Risk Assessment and Classification of Medical Device Software for the Internet of Medical Things

Challenges arising from connected, intelligent medical devices

Irina Brass*
University College London, London, UK
i.brass@ucl.ac.uk

Andrew Mkwashi
University College London, London, UK
a.mkwashi@ucl.ac.uk

ABSTRACT

Although the medical device industry operates within a stringent regulatory environment, the growing deployment of connected, intelligent medical devices (CIMDs) in the healthcare sector is challenging these established regulatory frameworks. CIMDs come in a variety of forms, from implantables, to specialist IoMT devices deployed at the point-of-care, to AI-based medical devices, and AI as a medical device (AIaMDs). These devices raise several cybersecurity, data management, and algorithmic integrity concerns for patient safety and the delivery of reliable, responsible healthcare. The purpose of this article is to focus on a particular characteristic of CIMDs: their changing risk profile, several times throughout their lifecycle, with limited awareness from users, manufacturers, and regulators. Looking at the implications of these often subtle yet meaningful software modifications for current medical device regulations and for critical stakeholders in the CIMD ecosystem, the article highlights three main challenges to: i) risk assessment, classification and management frameworks that underpin current medical device regulations; ii) current medical device compliance frameworks, especially the post-market surveillance of medical devices; and iii) the detection, categorization, and reporting of compromised devices that might not perform according to their intended purpose. The article brings empirical evidence from a qualitative research study conducted with critical stakeholders in the medical device sector.

CCS CONCEPTS

• **Internet of medical things**; • **artificial intelligence**; • **risk classification**;

KEYWORDS

Medical device software, regulation

ACM Reference Format:
Irina Brass and Andrew Mkwashi. 2022. Risk Assessment and Classification of Medical Device Software for the Internet of Medical Things: Challenges arising from connected, intelligent medical devices. In *Proceedings of the 12th International Conference on the Internet of Things (IoT '22), November 07–10, 2022, Delft, Netherlands.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3567445.3571104

*Corresponding author.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
IoT '22, November 07–10, 2022, Delft, Netherlands
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9665-3/22/11.
https://doi.org/10.1145/3567445.3571104

1 INTRODUCTION

The use of software in medicine and healthcare, including software as a medical device (SaMD[1]), has been growing over the years, having a variety of applications and purposes from clinical decision support, diagnosis, treatment, and assistance in complex medical interventions. Within the broad category of software-based medical devices and SaMDs, we have identified a category of healthcare technologies that is raising considerable concerns about patient safety and security, and the integrity of digital healthcare, which we call "connected, intelligent medical devices" (CIMDs). CIMDs are "medical devices that incorporate software and artificial intelligence tools, and use communication technologies and networks to transfer, manage, store, and analyze health data" [31:10]. These can be connected devices used in various healthcare settings such as smart CT scanners, as well as wearables or implantables such as heartrate monitors that collect patient data and can provide therapeutic options. They can also be AI-based medical devices or standalone AI as a Medical Device (AIaMD) that provide decision support or assistance to professional staff. Together, these devices form a connected, intelligent medical device ecosystem at the confluence of the Internet of Medical Things (IoMT) and artificial intelligence (AI) [22, 40, 42] – a connected infrastructure of smart medical devices, software applications, and communication systems and services that facilitate data collection, transmission, storage, management, analysis, and actuation in digital healthcare.

While CIMDs have undoubtable benefits – from remote management of heart failure in implantables to sophisticated machine learning software that provides considerable support in diagnosis or surgery – they also raise critical cybersecurity and algorithmic integrity concerns. These challenges are becoming more well documented in the specialist literature and practice [7, 8, 10, 25, 27, 29, 41, 44], highlighting their serious consequences for patient safety, their health outcomes and fundamental rights, as well as important consequences for medical professionals and the resilience of the healthcare infrastructure [31].

Because of their potentially life-threatening consequences for patients, medical devices are strictly regulated in most jurisdictions in order to evaluate and manage their safety and performance (Section 3). The regulation is generally structured on a risk-based medical device classification system, from low to moderate to high risk. Risk assessment and classification are conducted by the device manufacturer and reported to the regulator for review and/ or market authorisation, or to an approved body as part of conformity

[1]The IMDRF defines SaMD as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device" [21:6].

assessment and certification. In recent years, medical device regulations have tried to keep pace with innovations in digital healthcare, recognizing that medical device software and SaMDs fall within their scope as "active medical devices" [14, 54].

However, these regulatory reforms are also perceived as insufficient in keeping pace with the cybersecurity, data quality, and algorithmic integrity challenges raised by CIMDs [6, 31, 33, 39, 49, 55]. A particular concern that will be explored in this article is guided by the following set of questions:

1. To what extent does the risk profile of a connected, intelligent medical device – based on its risk classification declared by manufacturers at certification, regulatory review, or market approval – changes once the device is in use?

2. How often is a CIMD's risk profile modified throughout its lifecycle, and how transparent are those changes to critical stakeholders?

These are important questions given that the risk-based classification of medical devices is a fundamental pillar of current medical device regulations.

1.1 Key Findings

In this article, we highlight that some CIMDs can change their risk profile several times throughout their lifecycle with limited awareness from users, manufacturers, and regulators. These changes are attributed to the more dynamic nature of software in/as medical device than their conventional counterparts. The modifications can result from both the exploitation of cybersecurity vulnerabilities in the IoMT and from the growing use of unlocked, "adaptive" algorithms in medical device software based on continuous learning[2] [15, 43, 45]. Our scenario below (Section 2) shows that these dynamics can have serious consequences in the connected, intelligent medical device ecosystem where IoMT and AI medical devices are increasingly embedded and rely on each other's functionalities to create the smart healthcare environment our societies require. In our discussion (Section 4), we further highlight that the changing risk profile of CIMDs raises challenges for: i) the risk assessment, classification and management frameworks that underpin current medical device regulations; ii) current medical device compliance frameworks, especially the post-market surveillance of medical devices; and the iii) the detection, categorization, and reporting of compromised devices that might not perform according to their intended purpose.

