# Assessment Methods of Artificial Intelligence Supported Health Technologies. A Systematic Literature Review and Web-page Search Daugbjerg S<sub>1</sub>, Di Bidino R<sub>2</sub>, Fiore A<sub>1</sub>, Cicchetti A<sub>1</sub>







1 Alta Scuola di Economia e Management dei Sistemi Sanitari (ALTEMS), Università Cattolica del Sacro Cuore, Rome, Italy 2 Health Technology Assessment Unit, Fondazione Policlinico Universitario Agostino Gemelli Istituto di Ricovero e Cura a Carattere Scientifico, Rome, Italy

## BACKGROUND

- AI-supported technologies are rapidly developing, and potential to improve healthcare quality at reduced cos
- However, few examples exist of successfully deployed technologies in a real-world context that have been a assessed.

## OBJECTIVES

The objective of this research is to: 1) identify existing Technology Assessment (HTA) methods developed or assess AI supported health technologies, 2) identify r assessment topics or domains relevant for AI-technology 3) take the first step in developing a framework appli new challenges that emerge with the introduction of A

## METHODS

- A systematic literature review of studies describing r frameworks to assess AI supported health technologie performed on PubMed from January 2010 until Februa
- Web-page search of international HTA agencies and organizations such as the WHO, OECD, and the Europe Commission were performed to identify important as consider when implementing and assessing AI techno

## RESULTS

• No assessment frameworks for AI technologies were from the systematic literature review or web-page sea international HTA agencies.

#### Table 1. Topics for assessment of trustworthiness

	Domain	Identified topics
nd have the	Description of	<ul> <li>Full disclosure of sourcecode</li> </ul>
	technical	<ul> <li>Description of model, intention, meth</li> </ul>
ost.	characteristics	training and limitations
		<ul> <li>Algorithm open to critism</li> </ul>
		<ul> <li>Auditable codes</li> </ul>
u AI-	Safaty useuos	Dick mitigation strategies
doguataliz	Salety ussues	• RISK IIIIIgation strategies
laequalery		Onexpected outcome assessment
		Accountability
		<ul> <li>Human oversight measures</li> </ul>
		<ul> <li>Cyper security</li> </ul>
		<ul> <li>Accuracy assessment</li> </ul>
		<ul> <li>Reliability</li> </ul>
		<ul> <li>Data govenance</li> </ul>
g Health		<ul> <li>Transparency</li> </ul>
-	<b>Clinical effectiveness</b>	<ul> <li>Intended use and expected outcomes</li> </ul>
r adapted to		Main functions
		<ul> <li>Indirect outcomes</li> </ul>
new		<ul> <li>Limitation of validity and efficiency or</li> </ul>
		<ul> <li>Outcome and impact assessment</li> </ul>
ogy uptake		<ul> <li>Accuracy according to setting</li> </ul>
<b>3</b> 7 <b>1</b>	Detient and social	Accuracy according to setting
icable for	Patient and social	• Identification of direct and indirect us
	aspects	<ul> <li>Engagement and education of stakeh</li> </ul>
AI.		<ul> <li>Ability of patients to refuse use</li> </ul>
		<ul> <li>Impact on society at large</li> </ul>
	Legal aspects	<ul> <li>Human right issues</li> </ul>
		<ul> <li>Protection of privacy</li> </ul>
		<ul> <li>Product liability and personal liability</li> </ul>
nethods or		<ul> <li>Input liability and liability to data do</li> </ul>
		<ul> <li>Accountability</li> </ul>
es was	Ethical aspects	<ul> <li>Concont of uso</li> </ul>
	Ethical aspects	• Uuman rights
arv 2021		<ul> <li>Doduction of inconvolity and bios</li> </ul>
		<ul> <li>Reduction of inequality and bias</li> </ul>
intornational		Prevention of re-identification
IIIIeIIIdiiOildi		<ul> <li>Preservation and enhancement of his</li> </ul>
000		<ul> <li>Inclusion and diversity throughout the second second</li></ul>
Call		cycle
inoata to		<ul> <li>Identification of potential bias</li> </ul>
specis io		<ul> <li>Accessibility of technology</li> </ul>
ologiog	Transparency	<ul> <li>Traceability (is AI system procedures</li> </ul>
orogres.		Full disclosure of sourcecode
		Description of model, intention, met
		training and limitations
		<ul> <li>Algorithm open to critism</li> </ul>
id antifical		Auditable codes
ideniiied		
archag of	<b>Explainability</b>	<ul> <li>Of technical processess</li> </ul>
arches of		<ul> <li>Of the reasoning behind the decisio</li> </ul>
		system

#### hod of architecture,

over time

users of the tool olders

cy of decision makers ners

iuman autonomy he entire AI systems life

s properly documented)

ethod of architecture,

ons/predictions of the AI

 Reports from international organizations highlights limitations or inability of most AI technologies to 'explain' their decision-making process (Black box issue), leading to lack of trust in the technology and affect the adaptation.

Figure 1. Framework of association between trustworthy AIsolution and uptake (simplified)



- To create trust in the AI-solution an assessment of the technology trustworthiness should be included in the overall technology assessment (see Figure 1).
- Table 1 shows a preliminary list of identified topics to be included in the assessment of AI-supported technologies to address the trustworthiness of the technology.
- Overall, more emphasis should be put on assessing:

Transparency and Explainability of the AI solution

> Aspects related to safety, ethical, legal and social issues related to implementation and the development and training phase of the AI technology

## CONCLUSION

• The results from this study uncovers key gaps posed for performing a systematic and holistic assessment of AI in a real-world context of health care.



• The results will form the basis for the development of a framework to assist decisionmakers in assessing AI supported technologies in a holistic manner for a responsible deployment – the HTA AI Framework.

### ACKNOWLEDGMENT

This work was performed as a part of the European project AI-Mind. It has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964220



PhD Signe Daugbjerg Signeb.daugbjerg@unicatt.i