Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P5204-cd

Safety and security co-design with application to medical device industry

Roberto Filippini
PhD, Ind. cons. medical device safety and regulatory affairs, Italy. E-mail: filippini.consulting@gmail.com

John Eidar Simensen
Dep. Risk and Security, Institute for Energy Technology, Norway. E-mail: john.eidar.simensen@ife.no

Engineering is experiencing rapid changes in response to new needs, from embedding "intelligence on board" for communication, control and decisions, up to large-scale architecture such as systems of systems and Internet of Things. These systems rely on high degree of autonomy together with complex functional dependencies, and are exposed to endogenous and exogenous risks which cannot be understood and mitigated as a sum of their parts. Among these risks, those related to security are becoming a serious concern for safety and operability.
The healthcare sector recognizes the importance of integrating security and safety in the life cycle of a medical device. This is reflected in the MDR 2017/745 medical device regulations, which require manufacturers to address security in the life cycle of medical devices that incorporate software, or software that are medical devices in themselves. This paper discusses principles of safety and security for medical devices according to the state of the practice, and exemplifies their pros and cons with the intent of converging in a new state-of-the-art. This eventually becomes the "game changer" in favor of co-design. Design challenges and expected benefits are discussed based on security and risk management expertise from information technology and operational technology.

Keywords: Safety, Security, Co-engineering, Risk management

1. Introduction

For safety-critical digital systems, the requirement to consider security and cyber security in addition to safety is recent. Digital Instrumentation and Control (DI&C) systems, and systems comprising operational technology (OT), such as Programmable Logic Controllers (PLC) and sensors, were traditionally treated by industry as local systems, rarely connected to other systems and more rarely to the internet. In pursuit of effectiveness and profit, industry has connected industrial systems to business side systems, and indirectly or directly to the Internet. Existing systems, originally developed for safety, suddenly need to address security threats. without impeding the operation and profitability of the system. This has resulted in a practice in which security is added or appended to existing solutions rather than being implemented from the ground up. The notion that safety and security can have different cultures, goals, standards, and practices Simensen and Gran (2021), should forewarn the manufacturer that the problem is complex.

As many other sectors, which implemented the latest ICT innovations, the healthcare sector was not ready to face security and implement the necessary controls. Healthcare service disruptions can severely affect the society and because of that they have become one of the preferred targets of cyber-attacks Williams and Woodward (2015) ENISA (2024).

The goal of this paper is to review the design and manufacturing of a medical device, starting from the regulatory framework. A case study is taken as proof of concepts to identify the conceptual gaps of the state of practice against advanced security threats, in order to show the urgency of safety and security codesign for medical devices.

The paper consists of six sections. Section 2 introduces the regulatory and technical aspects of the design and manufacture of safe and secure medical devices. Section 3 identifies and analyzes security challenges in the design process. The conceptual gaps of the state of practice are exemplified in the case study of Section 4. Section 5 discusses requirements of co-design of safety and security, before conclusions in Section 6.

2. Medical device safety and security

The category of medical devices is very diverse in terms of applications and technologies. In this section, we focus on modern medical devices that are safety critical and depend on software and Information and Communication Technologies (ICT). Examples are medical devices that administer treatment (e.g. insulin pumps, radiotherapy systems), medical devices used for recovery or for sustaining life (e.g. defibrillators, ventilators), medical devices that support vital functions (e.g. pace makers), monitoring and diagnostic systems. These systems are critical to safety because their failure or misuse can cause serious harm to patients. And because failures may be caused by cyber-attacks, they are also security critical.

2.1. *The regulatory framework*

Medical Device Regulation MDR 2017/745 Council of the European Union (2025) establishes norms and requirements for medical devices to ensure their safe use and protection of patients, users and other persons' health. In particular, Annex I defines the General Performance and Safety Requirements for the design and manufacturing of the medical device and also includes clauses that cope with operations security (clauses 14.1, 14.2 and 17.1), information security (clause 17.2) and IT security (clause 17.4). The verification of the clauses in Annex I is by complying with applicable safety and security standards, which overall define the regulatory framework; see Figure 1.

