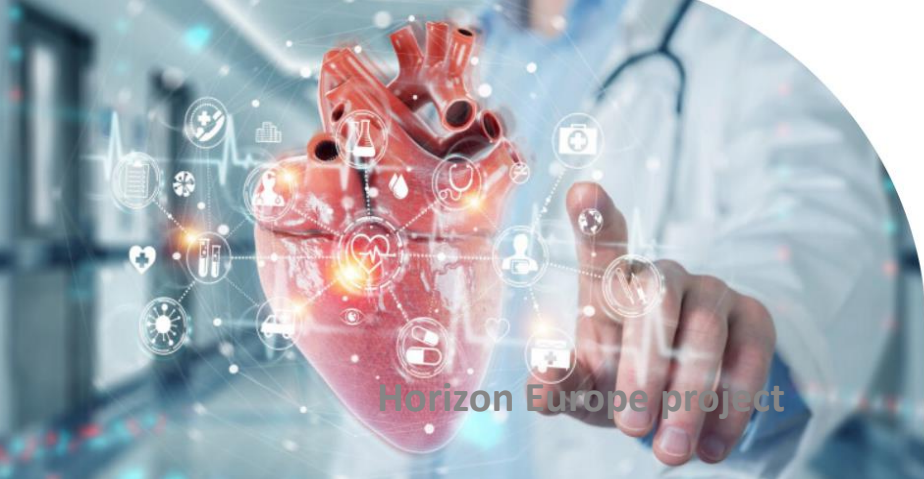


Clinical and patient requirements for trustworthy AI

**Results of co-creation workshops of Trustworthy Artificial Intelligence
for Personalised Risk Assessment in Chronic Heart Failure (AI4HF)
European Union program**

Carina Dantas | SHINE 2Europe

13 November 2024 | Zurich, Switzerland



Horizon Europe project



AI4HF

Trustworthy Artificial Intelligence
for Personalised Risk Assessment
in Chronic Heart Failure

Horizon Europe project



1 June 2023 → 31 May 2027

AI4HF harnesses the power of **Artificial Intelligence** to co-design, develop, and evaluate a new digital system for **personalized risk assessment and advice** for individuals living with **Chronic Heart Failure**.

It uses advanced AI **algorithms**, global **collaboration**, and a **patient-centered** approach to improve healthcare outcomes.

The problem to solve

The prevalence of Heart Failure (HF) is expected to increase by 46% by 2023 due to an ageing population and increases in unhealthy environments.

The challenge in Heart Failure care and management

- HF is complex as it has many risks, causes and outcomes
- Current medical models for HF follow one-size-fits-all guidelines



The challenge in current AI-based solutions

Existing AI models for predicting HF risk are promising, but far from being used in clinical settings. They often share common limitations:

- **Low levels of trust** between healthcare providers and patients
- **Limited real-world validation** of the AI models
- Though studies focus on the accuracy of prediction, **few focus on aspects of *Trustworthy AI***

The AI4HF solution



The foreseen system developed in AI4HF integrates the following features:

Patients & clinicians | a **clinical decision support tool** to enable shared decision-making throughout the care journey.

Citizens & patients | an educational **information and communication package** for increasing transparency and digital literacy.

Clinicians | a **multi-modal AI risk assessment tool** for predicting personalised outcomes in HF patients and improving quality of care.

Researchers | an **AI traceability technology** to help researchers to effectively monitor and adjust the AI tools over time.



Partners



ABOUT US



SHINE is a Portuguese SME - applied research - promoting inclusive communities for all citizens.



4



HORIZON

13



ERASMUS & OTHERS

23



Social Innovation



Ethics



Policy



Dissemination



Gamification

CHAIR COST ACTION



50 countries | +775 members

DIRECTOR



SHAFE Foundation

Smart Healthy Age-Friendly Environments

INCLUSIVE AND ACCESSIBLE COMMUNITIES



DIGITAL TECHNOLOGIES



SHAFE

SMART HEALTHY AGE-FRIENDLY ENVIRONMENTS



SUSTAINABLE BUILT ENVIRONMENTS



HEALTH AND SOCIAL CARE

COORDINATOR



HORIZON EUROPE
Social Innovation
Responsive Environments
NETwork

What do we do at SHINE?



