Implementation Criteria for AI Systems in Health

1 SCOPING

Standard 1: Establish a scientific and ethics committee

The provider of a health AI system should set up, prior to the design phase, a scientific and ethics committee composed of a representative panel of potential users of the AI system as well as people for whom the system will be used (patients, service users). This committee should accompany the project throughout its lifecycle and validate the ethical compliance of the solution at each stage.

2 DEVELOPMENTS

Standard 2: Transparency on the origin of training data

When health data derived from care delivery are used to train the AI system, the provider must ensure that patients were informed and able to exercise their rights regarding secondary use of their health data.

Standard 3: Data minimization vs patient singularity

The AI provider should analyse and limit the data used for training, production and operation of the AI system to what is strictly necessary and sufficient, while ensuring that patient's singularity is adequately represented.

Standard 4: Data minimization vs environmental impact

The AI provider should analyse and limit the data used for model training, production and operation of the AI to what is strictly necessary, with the goal of reducing environmental impact and promoting digital sobriety.

Standard 5: Reduce and eliminate AI bias

The AI provider should implement mechanisms to minimise, and ideally eliminate, system biases.

Standard 6: Explainability of the AI system

When the AI system provides decision support with a user interface, the provider should implement mechanisms enabling the system to explain its outputs upon user request. If explanations are not possible, the provider must transparently describe the learning method and reasons why outputs cannot be explained.

Standard 7: Prevent generative AI from returning personal data that was used for its training (G)

When a generative AI has been trained using personal data, the AI provider should implement mechanisms to ensure that the personal data used for training cannot be used by the AI in the text of its responses.

Standard 8: Guarantee the absence of plagiarism in generative AI (G)

The provider of a generative AI should implement mechanisms to ensure that the answers provided are not exact copies of texts from authors used for training, unless they explicitly state that they are quotations and indicate the sources, thus guaranteeing the absence of plagiarism in the answers produced.

Standard 9: User autonomy to deactivate a modular AI system

The provider of a modular healthcare AIS (i.e., integrated into a system consisting of different applications that can be deactivated separately) should implement mechanisms allowing the user, whether a healthcare professional or a patient, to independently deactivate the AIS without having to stop the operation of the other applications in the system.

Standard 10: Eco-design of the AI system and prevention of hardware obsolescence

The provider of a healthcare AI should apply the national eco-design framework for digital services (RGESN) in the AI design process and support circular-economy approaches, allowing the system to run on devices that are not state-of-the-art when possible.

Standard 11: Carbon footprint assessment of AI system development

The provider should perform a carbon footprint assessment of the AI system, including its training phase.

Standard 12: Transparency that the user is interacting with AI

The provider of a healthcare AI system with a user interface should inform users, whether healthcare professionals or patients, that they are interacting with AI at the time of their use of the AI system.

Standard 13: Verify a clear understanding of the interaction with AI

The provider of a healthcare AI system with a user interface should implement mechanisms to verify that users, whether healthcare professionals or patients, understand that when using the AI system, they are interacting with AI.

Standard 14: Labelling AI Results

The healthcare AI provider should implement technical mechanisms to enable the labelling of results generated by AI.

Standard 15: Traceability of IAS deactivation

The provider of a modular healthcare IAS (i.e., integrated into a system consisting of different applications that can be deactivated separately) should implement mechanisms to ensure that any IAS deactivation is tracked (identity, time, duration).

Standard 16: Ensuring User Decision-Making Autonomy

Users must always retain the ability not to follow the system's recommendations and maintain full decisional autonomy.

Standard 17: Transparency on following AI recommendations

The system should allow tracing, for each patient, that the AI was used, that its recommendations were followed, and that the patient was informed.

Standard 18: Transparency on not following AI recommendations and reasons

The system should allow tracing that the AI was used but the recommendation was not followed, including the stated reason, with patient information ensured.

Standard 19: Transparency on not using the AI at all

The user interface, which offers decision support for healthcare professionals, should enable tracing situations for each patient, that the AIS has not been used, and that the patient has been informed.