1.2 Methodology

The findings and discussion points highlighted in this article are based on research conducted in the Reg-MedTech Project [56] of the PETRAS National Centre of Excellence in IoT Systems Cybersecurity between October 2021 and August 2022. The project examines the critical regulatory and standardization challenges raised by CIMDs. The findings outlined below are derived from the following data collection methods employed in our research:

Semi-structured interviews with 12 stakeholders from different organizations involved in the manufacturing, development, regulation, and operation of CIMDs in the digital healthcare and medical device space. The categories of respondents included 4 device manufacturers, 2 software developers/ researchers, 1 security practitioner, 1 lawyer, 1 regulator, 1 standards-maker, 1 clinician, 1 academic researcher. The interviews were conducted with a consistent line of inquiry that was two-fold: i) understanding the existing process of regulating CIMDs; ii) understanding the critical challenges posed by CIMDs to current regulatory frameworks, manufacturing and clinical practices. The chosen line of inquiry was to facilitate a detailed analysis based on practitioner and expert knowledge in the field. Consistent with the practice of conducting semi-structured interviews, the duration of interviews and the number of questions varied across different participants and was determined by the knowledge and willingness of participants to discuss pertinent issues. All the interviews were digitally recorded, transcribed, and checked. On average, most interviews lasted one hour. These practitioner and expert elicitation interviews allowed the researchers to gain a deeper understanding of the benefits and limitations of existing regulatory frameworks and the obstacles faced by different stakeholders in developing and deploying CIMDs.

A roundtable (in workshop format) entitled "The Future of Medical Device Regulation and Standards: Dealing with Software Challenges", held on 27 April 2022, organized in collaboration with BSI (UK national standards body) and MHRA (UK medicines and healthcare products regulator). The roundtable was attended by 45 participants with representation across the following stakeholder categories: medical device manufacturers, software developers, regulators, standards-makers, researchers, lawyers, security practitioners, clinicians, industry association representatives. The roundtable included a plenary session with keynote talks that focused on the latest regulatory responses to software-based medical devices and how standards can best support these regulatory developments. The keynotes were followed by several rounds of small group discussions tackling the main hurdles that software developers and device manufacturers face pre- and post-market to demonstrate conformity and ensure an appropriate level of cybersecurity, data governance, and integrity of algorithmic decisional tools in the medical field. This approach facilitated a broad range of responses and the provision of critical details based on expert and practitioner experiences [23]. Participants also reflected on current gaps in regulatory guidelines and standards and discussed priority areas for future standards development. Full details of the workshop methodology and findings, including the workshop design and questions, can be found in the White Paper entitled "The Future of Medical Device Regulation and Standards: Dealing with Critical Challenges for Connected, Intelligent Medical Devices" [31], in Section 3 and Appendices A-C.

Desk-based analysis of relevant policy, legislative, and regulatory documents applicable to medical devices, including software-based and standalone medical device software. The study applied in-depth content analysis as a method to review policy papers. Full details of the policy and regulatory frameworks reviewed can be found in the White Paper entitled "The Future of Medical Device Regulation

[2]The terms "locked" and "adaptive" algorithms are used in relation to medical devices. According to the FDA, "locked algorithms are those that provide the same result each time the same input is provided. As such, a locked algorithm applies a fixed function (e.g. a static look-up table, decision tree, or complex classifier) to a given set of inputs. These algorithms may use manual processes for updates and validation. In contrast to a locked algorithm, an adaptive algorithm (e.g. a continuous learning algorithm) changes its behaviour using a defined learning process. The algorithm adaptation or changes are implemented such that for a given set of inputs, the output may be different before and after the changes are implemented" [16:5].

and Standards: Dealing with Critical Challenges for Connected, Intelligent Medical Devices" [31], in Section 2.

Ethical considerations were fundamental in the conduct of this research as it involved personal data collection and collection of viewpoints from individuals and organizations. Ethics approval was thus obtained from UCL Research Ethics. Necessary steps for ensuring privacy, anonymity, and confidentiality were designed into the research programme. Prior to the interviews and roundtable, informed consent to take part in the study was obtained from each participant.

In the sections below, we provide anonymized evidence from our interviews and roundtable event to substantiate our findings and discussion points.

2 CHANGING RISK PROFILE OF CONNECTED, INTELLIGENT MEDICAL DEVICES (CIMDS)

Let us imagine the relatively likely scenario where a new pandemic is affecting our global society. Like Covid-19, it is based on the airborne transmission of a respiratory virus that affects the lungs. The extent to which and how the virus affects the respiratory system varies considerably. Patients arrive in hospital with different degrees of lung damage, making it hard to identify critical early treatment. Because the condition is poorly understood and continuously evolving, several hospitals decide to employ a new medical imaging analysis software using an unlocked, adaptive algorithm that processes and analyses chest and lung images sent from several hospitals and offers clinical decision support to professional staff, such as whether an antiviral agent or mechanical ventilation should be administered.

The benefits of employing an adaptive algorithm using a deep learning architecture that pools data from various hospital sources and learns continuously are undoubtable: it supports medical decision-making in situations of public emergency, where data about the nature of the condition and patient outcomes are abundant but dispersed and evolving. However, our scenario presents some critical challenges. The X-ray images that the medical device software pools from various hospitals differ considerably in quality, being taken on different machines, which reduces image comparability. The likelihood of this situation occurring is supported by a recent study that pointed out "algorithm performance may degrade when applied to images generated by equipment from a different manufacturer or in a different clinical environment than those of the training set" [26:415]. Another study emphasised that "the rapid evolution of continuous learning AI models makes implementation of continuous learning AI quality control measures challenging" [36:11].

