

Transforming Breast Cancer Screening with AI

Dr Helen Frazer

State-wide Clinical Director - BreastScreen Victoria
Clinical Director – St Vincent's BreastScreen
Lead Investigator – BRAIx program

22 February 2025



Learning objectives

Breast cancer screening today

Accuracy of screening with AI assistance

Future role of AI determining risk and personalising screening

Breast cancer screening

Australia introduced breast cancer screening in 1991

- Targets 50-74 years, available from 40 years
- 1 m pa, interval every 2 years
- Ongoing quality assurance
- Digitised since 2012

Successful public health initiative

- 41-52% reduction in mortality for women who attend breast cancer screening

One of only 22 countries globally with a national population screening program

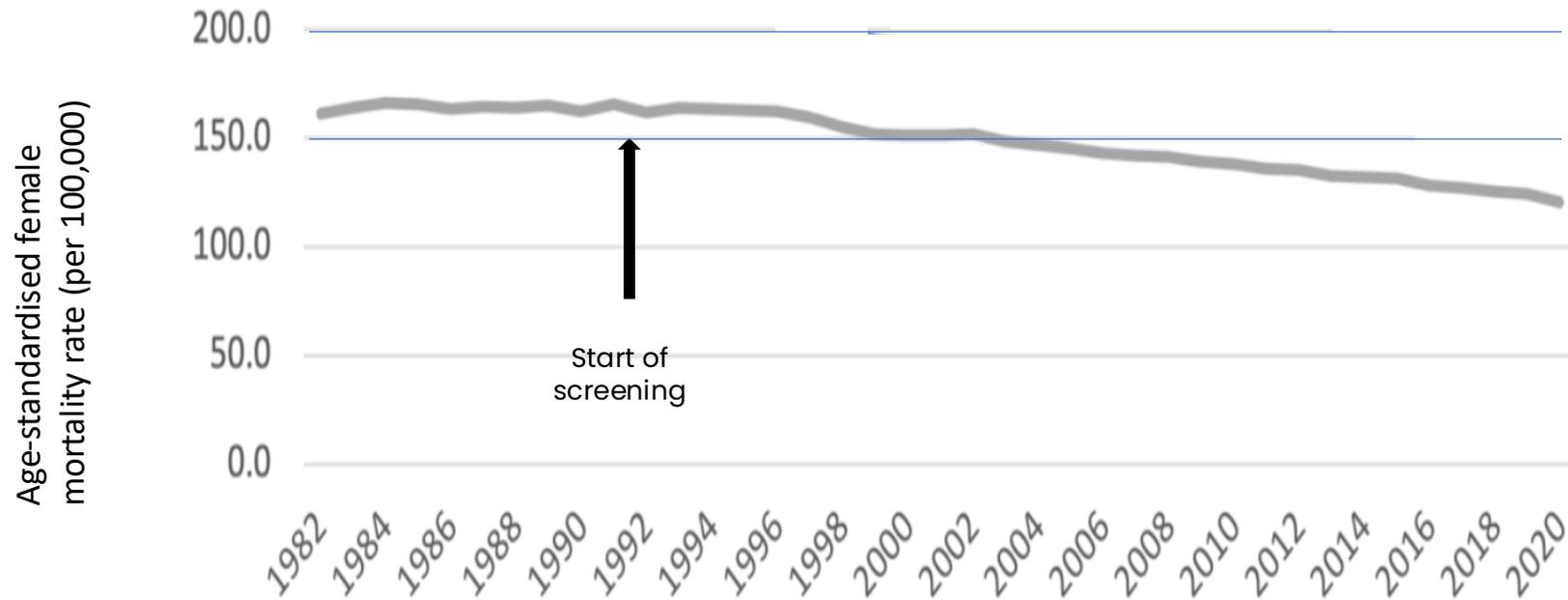


SOURCE:

Australian Institute of Health and Welfare 2020. BreastScreen Australia monitoring report 2020. Cancer series no. 129. Cat. no. CAN 135. Canberra: AIHW

Morrell, S., Taylor, R., Roder, D. & Dobson, A. Mammography screening and breast cancer mortality in Australia: an aggregate cohort study. *Journal of medical screening* 19, 26–34 (2012).