Annex I of MDR Council of the European Union (2025) constitutes the top level of the regulatory framework and points to IEC 60601-1 NEK/NK62 (2024), which covers the basic safety and essential performance of programmable electrical medical systems. The intermediate level of the framework includes the IEC 62304 ISO/TC 210 (2006) for the medical software life cycle and the IEC 80001-1 *Health software and health IT systems safety, effectiveness and security* for the integration of medical devices in IT networks. At the last level, the relevant standards are IEC TR 60601-5-4, derived from IEC 62443 ISA99 and IEC TC 65 WG 10 (2024), IEC 81001-5-1 for the security life cycle of medical device software, and ISO 27001 for the IT operating environment. The ISO 14971 standard *Medical devices – Application of risk management to medical devices* applies horizontally to each level. It is worth remarking that security requirements are addressed in the lower levels, and that verification is required for meeting safety requirements at the top level.

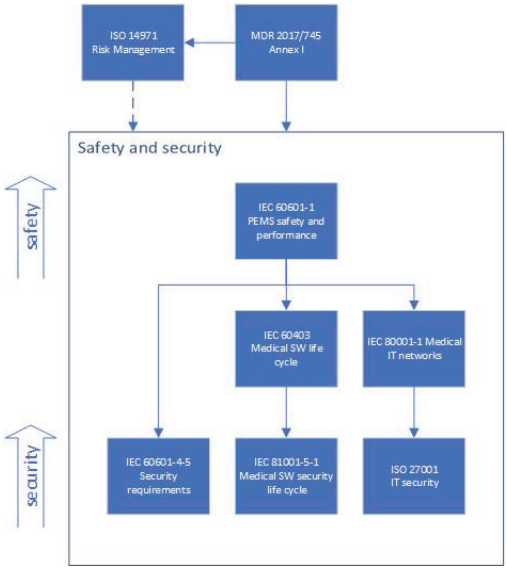

Fig. 1. The regulatory framework for security and safety of medical devices

2.2. *State of the practice*

EU member states have appointed the Medical Device Coordination Group, who has prepared the MDCG guidelines Medical Device Coordination Group (2019) with the objective of instructing the manufacturer to address security as an alternative to the implementation of security standards, in particular IEC 81001-5-1 *Health software and health IT systems safety, effectiveness and security*, and IEC 60601-5-4 NEK/NK62 (2024), which are not fully harmonized with EU laws. With similar intention, TUEV and Johner Institute have together prepared security guidelines to obtain conformity step-by-step M. Klinger (2021). Security requirements are arranged in four levels, from mandatory (level 0) to optional (level 3). These guidelines are not only for interpreting the

most controversial aspects of security. They also implicitly define the "state of the practice", by making it easier for a manufacturer to obtain certification and re-certification of medical devices. Having said that, the state of the practice is not state-of-the-art.

Meeting conformity by means of guidelines (instead of fulfilling all relevant standards) is an attractive short-cut to reach approval, however it is an approach that does not mandate additional efforts to secure the medical device. This also enforces the inappropriate habit of considering security at a later stage of the design. When applying the state of practice, it is essential to consider pros and cons and to look at the present situation as transitory.

3. Challenges with security for medical devices

The introduction of security in the medical device life cycle has increased the already significant efforts for medical device certification, which is comparable to and in certain cases higher than the effort to make a system safe. This situation is reflected in the definition of cyber-security as *a bull in a china shop* Filippini and Spiller (2024). In this respect, if the state-of-the-practice aims at simplifying the conformity assessment for the manufacturer, on the other hand, one may also wonder how much this is paid in terms of quality of designing and manufacturing a safe and secure medical device. In the following section, an example of safety and security assessment is chosen to revise the state of the practice and its potential gaps, omissions, and inaccuracy.

3.1. *An example of safety and security assurance*

The example takes clause 17.2 of Annex I MDR 2017/745 Council of the European Union (2025) and derives suitable claims and sub-claims to prove their verification in spite of the conceptual gaps in the state of the practice. Clause 17.2 says that *for devices that incorporate software or for software that are devices in themselves, the software shall be developed and manufactured in accordance with the state of the art taking into account the principles of development life cycle, risk management, including information security, verification, and validation*. The clause is structured as a safety and security assurance case, with claims for production and post-production phases.

Production claim: The manufacturer shall be able (1) to design safe and secure medical devices on the basis of (2) identified threats and vulnerabilities, and (3) to assess related risks with their mitigation.

Post-production claim: The manufacturer shall be able (4) to maintain safe and secure medical device along its life time against changes in the product and in the surrounding environment, and (5) until this becomes impossible or not economically convenient.

Each claim consists of sub-claims and assumptions, which are labeled by numbers.