3 KEY AREAS

RESEARCH

+ Knowledge

New approaches – **social innovation**

Multidisciplinary networks

POLICY

Local ecosystems

Evidence-based recommendations

Translate **research - policy** priorities

PEOPLE

Reliable information

Good tools

Promote **empowerment** and **engagement**



A model that can be applied by design, combining:



MULTIPLE STAKEHOLDERS

PARTICIPATION, CO-CREATION,
COMMITMENT & COLLABORATION



PARTICIPATORY ETHICS

ASSESSING SOCIETAL CHALLENGES AND
DISCUSSING POTENTIAL "SOLUTIONS"

PRODUCTS & SERVICES

PROMOTE EMPOWERMENT,
INCLUSIVENESS AND PARTICIPATION



OWNERSHIP

INCREASING LITERACY, EDUCATION,
NETWORKING & GOOD PRACTICES

Trustworthy A.I.



Clustering ETHICAL CHALLENGES on AI applied to health



1

INDIVIDUAL

- Low health and AI literacy of citizens
- Lack of adequate training on AI for healthcare professionals
- Personal convictions that lead to human biases
- Lack of trust on AI tools

TOUCHPOINTS

Education; training; prejudice; discrimination; trust



2

TECHNICAL

- Lack of scientific validation of the algorithms
- Lack of transparency and accuracy of algorithms and their predictions
- Lack of diversity in data, namely on underserved populations
- Lack of ethnicity-related data in datasets

TOUCHPOINTS

Algorithm design; training datasets; interpretation of results

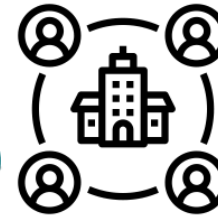
3

ORGANISATIONAL

- Lack of integration of new tools with existing clinical pathways & solutions
- Low diversity in working teams – on disciplines, culture, age, gender
- Lack of audit methods and tools to support organisations implementing ethical workflows

TOUCHPOINTS

Diverse teams; audit; ethics-by-design; supporting methods and tools



4

SOCIETAL

- Lack of clarity on professional liability with the use of AI
- Absence of studies analysing long-term impact of AI use in health
- Increase of inequalities due to financial reasons
- Need of cohesive regulations

