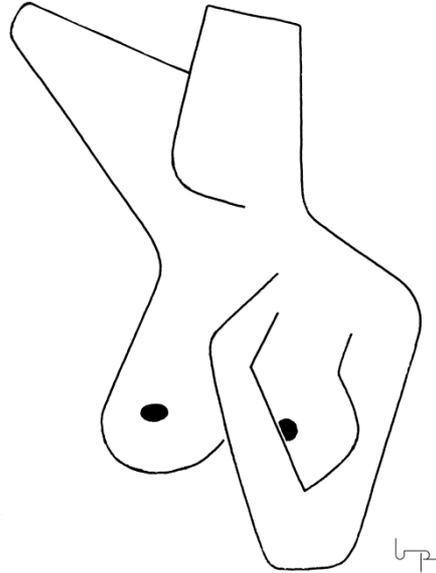


Can Artificial Intelligence (AI) completely replace human reader in mammography screening program?: A retrospective evaluation with Digital Mammography (DM) and Digital Breast Tomosynthesis (DBT)



***Esperanza Elías Cabot*, Sara Romero Martín*, José Luis Raya Povedano*,
Albert Gubern Mérida**, Alejandro Rodríguez Ruiz**, Marina Álvarez Benito*.***

**Reina Sofía University Hospital (Córdoba, Spain).*

***ScreenPoint Medical (Nijmegen, Holanda)*

PURPOSE

- To retrospectively evaluate an **AI** system as a **stand-alone reader** in a **breast cancer screening program** with **digital mammography (DM)** and **digital breast tomosynthesis (DBT)**.

HYPOTHESIS

- The **AI** system could achieve a **similar sensitivity to radiologists** with an **acceptable recall rate**.

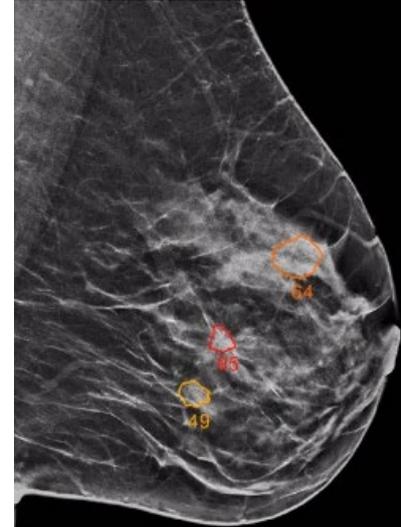
ARTIFICIAL INTELLIGENCE SYSTEM

The AI system (Transpara 1.7.0, ScreenPoint Medical BV) automatically detects lesions suspicious of breast cancer in DM and DBT.

- **REGION SCORE**: score from 1 to 100 in each region marked according to the probability of malignancy.

- **EXAM SCORE**: each exam is given an overall score from 1 to 100 which corresponds to the region assigned the highest score.

Exam score: 95



MATERIAL AND METHODS

Potentially eligible population:
16067 women
(Tomosynthesis Cordoba screening Trial)
113 cancers
(98 screen-detected cancers and 15 interval cancers)

Excluded exams:
(Images not retrieved from PACS)
68 exams
(no cancers)

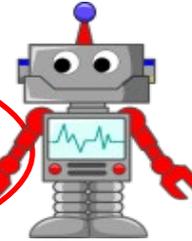
Study population:
15999 exams (DM y DBT)
113 cancers
(98 screen-detected cancers and 15 interval cancers)

European Radiology (2018) 28:2484–2491
<https://doi.org/10.1007/s00330-017-5219-8>

BREAST

Prospective study aiming to compare 2D mammography and tomosynthesis + synthesized mammography in terms of cancer detection and recall. From double reading of 2D mammography to single reading of tomosynthesis

Sara Romero Martin¹ · Jose Luis Raya Povedano¹ · Maria Cara Garcia¹ · Ana Luz Santos Romero¹ · Margarita Pedrosa Garriguet¹ · Marina Álvarez Benito¹



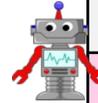
- ✓ All of the included exams were analyzed using our AI system.
- ✓ Only the cancers correctly localized and marked with the highest score were considered true positives.