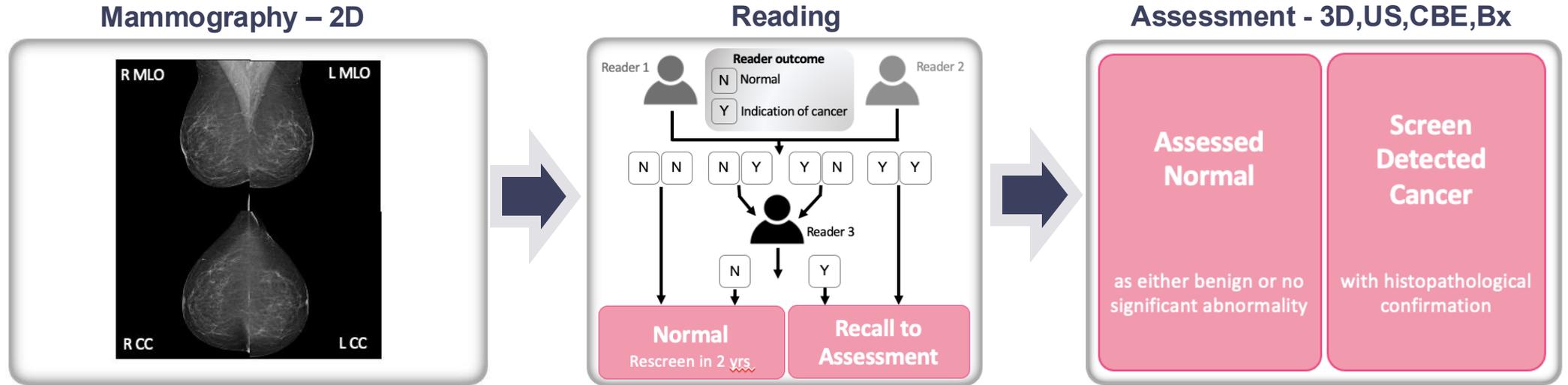
How do we take the next step change in mortality?



- Address modifiable causes (better prevention)
- Increase participation, reducing intervals (better screening)
- Improve survival (better treatments)

Screening has challenges with a “one size fits all” approach

PROCESS



CHALLENGES

ACCURACY

False positives (unnecessary assessment)

False negatives (Interval cancers)

