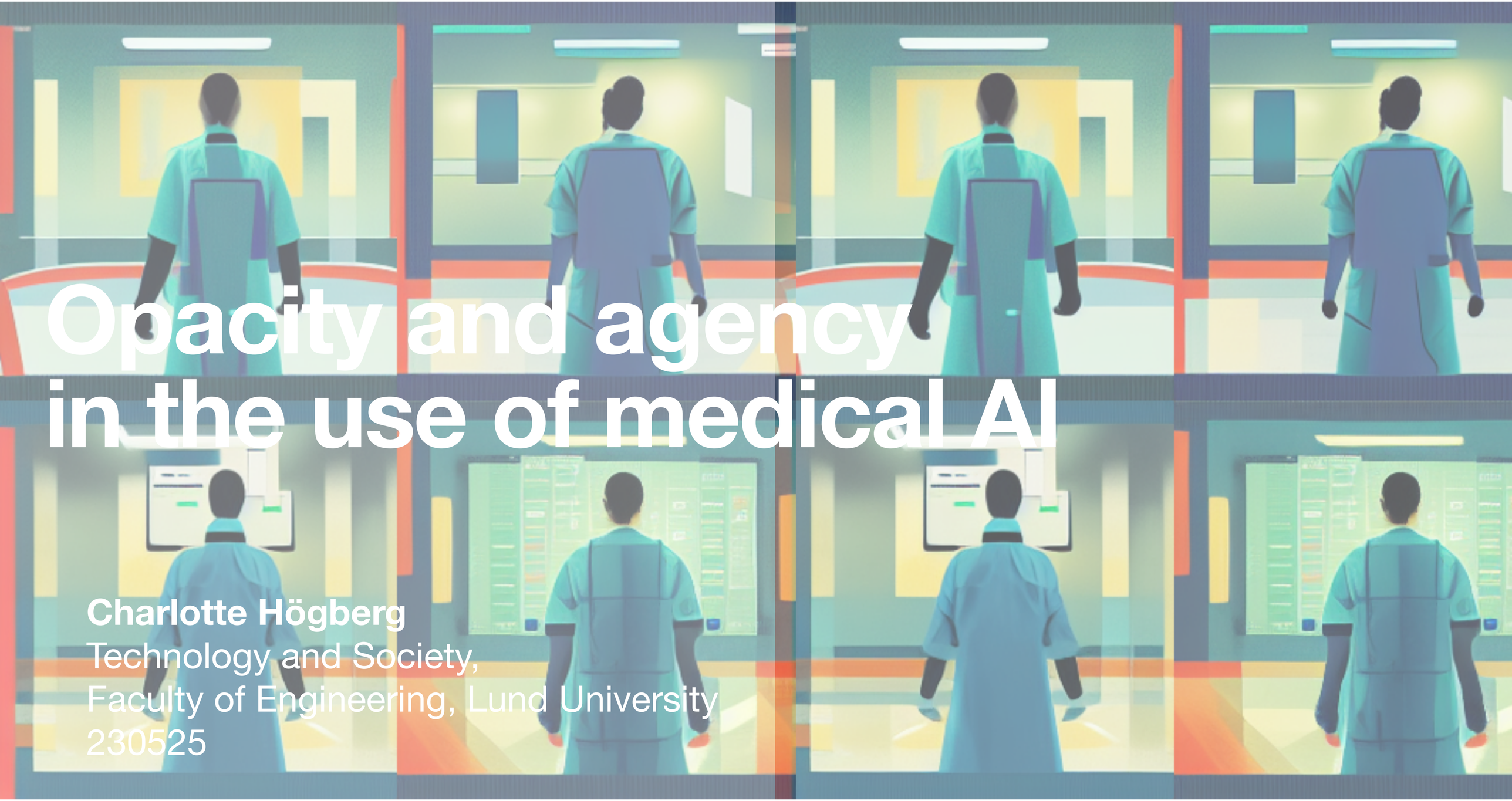


Opacity and agency in the use of medical AI

Charlotte Högberg
Technology and Society,
Faculty of Engineering, Lund University
230525

AIR Lund

Artificial Intelligent use of Registers

MASAI

Mammography Screening with AI

- Interdisciplinary
- My project:
    - Notions of information, and knowledge, needed about the use of AI (transparency & explainability in the use of applied AI)
    - Notions of information and knowledge made through the use of AI (explainability in knowledge production and AI as epistemic actant)
- Transparency, trust & the human judgement in healthcare
- Transparency & truths in knowledge production in medicine

AI transparency

EU High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI*, Transparency– the principle of explicability Encompasses transparency of: the data, the system and the business models (p. 18)

- Traceability
- Explainability
- Communication

AI transparency in healthcare

Key principle of World Health Organization's guidance on ethics and governance of AI for health: ensure transparency, explainability and intelligibility:

"AI technologies **should be intelligible or understandable** to developers, medical professionals, patients, users and regulators. Two broad approaches to intelligibility are to **improve the transparency of AI technology and to make AI technology explainable**."