Continuing with our scenario, we can easily imagine that, without the awareness of clinical staff, some of the smart X-ray machines used in hospitals to take and send images to the cloud have been compromised, exploiting a security vulnerability that allowed hackers to upload old chest X-rays from patients with already known conditions. As the decision support medical device software is based on a weakly supervised, deep learning model, it soon starts training on these images, slowly worsening its treatment recommendations. The changes are not immediately noticeable to professional staff, who continue to administer treatment informed by the outputs of the AI medical device software. Because staff use their medical discretion, some rely on the software outputs more than others, making it hard to identify which patient conditions have worsened due to what treatment input. It thus takes a while before hospital staff and administrators realize that something is going awry with the software tool. In addition, identifying which X-ray machines have been compromised and are uploading old images, and in which hospitals these are situated, is also proving difficult because only a particular model from a particular manufacturer had the compromise, and this model was distributed in several hospitals contributing to the imaging system, alongside other machines that were not compromised.

While this might seem like a long-term scenario, its fundamental pieces are very much present today. We are currently testing and deploying smart healthcare systems that present a multi-layered, connected, intelligent medical device ecosystem comprising of smart medical sensors, connected devices, cloud computing, and AI-based healthcare technologies [1, 28]. Adaptive algorithms based on deep learning models that are relatively inscrutable are already available on the market [46:16] and are considered for interventions similar to the healthcare crisis presented above [2, 34]. AI tools have already been considered for similar purposes during the Covid-19 pandemic, and can be used to distinguish chest X-rays of Covid-19 patients from other diseases like influenza pneumonia [19]. It is, thus, not impossible to image other scenarios where deploying these technologies can have life threatening consequences, such as aiding with brain mapping in surgery (helping surgeons decide which part of the brain to remove to treat a tumor) [37:544], helping with cancer screening and diagnosis, and even using implantable AI platforms to identify pathological patterns [9, 38].

A fundamental feature of the CIMDs described in this scenario is that their risk profile changes relatively subtly over time, yet the cumulative effects of the security vulnerability in the X-ray machines plus the adaptive learning properties of the imaging analysis algorithmic tool could cause substantial harm to patients, their health, and long-term wellbeing. In our roundtable, the "dynamic risk profile" of connected, intelligent medical devices was identified by both manufacturers and regulators as a critical challenge because it can affect the declared safety classification of the device without the manufacturer's, the healthcare professional's, or the regulator's knowledge. As a regulator explained during the roundtable, these subtle yet meaningful changes can be triggered in equal measure by exploited cybersecurity vulnerabilities or by dynamic algorithmic learning processes, even without continuous connectivity to the hospital network (Regulator, Reg-MedTech roundtable, 2022). This increases the opacity of the interactions in a connected, intelligent medical device ecosystem like the one described above that sits across several physical sites (hospitals), brings several digital technologies and processes together (IoT, AI, cloud computing), and involves both algorithmic and human decisions.

The extent of software vulnerabilities and, consequently, device failure modes in a CIMD ecosystem should not be taken for granted. The ECRI Top Ten Technology Hazards for 2022 report identifies both cybersecurity attacks and poorly performing AI tools in its list: "1. Cybersecurity attacks can disrupt healthcare delivery, impacting patient safety" and "7. AI-based reconstruction can distort images, threatening diagnostic outcomes" [11:2].

A recent study conducted by Cynerio in over 300 hospitals, looking at over 10 million IoT and IoMT devices, highlights that "a whopping 73% of IV pumps have a vulnerability that would jeopardize patient safety, data confidentiality, or service availability if it were to be exploited by an adversary" and that "more than a half of connected medical and other IoT devices in hospitals have a known critical vulnerability", with "a third of bedside healthcare IoT devices, the devices closest to patient care" having "an identified critical risk" [53:3]. Vulnerabilities in cardiac or drug infusion devices are also well known to the security community [3, 10, 24, 25]. Generally, vulnerabilities in hospital or patient-deployed connected medical devices stem from maintaining default passwords or settings, failing to update outdated software, long lifecycle of devices that are in continuous use as part of the healthcare infrastructure, and relatively insecure hospital networks [31:41].

The situation is equally challenging when it comes to AI software in standalone form or integrated into other medical devices or technologies. Recent examples include AIaMDs (with FDA clearance and a CE mark) aimed to detect diabetic retinopathy that performed as accurately as a human specialist during development, yet witnessed decreased performance over time in test clinics in rural Thailand [20]. Unlocked algorithms that are frequently fed new data and learn continuously from it using machine learning or deep learning models are particularly opaque and even the smallest change from the development lab, such as the image quality [43], can have important consequences to patients, from reinforcing harmful biases to affecting their safety and wellbeing [17, 18, 31:44–48, 48]. The consequences of deploying unlocked algorithms are already known in self-diagnosis apps, where they are shown to be worsening in results over time [5].

Lastly, several studies have shown that the number of software-related medical device recalls is increasing substantially, varying in application from decision support systems, to implants, to life-sustaining devices [37]. This raises critical concerns about the extent to which current medical device regulations are fit for purpose when it comes to medical software in standalone form or integrated in other products. Below, we briefly explore the main features of medical device regulations and highlight some of the key challenges that CIMDs raise for them.

3 THE REGULATION OF MEDICAL DEVICE SOFTWARE

To ensure patient safety, medical device regulations largely concern the "placing on the market, making available on the market or putting into service of medical devices for human use" [12:Art 1. para 1]. This generally means that without conformity assessment and certification by an approved body, or regulatory review and/or authorisation (depending on the jurisdiction and the device classification), medical devices cannot be placed on the market [12:Art 5. para 1]. Thus, the regulatory regime for medical devices can be imagined as having four stages: 1) device manufacturers have an obligation to specify the intended purpose of a device, conduct clinical trials and performance assessments, establish risk assessment, and classify their medical device following a low (class I), medium (II), or high (III/ IV) risk-based classification system set in place by the regulator; 2) this information is captured in technical documents

		Significance of Information provided by the MDSW to a healthcare situation related to diagnosis/therapy		
State of Healthcare situation or patient condition		**High** Treat or diagnose *~ IMDRF 5.1.1*	**Medium** Drives clinical management *~ IMDRF 5.1.2*	**Low** Informs clinical management *(everything else)*
	Critical situation or patient condition *~ IMDRF 5.2.1*	**Class III** *Category IV.i*	**Class IIb** *Category III.i*	**Class IIa** *Category II.i*
	Serious situation or patient condition *~ IMDRF 5.2.2*	**Class IIb** *Category III.ii*	**Class IIa** *Category II.ii*	**Class IIa** *Category I.ii*
	Non-serious situation or patient condition *(everything else)*	**Class IIa** *Category II.iii*	**Class IIa** *Category I.iii*	**Class IIa** *Category I.i*

