (EU) 2017/746–IVDR. European Commission. [Accessed 05 April 2023]. URL: https://ec.europa.eu/health/sites/health/files/md_topics-interest/docs/md_mdcg_2019_11_guidance_en.pdf

Medical Device Coordination Group (MDCG). 2020. Guidance on clinical evaluation (MDR)/performance evaluation (IVDR) of medical device software. [Accessed 05 November 2023]. URL: DocsRoom - European Commission (europa.eu)

Medical Device Coordination Group (MDCG) Working Groups 2023. [Accessed 15 October 2023]. URL: https://health.ec.europa.eu/medical-devices-dialogue-between-interested-parties/medical-device-coordination-group-working-groups_en

Melvin, T. & Torre, M. 2019. New medical device regulations: the regulator's view. EFORT Open Reviews, 4, 6, 351–356. https://doi.org/10.1302/2058-5241.4.180061

Melvin, T. 2022. The European Medical Device Regulation–What Biomedical Engineers Need to Know. IEEE Journal of Translational Engineering in Health and Medicine, 10, 1–5. https://doi.org/10.1109/jtehm.2022.3194415

Muehlematter, U.J., Daniore, P. & Vokinger, K.N. 2021. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. The Lancet Digital Health, 3, 3, e195–e203. https://doi.org/10.1016/s2589-7500(20)30292-2

Munro, D. 2017. Your EU Medical Device Regulation Guide. [Accessed 11 July 2023]. URL: https://ir.canterbury.ac.nz/items/1bbccee4-5c81-4a7a-9325-217b83a442d7

Negrouk, A., Lacombe, D. & Meunier, F. 2018. Diverging EU health regulations: The urgent need for co-ordination and convergence. Journal of Cancer Policy, 17, 24–29. https://doi.org/10.1016/j.jcpo.2017.05.007

Niemiec, E. 2022. Will the EU Medical Device Regulation help to improve the safety and performance of medical AI devices? DIGITAL HEALTH, 8, 205520762210890. https://doi.org/10.1177/20552076221089079

Näher, A.-F., Vorisek, C.N., Klopfenstein, S.A.I., Lehne, M., Thun, S., Alsalamah, S., Pujari, S., Heider, D., Ahrens, W., Pigeot, I., Marckmann, G., Jenny, M.A., Renard, B.Y., von Kleist, M., Wieler, L.H., Balzer, F. & Grabenhenrich, L. 2023. Secondary data for global health digitalisation. The Lancet Digital Health, 5, 2, e93–e101. https://doi.org/10.1016/s2589-7500(22)00195-9

OCEBM Levels of Evidence Working Group. 2011. " The Oxford 2011 Levels of Evidence." Oxford Centre for Evidence-Based Medicine.

Ordish, J., Hannah, M., & Hall, A. 2019. Algorithms as medical devices. PHG Foundation.

Pane, J., Francisca, R.D.C., Verhamme, K.M.C., Orozco, M., Viroux, H., Rebollo, I. & Sturkenboom, M.C.J.M. 2019. EU postmarket surveillance plans for medical devices. Pharmacoepidemiology and Drug Safety, 28, 9, 1155–1165. https://doi.org/10.1002/pds.4859

Pritchard, G. 2017. Clinical Evaluation Reports from the medical writer's perspective! Medical Writing, 26, 14-19

Pritchard, G. 2022. Clinical evaluation reports: 6 years after the introduction of MEDDEV 2.7/1 revision 4. Medical Writing, 31, 40-43

Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices. [Accessed 17 July 2023]. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745

Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices. [Accessed 17 July 2023]. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0746

Ravizza, A., Ciani, O., Ferrari, A., Grigolo, S., Perrone, F., Ravot, E., Santoro, E., Sternini, F. 2021. Clinical evaluation of digital therapeutics. Tendenze, 71. [Accessed 17 October 2023]. URL: Tendenze Nuove 1 2018 (passonieditore.it)

Sandelowski, M. 2004. Using Qualitative Research. Qualitative Health Research, 14, 10, 1366-1386. https://doi.org/10.1177/1049732304269672