World Health Organization, Ethics and governance of artificial intelligence for health: WHO guidance. p. xiii

Transparency

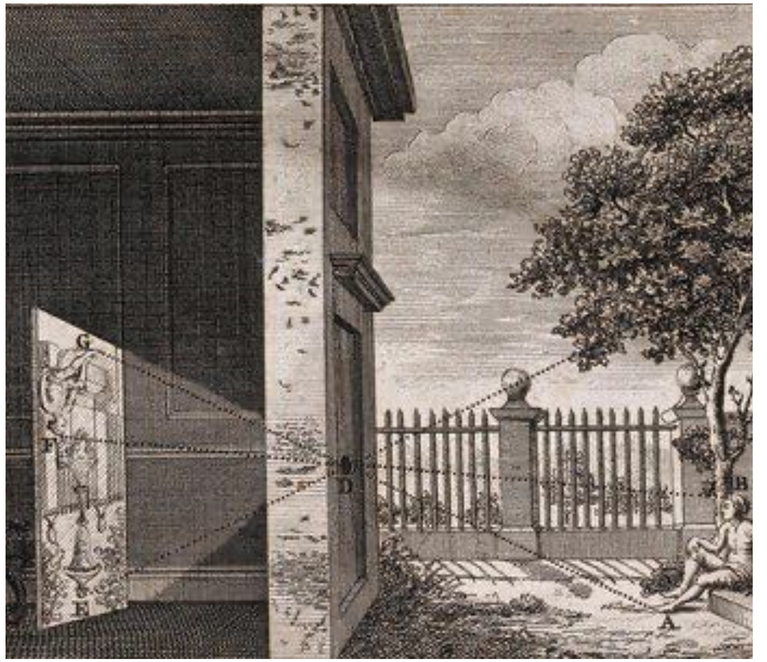
- What can be seen can be known (Ananny and Crawford, 2018; Larsson and Heintz, 2020).

- High hopes –
  fairness, accountability, robustness, etc.

- Mend opacity of the ”black box” (Burrell, Ferretti et al).
  Reveal structural & discriminatory bias,
  Empower individuals (de Vries 2021)

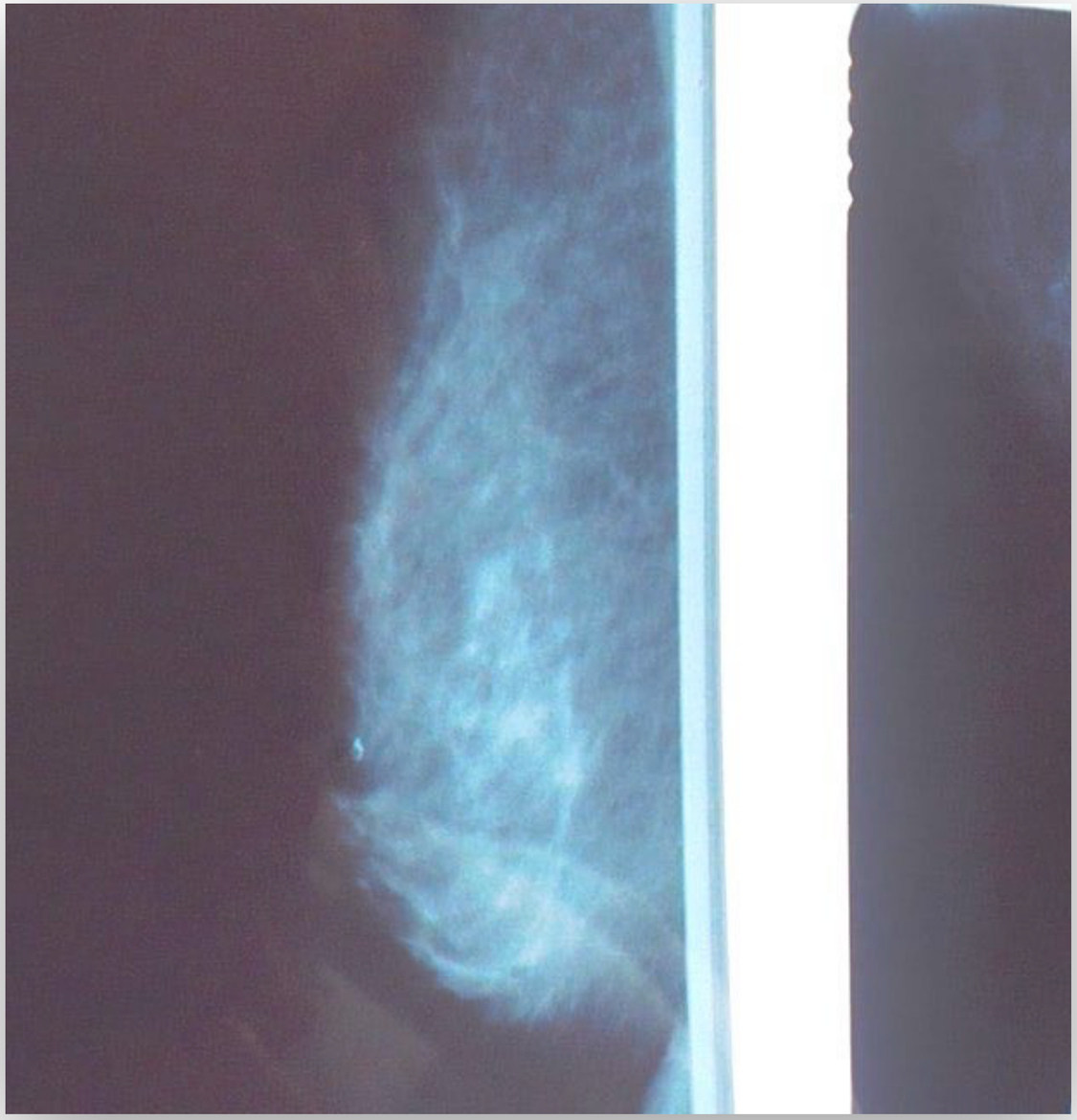
- Faulty binary, instead: Multiple, situated translucensies (Lee 2021)

- Management of visibilities (Flyverbom 2019)
  Not a state= Performative, processual (Cellard 2022)