TOUCHPOINTS

health inequities, ethical and legal; regulatory framework; policy measures

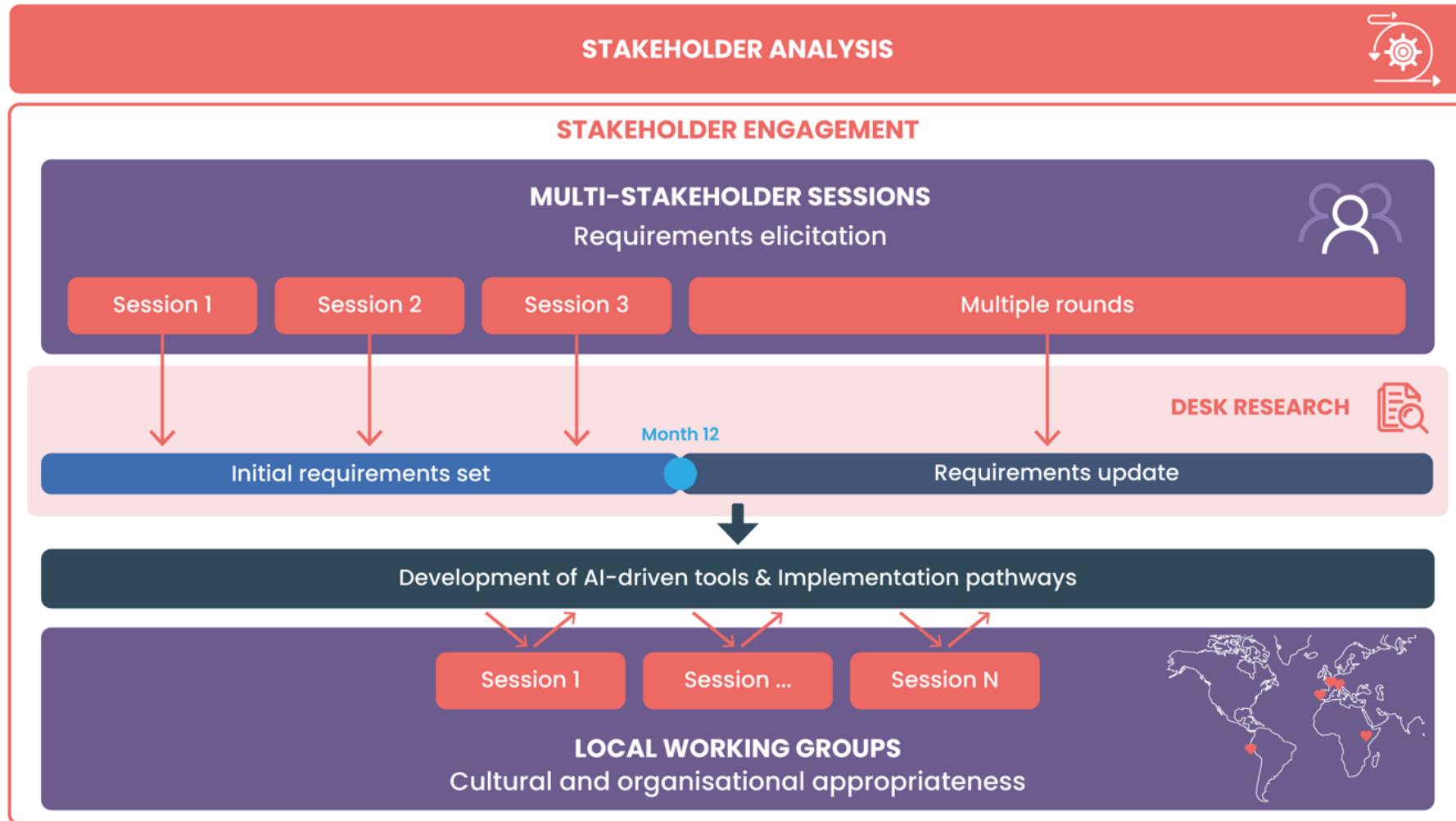


WP1: Multi-stakeholder engagement and social innovation (SHINE)

- Develop a **social innovation framework** to **engage relevant stakeholders**, including cardiologists, patients, AI technologists, data/IT managers, social scientists, policymakers and regulatory experts.
- Leverage the social innovation framework to identify **multi-disciplinary needs, requirements, obstacles** and **implementation pathways** for AI4HF's real-world adoption.
- Translate the multi-stakeholder requirements into a set of **DESIGNS, PROCEDURES AND SOLUTIONS** based on the FUTURE-AI guidelines for subsequent **trustworthy AI implementation and evaluation**.



Social Innovation Framework in AI4HF






















































Multi-stakeholder requirements

1. **Stakeholder mapping**
2. Multi-stakeholder **workshops**
3. **Local Working Groups** Clinical/Patient
4. **ELSI** requirements
 1. Literature review: alignment of AI4HF with key ethical principles
 2. Literature review: patient reidentification risks
 3. Desk research: accountability of AI-driven solutions in healthcare
5. Healthcare and **regulatory** requirements
 1. Interviews with HTA experts
 2. Review of relevant regulations
6. **Requirements specification**, including aligning to the **FUTURE-AI guidelines**

Stakeholder Analysis | an example of a Stakeholder Map



| | Steps of AI development process according to the FUTURE-AI guidelines | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Stakeholder | Clinical conceptualisation | End-user requirements gathering | Technical Design | Data selection, collection and/or preparation | AI implementation and optimisation | AI evaluation | AI deployment | AI monitoring |
| 1. AI Ethicists |   | |  |  |  |  |  |  |
| 2. Legal consultants |  | | |  |  |  |  |  |
| 3. Human rights advocates |  |  | | | | |  |  |
| 4. European Health Management Association | | | | | |  |  |  |
| 5. European Committee for Standardization |  |  |  | | |  |  |  |
| 6. National Medicine and Medical Devices Agencies | | | |  |  |   |   |   |
| 7. Data protection authorities | |  |  |   |  |  |   |   |
| 9. Social scientists |  |  | | | | |  | |
| 9. HTA experts | | | | | |  |  |  |



LISTENER

The person in this role is given information.