EXPERIENCE

Flat participation

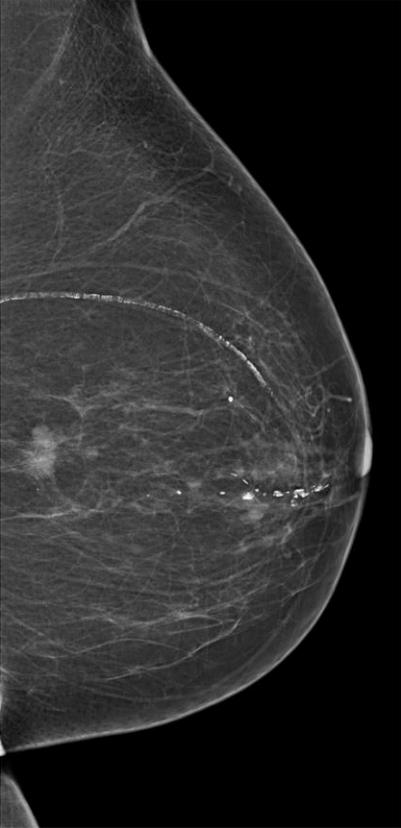
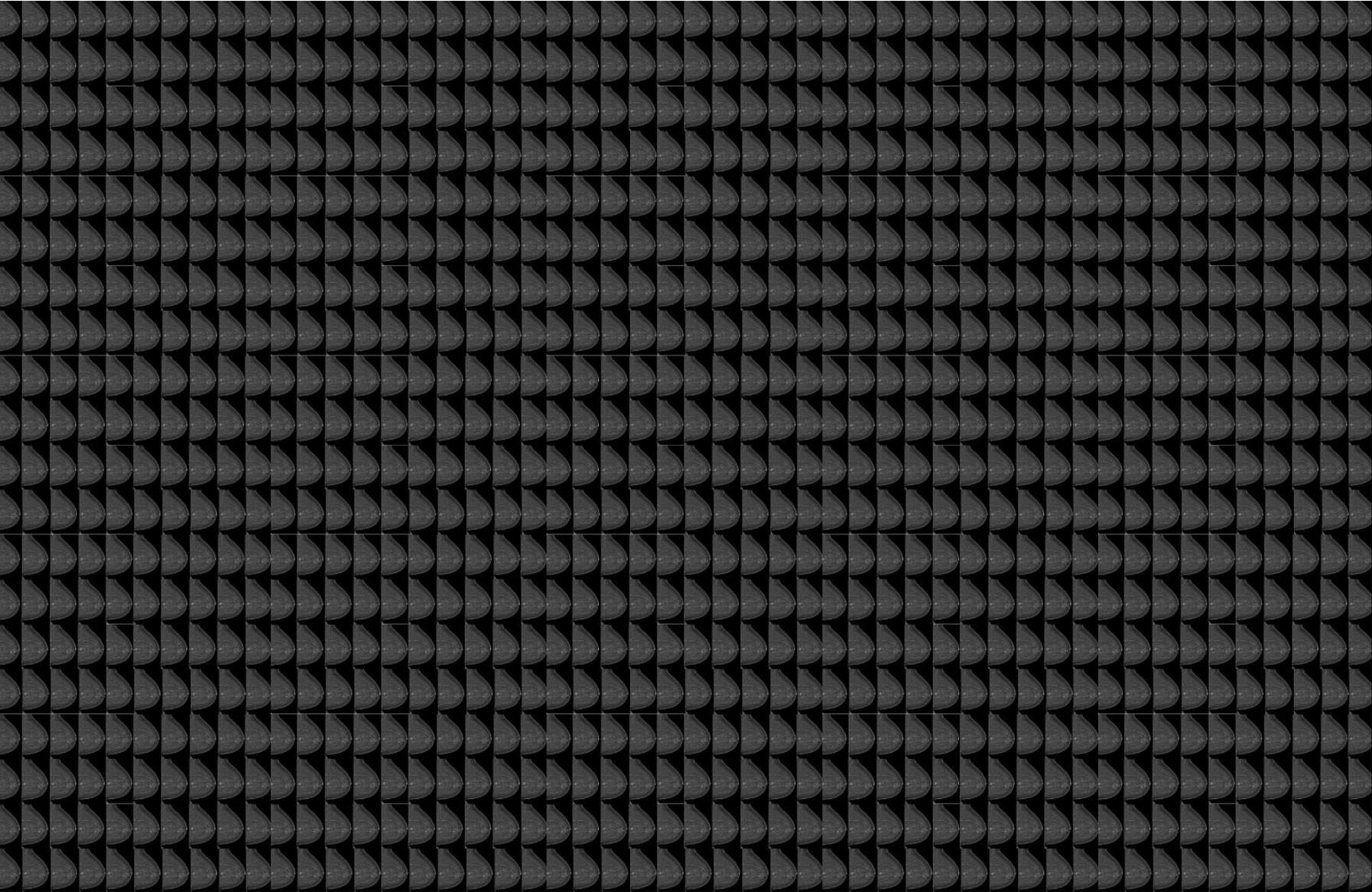
Service Level