Studies

- **Patients' rights**
    - Document analysis
- **Medical professionals**
    - Survey, interviews, observation
- **AI-developers**
    - Interviews & observations with AI-developers/AI-researchers/ Medical researchers developing AI for medical research/healthcare



Information-related principles potentially impacted by AI use

1. **The right to equal healthcare**
    - e.g., Non-discrimination
2. **The right to privacy and personal integrity**
    - e.g., Info about use of data
3. **The right to receive information**
    - e.g., Explanation
4. **The right to dignity and autonomy**
    - e.g., Informed consent
    - e.g., Right to refuse automatic processing

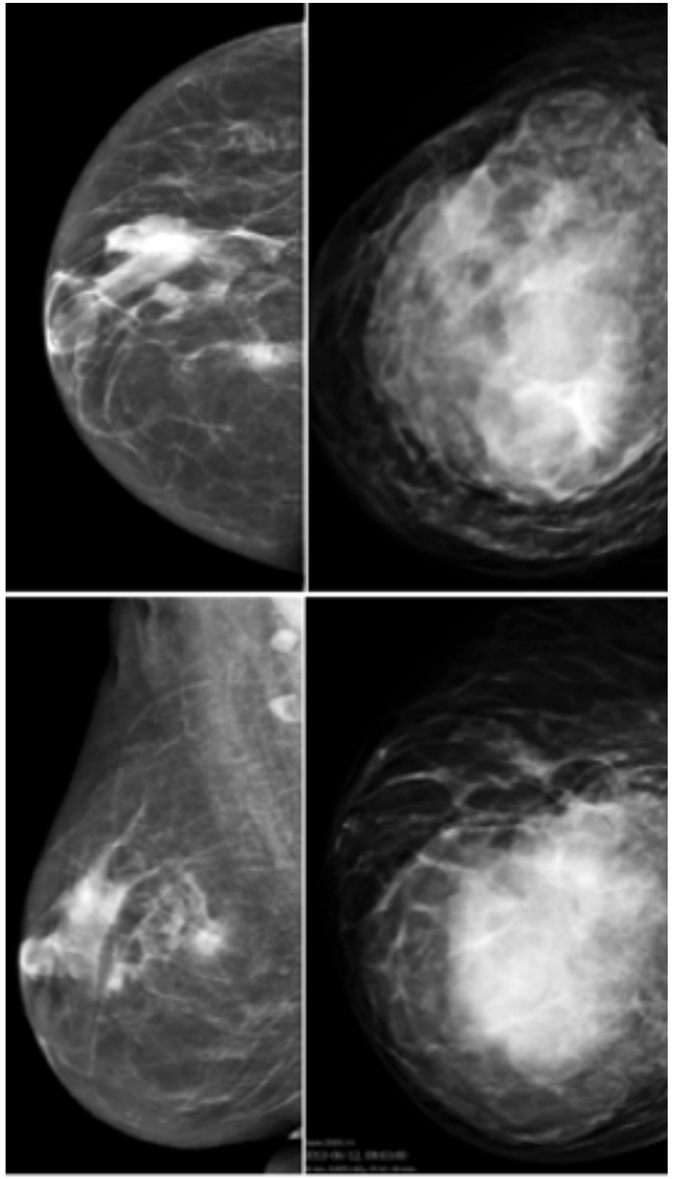
Analysis based on: Medical ethics (Declaration of Geneva), EU regulation (GDPR, MDR),
Swedish regulations (Health and Medical Service Act, Patient Act, Patient Safety Act and Patient Data Act)

- Some level of AI transparency is needed to safeguard ethical and legal principles of provision of healthcare and patients' rights
- Meaningful transparency – specific to context and situation
  - No fits all solution to AI transparency in healthcare
  - Incorporated into current information practices in healthcare

Högberg, C & Larsson, S 2022, AI and Patients' Rights: Transparency and information flows as situated principles in public health care. i K de Vries & M Dahlberg (red), De Lege – Yearbook Uppsala Faculty of Law 2021: Law, AI & Digitalization. De Lege, vol. 2021, Iustus förlag, Uppsala, s. 401-429.

