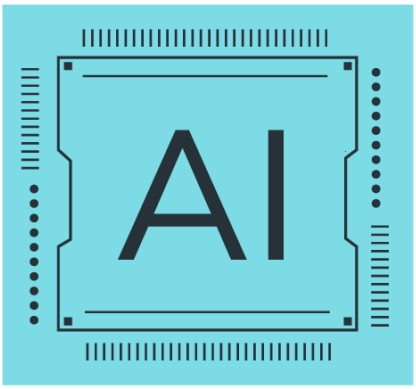
TÜV AI.LAB

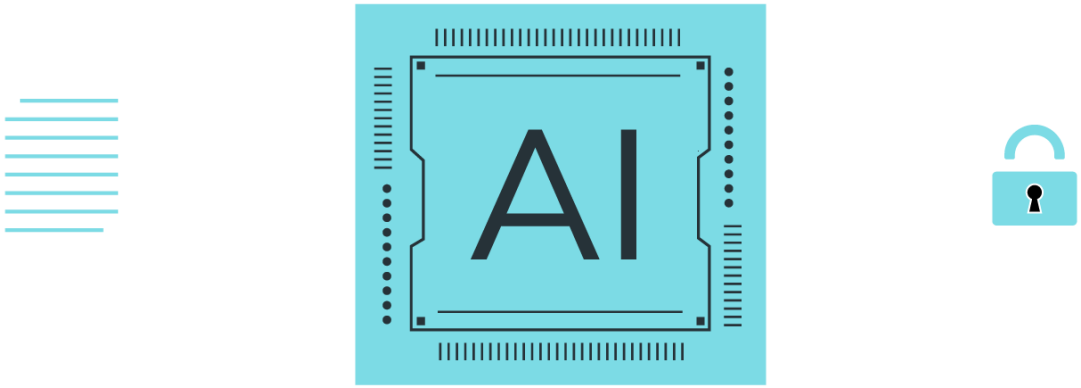
Certifying the future



Certifying AI medical devices: Challenges and recommendations

Summary

The integration of Artificial Intelligence in Medical Devices Software (AIMDSW) has the potential to radically transform the European healthcare sector. As Artificial Intelligence (AI) solutions mature and are increasingly deployed, ensuring these applications are of high quality, meet the necessary performance requirements and are safe in usage, is paramount. A key pillar in verifying this is independent third-party assessment, certifying compliance with legal requirements and validating whether AIMDSW applications are safe, secure and trustworthy. Even though this is already common practice for Notified Bodies (NBs) across Europe, certifying AIMDSW along the traditional pathway of the Medical Device Regulation (MDR) is challenging in central aspects. While the AI Act and additional standards will provide more guidance, their potential to alleviate the identified challenges remains to be determined. In general, certifying AI-systems poses significant challenges specific to this technology. Therefore, future adjustments to the certification process are necessary, favouring a.o., more agility, shorter time-to-market and allowing for continuous monitoring of ever-changing AI-systems.



AI In Healthcare

AI is expected to have a significant effect on medical practice. Already today, AI applications can perform certain tasks close to – or with even higher levels of accuracy than those of human experts. Moreover, AI may not only impact medical practice but the entire workflow from biomedical research to administrative tasks, resulting in an overall increase of access to high quality medical services on a broad scale. As the exact effect of AI will unfold over the coming years, making sure these innovations are of high quality as well as safe and secure is a central challenge we are facingtoday.

Multiple applications prove this already. One prominent example is the use of one-dimensional data (1-D) for the detection of different skin lesions to predict benign or malignant types of skin cancer. Depending on the exact learning model and data used, these applications can score up to 99.6% accuracy, 100% recall, 100% sensitivity and 99.2% specificity and thus outperform any human doctor[1]. Additionally, AI can help to alleviate some of the bigger, underlying problems of the European healthcare sector like an ageing population, overworked doctors and nurses due to a lack of health personal or may reduce healthcare inequalities overall by making quality treatments cheaper.

At the same time, however, key challenges remain as AI applications can produce different harms, which are difficult to track across the application's life cycle. Furthermore, there is still little mechanistic understanding of certain learning methods like deep neural networks, which are a popular framework for both predictive as well as generative AI. This makes it difficult to explain model behavior where necessary and detect sources of harm like unjustified bias early on.

Legal Background

When introducing an AIMDSW into the European Market, MDR compliance remains the central regulatory entry point. At the same time, the upcoming AI Act will fill regulatory gaps regarding AI-based applications that are not covered in the general provisions of the MDR or respective standards. This will provide more guidance to the deployment as well as the certification of AI-based application in the healthcare sector. However, it remains to be determined how well upcoming AI standards will directly translate to healthcare use-cases or whether sector specific standards need to follow.

1 Sikandar, S., Mahum, R., Ragab, A.E., Yayilgan, S.Y., Shaikh, S. (2023): *SCDet: A Robust Approach for the Detection of Skin Lesions*, In: Diagnostics 13, 1824, https://doi.org/10.3390/diagnostics13111824

Medical Devices Regulation

In force since May 2021, the MDR regulates all medical devices, including AIMDSW. The main goal of the MDR is to prove the safety and performance of medical devices that enter the European Market. It divides medical devices into four risk classes (I, IIa, IIb and III), which determine the assessment's depth to prove conformity of the respective device.[2] In short, Notified Bodies assess the product's conformity and issues a certificate if all relevant legal requirements are met. Following this, the manufacturer can affix the CE marking and sell the respective medical product in the European Market. A re-certification Is mandatatory at least every five years while surveillance audits and unannounced audits are in part carried out annually.