MATERIAL AND METHODS

ORIGINAL SCREENING SETTING

DM		Single reading of DM
DM		Double reading of DM
DBT		Single reading of DBT
DBT		Double reading of DBT

VS



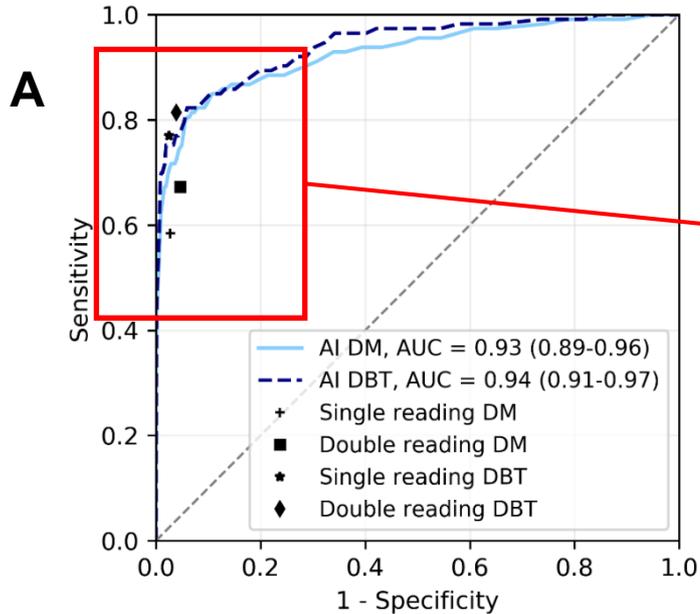
ARTIFICIAL INTELLIGENCE SYSTEM

DM	Cut-off 80*
DM	Cut-off 74*
DBT	Cut-off 65*
DBT	Cut-off 57*

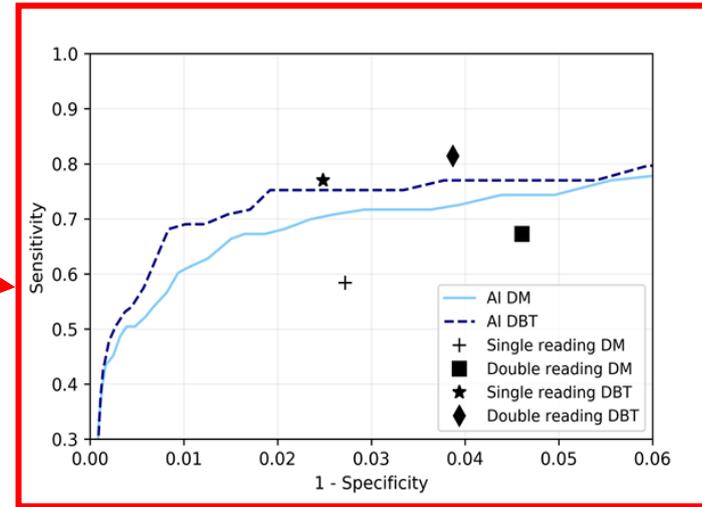
- We compared the AI system performance as a stand-alone reader for DM and DBT with the original human readings.
- AI stand-alone performance was measured using the area under the receiver operating characteristic curve (AUC). The highest region score found in the exam was used as exam score. 95% CIs were computed using bootstrapping.
- We investigated whether AI as a single independent reader in DM or DBT screening could achieve similar sensitivity with an acceptable recall rate. The recall rate of AI and the human readings were compared using a McNemar test.

(*)Different operating points selected to have non-inferior sensitivity as the human readings (non-inferiority margin 5%)

MATERIAL AND METHODS



B



A) Receiver operating characteristic curve (ROC) of AI system for DM and DBT exams. The AI system achieved an AUC of 0.93 (95%CI :0.89-0.96) for DM and 0.94 (95%CI :0.91-0.97) for DBT.

B) Zoomed image of the ROC curve of AI system shows the range when radiologist operated.

RESULTS

DM		 Cut-off 80	P
Sensitivity	58.4% (66/113)	62.8% (71/113)	0.458
Recall rate	3.1%	 1.7%	<0.001

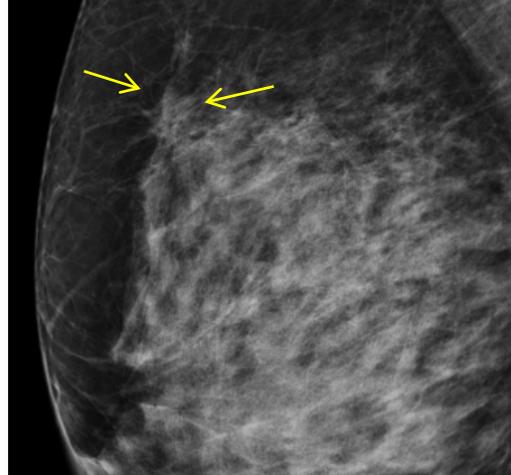
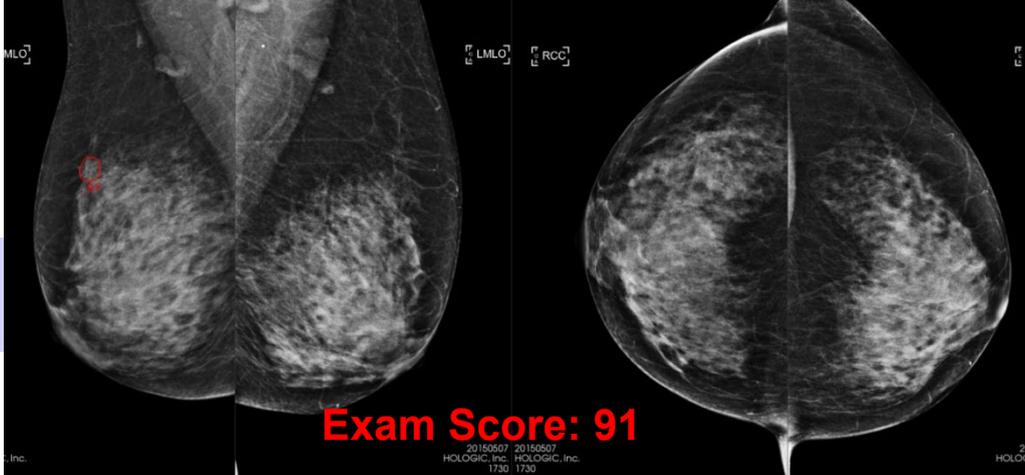
DM		 Cut-off 74	P
Sensitivity	67.3% (76/113)	70.8% (80/113)	0.523
Recall rate	5.1%	 3.1%	<0.001

