

1.3.1 – Validation of safety requirements

Practical guidance – healthcare

Author: SAM demonstrator project

The introduction of artificial intelligence (AI) and machine learning (ML) applications into clinical systems can create challenges for traditional design approaches that require clearly defined and precise specifications of the operating environment, operational scenarios and of the resulting safety requirements that bound the behaviour of the AI / ML system. Healthcare is a complex domain, and clinical systems are made up of many different actors and technologies all interacting with one another in ways that can be very dynamic and responsive to the requirements of a specific situation.

One way of both eliciting and validating safety requirements (though not the only one, nor even a standalone approach) is to seek input and feedback from stakeholders (e.g. through simulation or interviews). This approach is particularly appropriate where human-machine interaction and training are concerned.

For example, in the case of the design of an autonomous infusion pump to be used in intensive care, input from stakeholders might produce training requirements as shown below. Feedback from stakeholders on a design prototype can then provide information about the extent to which the requirements have been met.

Clinicians need to maintain core clinical skills	When an autonomous system fails or becomes unavailable, staff need to remain vigilant and be able to take over. They require training and exposure to maintain their clinical skills.
Clinicians need to build a baseline understanding of AI and its limitations	Clinicians will become users as well as supervisors of AI systems. They shall be provided with a baseline understanding of how AI works so that they are able to identify limitations and problems.
Training needs to address over-reliance on AI	Staff might rely too much on AI. They shall receive training in core clinical skills and education about limitations of AI to help address over-reliance.

Similarly, high level safety requirements relating to autonomy and control might be validated through feedback on a proposed interaction design.

Clinicians need to be able to maintain autonomy	Clinicians feel responsible for their patient and want to remain in control. Autonomous systems can challenge this sense of autonomy, and clinicians need to be
--	---

	allowed to remain in charge (e.g. through manual override options).
Feedback and alerts shall provide clinicians with an awareness of what the AI is doing	Feedback and alerts can help to maintain situation awareness and stay in control of the overall treatment and care for the patient. The design shall determine clearly when an alert is raised. The system shall avoid alert fatigue or overload.
Clinicians need to be able to build trust in AI	Clinicians have to trust AI in order to realise its benefits. The interaction design shall include training and feedback. The AI system shall be introduced gradually in low-risk areas over time.