COST

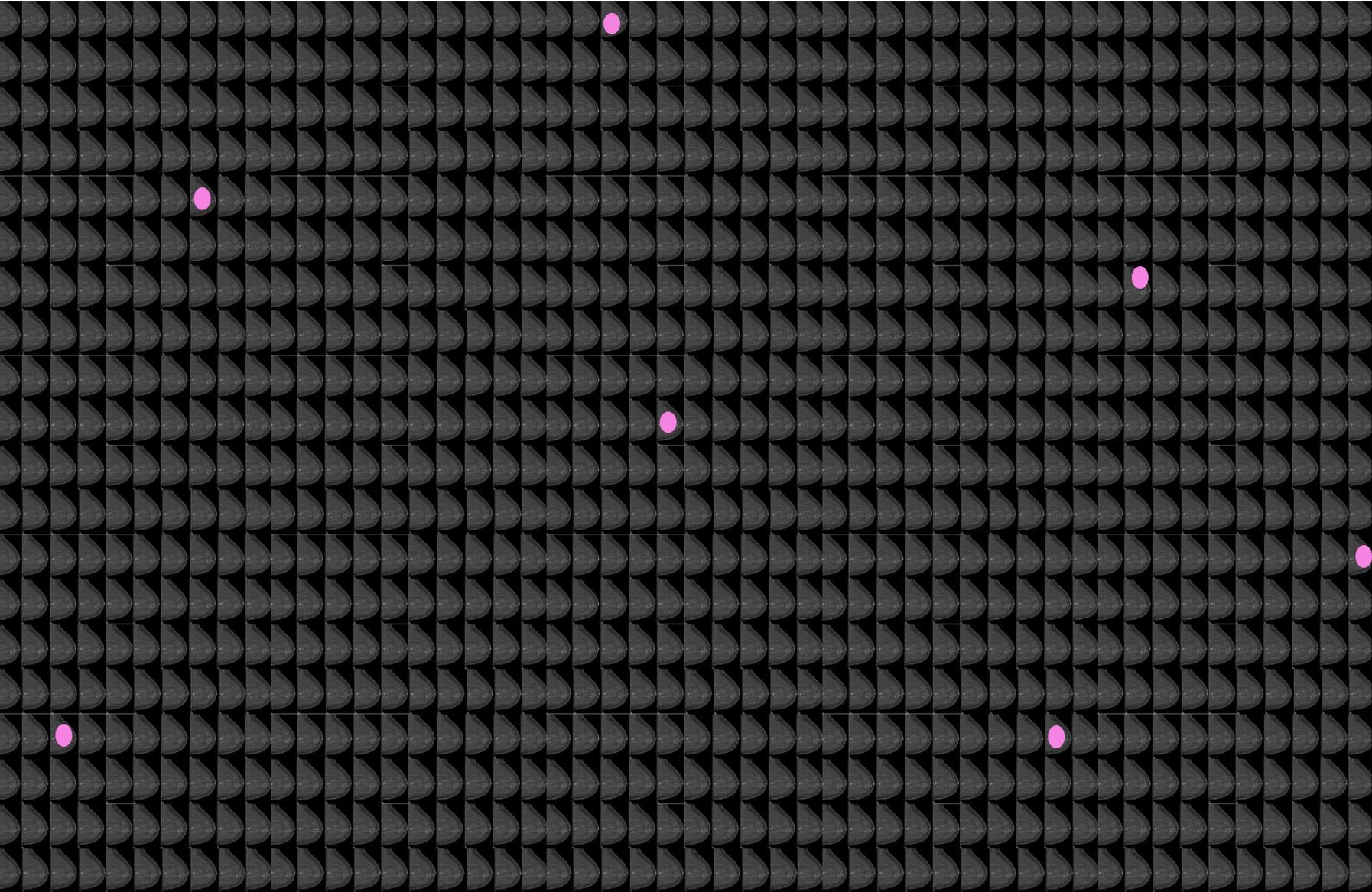
High Cost

Workforce availability

For every 1000 women screened

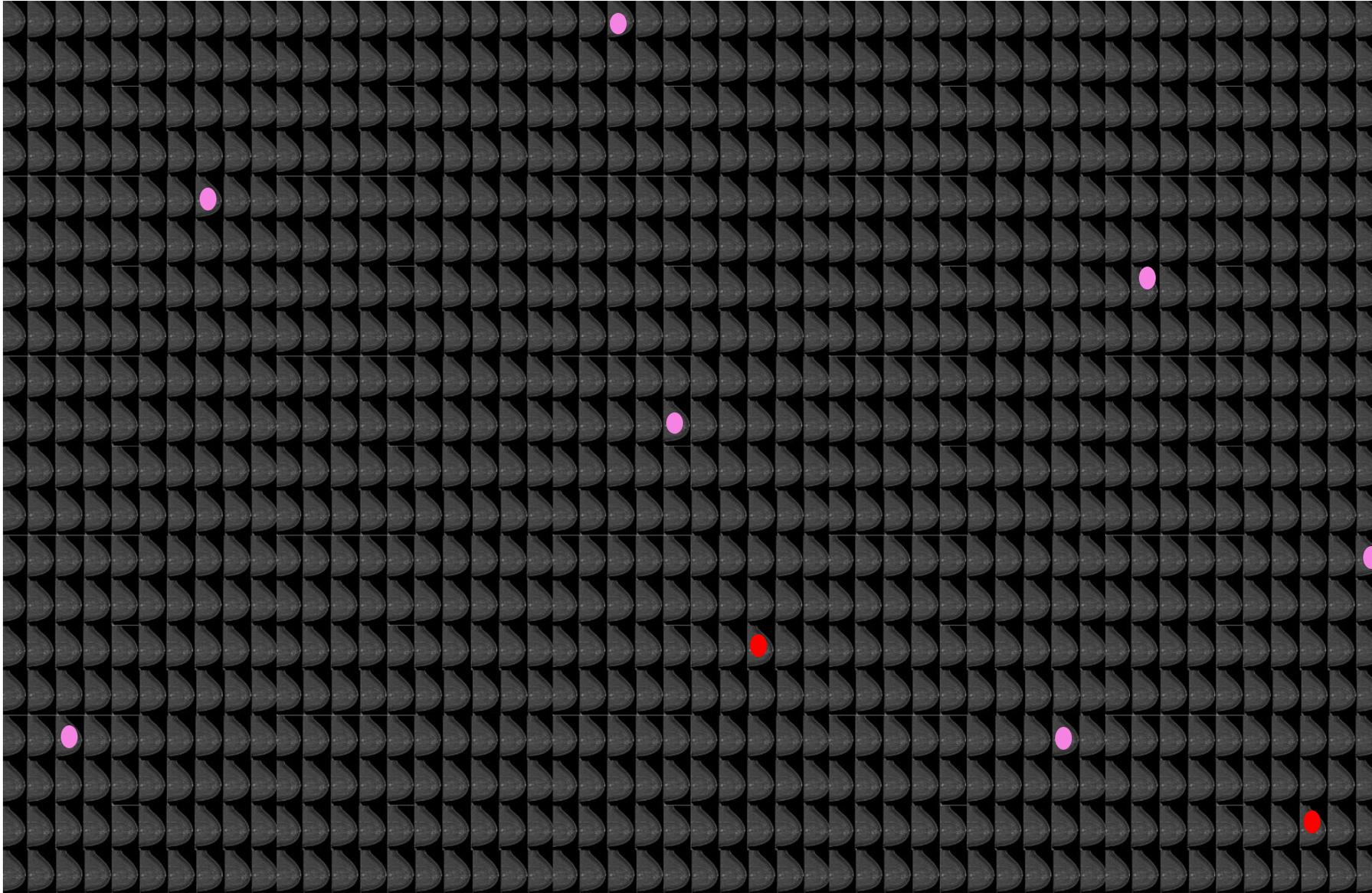


For every 1000 women screened



**7 cancers
are detected
(true positives)**

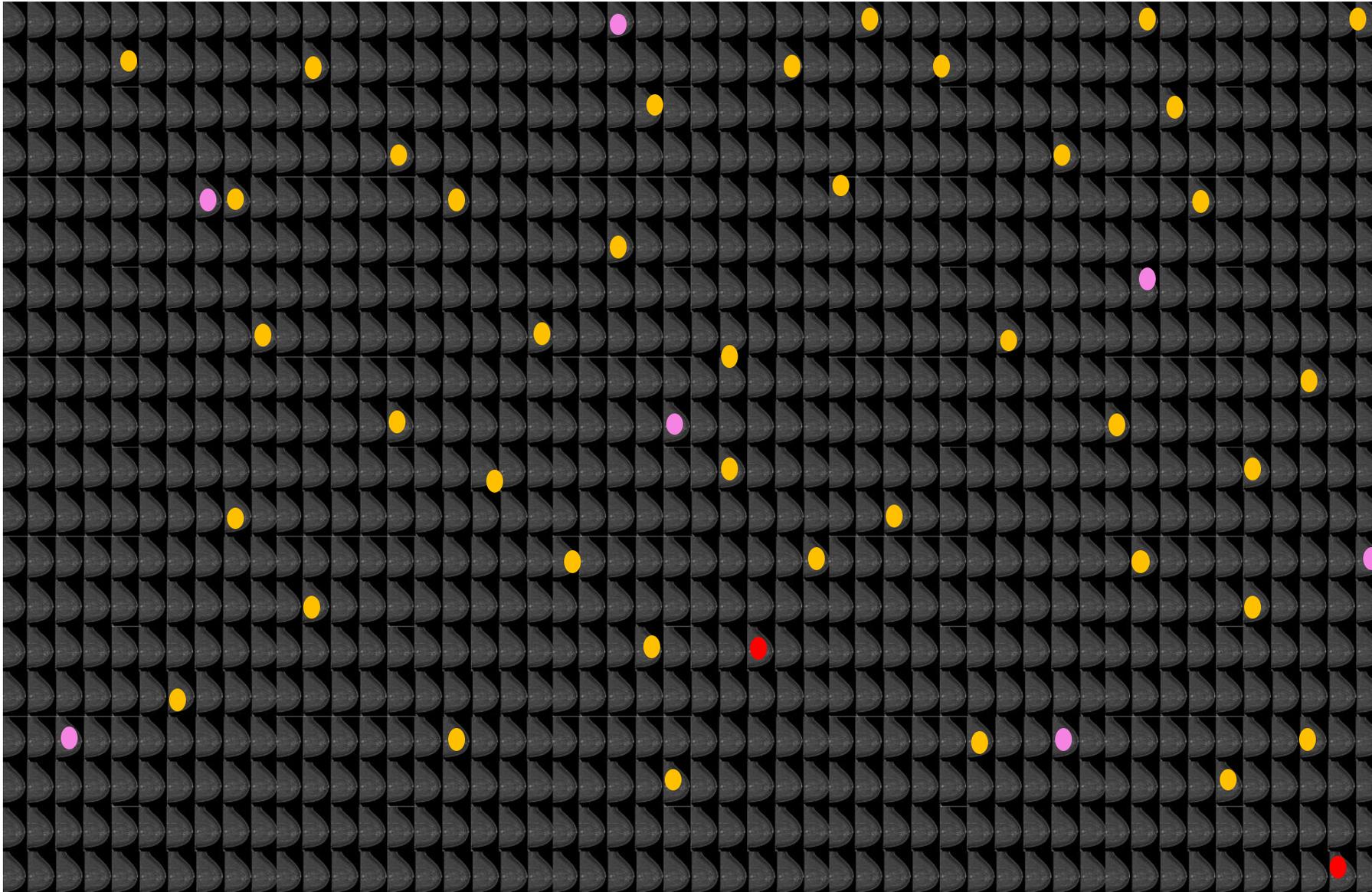
For every 1000 women screened



**7 cancers
are detected
(true positives)**

**1.8 cancers
diagnosed in interval
(false negatives)**

For every 1000 women screened



**7 cancers
are detected
(true positives)**

