

Assurance Challenges for Artificial Intelligence and Machine Learning in Healthcare



Photo Credit: Dr Nick Reynolds, Royal Derby Hospital

Shakir Laher and Mark Sujan discuss the challenges of assuring health IT systems that use Artificial Intelligence and Machine Learning technologies in delivering patient care.

The rate at which Artificial Intelligence (AI) and Machine Learning (ML) technologies are being developed in healthcare does not show any signs of passing by as just the latest fad. NHSX, recently set up by the government to drive NHS digital transformation and lead policy, published a report entitled "Artificial Intelligence: How to get it right". In it they surveyed 112 developers of AI products, finding 24 were very likely to deploy at scale in a year, with up to 82 in five years [1]. There will be many development challenges to overcome, which becomes the primary focus. However, consideration of potential patient harm resulting from this form of technology and the assurance requirements of the technology have never been more crucial.

"Medication errors present a significant and stubborn cause of patient harm, with as many as 237 million medication errors occurring in the NHS in England every year"

This article focuses on the potential assurance challenges faced by a specific instance of AI/ML technology in the form of delivering insulin autonomously to Intensive Care Unit (ICU) patients. Medication errors present a significant and persistent cause of patient harm, with as many as 237 million medication errors occurring in the NHS in England every year, causing around 700 deaths and contributing to around 1700 patient deaths [2]. Infusion errors are particularly error-prone and high-risk due to their complexity and the acuity of the patient. The findings have been extrapolated from the Safety Assurance of Autonomous Intravenous Medication Management Systems (SAM) project [3], which is part of the wider Assuring Autonomy International Program (AAIP). The project was tasked with identifying stakeholders' safety assurance requirements in the context of an autonomous intravenous medication management systems. In addition, NHS Digital's contribution was to conduct a case study at Royal Derby Hospital (RDH) to develop a preliminary hazard analysis and explore possible safety assurance strategies for different levels of an automated and autonomous intravenous medication management system. It is with this focus the article presents its findings of assurance challenges faced by autonomous healthcare systems.

The Regulatory Space

Technology that is driven by software and used in the health domain is generally classified into 1 of 2 types: Medical Device (MD) or Health Information Technology (HIT). If the technology is defined as a MD, it will be subject to a regulatory framework established under European law. Adhering to the framework establishes an acceptable level of safety (from a legal perspective) of the product to be placed on the market. If the technology is classed as HIT, it falls out of scope of European legislation. The NHS does however provide standards established under the Health and Social Care Act that need to be followed to assure technology that is placed in or accesses NHS IT infrastructure.

NHS Digital's clinical safety team utilises two standards to assure the safety of HIT in the NHS. DCB0129 [4] is used by manufacturers of HIT to evidence the clinical safety of their products. DCB0160 [5] is in place for healthcare organisations who deploy and use the HIT.

Safety Modelling, Assurance and Reporting Toolset (SMART)

Since the introduction of the standards, the clinical safety team have observed that organisations struggle with key concepts around hazard analysis, risk assessment and development of safety cases [6]. There is a variety of reasons for these difficulties, which are covered in the study at [6] and not reproduced here. However, if there were a software tool that guided organisations through the implementation of the standards, the safety cases produced would be of much richer quality. This is where SMART takes centre stage.

SMART is a software tool which has been developed in-house by NHS Digital's clinical safety team. It guides users through entering information that forms a safety argument based on the principles of a goal structuring notation (GSN) argument [7]. This technique allows users to concentrate on the intellectual work involved in the clinical risk management process and removes the burden of generating the documentation as the tool exports the documentation in the form of a clinical safety case report (CSCR).