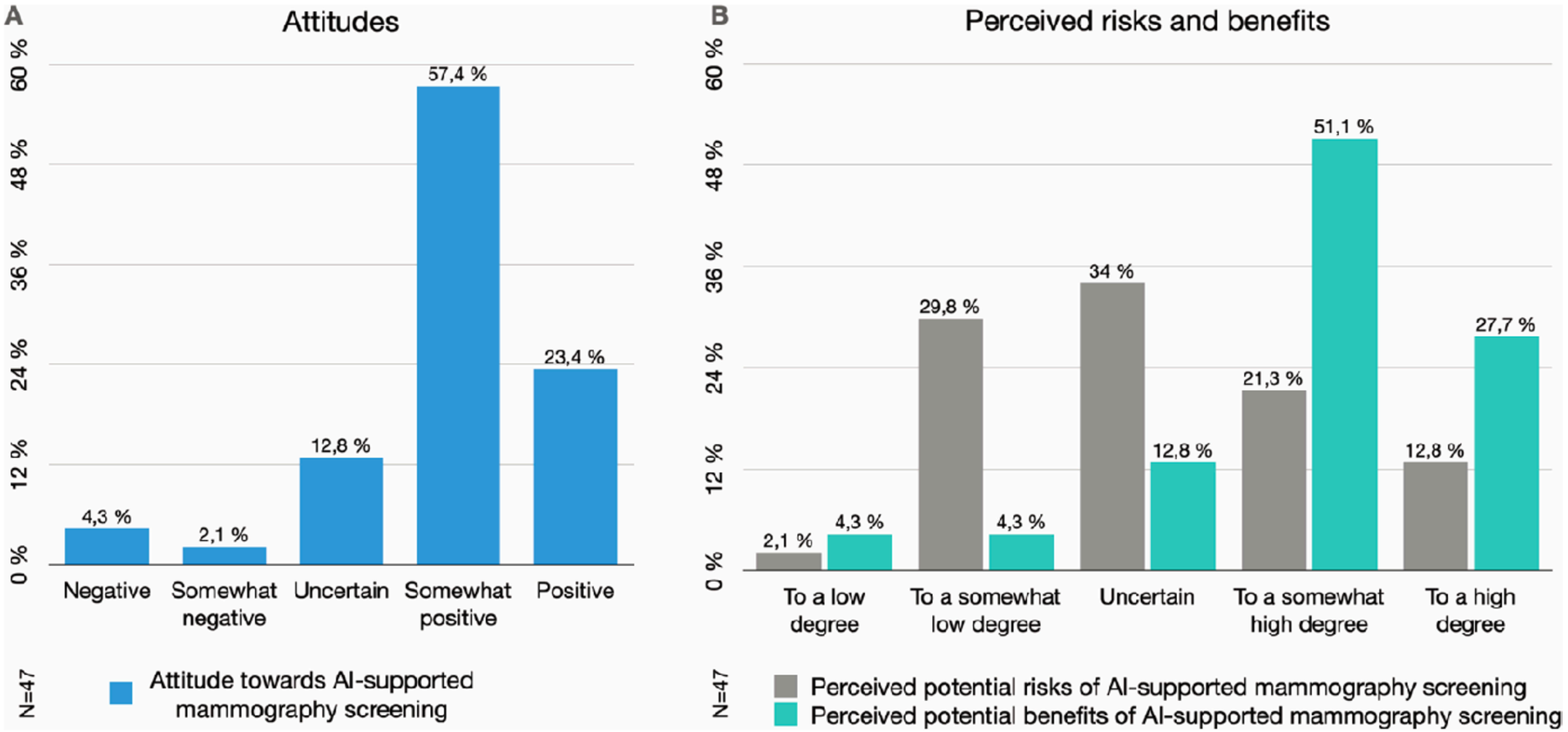
“A diagnostic technique never merely registers facts. It intervenes in the situations in which it is put to use” (Mol 2000, p.19).

- Introduced in sociotechnical relations and organizations
- Epistemic setting (Knorr-Cetina, K. 1999)