**1.8 cancers
diagnosed in interval
(false negatives)**

**40 women
are falsely recalled
(false positives)**

BRAIx developing foundational AI capabilities to address these challenges

Summary

- Developed globally unique AI dataset and digital twin - Value is in the data
- Demonstrated benefits retrospectively - Likely and easiest integration is second reader replacement
- Demonstrated AI reader result combined with other data is a better predictor of future risk than current measures - AI reader will also enable the basis for future risk communications and personalization
- Preparing RCT with AI reader as a second reader replacement in BSV and BSSA with significant ongoing stakeholder engagement - Ethics recently approved
- Developing a roadmap for deployment that incorporates vendor engagement and assessment, quality management and future personalisation

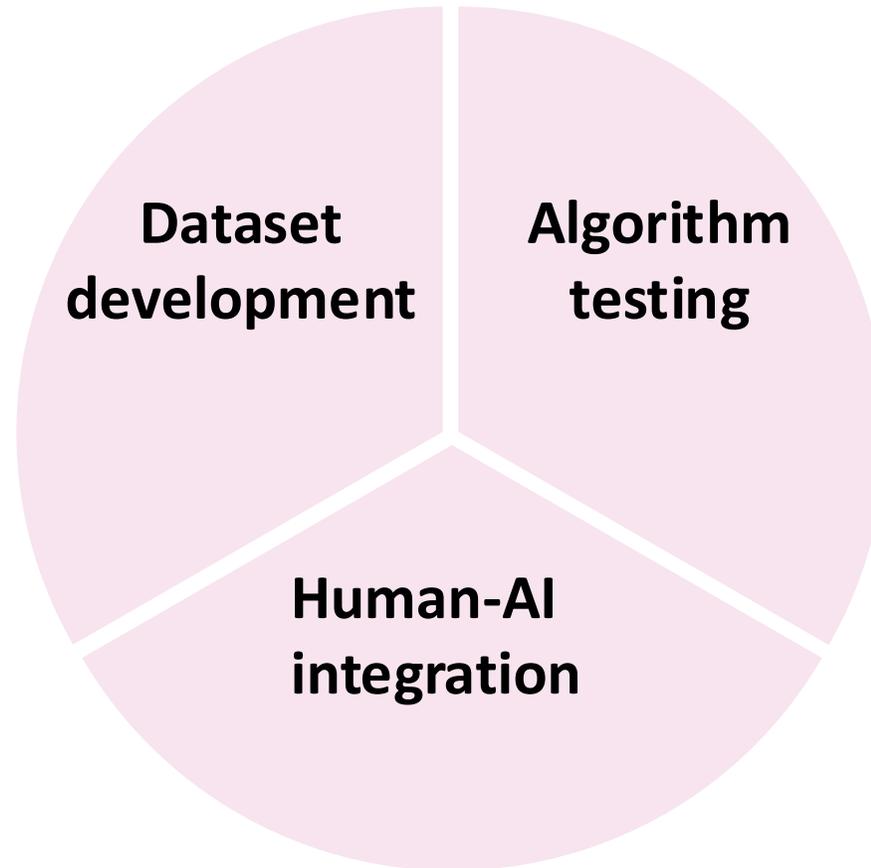
BRAIx is a key AI program of BSV and partners supported by MRFF and philanthropic grants