Sub-claim (1) assumes the know-how of the manufacturer to deal with safety and security by design of medical devices. The know-how constitutes state of practice that addresses the design of two attributes separately. Because of this, a comprehensive definition of safety and security critical medical devices does not exist and this is a conceptual gap that affects several design principles. The single fault safety principle in the IEC 60601-1, i.e., that the medical device remains free of unacceptable risk under single fault condition, turns out to be challenged when facing cyber-attacks. The same holds for the segregation principles of safety critical and non-safety critical software in IEC 62304 *Medical device software – software life cycle processes*. As long as safety and security are separated, the defense-in-depth principle may end up in a complex architecture with the several layers of defense, eventually interfering with each other. These conceptual gaps may result in a suboptimal design with respect to safety and security.

Sub-claim (2) requires the identification of threats and vulnerabilities, both constituent elements of a cyber-attack. Unlike safety risk analysis, for which independent double faults are excluded and not analyzed, security risk analysis does not have such rule of thumb. An example here is where a cyber-attack may elaborate a strat-

egy to exploit more combinations of vulnerabilities and threats. The applicable criteria is based on the exploitability of a cyber-attack. In other words, the border between the analyzed security risks and those that are acknowledged but not analyzed (i.e. the "known unknown") is subtler and subjected to the experience of the security analyst.

The assessment of security risks and their mitigation, namely sub-claims (3), is done separately. While this is a consolidated practice, in the scientific community new analysis methods are proposed with the intent of enriching the description of the hazardous situations that are caused by random failure scenarios and cyber-attacks. IT security measures (e.g. access control, firewall, intrusion detection, etc.) are often considered sufficient to protect the medical device. This is a misinterpretation of the As Far As Possible (AFAP) principle of ISO 14971. As a consequence, a cyber-attacker may gain full control of the medical device after breaching IT security measures.

The sub-claims (4) and (5) apply to post-production. Any changes in the medical product and in the operating environment are expected to trigger reassessment of risks and their mitigation. Changes in the operating environment depend on the surveillance of vulnerabilities and threats, which is a big concern. Due to a fast changing threat landscape, all claims on security are impermanent. In this respect, a continuous surveillance is essential to ensure that the level of safety and security does not decrease. This problem is also related to the obsolescence of software. Examples are legacy software, some of which provide exploitable back-doors for cyber-attackers. Again, the meeting of the requirements depends on the experience of the manufacturer.

3.2. *Threat landscape paradigm shift*

In cybersecurity, reality is that adversary skill is beyond that of the good guys (i.e., us). When factoring in zero-day vulnerabilities, advanced tools supporting hackers, and recent development in artificial intelligence (AI) supporting both skills and tools of the adversaries, its a losing game. AI-tools are already powerful tools in developing advanced cyber-attacks, and they can empower low-skill actors to better understand the steps of attacks and develop their overall knowledge and strategy.

Current practice is to protect assets according to perceived value of the asset against realistic adversaries. The National Institute of Standards and Technology (NIST) provides the following main adversary types: state actors, terrorists, organized crime, hacktivists, and hackers. Where home-schooled hackers use simple tools and tactics to obtain smaller gains in a short time, state-backed actors with access to advanced knowledge and tools are playing the long game. Depending on the relevant adversary capabilities, existing system vulnerabilities are mitigated to reach an acceptable level of risk. Most civilian industries do not consider themselves to be a target of state-actor interest, hence the need to protect against advanced techniques and tactics is not seen as relevant. Mostly, industry protects against hackers and hacktivist activities, in a manner suggested by best-practice standards and regulations, where security controls often can successfully be appended to existing systems, e.g., firewalls, simple encryption and authorization/authentication.

Where previously e.g., a hacktivist would be stopped by firewalls and simple encryption, an AI-empowered hacktivist might not be. If this AI-empowerment is true for different adversary types, it can render current protection strategies obsolete. Safeguarding the borders of a system might no longer be sufficient.

4. A case study

4.1. *System description and analysis*

The case study makes it possible to discuss technical safety and security requirements, according the present state of practice, and although it is based on a real system, it is generalized as to not directly represent any specific implementation. The case study is a medical system that provides remote supervision and control of radiotherapy sessions for eye cancer treatment. Remote supervision and control are required in several clinical applications and in this case prevent exposure of medical personnel to radiation from the particle beams as well as secondary effects of scattered

particles.