In general, MDR certification is a lengthy process. According to a recent study by the European Commission[3], initial certification can take between 12–24 months, while re-certification can take 6–24 months, depending on the complexity of the device in question. For many manufacturers this led to a gradual shift away from the European Market, leaving it at best a secondary market[4].

European AI Act

AI adds an additional layer of complexity and uncertainty to certification processes and MDR compliance. To fill regulatory gaps with respect to AI across all relevant sectors, the European Commission introduced the AI Act. The Act applies horizontally and divides AI applications into four risk categories: unacceptable risk, high-risk, low-risk and no risk. Only high-risk applications are required to comply with the provisions of the Act, which demands for example the establishment of a quality management system, a risk management system, registration in a public EU database and an internal conformity assessment. While this mirrors many of the requirements already laid out in the MDR, the Act further specifies AI-specific requirements around data quality, documentation and traceability, transparency, human oversight, accuracy, cybersecurity and robustness. It also expands the MDR in aspects like dynamic and continuous improvements for example.

2 Class I medical devices are further divided into I, Is, Ir and Im. However, for medical devices in class I no third-party conformity assessment needs to be conducted.
3 European Commission (2023): *Notified Bodies Survey on certifications and applications*, available at: https://health.ec.europa.eu/system/files/2023-11/md_nb_survey_certifications_applications_en.pdf
4 Pedersen (2022): *Gone are the Days of Europe-First Medical Device Innovation*, available at: https://www.mddionline.com/regulatory-quality/gone-are-the-days-of-europe-first-medical-device-innovation

Challenges of AIMDSW certification

Certifying compliance of AIMDSW with the MDR is already common practice in many Notified Bodies across Europe. However, challenges remain even now and will greatly increase with more innovative and advanced AI-based systems being integrated into medical devices as well as growing market-penetration of AI-based systems.

Challenge 1: Lengthy certification process meets fast iteration cycles

As outlined above, obtaining MDR certification is already a lengthy and tedious process. This, however, is less of a problem for physical medical devices where the purpose of the device, the device itself or the target population changes on a much slower basis. AI-based applications on the other hand, are developed in quick iterations, improving the product step by step. As such, their innovation cycles are much faster than those of physical devices or even traditional software. Consequently, finding efficient certification processes within the existing MDR pathway becomes crucial for ensuring the timely integration of the most advanced models into the European Market.

Challenge 2: Small room for continuous adjustments

Under the current framework, medical devices need to be frozen during certification and can only be adapted within a very specific range without requiring re-certification. The power of learning models like neural networks, however, lies in constant improving based on new input data they receive after deployment. As such, AI-based models and products can change even after they have been introduced to the market. This type of adjustments is common in Large Language Models (LLMs) like ChatGPT for example. A change in the foundational model results in a change of behaviour of the application despite it already being in use. These types of dynamical adjustments poses a great challenge for MDR certification and highlight the need for finding ways that allow more possibilities for dynamic adjustments.

Challenge 3: No standardised approaches for AIMDSW certification

The certification of AI presents distinct new technical challenges. The problem of unjustified bias, the challenge of inscrutable models and the need for transparency and traceability are just a few. While certifying AIMDSW is possible and already a reality, there are no harmonised standards or common criteria manufacturers can follow and NBs can assess against. The lack of these hinders the ability to systematically monitor the compliance of evolving models and impedes the facilitation of swift iterations.

The way forward – initial recommendations

With respect to AI certification and governance, the medical sector serves as a pivotal example. It stands as one of the few domains where AI is already being certified, highlighting the relevance of effective certification practices to ensure compliance of future innovations. Finding ways to efficiently certify AI-based applications of today and tomorrow holds an immense potential for the healthcare sector and ultimately benefits patients in general. To alleviate some of the above-mentioned challenges, we propose the following set of recommendations:

Recommendation 1: Speeding up the existing certification processes through digitalisation

To account for fast iteration cycles of dynamic systems, we propose a platform-based approach. Building on standardised, open-source data formats, connectivity and portability, we can transform the approval process from a document-heavy to a digital one. Digitalised documentation has multiple advantages. First, it makes information layerable, allowing the separation of content and carrier. Second, it makes information referenceable, ensuring the real-time synchronisation of information across devices. Finally, information becomes actionable, meaning that certain tasks can be triggered automatically. Such a platform not only supports the work of NBs but also contributes to manufacturers' submission of required information in the correct format.

Recommendation 2: Continuous monitoring and agility for faster compliance

While digitalisation is an important first step, one should not stop here. Rather, the certification process as a whole needs to be reshaped to secure compliance even post deployment and make dynamic systems easier to certify. Such an agile certification process needs to be designed from end-to-end, taking today's fast-paced developments and changes of product iterations into account. This ensures ongoing compliance of all iterations without any shortcomings with regard to safety, security or trustworthiness. This lays the basis to certify even the most complex models like deep neural networks or LLMs on a continuous basis.

Recommendation 3: Standardised approaches and common benchmarks

New approaches are needed that standardise the evaluation of AIMDSW. Without these recognised standards, the task of reliably certifying medical devices remains challenging. Furthermore, a unified set of standards will not only facilitate NBs in their conformity assessments but will also help manufacturers in increasing reliability and patient benefit of their applications. Moreover, it streamlines the tracking of the application's most critical components along a pre-defined corridor of change across multiple iterations. This will facilitate the validation of system's trustworthiness, reliably and reproducibly.

Author and Contact

Alexander von Janowski
Advisor AI Certification
alexander@tuev-lab.ai
www.tuev-lab.ai

Herausgeber
TÜV AI.LAB GmbH
Max-Ulrich Straße 3
13355 Berlin