THINKER

The person in this role thinks along with the research team.



DECISION-MAKER

The person in this role takes the initiative and/or makes the necessary (final) decisions. Key player.

Multi-stakeholder sessions



AI4HF

M3 Workshops: Clinical Conceptualization

On-site meeting in Brussels with 14 patients

On-site meeting in Tanzania with 22 patients

On-site meeting in Peru with 15 patients

Series of online meetings with healthcare professionals from AI4HF consortium



M6 Workshop: Ethical considerations

Online meeting with patients, healthcare professionals, and social scientists (17 participants)



M12 Workshop: Trustworthy AI

On-site meeting in Lisbon (Portugal) with patients, healthcare professionals, social scientists, and technologists (21 external participants)



PART I

| | Phase I - Pre-diagnosis | | | Phase II - Diagnosis | Phase III - Management | Phase IV - Hospitalization |
|--|--|--|---|--|--|--|
| <p>Scenario Patient journey of a patient with heart failure</p> | <p>Primary care Entry point through primary care</p> | <p>Secondary Care Entry point through medical specialist</p> | <p>Tertiary Care Entry point through hospitalization</p> | <p>Diagnosis What happens at diagnosis?</p> | <p>Regular follow-ups What happens during regular follow up?</p> | <p>Hospitalization after diagnosis What happens from the moment of hospitalization until discharge?</p> |
| <p>Steps Write down a short story of what is happening at each step.</p> | <p>First in primary care, as a gatekeeper. (entry-point) Stable and for continuous follow-up to GP. (management)</p> | <p>hospitalisation can be entry point, most often via GP (not-academic hospital) referral chronic HF GP or consultation (internal medicine/respiratory care)</p> | <p>hospitalisation can be entry point, most often via GP (academic hospital) sometimes second opinion or referral chronic HF GP or consultation (internal medicine/ respiratory care)</p> | <p>Clinical work-up echocardiogram, lab (BNP/nt-pro-BNP), ECG. Signs and symptoms of heart failure, patient history, complaints.</p> | <p>Classification during follow-up (preserving, reduced, mildly-reduced, preserved EF), asymptomatic (volume overload reduction, initiate guideline directed medical therapy (GDMT) guidelines, but there may be local variation via the doctor's in guideline adherence), same here for the heart team and cause of HF as during hospitalization.</p> | <p>medication optimization, reduce volume overload (IV-diuretic or oral medication) and catheterization lab (intervention) heart team discussion if it is caused by valvular disease, ischemic (revascularization) or other causes</p> |
| <p>People involved Write down all the actors that are involved in this step. If there is more than one clinician involved please specify the type of clinicians.</p> | <p>General practitioner (cardiologist as consultation partner for GP), also as video consultation</p> | <p>1.5 care cardiologist, cardiologist, nurses</p> | <p>cardiologists, nurses</p> | <p>cardiologists</p> | <p>cardiologist, nurses, cardio-surgeons</p> | <p>cardiologists, nurses, cardio-surgeons</p> |
| <p>Where does this step take place? Write down all physical locations where this step can take place. E.g., healthcare center, hospital, GP clinic, etc.</p> | <p>GP office, at-home video consults, GP-cardio video consults</p> | <p>Hospital</p> | <p>Hospital</p> | <p>Hospital (both secondary and tertiary) or 1.5 care units</p> | <p>hospital</p> | <p>hospital</p> |
| <p>Time duration How much time does this step usually takes until the patient moves to the next step?</p> | <p>within 6 weeks</p> | <p>Within the same week (1.5 partners) Within six weeks</p> | <p>Within six weeks</p> | <p>[How much time does this step usually takes until the patient moves to the next step?]</p> | <p>[How much time does this step usually takes until the patient moves to the next step?]</p> | <p>[How much time does this step usually takes until the patient moves to the next step?]</p> |
| <p>Use of AI-driven prediction tools Write the name of the tool, good points and bad points. Please mention tools that you currently use, have used in the past, or that you heard about.</p> | <p>none</p> | <p>SCORPIO System (AI-based)</p> | <p>none</p> | <p>none used in clinical practice</p> | <p>none used in clinical practice</p> | <p>none used in clinical practice</p> |