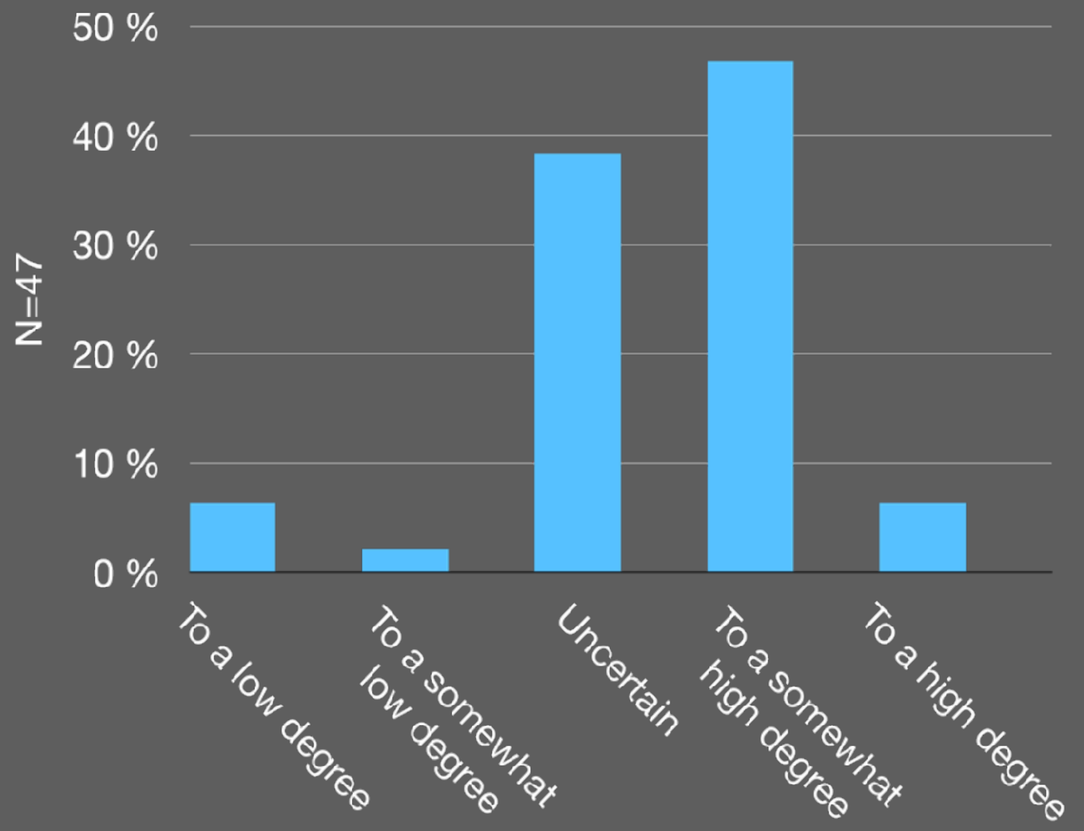
AI in mammography screening

Swedish breast radiologists: Positive, but in need of clarifications

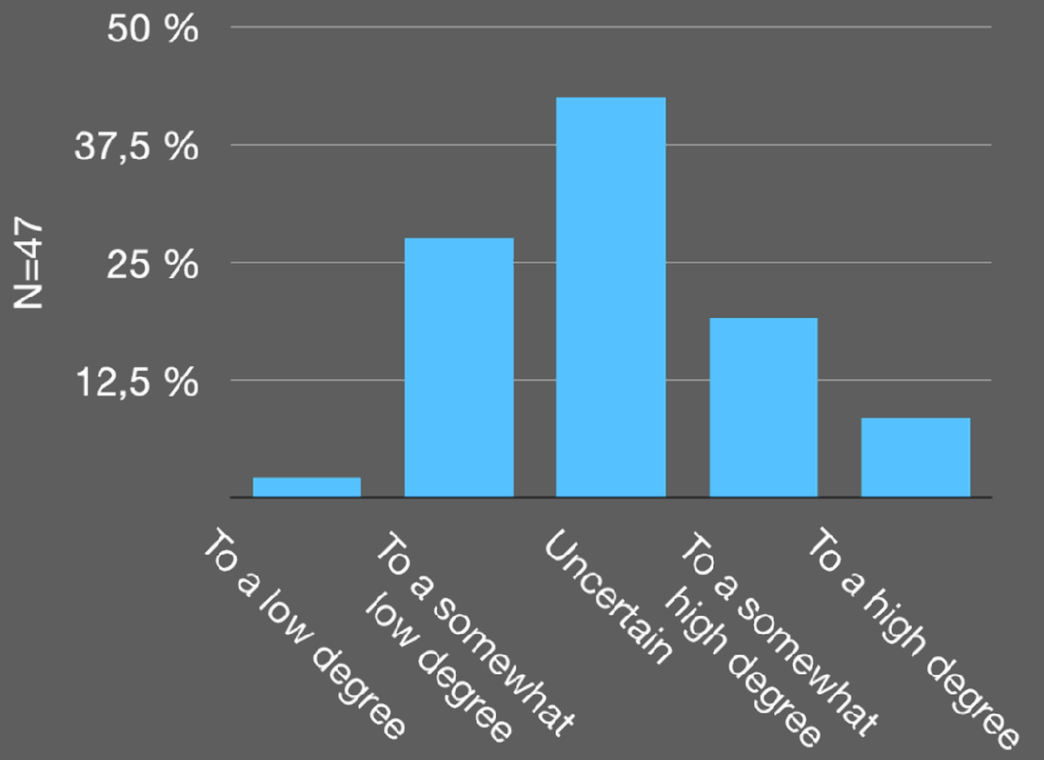


Högberg C, Larsson S, Lång K. Anticipating artificial intelligence in mammography screening: views of Swedish breast radiologists. BMJ Health and Care Informatics 2023;30(1) doi: 10.1136/bmjhci-2022-100712

Trust in assessments made by AI-systems



Perceived risk of radiologists having overconfidence in assessments by AI-systems