Welcome to the Safety Modelling, Assurance and Reporting Toolset

SMART METHODOLOGY

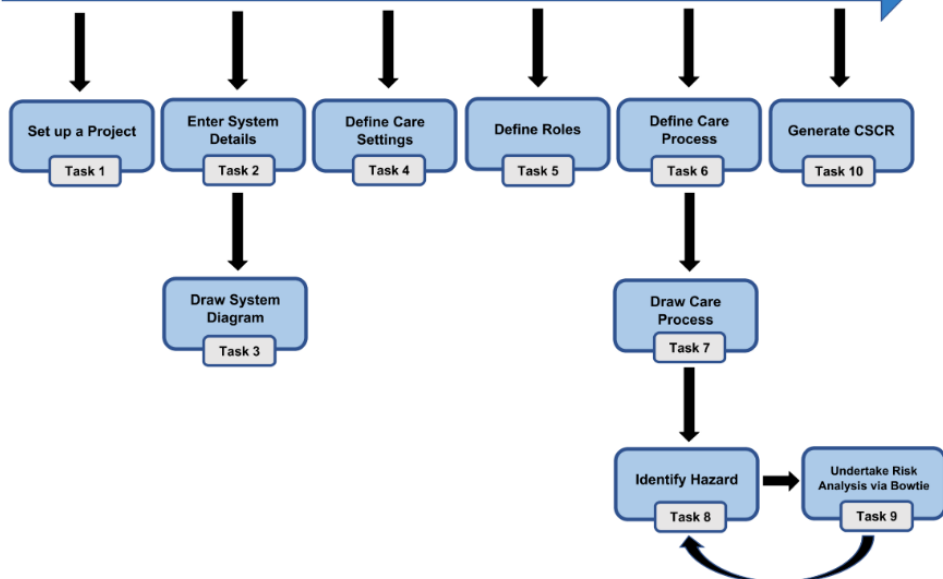
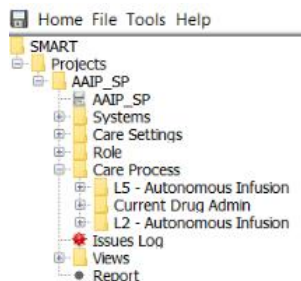


Figure 1: SMART Dashboard

Figure 1 graphically demonstrates the sequence of tasks the user needs to follow to complete the safety argument. The content can be navigated through the tree in the left-hand panel as shown in the figure opposite. The clinical risk management process for the case study was completed using SMART. It is envisaged SMART could become a standardised tool for clinical risk management.



Automated to Autonomous

To understand how the technology within the SAM project could be implemented, an approach was taken to describe the level of autonomy through levels similar to the automotive domain. Figure 2 illustrates different levels of capabilities of the technology, moving from automated to an autonomous implementation. This aided the preliminary hazard analysis and safety assurance argument as the technology took greater control of the process.

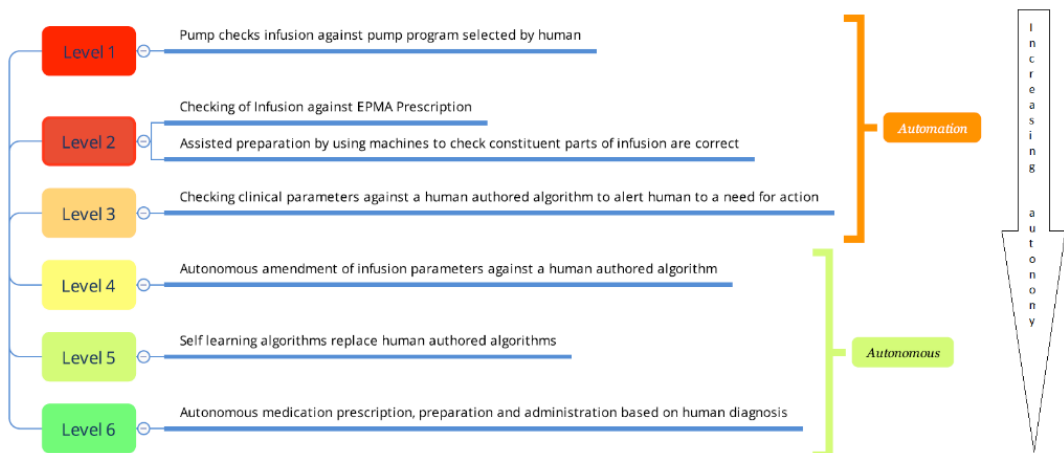


Figure 2: Automated to Autonomous

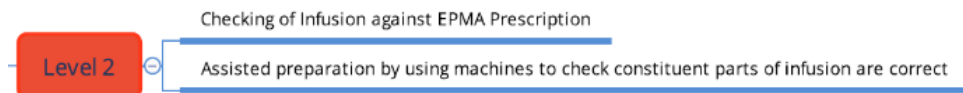
Using SMART's graphical process editor, three process maps were drawn to fully understand the current drug-administration process compared to the future processes at Level 2 (L2) and Level 5 (L5).