Table 1: Classification Guidance on Rule 11

Figure 1: Example of IMDRF's SaMDs Risk Categorization Interfacing with Regulation (EU) 2017/745 – Medical Device Regulation [54:26]. A similar approach can be seen in the US [16:5].

that are put forward to the regulator or other approved bodies for certification, review, or market approval (market approval is needed for devices classed as high or highest risk); 3) once the device can be placed on the market, the manufacturer has a responsibility to conduct post-market surveillance to ensure the device continues to be safe and is performing according to its intended purpose, as stated in its original documentation; 4) manufacturers are required to maintain records and report the results of their post-market monitoring and surveillance to the regulator on a regular basis. If a device becomes faulty or new risks to patient safety are discovered or disclosed, the regulator can recall the device from the market. While jurisdictional differences exist, the foundational principles of medical device regulations are largely aligned internationally and voluntary bodies such as the International Medical Device Regulators Forum (IMDRF) provide an opportunity for rule harmonization in this space.

Initially, when software become recognized in the medical device regulation, it was classed as low risk due to its largely assistive role in clinical management. As software become more complex and increasingly used in treatment and diagnosis, regulators exhibited concerns about its impact on patient safety and its performance over time. In 2014, the IMDRF put forward a risk categorization framework for SaMDs, which interfaces with national regulatory frameworks and their device classification systems, as seen in 1 below.

As noted above, products expected to pose less risks to patient safety do not go through a full regulatory market authorisation process. Instead, in jurisdictions such as the US or the EU, an "equivalence" pathway can be used, which allows manufacturers to submit clinical evaluation documentation for their products based on equivalence to an already marketed device [47:15, 52]. In the US, this is known as the 510(k) review pathway.

Growing evidence from the specialist literature shows that medical device software is placed on the market with a predominantly

medium (even low) risk classification and through the equivalence review process, which essentially asks manufacturers to demonstrate that their products are similar to others already approved on the market, before they are cleared. For instance, in their study on medical device software recalls in the US, Ranquillo and Zuckerman note: "Of the 11 software devices reviewed through the 510(k) process, none was tested in clinical trials to determine safety or effectiveness. In their publicly available 510(k) application summaries, the manufacturers stated that they performed nonclinical testing (eg, simulation, performance testing, software validation) to determine substantial equivalence to software devices already on the market" [37:544]. In addition, Vokinger et al note that: "A majority of AI/ML-based medical devices are cleared through the 510(k) pathway. [. . .] The 510(k) clearance can lead to chains of medical devices that claim substantial equivalence to each other, but over the years or even decades, may diverge substantially from the original device" [47:15]. They further note that, in the manufacturer submitted documentation, "only rarely does the device description state whether a medical device contains an AI/ML component", while in practice several devices are entering the market being advertised as "adaptive" algorithms on the manufacturers' websites [47:16–17]. This is confirmed by further studies, which demonstrate that a considerable proportion of FDA-cleared AI/ ML medical device software is not declared as such in official documentation [4] and that, at the moment, there are clear limitation with the evaluation process for approved medical AI devices [50].

We are, thus, witnessing important challenges with the current processes for evaluating, reporting, classifying, reviewing, and approving CIMDs on the market. While the regulation has been trying to catch up with digital technology innovation in this space, we are also seeing that it is not always possible to capture the appropriate risk profile of these devices at the pre-market and review stage. The situation becomes even more complex once the devices are in use, as discussed below.

4 DISCUSSION: IMPACT OF CIMDS ON RISK ASSESSMENT, CLASSIFICATION, AND MANAGEMENT

If we return to our scenario, we see how easy it is for CIMDs to change their risk profile and interact with each other in less transparent ways, increasing vulnerabilities, risks, and potential harm to patients. Below, we focus on what the changing risk profile of CIMDs, once deployed and in use, means for the fit-for-purpose of current medical device regulations.

4.1 Risk assessment and classification challenges

The procedural limitations discussed above are critical in the context of CIMDs, because a large proportion of these devices rely on general purpose technologies. Thus, it is important to note the "user and context-specific nature of AI applications" [35:326] and IoMT systems. As noted by Park et al, "AI solutions for healthcare differ from drugs or [traditional] medical devices in that they are designed to affect human decision-making. The utility of conveyed information is determined by perception, comprehension, and subsequent actions of the user" [35:327].

This implicitly means that the profile of CIMDs is very likely to change based on its use, which often makes it hard for manufacturers to risk assess and classify their devices with complete understanding of how they might change over time and who might affect this change. This is confirmed by a participant in our study, who noted: "The risks of artificial intelligent software are not well understood because a lot of times the algorithms are learning upon themselves without necessarily an oversight right away from an expert or a clinician" (Manufacturer, Interview-004, 2022). Furthermore, a regulator present in our roundtable highlighted that, as more innovations enter this space – such as the use of foundation models [32] and meta learning - the qualification and classification of AI medical devices will become increasingly difficult and maintaining the initial device risk profile is almost impossible (Regulator, Reg-MedTech roundtable, 2022).