Sarlio-Siintola, S., Nikula, K., Aholaakko, T-K., Alapuranen, N., & Kartsidis, P. et.al 2019. Smart and Healthy Ageing through People Engaging in Supportive Systems.D8.2 – Baseline for the project ethics. [Accessed 17 October 2023]. URL: https://shapes2020.eu/deliverables/

Scholtes, M., Behrend, A., Buedenbender, S., Volker, G. & Keywan, S. 2018. Clinical Evaluation of Mobile Medical Apps. Current Directions in Biomedical Engineering, 4, 1, 173-176. https://doi.org/10.1515/cdbme-2018-0043.

Seidel, K., Labor, M., Lombard-Vance, R., McEvoy, E., Cooke, M., D'Arino, L., Desmond, D., Ferri, D., Franke, P., Gheno, I., Grigoleit, S., Guerra, B., Krukowski, A., Pešoutová, M., Pietri, I., Prendergast, D., Maguire, R., Manso, M., Saris, A.J., Sarlio-Siintola, S., Silva, T., Zarogianni, E. & MacLachlan, M. 2022. Implementation of a pan-European ecosystem and an interoperable platform for Smart and Healthy Ageing in Europe: An Innovation Action research protocol. Open Research Europe, 2, 85. https://doi.org/10.12688/openreseurope.14827.1

Singhal, A. & Cowie, M.R. 2021. Digital Health: Implications for Heart Failure Management. Cardiac Failure Review, 7. https://doi.org/10.15420/cfr.2020.28

Spargo, M., Goodfellow, N., Scullin, C., Grigoleit, S., Andreou, A., Mavromoustakis, C.X., Guerra, B., Manso, M., Larburu, N., Villacañas, Ó., Fleming, G. & Scott, M. 2021. Shaping the Future of Digitally Enabled Health and Care. Pharmacy, 9, 1, 17. https://doi.org/10.3390/pharmacy9010017

Spargo, M., Goodfellow, N. 2023. D6.4 Medicine Control and Optimisation Pilot Activities Report. EU SHAPES project. [Accessed 17 November 2023]. URL: 857159_Deliverable_62_Medicine-Control-and-Optimisation-Pilot-Activities-Report.pdf (shapes2020.eu)

Software and AI as a Medical Device Change Programme 2023. [Accessed 17 October 2023]. URL: https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme

Starks, H. & Brown Trinidad, S. 2007. Choose Your Method: A Comparison of Phenomenology, Discourse Analysis, and Grounded Theory. Qualitative Health Research, 17, 10, 1372–1380. https://doi.org/10.1177/1049732307307031

Sutton, R.T., Pincock, D., Baumgart, D.C., Sadowski, D.C., Fedorak, R.N. & Kroeker, K.I. 2020. An overview of clinical decision support systems: benefits, risks, and strategies for success. npj Digital Medicine, 3, 1. https://doi.org/10.1038/s41746-020-0221-y

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H. & Connelly, A. 2019. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. NeuroImage, 202, 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

Tzintoli, A. 2022. Data Act: MedTech Europe calls for a data-sharing ecosystem that supports innovation in healthcare - MedTech Europe. [Accessed 27 July 2023]. URL: https://www.medtecheurope.org/news-and-events/press/data-act-medtech-europe-calls-for-a-data-sharing-ecosystem-that-supports-innovation-in-healthcare/

Unsworth, H., Dillon, B., Collinson, L., Powell, H., Salmon, M., Oladapo, T., Ayiku, L., Shield, G., Holden, J., Patel, N., Campbell, M., Greaves, F., Joshi, I., Powell, J. & Tonnel, A. 2021. The NICE Evidence Standards Framework for digital health and care technologies – Developing and maintaining an innovative evidence framework with global impact. DIGITAL HEALTH, 7, 205520762110186. https://doi.org/10.1177/20552076211018617

Valla, V., Tzelepi, K., Charitou, P., Lewis, A., Polatidis, B., Koukoura, A., Karapatsia, A., Antonopoulou, K., Prodromidou, K., Papadaki, E. & Vassiliadis, E. 2023. Use of Real-World Evidence for International Regulatory Decision Making in Medical Devices. International Journal of Digital Health, 3, 1, 1. https://doi.org/10.29337/ijdh.50

Zapata, K. A. C., Patil, R., Ward, T., Loughran, R., & McCaffery, F. 2023. Analysis of the Classification of Medical Device Software in the AI Act Proposal. Workshops at the Second International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 26—27, 2023, Munich, Germany.

