2.6 – Handling change during operation

Practical guidance – healthcare

Author: SAM demonstrator project

Healthcare is delivered in a highly dynamic and non-deterministic environment, and successful outcomes are dependent on actions and decisions made by humans. However, humans are fallible and unintentional errors and mistakes have led to unsafe care outcomes. The application of artificial intelligence (AI) offers great potential in this domain by: automating routine tasks that are susceptible to human error (e.g. transcribing prescriptions and printing syringe labels); and working autonomously to monitor and manage care scenarios (e.g. optimising insulin infusion regime).

However, by removing the healthcare professional (HCP) from the real-time, closed loop, care pathway there is a significant risk that they will lose their situational awareness and their ability to deliver effective care could be compromised. So, whilst there is a great opportunity to improve efficiency within healthcare, careful consideration needs to be given to monitoring and handover protocols such that effective human intervention occurs should the AI's behaviour exhibit characteristics that could cause or contribute to patient harm.

The following guidelines provide a framework which will support effective monitoring and handover between AI technology and HCPs.

Upskill HCPs	HCPs will need to establish an understanding of the technology, its capabilities and weaknesses so they are better placed to recognise anomalous behaviour.
Baseline and understand care pathway	Care pathways need to be defined and baselined (representing work as is done) so that the contribution and authority of AI is clearly expressed and understood within the clinical team.
Define AI capability	The specific capability that the AI is providing needs to be defined and characterised in the context of supporting the care pathway. This must consider the interaction between the AI and the HCP both as a user and also a supervisor. This must consider the authority limits autonomous AI can have. This must consider the monitoring and alerting mechanisms and whether these are undertaken by the AI itself or independently by another element of the care system.
Conduct pre-emptive hazard analysis	Need to understand the potential patient-level harm effects that could occur in the care-pathway and the specific contributions AI could make. The severity of harm outcome

	and the significance of the AI contribution will impact the definition of key activities.
Develop monitoring standard operating procedure (SOP)	Need to develop a regime within the care pathway that will ensure the continued safe operation of the AI. This will be dictated by the capability that the AI is providing (automation and/or autonomy) but typically would need to consider: • Pro-active or re-active: HCP routinely monitors behaviour of AI or responds to an alert or alarm. • Frequency: sufficient to maintain situational awareness but not so frequent that it compromises efficiency. • Trends: does the monitoring indicate progression to an unsafe state or imminent point of handover? • Escalation: is there a need for a monitoring HCP to seek second opinion or high authority before initiating any action? • Monitoring architecture: this needs to be defined and consideration given to whether the AI simply monitors itself, whether the AI is monitored independently by another element of the care system or whether it's a combination of both.
Develop handover SOP	Need to develop a regime within the care pathway that will ensure timely HCP intervention when required. This will be dictated by the capability that the AI is providing (automation and/or autonomy) but typically would need to consider: • Definition of safe limits: the limit of authority the AI can have before HCP intervention is required needs to be defined. Definition should consider the need for soft limits (i.e. those that can be transgressed but signify an impending handover requirement). Hard limits (i.e. those that must never be exceeded) need to be defined. The protocol needs to consider the degree of authority the AI has; does it have the same as a HCP or is it restricted to a lower level? • Definition of transfer state: the AI's behaviour whilst a handover is being determined needs to be specified. Does the AI continue to perform its function that may result in a change in outcome, does it maintain a steady state of output or default to the previous known good (safe) output? The period of

	time that the transfer state can persist needs to be defined.
	• Definition of safe state: the AI's behaviour once authority has been relinquished needs to be defined. Does the AI revert to an "off" state and dissociate itself from the care pathway or does it continue to function in "hot standby" in order to support a subsequent transfer of authority back from the HCP?
	• Definition of re-engagement criteria: the criteria and process for re-engagement of the AI needs to be defined.
	• Definition of audit log: an audit log may be needed to support informed and safe handover of authority to an HCP. It will be necessary to identify those clinical variables, environmental conditions and system parameters that influenced the learning. The HCP will need to be able to quickly assimilate the clinical scenario and take effective mitigating action.
Simulation and dry-runs	Need to train HCPs in the execution of the SOPs potentially through simulation (outside of the care environment) and dry-runs (inside the care environment) of hazard scenarios. This needs to verify the effectiveness of the SOP and the ability of the organisation to follow it in real-life scenarios.
Re-active incident management	There is a need for an organisation to recognise when handover between the AI and HCPs has resulted in an incident or near-miss. This needs to be accommodated in the organisation's existing service/safety management process. Such events need to be reviewed and the impact on the organisation's safety case and understanding of AI technology considered.