Summary Results Workshop: “Towards trustworthy AI-driven tools for personalized treatment of heart failure” | Lisbon, 13th of May 2024

The results of this workshop led to 22 new stakeholder requirements.

Alignment of AI HF solutions with key ethical principles

Umbrella review with 25 articles until September 2023



Right to Autonomy

- AI systems should **support**, not replace, human judgment. Clinicians must remain central to decision-making, preserving empathy and managing complex conditions. Establishing **clear ethical standards** and **accountability mechanisms** will protect patient autonomy.

Confidentiality

- Address privacy concerns in **remote monitoring** and **data transmission**. Implement robust data security measures, such as encryption and blockchain technology.

Information privacy

- Tackle ethical concerns about data sharing and **potential biases in AI algorithms**. Develop a **code of conduct** and comply with **GDPR guidelines** to protect patient privacy. Ensure AI systems are explainable to avoid hidden biases.

Equal treatment

- Prevent discrimination from **biased training data**. Implement transparent accountability for AI errors. Techniques like **semi-supervised** and **federated learning** can enhance data representativeness and collaboration. More randomized controlled trials will evaluate clinical suitability of ML systems.

The results of the desk research on key ethical principles led to **15 new requirements**.

Patient Reidentification Risks

- Attack models and solutions preventing reidentification found across numerous healthcare sub-fields and application domains. However,
- **No solution identified fully protects** patients from reidentification
- Encouraging **multi-modality in privacy-enhancing frameworks** can lead to decreased reidentification risks
- A large **shift towards synthetic data** processing algorithms is observed
 - Requires novel solutions and (evaluation) measures
- FEDERATED LEARNING is **not vulnerable** to patient reidentification attacks
 - Aims to mitigate such attacks → **can be recommended** to be used for mitigating privacy issues within AI4HF

Accountability of AI-driven solutions in healthcare



- 3 types of liability: fault-base liability, strict liability, and product liability
- Relevant legislative initiatives:
 - EU Artificial Intelligence Act (AI Act)
 - AI Liability Directive (AILD)
 - Product Liability Directive (PLD).
- Questions discussed with stakeholders:
 - How can **responsibility be clearly assigned** in the complex ecosystem of AI-driven healthcare? What legal reforms are necessary to address this issue?
 - How can regulatory frameworks be designed to **promote innovation while ensuring safety and accountability**?
 - What **standards** should be established for the transparency of AI algorithms in healthcare? How can these standards be enforced?
 - What strategies are most effective in **building public trust in AI-driven treatments**?
 - How can **educational** initiatives be designed to enhance understanding and acceptance?

Interviews with HTA experts: towards healthcare and regulatory requirements



Interviews with **6 experts** in Health Technology Assessment regarding key issues such as:

- Implementation
- Reimbursement
- Potential barriers
- Ensure HTA requirements are met