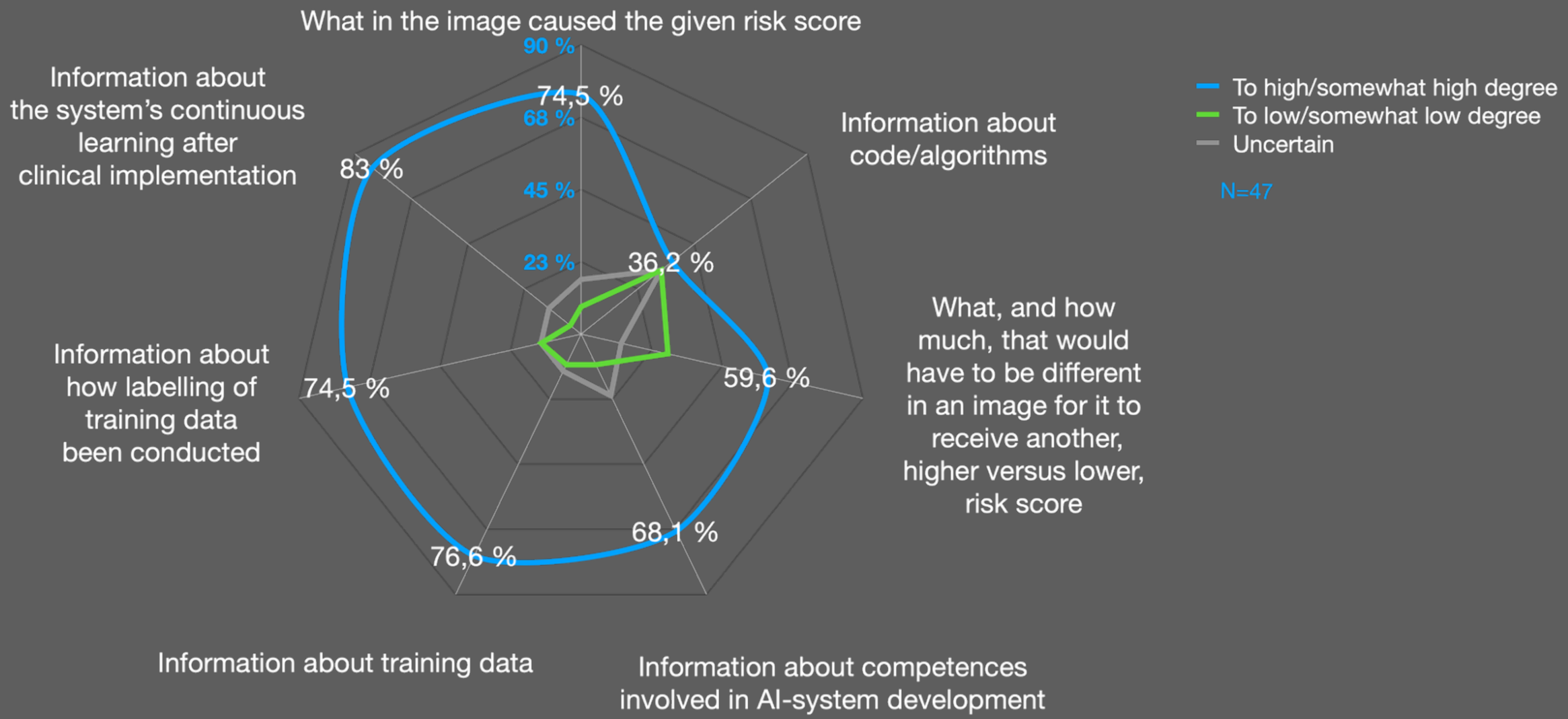


“My experiences are that you can to somewhat high degree trust AI-assessments in simple easily assessed cases. But when the breast tissue is a little more difficult to assess, you can to somewhat low degree trust the AI-assessment.”

“Need to see with my own eyes…”

“you need to test the system on your own material first”

Would the following information support your evaluation of trust in an AI-system's assessments in mammography screening?



- How to evaluate AI in medical decision-making?
- Not something to consider after certification?

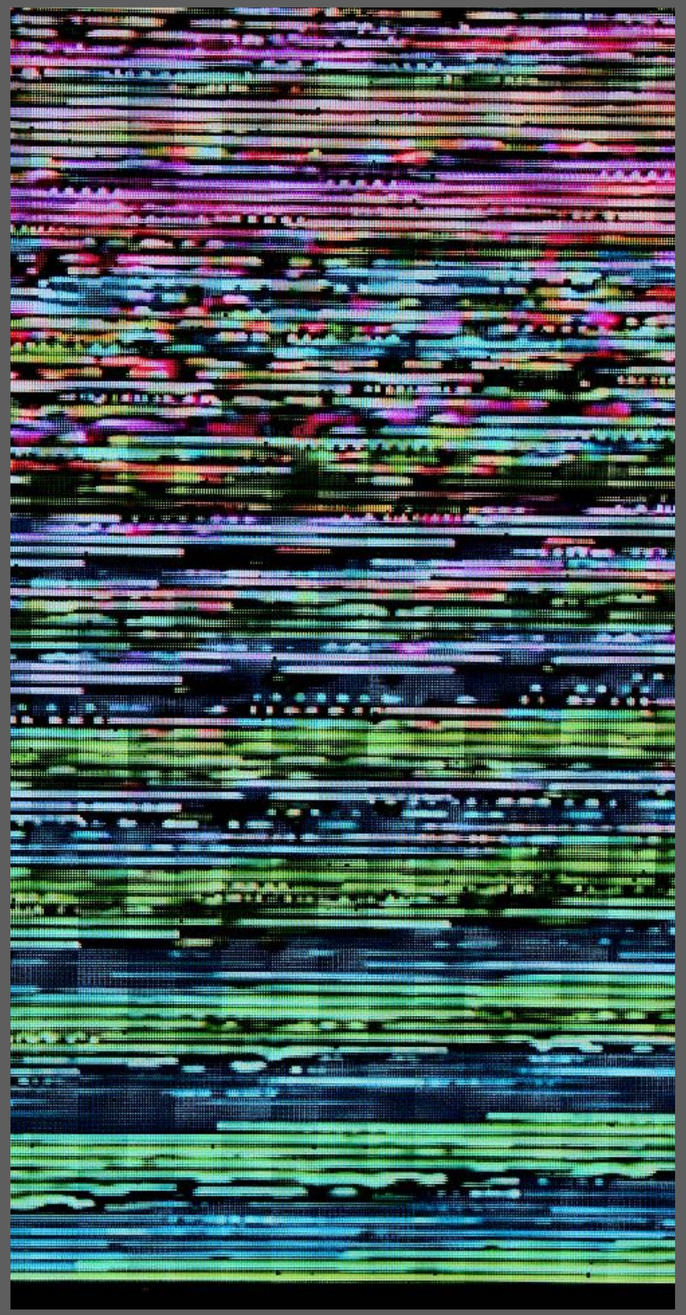
"It is not really the radiologist's task, as little as there are few radiologists who understand the physical foundation and the algorithms that are used for example when producing MR-images."

- Different notions of the role of human expertise

Introduction of x-rays

”Thus technology gradually became part of the art; it became a 'blackbox' both in radiological practice and in the publications of X-ray workers. It became a controllable part of the heterogeneous mixture that made the pictures diagnostic”

(Pasveer, 1989, p. 376)



Introduction of
CT-scanners

"the scanners' technical complexity and the radiologists' lack of familiarity with CT's diagnostic signs threatened the in-experienced radiologists' authority and forced them to rely more heavily on the technologists. However, […] complexity and uncertainty are functions of how the machine merged with the social system; they are not attributes of the machine itself.”

(Barley, 1986, p. 106)