Regulators are already thinking about how to address this challenge, but much more is needed in terms of procedural clarity and guidance provision. For instance, in the US, software modifications through the lifecycle of a medical device are already tackled in regulatory guidance, with manufacturers having to submit and obtain FDA clearance of a new premarket notification (510(k)) if their software has changed in a manner that "could significantly affect the safety or effectiveness of the device", or has seen major changes in its intended use [13]. In addition, The FDA is also proposing a new regulatory approach to better capture modifications to AI/ML-based software as a medical device throughout their lifecycle [15]. Through a "Predetermined Change Control Plan" in premarket submissions and a continuous algorithmic performance monitoring plan, the FDA is hoping to address the changing risk profile of AI in and as a medical device. Regulatory plans to mitigate the risks associated with adaptive algorithms in the medical device sector are also considered in the UK [30]. While these initiatives are still under development, several challenges remain in this space. A manufacturer at our roundtable noted: "There is lack of regulatory guidance on what constitutes a major change (i.e. change in input, architecture) and how to keep up with the pace of innovation" (Manufacturer, Reg-MedTech roundtable, 2022). Another stressed the lack of clarity for when regulatory documents need to be updated if software updates are performed ((Manufacturer, Reg-MedTech roundtable, 2022). A start-up developer added: "We have a physical device and then software device which goes on top of that. Issue: how do you rate the classification of the first device, when you have lots of different software on it?" (Software developer, Reg-MedTech roundtable, 2022).

4.2 Risk management and post-market surveillance

Due to the critical nature of medical devices, regulatory compliance in this field is complex and, as noted in Section 3, extends beyond review and market approval, to post-market surveillance of the product's safety and performance through established quality and risk management processes. This is particularly challenging for CIMDs. In our study, several roundtable participants reflected on issues with measuring CIMDs performance and identifying when things go wrong, especially for AIaMDs. This opacity was revealed in our scenario too (Section 2), where it is not clear whether the AI

imaging analysis software was performing poorly because it was fed images of different quality from different machines, old images from the hacked X-ray devices, or both. This makes the issue of the evidence requirement for risk assessment of post-market device performance a particularly thorny one. A start-up developer in our study noted that even if you plot a pathway for the algorithm, it is often difficult, if not impossible, to map out what your code is doing at all times, especially if the algorithm is adaptive (Software developer, Reg-MedTech roundtable, 2022).

Currently, regulators are proposing increased post-market device surveillance as a potential solution for dealing with the changing nature of CIMDs in use. For instance, in its AI/ML-based Software as a Medical Device Action Plan discussed above, the US regulator is proposing an "Algorithm Change Protocol" based on "transparency and real-world performance monitoring by manufacturers that could enable FDA and manufacturers to evaluate and monitor a software product from its premarket development through post-market performance" [15:1]. This process is intended to improve existing quality management systems used in device surveillance and to address the challenges of conducting postmarket clinical follow-ups on adaptive algorithms [18:3–5]. This approach would entail that a substantial software modification could lead to the AI medical device undergoing a new conformity assessment or market review (depending on the jurisdiction). Yet, adaptive algorithms are highly dynamic, and the change might be drastic but opaque and temporary, hence difficult to spot, measure, and evaluate. Another consideration is that the kind of continuous post-market surveillance needed for adaptive algorithms will inevitably favour established manufacturers with the resources to do so, rather than start-ups. A participant at our roundtable noted that, in order to address some of these challenges, "we need to define performance criteria for AI, depending on medical device characteristics. Also, we need methods for evaluation of performance criteria. This is where standards would also come in to evaluate the performance and safety of AI devices" (Product assessor, Reg-MedTech roundtable, 2022).

4.3 Detecting, categorizing, and reporting compromises

A final consideration in our analysis is the difficulty of detecting, categorizing, and reporting device compromises if exposed to a cybersecurity incident or failure mode in the case of AI medical devices. This goes back to the user and context-specific nature of CIMDs and, subsequently, their contextual risk profile. In our scenario (Section 2), we demonstrate that CIMD ecosystems can become incredibly complex, with marginal changes over time resulting in disruptive systemic risks with serious consequences to patient safety and wellbeing.

In our scenario, the human-machine relationship at the point of care is critical, yet highly opaque. What triggered the suboptimal outcome we saw in the scenario? Is it the fact that the compromised smart X-ray equipment was not updated or had other security vulnerability that should have been addressed by the manufacturer and/or the hospital staff? Is it that the algorithm was not trained on different data quality in development (X-rays taken by different devices) and, once exposed to that, started deviating from its expected performance? Or was it that the algorithm itself performed absolutely fine, but by deliberately being fed inappropriate data (chest X-rays from previously known illnesses) via the compromised machines, it ultimately produced the outcomes it did? How can healthcare professionals and manufacturers be aware of these dynamics?

The critical aspect of the "human factor" or "human oversight" was addressed in our roundtable, with a participant asking: "Is it reasonable and realistic to expect clinicians to remember all the questions AI can ask? Do we end up building software that even experts can't understand the outputs of?" (Standards-maker, Reg-MedTech roundtable, 2022). Another participant noted that the question around explainability of AI medical devices is a recurring one with clinicians, who make the main customer base for advanced medical device software used, for instance, in diagnostics. They highlighted that, when questioned why an algorithm produced the output it did, which the clinician considered false, they went back and created a heatmap technology to show what the algorithm was "seeing", providing further explainability for their tool (Software developer, Reg-MedTech roundtable, 2022). Such measures can help healthcare professionals better understand why CIMDs are behaving in the way that they are and how to distinguish between expected and abnormal device performance. However, as uncovered in our study so far, the awareness and preparedness of medical professionals dealing with CIMDs has to be substantially strengthened [31]. Thus, it is vital for qualified users to have information about recent updates and current performance metrics. Testing AI-based medical devices "performance is a necessary and important task, primarily to control a sufficient level of efficiency and minimize false results that may affect medical decisions" [51:1974].

5 CONCLUSION

In this article, we investigated some of the risk assessment, classification, and management challenges emerging from the growing deployment of connected, intelligent medical devices in healthcare settings. CIMDs come in a variety of forms, from implantables, to specialist IoMT devices deployed in hospitals, to AI-based medical devices and AI as a medical device (AIaMDs). These devices raise several cybersecurity, data management, and algorithmic integrity concerns and, when interacting with each other in complex CIMD ecosystems, can have serious effects on patient safety and on the delivery of reliable, responsible healthcare.