Standard 20: Compliance with digital accessibility standards (RGAA)

Any AI system with a user interface must comply with the national accessibility framework (RGAA).

Standard 21: Intelligibility and usability of interfaces

The provider must ensure that interfaces support intuitive, easy, and safe use for all user profiles.

Standard 22: Consistency of answers from a generative AI system (G)

The provider should ensure that identical queries submitted at the same time produce consistent responses.

Standard 23: Quality of responses from a generative AI system (G)

The provider must ensure that responses are reliable and accompanied by warnings regarding potential AI hallucinations.

Standard 24: Ensure patient-facing generative AI prompts referral to a health professional when needed (G)

The provider of an AI system equipped with a user interface, which offers decision support for patients, should put in place technical mechanisms to ensure that the SIA does not exceed its capabilities and effectively encourages patients to consult a healthcare professional at the slightest sign of a health emergency for the patient.

Standard 25: Equal performance across all groups within the target population

The provider of a health AI system that offers decision support, diagnostic or therapeutic, for patients should ensure that the decision support produced is fair and equitable for each patient in the target population for which the solution can be used.

3 DEPLOYMENT, TRAINING & USE

Standard 26: Support users/organizations on impacts of deployment of a SAI

The provider of a health AI system must support organizations in anticipating and analysing impacts on care delivery, professional relationships, and patient relations.

Standard 27: Co-construct training resources with users

Training materials on AI-related issues for healthcare professional or patient must be co-designed with a representative user panel.

Standard 28: Train users on AI issues (bias, limits, performance)

The provider should train users (professionals and patients) on AI-related risks, bias, and performance limits. He should inform about the way he acted to reduce or to erase those risks and bias.

Standard 29: Verify user understanding of AI issues

The provider must ensure users (professionals and patients) fully understand the system's limits and can make informed decisions about AI use and AI proposals following.

Standard 30: Train users before deployment

The provider of AI solutions is responsible for the training of relevant personnel (those responsible for roll-out, or if relevant, the healthcare professional users) for the solution, which use must occur before the system is introduced into practice.

Standard 31: Provide easy and intuitive access to documentation

All documentation must be readily accessible and understandable.

Standard 32: Transparency on types of AI used

Users (professionals and patients) must be informed by the documentation and training sessions about the types of AI employed in the system.

Standard 33: Transparency on system performance

The provider should clearly indicate the specifies of the target population and the performance levels expected for each segment of the target population.

Standard 34: Detect user dependence on the AI system

Mechanisms should be implemented to detect both professional and patient users' signs of dependence on AI-generated recommendations.

Standard 35: Alert in case of detected dependence

The provider of a health AI system should ensure that users' dependence patterns to AI can be identified.

Standard 36: Maintain critical thinking and clinical expertise

The training provided to health professionals for the use of AI solutions must emphasise the need for human critical thinking about AI proposals and trust in clinical judgement.

Standard 37: Remind users of their responsibility for decisions

The provider of a health AI system that offers decision support should remind users (professionals or patients) that they remain fully responsible for their decisions.

Standard 38: Promote environmentally responsible use of AI

AI training should include clear and accessible guidance on environmentally responsible use; complete and didactic documentation should be provided.

Standard 39: Data sovereignty for the deploying organisation

The AI provider should offer technical solutions to ensure data is not subject to extra-European regulations.

4 MONITORING, UPDATE & CONTINUOUS IMPROVEMENT

Standard 40: Continuous evaluation of the AI system

IA provider should carry out regular evaluations of system outputs.

Standard 41: Continuous improvement process

The IA provider must engage in continuous improvement based on regular evaluation results, including drift detection and human oversight.

Standard 42: Transparency about system updates and consequences

Users (professionals and patients) must be informed of IA system updates that may change system usage or impact judgement.

Standard 43: Verify user understanding of update consequences

The AI system provider should ensure users' (professional and patient) understanding of the impact of system updates, particularly those affecting usage or clinical decision-making.