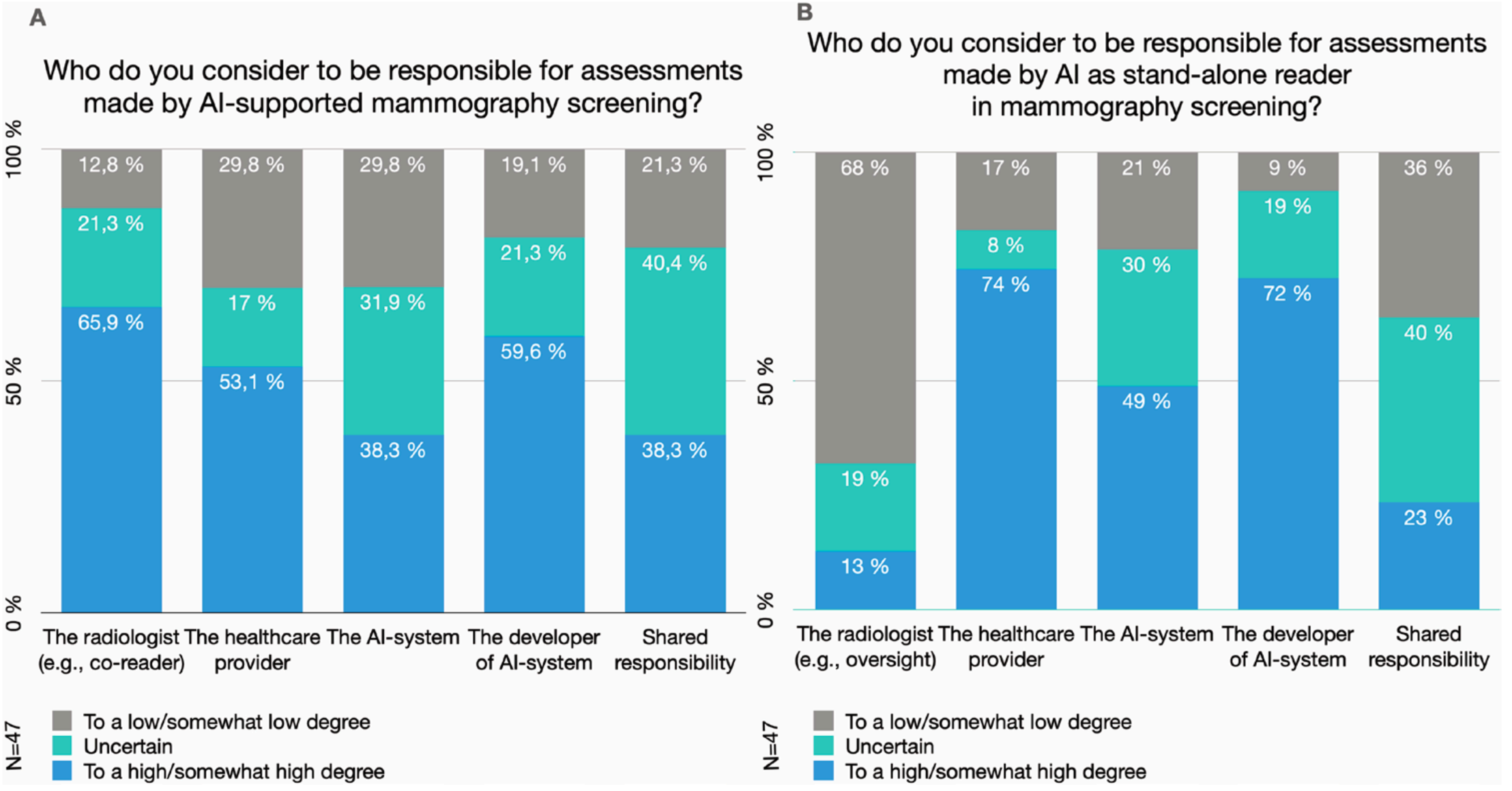
1960-1970s future visions

”Doctors spoke of the ***creeping amoeba of automation***, of the electronic brain that threatened their status, even their livelihood. [...] He [Jerold Maxmen] predicted [in 1976] that by about the first quarter of the twenty-first century, ’doctors would be rendered obsolete,’ ***replaced by a ’medic-computer symbiosis.’*** His model assigned to computers most diagnostic and therapeutic decisions...” (Reiser, 1978, p. 224-225).

AI – a big difference?

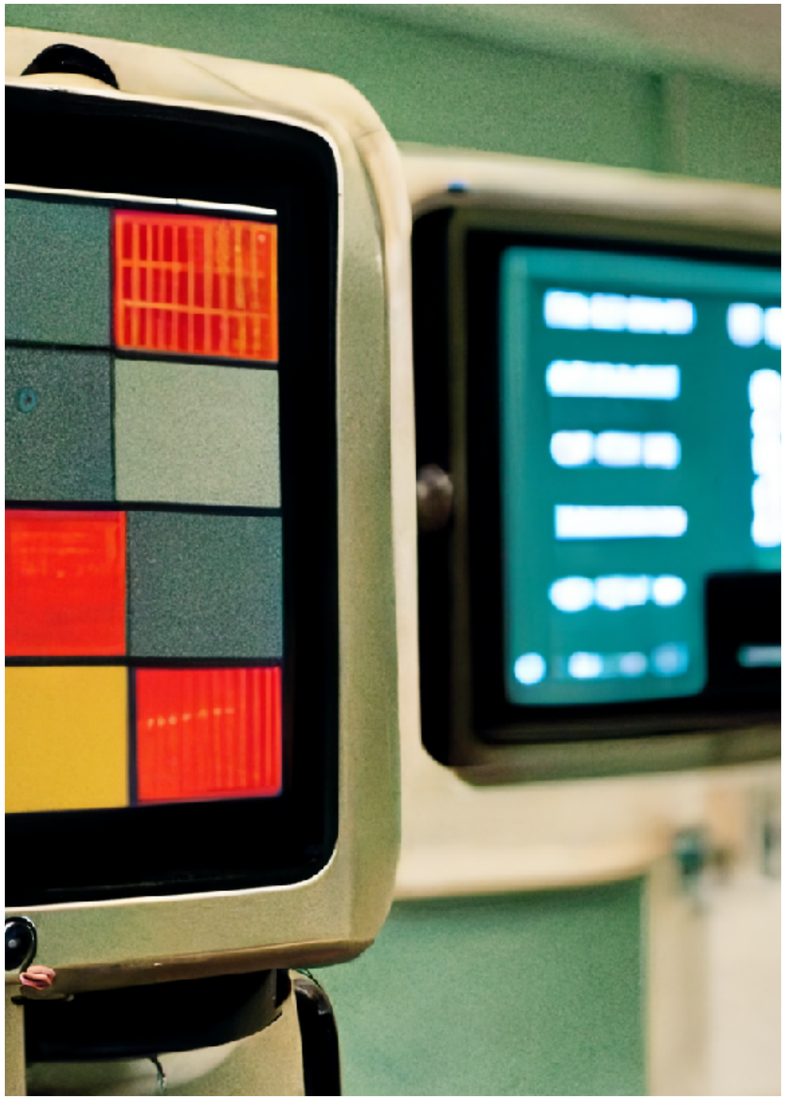
- Nearly half of the breast radiologists (n=21, 44.7%): integrating AI in mammography screening = substantial difference, in comparison to previously introduced technologies (such as digital mammography and tomosynthesis), to a high/somewhat high degree.
- more than one-third (n=17, 36.2%) were uncertain.