Figures

Figure 1: MDR Annex VIII, section 6.3, Rule 11 14

Tables

Table 1: List of some use cases of the SHAPES pilot project (Grigoleit 2020) 20
Table 2: Informants (N=10), their ID, position, professional field, and number of words in transcript 24
Table 3: PICO model applied to conduct the study within the EU SHAPES pilot project 27
Table 4: Themes and subthemes derived from thematic analysis 29

Appendices

Appendix 1: Participant information sheet

Study title: EXPLORING THE CHALLENGES IN THE HARMONIZATION OF CLINICAL EVALUATION OF MEDICAL DEVICE SOFTWARE ACROSS EU MEMBER STATES

This information sheet describes the study and Your role in it. Before you decide, it is important that You understand why the research is being done and what it would involve for You. Please take time to read this information and discuss it with others if You wish. If there is anything that is not clear, or if You would like more information, please ask us. After that we will ask You to sign a consent form to participate in the study.

Voluntary nature of participation

Participation in this study is voluntary. You can withdraw from the study at any time without giving any reason and without there being any negative consequences. If You withdraw from the study or withdraw Your consent, any data collected from You before the withdrawal can be included as part of the research data.

Purpose of the study

The purpose of the research is to explore challenges in the harmonization of clinical evaluation of MDSW/SaMD among EU member states in order meet requirements of the EU regulations for safety and performance.

Who is organizing and funding the research?

The research is organized within pan-European SHAPES pilot project and conducted by Anna Polishchuk as a Master's thesis project. The study does not have funding.

What will the participation involve?

Semi-structured online interviews using "Microsoft Teams" and video conferencing technology will be used to gather data. To match the inductive and exploratory nature of this study, the interview guide is built on the research questions. With the interviewee's permission, the interviews will be taped and then verbatim transcribed. The participants are contacted to seek clarification from them if they noticed any uncertainty in the verbatim interview transcripts so that their statements could be properly understood.

The target audience will only be senior professionals spread out across the EU with likely busy schedules; therefore, Microsoft Teams video conferences over the internet were chosen to make the best use of their time. Participants will be sought out by email and will be found through a range of sources.

All the interviews will take place in English. After confirming that they read the information sheet and consent form during these calls, people will be given the option to give their informed consent (Appendix 2). After that, the recording started, and their permission will be verified once more for the record. The participants will have the option to stop the recording on their own, and they will be made aware that they can withdraw from the interview and the research at any time, as well as the information they gave.

Possible benefits of taking part

Indirect benefits to the participants in terms of improving safety and performance of Software as a Medical Device in the EU.

Possible disadvantages and risks of taking part

There is no reasonably foreseeable discomforts, disadvantages and risks for the participants.

Financial information

Participation in this study will involve no cost to You. You will receive no payment for Your participation.

Informing about the research results

The results of the study will be made available to the participants in the form of a Master's Thesis or scientific publication. The participants will not be identified from any report, publication or Master's Thesis of Anna Polishchuk placed in the public domain.

Termination of the study

The researcher conducting the study can also terminate the study due to some force major reasons.

Further information

Further information related to the study can be requested from the researcher / person in charge of the study.