**The results of the interviews with HTA experts led to
15 new requirements.**



AI4HF

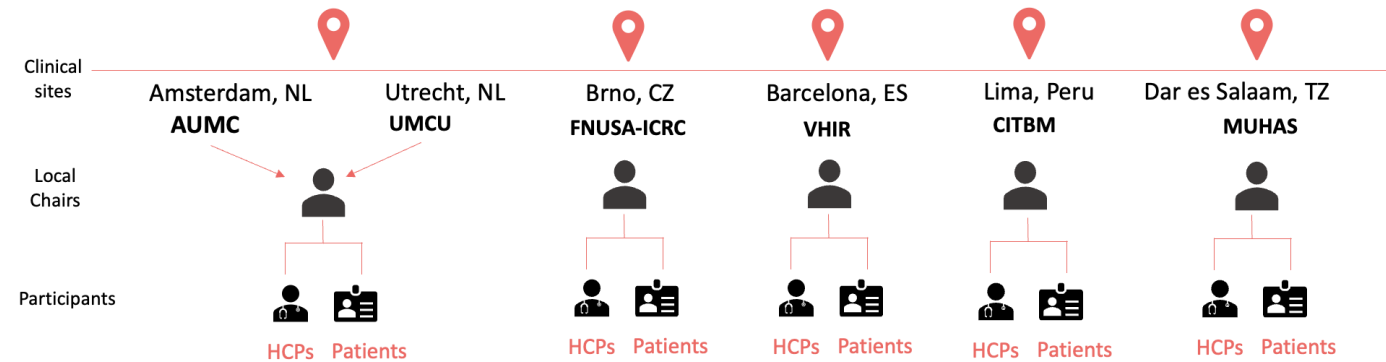
SHINE DEVELOPED

1. Workshop protocol with step-by-step guide
2. Workshop planning checklist
3. Reporting template

Provided in EN and translated to local languages:

1. Invitation letter
2. Informed consent
3. Workshop presentation
4. Material for participants during the workshop

Local Clinical/Patient WGs



Funded by
the European Union

First Local WG workshops: Spring 2024

Highlights

- Patients and healthcare professionals involved have good access to technology.
- Neither healthcare professionals (NL), nor patients (CZ) see the need for *yet another risk prediction tool*.
- Patients highlight the wish for **support in the daily management** of the disease (CZ) and **choosing best treatment** (ES).
- Need for **personalized solutions** (no “typical” patient, NL).
- High lack of trust in AI from patients. Need for AI literacy (PE).
- Positive experience from patients and HCP in participating in the local groups.