- Previous tech dev about improving image quality to support radiologists' assessments, AI about delegating assessments and decisions to technology



Högberg C, Larsson S, Lång K. Anticipating artificial intelligence in mammography screening: views of Swedish breast radiologists. BMJ Health and Care Informatics 2023;30(1) doi: 10.1136/bmjhci-2022-100712

Informed AI trust assessment work			
Matter of possibilities	Explainability (UI, risk score, visual markings, confidence levels etc.)	Competence, literacy (competence-enhancing initiatives, AI-specific knowledge)	Transparency, information seeking and source evaluation (active practice, info. attainable)
Matter of responsibilities	Assessment of screening exams (images, AI-findings, patient info.)	Critical evaluation of technological support and outputs	Liability in medical decision-making and time management



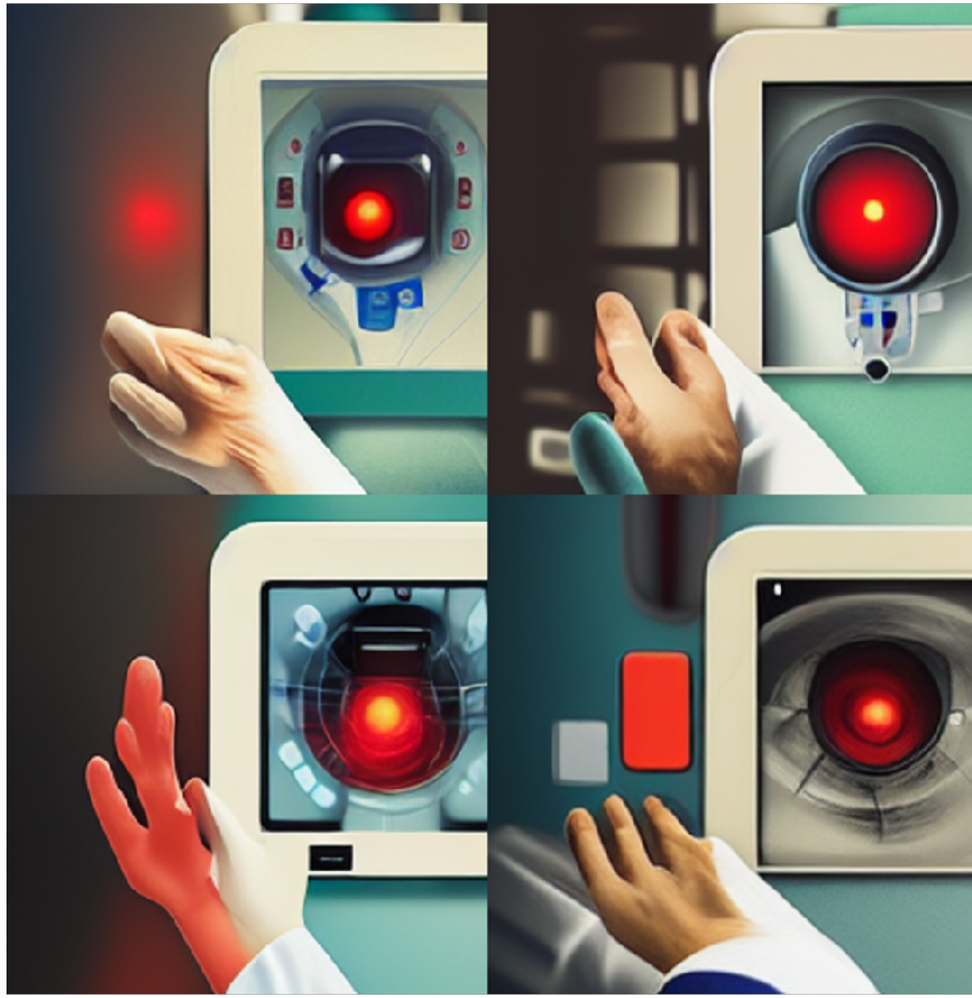
What is the role of the human? What does she need to know?

- What is expected of the human-in-the-loop? Human as critical safeguard
  (Enarsson, T., Enqvist, L. & Naarttijärvi, M. (2022)

- What transparency/explainability of AI is wanted/needed? For whom?

- Unclear what transparency of AI in healthcare is supposed to be, what purpose it is serving, and what level of critical engagement with AI output that is expected

- If we say that trust is missing – what do we want trust to contribute with?



Comments? Thank you!

charlotte.hogberg@lth.lu.se

Swedish Research Council

CANCERFONDEN

REGIONALA CANCERCENTRUM I SAMVERKAN

WASP—HS