European Partnership on Metrology
Call 2022 – Digital Transformation, Health, Integrated European Metrology, Normative and Research Potential

Selected Research Topic number: **SRT-h04**
Version: 1.0

Title: Developing a metrological framework for assessment of image-based Artificial Intelligence systems for disease detection

Abstract

Image-based artificial intelligence (AI) systems for disease detection are increasingly being developed, and it is vital that these tools are robust and effective in heterogeneous clinical settings. To date, performance has been assessed in an ad hoc manner as there are no approved guidelines for evaluation. Most studies have methodological weaknesses and results that are not comparable. A standardised and impartial framework for performance, generalisability, and suitability assessment of AI tools will address these needs and enable more efficient, reliable and reproducible validation of image-based AI systems for disease detection.

Keywords

Image-based AI systems, explainable AI, disease detection, clinical implementation, metrological framework, assessment of AI, cancer

Background to the Metrological Challenges

AI has the potential to transform healthcare systems significantly and play a key role in future clinical decision-making. The exponential increase in healthcare data over the last decade and fast-paced technology developments have resulted in new AI approaches for diagnostic applications and risk prediction developing rapidly, yielding promising outcomes. However, the implementation of AI in clinical settings remains limited, mostly due to a lack of robust validation procedures and trustworthiness as well as a lack of clear understanding from the national bodies on steps for adoption.

There are many challenges to accessing suitably anonymised healthcare data, due to necessary information governance and complex data structures. These difficulties may partially explain the limitations of previous evaluation studies on small scale datasets and populations. The development of accessible clinical data warehouses will provide sufficient data and ensure future validations are robust.

Thorough and consistent validation of image-based AI systems for disease detection is essential, not only in terms of sensitivity and specificity, but also in how they will be intended to be used in the clinical environment. Crucially, the AI systems must be sufficiently generalisable and fit for purpose for the population they are to be applied to. Well-designed comparative test accuracy studies and cohort studies in large and diverse populations are urgently needed, as well as evidence from direct comparison of different AI systems; the effect of different imaging machines on the accuracy of AI systems; the effect of differences in screening programmes on disease detection with AI; and the effect of making available additional information to AI systems for decision making.

Objectives

Proposers should address the objectives stated below, which are based on the PRT submissions. Proposers may identify amendments to the objectives or choose to address a subset of them in order to maximise the overall impact, or address budgetary or scientific / technical constraints, but the reasons for this should be clearly stated in the protocol.

The JRP shall focus on metrology research necessary to support standardisation in image-based AI systems for disease detection.

The specific objectives are

1. To develop technical infrastructure to be able to query and extract the relevant data from various databases (such as the OPTIMAM Mammographic Image Database).
2. To identify the key factors differentiating clinically relevant subgroups in the data and use these to categorise the data into subsets based on these key factors, identifying where there is sufficient data for training and validation. To develop a methodology to generate data derived from physics-informed models based on measurement knowledge.
3. To develop an explainable and traceable AI tool for disease screening, providing the capability to train and re-train the tool as necessary, to evaluate performance under different scenarios, including processed versus unprocessed data, low versus high image quality data, validation for specific patient demographics, etc. To develop and validate methods for the interpretation of the behaviour of the trained AI tool.
4. To summarise the performance testing evaluations and to provide metrology recommendations for the assessment of AI tools for disease screening- with a focus on understanding their generalisability and sensitivity to varying populations, manufacturers, image processing and acquisition techniques. To use the recommendations to design a global, standardised, and impartial AI assessment framework.
5. To facilitate the take up of the technology and measurement infrastructure developed in the project by the measurement supply chain, standards developing organisations (British Standards Institution, ISO/IEC JTC 1/SC 42 – Artificial intelligence), and end users (e.g. clinical stakeholders, manufacturers of medical and healthcare products, regulators).

These objectives will require large-scale approaches that are beyond the capabilities of single National Metrology Institutes and Designated Institutes, and it is expected that multidisciplinary teams will be required. To enhance the impact of the research, the involvement of the appropriate user community such as medical practitioners, medical (academic) hospitals and industry is strongly recommended, both prior to and during methodology development.

Proposers should establish the current state of the art and explain how their proposed project goes beyond this.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 1.9 M€ and has defined an upper limit of 2.6 M€ for this project.

EURAMET also expects the EU Contribution to the external funded beneficiaries to not exceed 35 % of the total EU Contribution across all selected projects in this TP.

Any industrial beneficiaries that will receive significant benefit from the results of the proposed project are expected to be beneficiaries without receiving funding or associated partners.

Potential Impact

Proposals must demonstrate adequate and appropriate participation/links to the ‘end user’ community, describing how the project partners will engage with relevant communities during the project to facilitate knowledge transfer and accelerate the uptake of project outputs. Evidence of support from the ‘end user’ community (e.g., letters of support) is also encouraged.

You should detail how your JRP results are going to:

- Address the SRT objectives and deliver solutions to the documented needs,
- Feed into the development of urgent documentary standards through appropriate standards bodies,
- Transfer knowledge to the medical sector.

You should detail other impacts of your proposed JRP as specified in the document “Guide 4: Writing Joint Research Projects (JRPs)”

You should also detail how your approach to realising the objectives will further the aim of the Partnership to develop a coherent approach at the European level in the field of metrology and include the best available contributions from across the metrology community. Specifically, the opportunities for:

- improvement of the efficiency of use of available resources to better meet metrological needs and to assure the traceability of national standards
- the metrology capacity of EURAMET Member States whose metrology programmes are at an early stage of development to be increased
- organisations other than NMIs and DIs to be involved in the work.

Time-scale

The project should be of up to 3 years duration.