Our analysis focused on the changing risk profile of CIMDs and the implications of these often subtle yet meaningful software modifications for current medical device regulations and for critical stakeholders such as device manufacturers, software developers, and healthcare professionals. The discussion points above show that much more regulatory guidance and standards need to be provided so that critical stakeholders can confidently develop, place on the market, and use CIMDs, especially those based on adaptive algorithmic learning. This point has been highlighted in our research study, when several roundtable participants noted that "the state of the art in this field is not yet settled" and changes need to occur to the way regulatory obligations and expectations are set (Device

Manufacturer, Reg-MedTech roundtable, 2022). Stakeholders consulted throughout our research highlighted some immediate steps that can be taken to address these challenges: providing clear guidance and standards for the evaluation, classification, and software modification of adaptive AI-based medical devices and AIaMDs; the importance of developing continuous performance monitoring of devices in the post-market surveillance stage; and the need to provide further explainability of device performance and safety parameters for medical staff and healthcare professionals.

ACKNOWLEDGMENTS

This research is funded by the PETRAS National Centre of Excellence in IoT Systems Cybersecurity (EPSRC grant number EP/S035362/1). The authors would like to express their appreciation to all stakeholders who took part in our primary research and provided us with a wealth of information on this topic. Special thanks go to our project partners at BSI for their guidance and support throughout this project, especially to Rob Turpin, Paul Sim, Emma Glass, and Matthew Chiles.

REFERENCES

[1] Imran Ahmed, Gwanggil Jeon, and Abdellah Chehri. 2022. An IoT-enabled smart health care system for screening of COVID-19 with multi layers features fusion and selection. Computing (January 2022). DOI:https://doi.org/10.1007/s00607-021-00992-0

[2] Ken Asada, Masaaki Komatsu, Ryo Shimoyama, Ken Takasawa, Norio Shinkai, Akira Sakai, Amina Bolatkan, Masayoshi Yamada, Satoshi Takahashi, Hidenori Machino, Kazuma Kobayashi, Syuzo Kaneko, and Ryuji Hamamoto. 2021. Application of Artificial Intelligence in COVID-19 Diagnosis and Therapeutics. Journal of Personalized Medicine 11, 9 (September 2021), 886. DOI:https://doi.org/10.3390/jpm11090886

[3] Adrian Baranchuk, Bryce Alexander, Debra Campbell, Sohaib Haseeb, Damian Redfearn, Chris Simpson, and Ben Glover. 2018. Pacemaker Cybersecurity. Circulation 138, 12 (September 2018), 1272–1273. DOI:https://doi.org/10.1161/CIRCULATIONAHA.118.035261

[4] Stan Benjamens, Pranavsingh Dhunnoo, and Bertalan Meskó. 2020. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. npj Digit. Med. 3, 1 (September 2020), 1–8. DOI:https://doi.org/10.1038/s41746-020-00324-0

[5] A. Ćirković. 2020. Evaluation of four artificial intelligence-assisted self-diagnosis apps on three diagnoses: Two-year follow-up study. Journal of Medical Internet Research 22, 12 (2020). DOI:https://doi.org/10.2196/18097

[6] I. Glenn Cohen, Timo Minssen, W. Nicholson Price II, Christopher Robertson, and Carmel Shachar (Eds.). 2022. The Future of Medical Device Regulation: Innovation and Protection. Cambridge University Press, Cambridge. DOI:https://doi.org/10.1017/9781108975452

[7] Michael Da Silva, Colleen M. Flood, Anna Goldenberg, and Devin Singh. 2022. Regulating the Safety of Health-Related Artificial Intelligence. Healthc Policy 17, 4 (May 2022), 63–77. DOI:https://doi.org/10.12927/hcpol.2022.26824

[8] Stephane Doyen and Nicholas B. Dadario. 2022. 12 Plagues of AI in Healthcare: A Practical Guide to Current Issues With Using Machine Learning in a Medical Context. Front. Digit. Health 4, (May 2022), 765406. DOI:https://doi.org/10.3389/fdgth.2022.765406

[9] Tu Dresden. 2021. Implantable AI System Developed for Early Detection and Treatment of Illnesses. SciTechDaily. Retrieved September 1, 2022 from https://scitechdaily.com/implantable-ai-system-developed-for-early-detection-and-treatment-of-illnesses/

[10] Chuek Easttom and Nagi Mei. 2019. Mitigating Implanted Medical Device Cybersecurity Risks. In 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 0145–0148. DOI:https://doi.org/10.1109/UEMCON47517.2019.8992922

[11] ECRI. 2022. Top 10 Health Technology Hazards for 2022. Retrieved August 29, 2022 from https://www.ecri.org.uk/wp-content/uploads/2022/05/ECRI_2022_Top_10_Hazards_Executive_Brief.pdf

[12] European Commission. 2017. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. (2017). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745

[13] FDA. 2020. Deciding When to Submit a 510(k) for a Software Change to an Existing Device. U.S. Food and Drug Administration. Retrieved September 2, 2022 from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/deciding-when-submit-510k-software-change-existing-device

[14] FDA. 2020. Software as a Medical Device (SaMD). UCl Food and Drug Administration. Retrieved September 2, 2022 from https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd

[15] FDA. 2021. Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan. US Food and Drug Administration. Retrieved September 1, 2022 from https://www.fda.gov/media/145022/download

[16] FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback. US Food and Drug Administration. Retrieved September 1, 2022 from https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf

[17] Paul Festor, Yan Jia, Anthony C. Gordon, A. Aldo Faisal, Ibrahim Habli, and Matthieu Komorowski. 2022. Assuring the safety of AI-based clinical decision support systems: a case study of the AI Clinician for sepsis treatment. BMJ Health Care Inform 29, 1 (July 2022), e100549. DOI:https://doi.org/10.1136/bmjhci-2022-100549

[18] Stephen Gilbert, this link will open in a new window Link to external site, Matthew Fenech, this link will open in a new window Link to external site, Martin Hirsch, this link will open in a new window Link to external site, Shubhanan Upadhyay, this link will open in a new window Link to external site, Andrea Biasiucci, this link will open in a new window Link to external site, Johannes Starlinger, and this link will open in a new window Link to external site. 2021. Algorithm Change Protocols in the Regulation of Adaptive Machine Learning–Based Medical Devices. Journal of Medical Internet Research (October 2021), e30545. DOI:https://doi.org/10.2196/30545