A radiotherapy treatment session consists of the delivery of a proton beam of required energy and intensity, which is generated by the particle therapy accelerator. The proton beam is steered and collimated along the beam line towards the target tumor area. In order to maintain the alignment with the beam, the patient shall gaze a LED light during the treatment session. A radiotherapist in the control room (i.e. the user of the medical device) supervises the treatment through a monitor, showing a live image of the patient. The live image is captured by camera and sent to a server which formats the image and sends it to the client in the control room. The radiotherapist makes decisions on the basis of the live image. If the patient moves the eye out of the target area, the alignment with the beam is lost, and the radiotherapist shall stop the beam by pushing the interrupt button. A delayed live image, a frozen image, or a low-quality image will compromise the effectiveness of the supervision activity, resulting in late or missed beam stop. Worst case, surrounding healthy tissues are irradiated and harm to the patient is severe.

The client software displays additional information related to the machine state (i.e., preparatory, ready, irradiation) as well as warning and error messages. Error messages are generated by either the server or client in response to violation of the real-time property, substandard quality of the live image, or from errors in the software, operating environment, and communication channels. The acquisition, framing, transmission and display of the live image and the supervision activity by the user are shown in the swim lane diagram in Figure 2. The server and the client are safety critical processes and they are developed as class C medical device software according to IEC 62304. Several risk control measures are implemented to mitigate risks in the server and client; frozen image detection, live image check (image time stamp and timeouts). In addition user takes care of the image quality supervision.

The medical device software has been analyzed and designed to guarantee a safe treatment with respect to identified failure scenarios and foreseeable misuse. Security was addressed at a later stage and separately from safety. According to the state of practice, STRIDE methodology together with Failure Modes and Effect Analysis assessed risks of cyber-attacks that may compromise integrity and availability of live image. New risk control measures were identified to secure the operating environment such as access control, firewalls, and resource monitoring, as requirements for IT. In conclusion, the medical software was evaluated safe and secure for its intended use under identified failure conditions and foreseeable misuse.

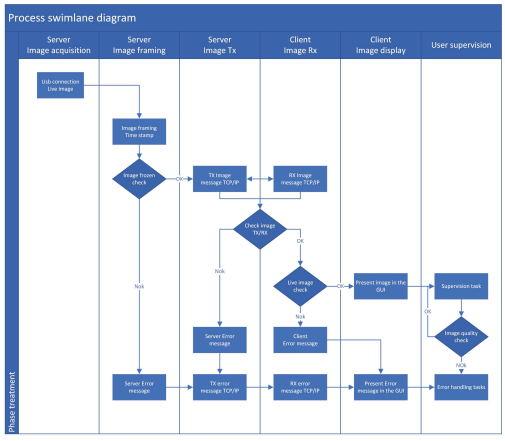

Fig. 2. Swim lane diagram of the case study

4.2. *Use case vulnerability assessment when considering advanced threats*

When assessing the swim-lane diagram for potential attack vulnerabilities that can allow for actions on objectives, the importance of an up-to-date image is critical. There are functions on both server- and client-side checking for potential delays in image acquisition and transmission. The image acquisition check verifies the time stamp to confirm that the image is recent and correct. Image transmission and receive functions ensure that the transmission of the image from server to client does not introduce delays that render the image out-of-date. The 'time stamping' imprints the time off image capture into the picture and it is this image-imprinted information that is checked at several points in the process to ensure that the

picture is sufficiently recent.

When considering cybersecurity risks, structured adversarial thinking can be supported with e.g., the Lockheed Martin cyber kill chain model Eric M. Hutchins and Michael J. Cloppert and Rohan M. Amin (2011). Considering the models steps; Weaponization and Delivery, can each provide an attacker with access. With access, the step Exploitation can consist of acquiring a better understanding of how the system actually works; including identifying further weaknesses, vulnerabilities, the potential to escalate rights and access, and identifying potential delivery mechanisms for payloads, commands, and other actions to reach objectives. For sake of example we identified two potential effects of interest; 1. compromising the time-stamp process in such a way that the system is stopped, and 2. compromise the time-stamp process so that a wrong image is used without being flagged. The first can be achieved in several ways, e.g., denial of service attack on server side that slows down image and denial of service attack on the network to introduce delays in transmission. The second requires more than one of the checking functions to be manipulated in order to not throw an error, or it requires breaking the error checking functions itself. Which of these are more probable depends on the goal of the attacker, if it is to stop the process and prevent treatment, or if it is to harm a patient through administering treatment based on wrong information.