DBT		 Cut-off 65	P
Sensitivity	77% (87/113)	80.5% (91/113)	0.648
Recall rate	3%	 9.2%	<0.001

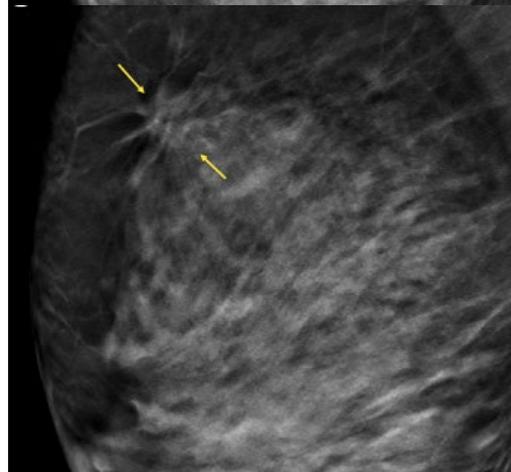
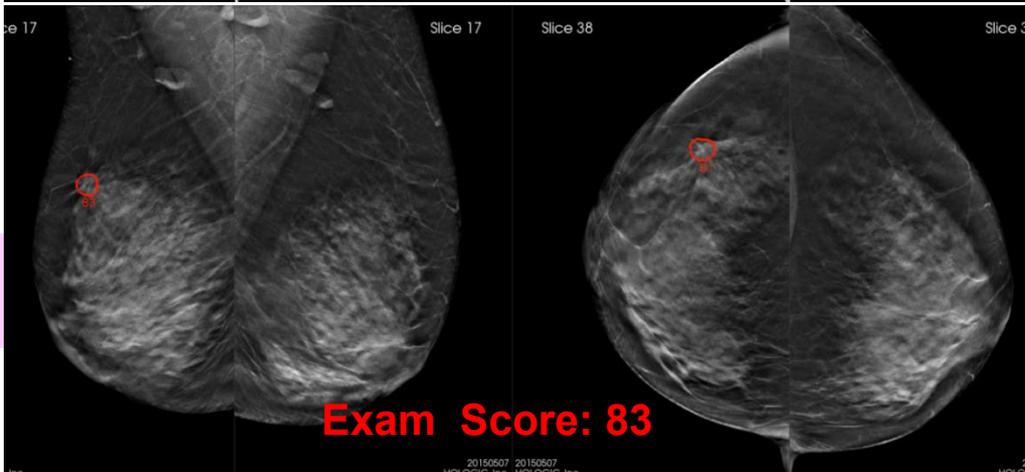
DBT		 Cut-off 57	P
Sensitivity	81.4% (92/113)	85% (96/113)	0.481
Recall rate	4.4%	 16.7%	<0.001

- ✓ AI in DM achieved non-inferior sensitivity as a single or double reading with a significantly reduction in recall rate.
- ✓ AI in DBT achieved non-inferior sensitivity as a single or double reading but with a higher recall rate.

DM



DBT



**A 67-year-old woman recalled only by the original DBT readings.
This study would be recalled both in DM and DBT by the AI. Grade II invasive lobular carcinoma of 18mm was diagnosed at percutaneous biopsy.**

CONCLUSIONS

- In breast cancer screening with digital mammography, artificial intelligence could replace human readings, reducing the recall rate with non inferior sensitivity.

- In breast cancer screening with digital breast tomosynthesis, however, to maintain the sensitivity of radiologists, the recall rate would be substantially increased.

THANK YOU FOR YOUR ATTENTION