[19] Ilana Harrus and Jessica Wyndham. 2021. Artificial intelligence and COVID-19: applications and impact assessment. AAAS AI Report. Retrieved October 26, 2022 from https://www.aaas.org/sites/default/files/2021-05/AIandCOVID19_2021_FINAL.pdf

[20] Will Heaven. 2020. Google's medical AI was super accurate in a lab. Real life was a different story. MIT Technology Review. Retrieved September 1, 2022 from https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease/

[21] IMDRF (International Medical Device Regulators Forum). 2013. Software as a Medical Device (SaMD): Key definitions. Retrieved from https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf

[22] Kourosh Kakhi, Roohallah Alizadehsani, H. M. Dipu Kabir, Abbas Khosravi, Saeid Nahavandi, and U. Rajendra Acharya. 2022. The internet of medical things and artificial intelligence: trends, challenges, and opportunities. Biocybernetics and Biomedical Engineering 42, 3 (July 2022), 749–771. DOI:https://doi.org/10.1016/j.bbe.2022.05.008

[23] G Kamberelis and G Dimitriadis. 2018. Focus Groups: Strategic Articulations of Pedagogy, Politics, and Inquiry. In The SAGE handbook of qualitative research / edited by Norman K. Denzin, Yvonna S. Lincoln. (Fifth edition.). SAGE, Thousand Oaks, California, 887–902.

[24] Aditya Kapoor, Amit Vora, and Rakesh Yadav. 2019. Cardiac devices and cyber attacks: How far are they real? How to overcome? Indian Heart Journal 71, 6 (November 2019), 427–430. DOI:https://doi.org/10.1016/j.ihj.2020.02.001

[25] David Klonoff and Julia Han. 2019. The First Recall of a Diabetes Device Because of Cybersecurity Risks. J Diabetes Sci Technol 13, 5 (September 2019), 817–820. DOI:https://doi.org/10.1177/1932296819865655

[26] David B. Larson, Hugh Harvey, Daniel L. Rubin, Neville Irani, Justin R. Tse, and Curtis P. Langlotz. 2021. Regulatory Frameworks for Development and Evaluation of Artificial Intelligence-Based Diagnostic Imaging Algorithms: Summary and Recommendations. J Am Coll Radiol 18, 3 Pt A (March 2021), 413–424. DOI:https://doi.org/10.1016/j.jacr.2020.09.060

[27] Daniela Luzi and Fabrizio Pecoraro. Medical Device Software: A new Challenge. 6.

[28] Pandiaraj Manickam, Siva Ananth Mariappan, Sindhu Monica Murugesan, Shekhar Hansda, Ajeet Kaushik, Ravikumar Shinde, and S. P. Thipperudraswamy. 2022. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. Biosensors 12, 8 (August 2022), 562. DOI:https://doi.org/10.3390/bios12080562

[29] Jon B. Martinez. 2018. Medical Device Security in the IoT Age. In 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 128–134. DOI:https://doi.org/10.1109/UEMCON.2018.8796531

[30] MHRA. 2021. Software and AI as a Medical Device Change Programme. Medicines and Healthcare Products Regulatory Agency. Retrieved April 24, 2022 from https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme

[31] Andrew Mkwashi and Irina Brass. 2022. The Future of Medical Device Regulation and Standards: Dealing with Critical Challenges for Connected, Intelligent Medical Devices. PETRAS National Centre of Excellent in IoT Systems Cybersecurity. Retrieved from https://petras-iot.org/wp-content/uploads/2021/06/White-Paper-The-Future-of-Medical-Device-Regulation-and-Standards.pdf

[32] Mike Murphy. 2021. What are foundation models? IBM Research Blog. Retrieved September 2, 2022 from https://research.ibm.com/blog/what-are-foundation-models

[33] Emilia Niemiec. 2022. Will the EU Medical Device Regulation help to improve the safety and performance of medical AI devices? DIGITAL HEALTH 8, (January 2022), 20552076221089080. DOI:https://doi.org/10.1177/20552076221089079

[34] Roseline Oluwaseun Ogundokun, Sanjay Misra, Mychal Douglas, Robertas Damaševičius, and Rytis Maskeliūnas. 2022. Medical Internet-of-Things Based Breast Cancer Diagnosis Using Hyperparameter-Optimized Neural Networks. Future Internet 14, 5 (May 2022), 153. DOI:https://doi.org/10.3390/fi14050153

[35] Yoonyoung Park, Gretchen Purcell Jackson, Morgan A Foreman, Daniel Gruen, Jianying Hu, and Amar K Das. 2020. Evaluating artificial intelligence in medicine: phases of clinical research. JAMIA Open 3, 3 (September 2020), 326–331. DOI:https://doi.org/10.1093/jamiaopen/ooaa033

[36] Oleg S. Pianykh, Georg Langs, Marc Dewey, Dieter R. Enzmann, Christian J. Herold, Stefan O. Schoenberg, and James A. Brink. 2020. Continuous Learning AI in Radiology: Implementation Principles and Early Applications. Radiology 297, 1 (October 2020), 6–14. DOI:https://doi.org/10.1148/radiol.2020200038

[37] Jay G. Ronquillo and Diana M. Zuckerman. 2017. Software-Related Recalls of Health Information Technology and Other Medical Devices: Implications for FDA Regulation of Digital Health. Milbank Quarterly 95, 3 (September 2017), 535–553. DOI:https://doi.org/10.1111/1468-0009.12278

[38] Jeff Rowe. 2021. How AI "implantables" might monitor your data from the inside. AI Powered Healthcare | Healthcare IT News. Retrieved September 1, 2022 from https://www.healthcareitnews.com/ai-powered-healthcare/how-ai-implantables-might-monitor-your-data-inside