Contact details of the researchers

Researcher / Student

Name: Anna Polishchuk

Tel. number: [redacted]

Email: anna.polishchuk@sudent.laurea.fi

Supervisor

Name: Teija-Kaisa Aholaakko, Principal Lecturer, PhD (Medicine)

Name of the organisation / Faculty: LAUREA University of Applied Sciences

Tel. number: [redacted]

Email: teija-kaisa.aholaakko@laurea.fi

Appendix 2: Participant consent form

Title of the study: EXPLORING THE CHALLENGES IN THE HARMONIZATION OF CLINICAL EVALUATION OF MEDICAL DEVICE SOFTWARE ACROSS EU MEMBER STATES

Location of the study:

Organization conducting study: Laurea University of Applied Sciences, Finland

Researcher: Anna Polishchuk, Global Health and Crisis Management Master program student, anna.polishchuk@student.laurea.fi, Mobile: +358 468114254

Supervisor: Teija-Kaisa Aholaakko, Principal Lecturer, PhD (Medicine), LicSc (Education), MSc (Healthcare), RN (Spec.), Program leader of Masters of Global Health and Crisis Management – Degree, Laurea Development Unit (Education), Laurea Tikkurila, Ratatie 22, 01300 VANTAA, Finland, Mobile: +358468567348, teija-kaisa.aholaakko@laurea.fi

I (NAME) have been invited to participate in the above research study. The purpose of the research is to explore challenges in the harmonization of clinical evaluation of Software as a Medical Device among EU member states in order meet requirements of the EU regulations for safety and performance.

I have read and understood the written participant information sheet. The information sheet has provided me with sufficient information about the above study, the purpose and execution of the study, about my rights as well as about the benefits and risks involved in it. I have had the opportunity to ask questions about the study and have had these answers satisfactorily.

I have had sufficient information of the collection, processing and transfer/disclosure of my personal data during the study and the Privacy Notice has been available.

I voluntarily consent to participate in this study. I have not been pressurized or persuaded into participation.

I have had enough time to consider my participation in the study.

I understand that my participation is entirely voluntary and that I am free to withdraw my consent at any time, without giving any reason. I am aware that if I withdraw from the study

(I can continue it later), any data collected from me before my withdrawal, can be included as part of the research data.

By signing this form, I confirm that I voluntarily consent to participate in this study.

The legal basis of processing personal data within this study is a consent granted by me as the data subject, by signing I grant the consent for process my personal data. I have right to withdraw the consent regarding processing of personal data as described in the Privacy Notice.

Date

_____

Signature of Participant

The original consent signed by the participant and a copy of the participant information sheet will be kept in the records of the researcher. Participant information sheet, privacy notice and a copy of the signed consent will be given to the participant.

Appendix 3: Questions used for semi-structured interviews

1. What in your opinion are the challenges to the harmonization of clinical evaluation of SaMD/MDSW among EU member states?

2. May some national solutions to the issues increase legal differences? What problems do you see and what changes need to be done?

4. Can you think of any specific approaches that could be made to begin implementing these changes?

5. What in your opinion the EU SHAPES pilot project should consider for clinical evaluation of MDSW?

Appendix 4: Extract from the row data of the transcript

0:6:31.380 --> 0:6:52.950
P1
Project related softwares have been in class one. Maybe the decision the classification was not right. But according to the new rules it is absolutely clear that the minimum classification will be 2A and that requires at least the quality management system and an assessment of the common from the notified body.

0:6:53.830 --> 0:7:8.650
P1
And as you know, we have changed a little bit, but that is already something that we have done in 2007.

0:7:10.40 --> 0:7:36.540
P1
Where we decided that even if she has only the requirement for quality management system, we are asking the notified body to have an assessment of a technical documentation of at least one product of the portfolio of the manufacturer. Yeah, and that means for many of those software companies that they also have to have a kind of assessment of the full technical documentations. That is kind of Class 3 device.

0:7:36.620 --> 0:7:44.690
P1
And conformity secure. That means the notified body, but not only check the quality management system, but also.

0:7:45.450 --> 0:8:3.620
P1
I'm check see technical file. The technical documentation including the clinic evolution and that is something that you have asked for clinical evaluation for software that is special, special issue and maybe and special challenge.