Workshop on the C&I Package

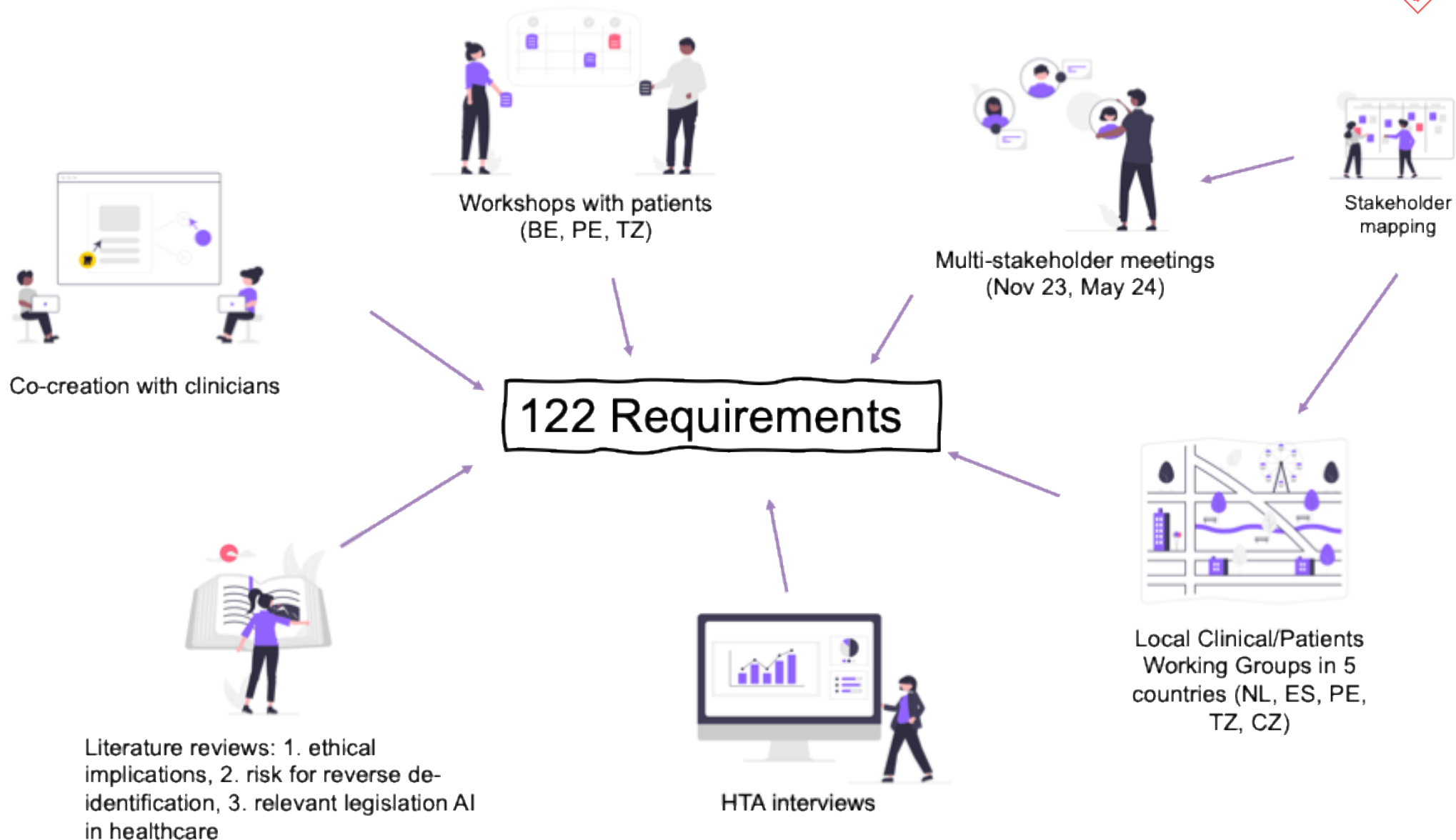


Date | 5 December 2023, online meeting | 7 EHN patient consultants (from 4 European countries), 5 project partners

| Knowledge Gap | Information Source | Trustworthiness |
|---|--|--|
| <ul style="list-style-type: none">• 5/7 did personal research right after HF diagnosis• Helpful info about HF: symptoms, recommendations for self-management, psychological impact and support• What people should know about living with HF: it is an invisible disability, physical limitations, can still live well• 6/7 <u>did not know that AI could be used in HF care when first diagnosed</u>• Discussed what they would like to know about the use of AI in HF and exiting knowledge gaps for patients in this area. | <ul style="list-style-type: none">• 7/7 would prefer <u>information disseminated by cardiologist</u>, followed by patient organisations (5/7) and other healthcare professions (4/7)• Shared different online sources for accessing information (i.e: chat GPT, medical journals, ESC, patient organisation website)• The consensus was that <u>multiple formats</u> for disseminating information to patients/the public is needed (i.e. video, infographics, text, podcast). | <ul style="list-style-type: none">• Majority would trust to learn this information from a <u>healthcare professional</u>, however, <u>patient-led support groups and patient organisations</u> were also described as trusted and important sources• Mixed response on how much information they would like to know about if AI were integrated into their care – suggestion to have <u>varying levels of information</u>• All respondents believed that there are <u>differences between EU countries</u> in access to trustworthy information on this topic. |

Stakeholder engagement M1-M15

| | M3 Workshops | | | | M9 Workshop* | M12 onsite | Local WGs: workshop #1 | | | | | Interviews* | Total |
|------------------------------------|-----------------|-----------|-----------|-----------|--------------|------------|------------------------|-----------|-----------|-----------|-----------|-------------|------------|
| | International * | Brussels | Peru | Tanzania | | | NL | CZ | ES | PE | TZ | | |
| Patients & caregivers | 0 | 14 | 15 | 22 | 5 | 6 | 2 | 5 | 9 | 5 | 9 | 1 | 93 |
| Healthcare professionals | 9 | 0 | 0 | 7 | 3 | 9 | 4 | 5 | 4 | 6 | 7 | 0 | 54 |
| Ethicists, regulators | 0 | 0 | 0 | 0 | 9 | 4 | 0 | 0 | 0 | 0 | 0 | 5 | 18 |
| AI developers and industry | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| Policy-makers / health authorities | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| Hospital administration | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 4 |
| Payors | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 9 | 14 | 15 | 29 | 17 | 21 | 6 | 10 | 14 | 11 | 19 | 7 | 172 |