[39] Michael Da Silva, Colleen M. Flood, and Anna Goldenberg and Devin Singh. 2022. Regulating the Safety of Health-Related Artificial Intelligence. Healthcare Policy 17, 4 (May 2022). Retrieved August 18, 2022 from https://www.longwoods.com/content/26824/healthcare-policy/regulating-the-safety-of-health-related-artificial-intelligence

[40] J. Song, D. Lyu, Z. Zhang, Z. Wang, T. Zhang, and L. Ma. 2022. When Cyber-Physical Systems Meet AI: A Benchmark, an Evaluation, and a Way Forward. 343–352. DOI:https://doi.org/10.1109/ICSE-SEIP55303.2022.9794128

[41] Aliya Tabasum, Zeineb Safi, Wadha AlKhater, and Abdullatif Shikfa. 2018. Cybersecurity Issues in Implanted Medical Devices. In 2018 International Conference on Computer and Applications (ICCA), 1–9. DOI:https://doi.org/10.1109/COMAPP.2018.8460454

[42] Karen Taylor, Amen Sanghera, Mark Steedman, and Matthew Thaxter. 2018. Medtech and the Internet of Medical Things: How connected medical devices are transforming health care. Deloitte UK Centre for Health Solutions. Retrieved August 31, 2022 from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-lshc-medtech-iomt-brochure.pdf

[43] The Medical Futurist. 2022. Locked And Adaptive Algorithms In Healthcare: Differences, Importance And Regulatory Hurdles. Retrieved August 25, 2022 from https://medicalfuturist.com/locked-and-adaptive-algorithms-in-healthcare-differences-importance-and-regulatory-hurdles/

[44] Nicole M. Thomasian and Eli Y. Adashi. 2021. Cybersecurity in the Internet of Medical Things. Health Policy and Technology 10, 3 (September 2021), 100549. DOI:https://doi.org/10.1016/j.hlpt.2021.100549

[45] Rob Turpin, Emily Hoefer, Joe Leweling, and Pat Baird. 2020. Machine Learning AI in Medical Devices: Adapting Regulatory Frameworks and Standards to Ensure safety and Performance [White Paper]. Association for the Advancement of Medical Instrumentation (AAMI), British Standards Institution (BSI). Retrieved from https://www.bsigroup.com/en-US/medical-devices/resources/Whitepapers-and-articles/machine-learning-ai-in-medical-devices/

[46] Kerstin N Vokinger, Thomas J Hwang, and Aaron S Kesselheim. 2022. Lifecycle Regulation and Evaluation of Artificial Intelligence and Machine Learning-Based Medical Devices. In The Future of Medical Device Regulation: Innovation and Protection. ed./I. Glenn Cohen; Timo Minssen; W. Nicholson Price II.; Christopher Robertson; Carmel Shachar. Cambridge University Press. DOI:https://doi.org/10.1017/9781108975452

[47] Kerstin N. Vokinger, Thomas J. Hwang, and Aaron S. Kesselheim. 2022. Lifecycle Regulation and Evaluation of Artificial Intelligence and Machine Learning-Based Medical Devices. In The Future of Medical Device Regulation: Innovation and Protection, Carmel Shachar, Christopher Robertson, I. Glenn Cohen, Timo Minssen and W. Nicholson Price II (eds.). Cambridge University Press, Cambridge, 13–21. DOI:https://doi.org/10.1017/9781108975452.002

[48] D. Wilhelm, R. Hartwig, S. McLennan, S. Arnold, P. Mildner, H. Feußner, T. Neumuth, and R. Bieck. 2022. Ethical, legal and social implications in the use of artificial intelligence-based technologies in surgery: Principles, implementation and importance for the user. Chirurg 93, 3 (2022), 223–233. DOI:https://doi.org/10.1007/s00104-022-01574-2

[49] Beau Woods, Andrea Coravos, and Joshua David Corman. 2019. The Case for a Hippocratic Oath for Connected Medical Devices: Viewpoint. Journal of Medical Internet Research 21, 3 (March 2019), e12568. DOI:https://doi.org/10.2196/12568

[50] Eric Wu, Kevin Wu, Roxana Daneshjou, David Ouyang, Daniel E. Ho, and James Zou. 2021. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med 27, 4 (April 2021), 582–584. DOI:https://doi.org/10.1038/s41591-021-01312-x

[51] Victoria Zinchenko, Sergey Chetverikov, Ekaterina Akhmad, Kirill Arzamasov, Anton Vladzymyrskyy, Anna Andreychenko, and Sergey Morozov. 2022. Changes in software as a medical device based on artificial intelligence technologies. Int J CARS 17, 10 (October 2022), 1969–1977. DOI:https://doi.org/10.1007/s11548-022-02669-1

[52] 2020. MDCG 2020-5 Clinical Evaluation - Equivalence. A guide for manufacturers and notified bodies. European Commission. Retrieved September 2, 2022 from https://ec.europa.eu/docsroom/documents/40903

[53] 2022. The State of Healthcare IoT Device Security. Cynerio. Retrieved July 13, 2022 from https://uploads-ssl.webflow.com/5d2ad783e06f4c19469d363a/61e70fd9286e1d6d68a86ba8_A%20Cynerio%20Report%20-%20The%20State%20of%20IoMT%20Device%20Security%202022.pdf?__hstc=40021757.f737178f25dbf3824b1345cefb284f49.1656863593379.1656863593379.1657728666011.2&__hssc=40021757.1.1657728666011&__hsfp=2600066663

[54] Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR. Retrieved August 31, 2022 from https://ec.europa.eu/docsroom/documents/37581

[55] Kioskli *et al.* - 2021 - The landscape of cybersecurity vulnerabilities and.pdf. Retrieved August 29, 2022 from https://dl-acm-org.libproxy.ucl.ac.uk/doi/pdf/10.1145/3465481.3470033

[56] Regulatory and Standardization Challenges for Connected and Intelligent Medical Devices (REG-MEDTECH). The PETRAS National Centre of Excellence for IoT Systems Cybersecurity. Retrieved September 1, 2022 from https://petras-iot.org/project/regulatory-and-standardization-challenges-for-connected-and-intelligent-medical-devices-reg-medtech/