0:8:5.120 --> 0:8:15.320

Appendix 5: Example of thematic analysis of the transcript texts

Research question 1. What are the challenges in the harmonization of clinical evaluation of MDSW across EU member states?

Quote from transcript	Sub-theme	Theme
A1: And I guess the second challenge then thinking of clinical evidence generation is that and the rules that we have in MDR, um, I guess they are designed with physical products in mind. And the way that software developers think of product development is often very different to the way that engineers or clinicians think about how you'd make a new stent or implant. A1: Yeah, I think we've covered most of the things that will be needed and I guess the other thing from a development perspective is that it's a real challenge because you have MDR, then you have GDPR. Then in the future you have the artificial intelligence regulation, European health data space, there's be a Cyber Resiliency Act. I think cyber security rules and for developers, they just see all regulation is not something they usually comfortable with because they're programming People and are very good at the technology side of things, but and often small organisations, but on the regulatory side it can be a real challenge and then when you get into the amount of detail you need for MDR compliance, now you need a full risk management file. P2: The move from the directive to the regulation, especially with regards to software, is a real implication, because previously under the directives in most software was Class I self declared, so you didn't need to go to a notified body and now it's the complete inverse. So 99% of software is Class II and above, and you always need to go to a notified body.	EU Legislation relevant to MDSW	EU-level challenges
A1: So transparency is a real challenge in the system. What they changed was MDR is they said that, um, there's a number of cases where there will be more information, but it's highly prescribed. And the information that will be released is basically copy and pasted from the technical file. So it could be very generic wording for how the evidence is described. It's probably going to be quite different to, you know, this summary of safety and effectiveness data that they have in the United States.	Clinical evaluation of MDSW	EU-level challenges

P2: So as I mentioned, our system is a little bit different. **The Member States themselves, they don't assess the software we have assess we, yeah, we have a system of designation of notified bodies.** Yes, they're independent organizations. Yeah, they're independent expert organizations who work with the national authorities, but they are mandated to do the technical assessments of documentation. A2: Challenges and harmonization within Member States in the EU on the basis of clinical evaluation, well to some small degree, **there may not be perfect harmonization because the Notified bodies are based in individual Member States, and you can choose your own notified body.** It can be based in any Member State, so it doesn't have to be with your company's based. You can choose any EU. Thus is getting notified body to be your notified body and **that notified body is subject to its own National Assessment Authority which audits another body and there are small differences of approach from those national designating authorities.**	**Notified Bodies**	National level challenges
P1: Let me see from the medical device regulatory framework, there are no differences. There might be some differences with regard to privacy data. Yeah, so that this could be. Sometimes the difficulty, even for the postmodern clinical follow up they have. You have the software and you are using data from patients and you can become an issue with regard to privacy and data protection. However, even I think we have said in the. General Data Protection Regulation that helps director are excluded from this. So there is some possibility even to sampling and acquire Health data from patients. But.So we think sometimes it could happen that some hospitals would say, uh, we cannot share those data with you because we do not want and we are not able to delete the names or the birth dates etcetera, yeah.	RWD	National level challenges

Research question 2. What are the challenges in the EU SHAPES project related to clinical evaluation of the MDSW developed within the project?

Quote from transcript	Sub-theme	Theme
P1: Take note that we have changed the classification of software. Yeah, and there's a kind of general rule that there is no more class I medical device software that means and I guess many of your or of the SHAPES project related softwares have been in class I. Maybe the decision the classification was not right. But according to the new rules it is absolutely clear that the minimum classification will be IIa and that requires at least the quality management system and an assessment of the common from the notified body.	Classification of MDSW	EU SHAPES pilot project
S1: ...Uh, no, it will be done by the health care provider, not only the software. **The software can predict some cases of abnormal vital signs.** But in general, the last word will be from the healthcare provider. **It will inform the patients if there is some abnormal cases where they should change their medication** and that will be done the future in it's not as through the pilot activities because the pilot activities and we cannot bring the patients into a any risk. Of course, we're not healthcare providers.	**Qualification of software**	EU SHAPES pilot project