Mapping of the requirements to the FUTURE-AI guidelines



| ID | Type | Requirement | Rationale | Corresponding FUTURE-AI principle | Corresponding (prospective) system requirement aligned with the selected principle | Source |
|-------|----------|---|--|-----------------------------------|--|------------------------------------|
| SI.17 | StakeReq | Patients who opt for the treatment using the AI-driven tool want to be able to decide whether their data can be used for research purposes | The patient data can be used for different purposes. Patients must be well informed of this. In this line, we should separate informed consent (related to treatment) from data protection consent (data for research). | TRACEABILITY 6 | The system shall request a data protection consent from the patient prior to processing patient's data. The system shall process the patient's data only if such a consent is obtained. | SI-Workshop-M6-breakout-ELSI |
| SI.8 | StakeReq | Patients do not want that AI-tools are the sole item on which clinical decision making is based. | Fear mentioned by a patient after the M6 workshop: "Implementing AI may shape and narrow the treatment options available to the patient, there will be no outside-of-the-box thinking as that will not be part of the AI programming." | USABILITY 1 | Same as for SI.3: The system shall provide the given clinician and patient, as part of the output, with a notification that the system is only used as an auxiliary tool in the decision-making process and that the clinician's decision does not only rely on the system's output | SI-Workshop-M6-followup-patients |
| SI.19 | StakeReq | Clinicians want that the clinical AI-tool is integrated within a specific clinical process based on the intended-use of the novel algorithm. | The use and application of the clinical AI tool should fit within specific work-flow and provide the necessary risk-prediction at required point of care (real-time versus non-realtime). In this sense, the risk-prediction is provided at the time-instance in the care pathway when it is required. | USABILITY 1 | The system shall provide the risk prediction that is coherent with the actual state of the patient. | SI-Workshop-M6-breakout-clinicians |
| LR.13 | StakeReq | The AI tool shall undergo rigorous validation and refinement processes to enhance its predictive performance in comparison to statistical models. | ML models did not achieve a significant advantage in predicting events, and their clinical feasibility and reliability were worse when compared with statistical models | TRACEABILITY 4 | The system's accuracy shall be compared with that of the state-of-the-art applications that only make use of statistical predictive models. | Desk research |

Horizon Europe EU funding for Trustworthy AI



Trustworthy AI is lawful, ethical, and technically robust. It is when trust in AI models can be established in each stage of its lifecycle, from design to development, deployment and use.

AI4HF uses the **FUTURE-AI guidelines**, which were developed based on 6 guiding principles:



FAIRNESS

Keeps the same quality of performance for different individuals and populations - developed to minimise potential bias



UNIVERSALITY

Can be successfully used in settings outside the study environment



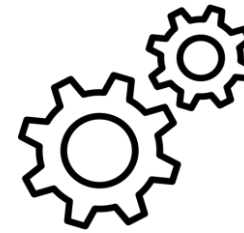
TRACEABILITY

Developed with measures for documenting and monitoring the tool from development to use



USABILITY

End users should be able to use the tool efficiently, easily and safely in real-world settings



ROBUSTNESS

The tool can maintain the same performance and accuracy when there are unexpected changes in the data it receives.



EXPLAINABILITY

Should provide clinically useful information about the logic behind the AI decisions it makes (i.e. no hidden processes)

Analysing challenges in AI for health



**FUTURE-AI
guidelines**

Fairness

Universality

Traceability

Usability

Robustness

Explainability

**ALTAI
GDPR**

Human agency and oversight

Data
privacy

**SOCIAL
INNOVATION**

Pathways to implementation



AI4HF

THANK YOU!



Funded by
the European Union