Worst-case scenario is harmful treatment. A potential way to achieve harmful effects can be through a side-channel attack compromising the premise of the time-stamp approach, e.g., tampering with system clocks. To ensure that different systems are aligned time-wise, network time protocols (NTP) are commonly used to adjust system time with a centralized 'time'. Although both server and client sides have internal battery back-up clocks, setting the time and adjusting for potential drifting is done with NTP. Existing time stamp checks do not consider system time manipulation. Without going to much into details, checking for time manipulation can be done in different ways such as checks to ensure that no commands that adjust or change time attributes or time-services are run when the system is operated, and checks that compares NTP time with internal clocks, and/or against other NTP services. The same acceptance that an adversary would be able to pass 'border security' protecting the system, supports accepting the possibility that other necessary services can be compromised as well, which means in the case of NTP that both protection measures should be identified and implemented for how the service is used, as well as checking mechanisms for cases where protection mechanisms have failed. In fact, should the NTP fail for non-security reasons, the checking mechanisms would still be needed. Regardless, both protection and checking mechanisms cannot be achieved without system redesign and implementation.

5. Towards Safety and Security Co-design

The case study is a paradigmatic example of how the implementation of the state of practice can be insufficient to handle more sophisticated security threats. In the use case example a vulnerability was identified that allows an attacker to achieve unwanted behavior of the medical device, compromising safety and challenging the effectiveness of the risk control measures.

In the following we discuss the reasons why this vulnerability remained after design, by root cause analysis. The existence of an unprotected software vulnerability has its causes in a few conceptual gaps with the application of the defense-in-depth principle and the AFAP principle. All safety risk control measures are within the medical device, while security risk control measures are in the surrounding IT operating environment, hence external to the medical device. This leads to an architecture that, in spite of being able to reduce risks, does not respond to the same defense-in-depth strategy. A risk analyst would also recognize a deviation from the AFAP principle. The AFAP applies to the manufacturing of the medical device and not to the IT environment, which is considered external. Therefore, the predominance of security measures in the IT operating environment over security measures in the medical device goes against the correct application of the AFAP. What is the

consequence of that? Simply, an attacker with proficient skill to breach IT security measures will be able to take full control of the medical device, including impacting safety measures. As it cannot be ruled out that an attacker would not be able to get through the 'border' security measures, internal security measures are needed with regards to defense-in-depth.

Why was this potential vulnerability overlooked? Maybe this is not just negligence. One might wonder if this was done to avoid a software design review to implement new security software requirements, which have been identified at a later stage. And why are security requirements introduced at a later stage? This is the consequence of the state of the practice, which allows safety and security to be designed separately, the result of which is a suboptimal engineering process. It seems we have found the root cause.

While best practice suggests addressing safety and security together in early life cycle phases, standards are not explicit in this respect. The integration of safety and security for medical devices is not reflected in a single standard. For example, IEC 62304 ISO/TC 210 (2006) addresses safety but does not include security, while IEC 62443 ISA99 and IEC TC 65 WG 10 (2024) considers security but not safety. Moreover, there is no risk management process that addresses safety and security altogether. Standardization committees have in part acknowledged this problem. Because cybersecurity is ubiquitous, it is expected that these obstacles will be addressed and removed one by one. But until then, instead of proceeding blindfolded, it is more beneficial to acknowledge the gaps, address them, and identify corrective measures.

In this respect a valuable contribution comes from the scientific community. The work of Schoitsch (2005) demonstrates that safety and security cannot be addressed separately, through examples where safety critical systems are accessible from 'outside'. Qi and Sangiovanni-Vincentelli (2018) defines co-design for cyber-physical systems as a multi-objectives optimization problem, in which security and safety are among the design attributes. Concerning risk management, Sango et al. (2019) suggests introducing more activities in between the RM process of safety and cybersecurity. Schmittner et al. (2014) developed new analysis methods combining safety and security by practically extending FMEA.

To summarize the discussion, it can be useful to list what we consider essential features of safety and security co-design of a medical device:

- Introduce the definition of safety and security critical system for medical devices
- Develop defense in depth principles that consider safety and security together
- Revise software segregation principles with respect to safety and security
- Reinforce the compliance with the AFAP principle for security
- Incorporate the threat landscape in the development and design review phases
- Design safety and security measures in order to minimize their interference
- Integrate safety and security in a single risk management process
- Revise the single fault safety concept in light of security threats
- Address safety-security trade-offs during the design phase.