Care Processes: Through the Levels

The component parts of the drug administration process are:

- Prescriber – Fills in the type, amount and dose of the medication
- ICU Nurse – Gathers, administers and monitors medication
- Second checker – Second checks at set intervals
- HIT – The ICU ward's computer system. Houses patient data
- Infusion pump – Delivers insulin into the patient's body

Current drug administration The current process is driven by a paper-based prescription chart that is passed around the staff. The consultant initiates the process by prescribing the medication on a prescription chart. The nurse assumes the responsibility of identifying the patient, gathering the medication, administering and monitoring infusion. The use of the pump in this scenario is limited to administering the medication with no decision-making functionality.



L2 This process moves the reliance away from the paper-based prescribing approach to a fully technological solution. The prescribing here is done using an Electronic Prescribing and Medicines Administration (EPMA) system, which is a component of the ward's HIT. The nurse is still responsible for gathering and administering the medication. However, the identification of the patient, viewing prescriptions and second checking the right syringe is placed in

the correctly programmed pump are all done by cross-checking against the HIT. The HIT prints a prescription barcode that has patient and prescription specific information. This barcode can be scanned by any device which can then communicate with the HIT to confirm any details. For example, scanning the barcode into an infusion pump will cross-reference whether it has been programmed correctly. At this level, much of the prescription second checking is now done by communicating with the HIT. This removes some of the duplication and double-checking tasks, from the nurse, which are error-prone and can lead to drug-administration errors.

Level 5

Self learning algorithms replace human authored algorithms

L5 This level extends level 2 significantly by making use of AI & ML. The HIT is monitoring the patient's blood sugar levels with sensors. This gives it the ability to make the same decisions a nurse would make in terms of smaller adjustments to the drug administration based on established clinical guidelines. The technology could also assume greater responsibility and go beyond the initial clinical guideline by learning from the patient's physiological response to develop a personalised drug administration protocol. To maintain a level of human supervision, the HIT has certain hard limits (e.g. upper limit on drug dose and rate of administration, which are known to have fatal outcomes). If it decides to act outside of guidelines, authorisation is required from a healthcare professional. It is within these bounds that the technology is considered to be autonomous.

Hazard Analysis

Hazard analyses were conducted for each level by firstly identifying the hazards. The causes and subsequent harm events were added. Where possible, mitigations were included. This was documented using bowtie diagrams.

Assurance Challenges

A fundamental component of this work was to assess if the clinical risk management process outlined in DCB0160 would be applicable when assessing autonomous technology. The process did not present significant difficulties and appeared relevant and applicable. The analysis

“Hazard scenarios and contributory factors changed, with a shift away from people towards the technology”

considered the different levels of automation. The major hazards remained the same across the levels of automation (e.g. patient harm resulting from overdose / underdose), but the hazard scenarios and contributory factors changed, with a shift away from people towards the technology. This seems a logical shift as the technology assumes more responsibility. However, it is also worth considering that novel human factors challenges are introduced, such as automation bias and handover between the technology and people [8]

The role of the human, mainly the nurse, in the baseline process is of someone who does the actual tasks of double-checking and drug administration. In L5, the technology undertakes these tasks and the role of the human becomes one

of providing supervision and expert oversight.

It is anticipated that this approach would lead to fewer drug administration errors. However, what assurances are needed to have confidence that the technology is functioning as it should be?

The technology at some point will fail or need to hand over to a clinician due to encountering a situation where it cannot continue. A protocol for handover will need to be established. Current nursing skills do not encompass handover from technology to humans, especially when there is an unplanned handover at L5. To assure safe operation of the service, the staff would need additional and new forms of training. It may be the case that a new specialised role might have to be established, such as an AI Operative Nurse who provides expert oversight and who would be present to give assurance that an appropriately trained human is in place.