Most of the features are process-related, and only a few are product-related. Process-related features shall be given the priority, because they establish the"co-design principle".

6. Conclusions

Security and safety have been considered as separated domains until the revolution of the internet and more recently the ubiquitous computing for which access from outside to critical systems, or even interaction between critical components or subsystems is part of the intended use Schoitsch (2005) . From this point in time, safety and security have stopped to be independent system attributes. The essence of the problem is how to integrate security and safety in the product life cycle because it is evident that a system cannot be secure and safe by pursuing the two goals independently. The statement of Bloomfield Bloomfield et al.

(2013) *a system is not safe if it is not secure*, pinpoints the core issue.

The healthcare industry needs significant improvements in protecting medical devices from cyber-attacks. Securing communication channels and network schema and adopting secure software practices is largely agreed, while fulfilling these objectives is not easy and is still debated. The compliance with applicable security standards does not seem to provide the solution and the current state of the practice may even enforce the misleading belief about safety and security of medical devices as two design processes that can be addressed separately. This conceptual gap has several drawbacks. For example, it reduces the coverage and effectiveness of the response to emerging security threats and the discovery of vulnerabilities. Security vulnerabilities in particular may undermine safety because they not only compromise confidentiality and integrity of data but also offer the possibility of taking control of the medical device.

This paper has identified the conceptual gaps of the state of the practice, and the reason why the design of safe and secure medical devices in best case is suboptimal. The case study showed that by considering safety and security as separate, realistic security vulnerabilities are not identified, needing additional corrective measures to improve security by design. This set of corrective measures constitutes essential features for improving the state of practice towards co-design of safety and security.

References

Bloomfield, R., K. Netkachova, and K. Stroud (2013). Security-informed safety: If it's not secure, it's not safe. *5th International Workshop on Software Engineering for Resilient Systems*.

Council of the European Union (2025, January). Regulation (eu) 2017/745 of the european parliament and of the council of 5 april 2017 on medical devices. Technical report, Official Journal of the European Union.

ENISA (2024, September). Enisa threat landscape 2024. Technical report, European Union Agency for Cyber Security.

Eric M. Hutchins and Michael J. Cloppert and Rohan M. Amin (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Technical report, Lockheed Martin.

Filippini, R. and S. Spiller (2024). Cybersecurity and medical devices: a bull in a china shop. *EnCyCriS/SVM 2024 workshop - Association for Computing Machinery*.

ISA99 and IEC TC 65 WG 10 (2024). IEC-62443 - Technical standards for requirements and processes for implementing and maintaining electronically secure industrial automation and control systems. Technical report, International Society of Automation.

ISO/TC 210 (2006). IEC 62304:2006 - Medical device software — Software life cycle processes. Technical report, International Society of Automation.

M. Klinger (2021). IT Security Guideline for Medical Devices. `https://github.com/johner-institut/it-security-guideline/blob/master/Guideline-IT-Security_EN.md`.

Medical Device Coordination Group (2019). Guidance on cybersecurity for medical devices. Technical report, European Commission.

NEK/NK62 (2024, January). IEC-60601 - Technical standards for the safety and essential performance of medical electrical equipment. Technical report, International Electrotechnical Commission.

Qi, Z. and A. Sangiovanni-Vincentelli (2018). Codesign methodologies and tools for cyber–physical systems. *Proceedings of the IEEE 106*(9), 1484–1500.

Sango, M., J. Godot, A. Gonzalez, and R. R. Nolasco (2019). Model-based system, safety and security co-engineering method and toolchain for medical devices design. *DMD2019 Design of Medical Devices Conference*, 16–18.

Schmittner, C., T. Gruber, P. Puschner, and E. Schoitsch (2014, 09). Security application of failure mode and effect analysis. In *International Conference on Computer Safety, Reliability, and Security*, Volume 8666, pp. 310–325.

Schoitsch, E. (2005). Design for safety and security of complex embedded systems: A unified approach. *Cyberspace Security and Defense: Research Issues*, 161–174.

Simensen, J. and B. Gran (2021). Information- and cyber-security practices as inhibitors to digital safety. *Proceedings of the 31st European Safety and Reliability Conference*, 1584–1590.

Williams, P. A. H. and A. Woodward (2015). Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Medical Devices*, 305–316.