“There would need to be a fundamental rethink of how the service is provided to the patient”

It is not anticipated that introducing AI/ML technology into a ward setting, while keeping all the other processes as they are, would be a wise approach. There would need to be a fundamental rethink of how the service is provided to the patient. It may be the case that the staff will need to work around the AI/ML product and fundamentally change their working practice. From an assurance perspective, AI/ML solutions will require certain staff and working practices to be in place.

Regulation will play a pivotal role in how AI/ML technology develops in this area. Based on the experiences within the SAM project, it appears that existing standards are reasonably well-suited, with some modifications, to address the technical challenges of assuring the safety of AI and autonomous systems in healthcare. However, further consideration needs to be given to the human factors aspects of using AI/ML in healthcare safely, and to the organisational and institutional context. For example, at present there is no communication between the current HIT and infusion pump.

At L5, as they begin to work in tandem, one of the two, or even both would be classed as a MD. This raises questions for the manufacturers, such as, will they want to take on the additional regulatory compliance and investment? The regulatory process under MD Regulation is far more complicated to navigate as opposed to going down the route of NHS standardisation. This will ultimately impact on the architectural design of the system, which in turn, will lead towards the types of assurance requirements that will be sought.

Conclusion

AI/ML technologies being developed in health domain show no signs of dissipating. As more of these technologies make their way on to the market, assuring them is, and will be a significant challenge. As observed from the case study, existing approaches to regulation and assurance still apply. However, deeper analysis is required in how humans and AI/ML technology collaborate to provide a service to the patient without harming them.

References

- [1] Joshi, I., Morley, J.,(eds). Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care. 2019. London, United Kingdom: NHSX
- [2] Elliott, R A, Camacho, E, Campbell, F, Jankovic, D, St James, M M, Kaltenthaler, E, Wong, R,Sculpher, M J, and Faria, R, "Prevalence and economic burden of medication errors in the NHS in England". 2018. Sheffield: Policy Research Unit in Economic Evaluation of Health & Care Interventions.
- [3] Sujan, M, Furniss, D, Embrey, D, et al. "Critical barriers to safety assurance and regulation of autonomous medical systems". 22–26 Sept 2019. In: Proceedings of the 29th European safety and reliability conference (ESREL 2019) (ed Beer, M, Zio, E), Hannover.
- [4] NHSDigital, "DCB0129: Clinical Risk Management: its Application in the Manufacture of Health IT Systems – Specification NPFIT-FNT-TO-TOCLNSA-1792.06, V4.2, 02.05.2018", NHS, Jun 2018 [Online] Available: <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture-of-health-it-systems> [Accessed: 18.11.2019]
- [5] NHSDigital, "DCB0160: Clinical Risk Management: its Application in the Deployment and Use of Health IT Systems – Specification NPFIT-FNT-TO-TOCLNSA-1793.05, V3.2, 02.05.2018 ", NHS, Jun 2018 [Online] Available <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0160-clinical-risk-management-its-application-in-the-deployment-and-use-of-health-it-systems> [Accessed: 18.11.2019]
- [6] Habli, I., White, S., Sujan, M., Harrison, S. and Ugarte, M., "What is the safety case for health IT? A study of assurance practices in England". 2018. Safety Science, 110, pp.324-335
- [7] , SCSC Assurance Case Working Group (ACWG), "Goal Structuring Notation Community Standard Ver 2", 2018, <https://scsc.uk/SCSC-141B>
- [8] Sujan M, Furniss D, Grundy K, et al, "Human factors challenges for the safe use of artificial intelligence in patient care". BMJ Health & Care Informatics 2019;26:e100081. doi: 10.1136/bmjhci-2019-100081

Shakir Laher, Safety Engineer, NHS Digital

Shakir Laher is a Safety Engineer currently working at NHS Digital. His role involves developing a software tool that generates safety cases, clinical risk management activities that support the safe development/deployment and use of Health Information Technology (HIT) within the NHS and writing academic research papers that explore the assurance requirements of HIT.

Mark Sujan, Senior Consultant, Human Reliability Associates

Mark Sujan, PhD, is Senior Consultant at Human Reliability Associates located in the North West of England, and Associate Professor of Patient Safety at the University of Warwick. He has more than 20 years of experience developing, applying and teaching human factors and safety engineering methods in both safety-critical industries and healthcare.