Three critical capabilities for AI translation

Capabilities for multi modal data development and management

- Establishing high quality ground truths
- Establishing dataset size and diversity to underpin generalisability



Capabilities for algorithm development, training and testing

- Testing of standalone performance
- Retrospective screening system simulation
- Classification and risk score

Capabilities for prospective study and ongoing quality assurance, human in control

- RCT and Quality Management
- Explainability, ethics and client/clinician engagement

Dataset development – globally unique

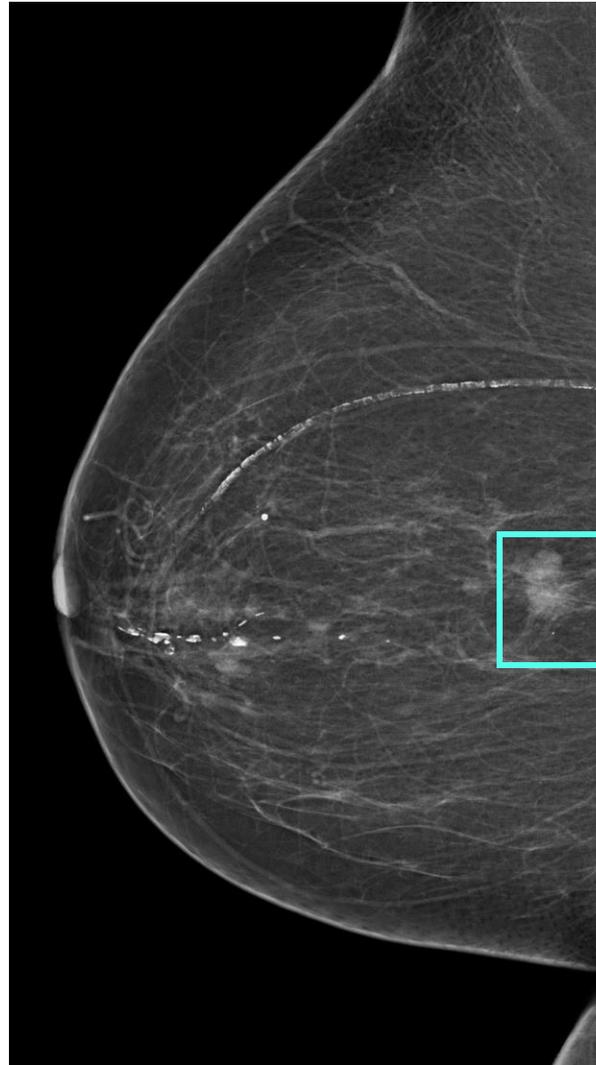
DATA DEVELOPMENT

We have built a globally unique dataset for AI development with 6.2 million+ images from almost 1.5-million screening episodes and 692,637 women during 2013-2021

- 29,363 screen detected cancer images
- 7,463 interval cancer images

Complete sequential 2016-2021 population data

Helen M. L. Frazer, et.al.. ADMANI: Annotated Digital Mammograms and Associated Non-Image Datasets. (2022) Radiology AI, RSNA, Published Online Dec 21, 2022.



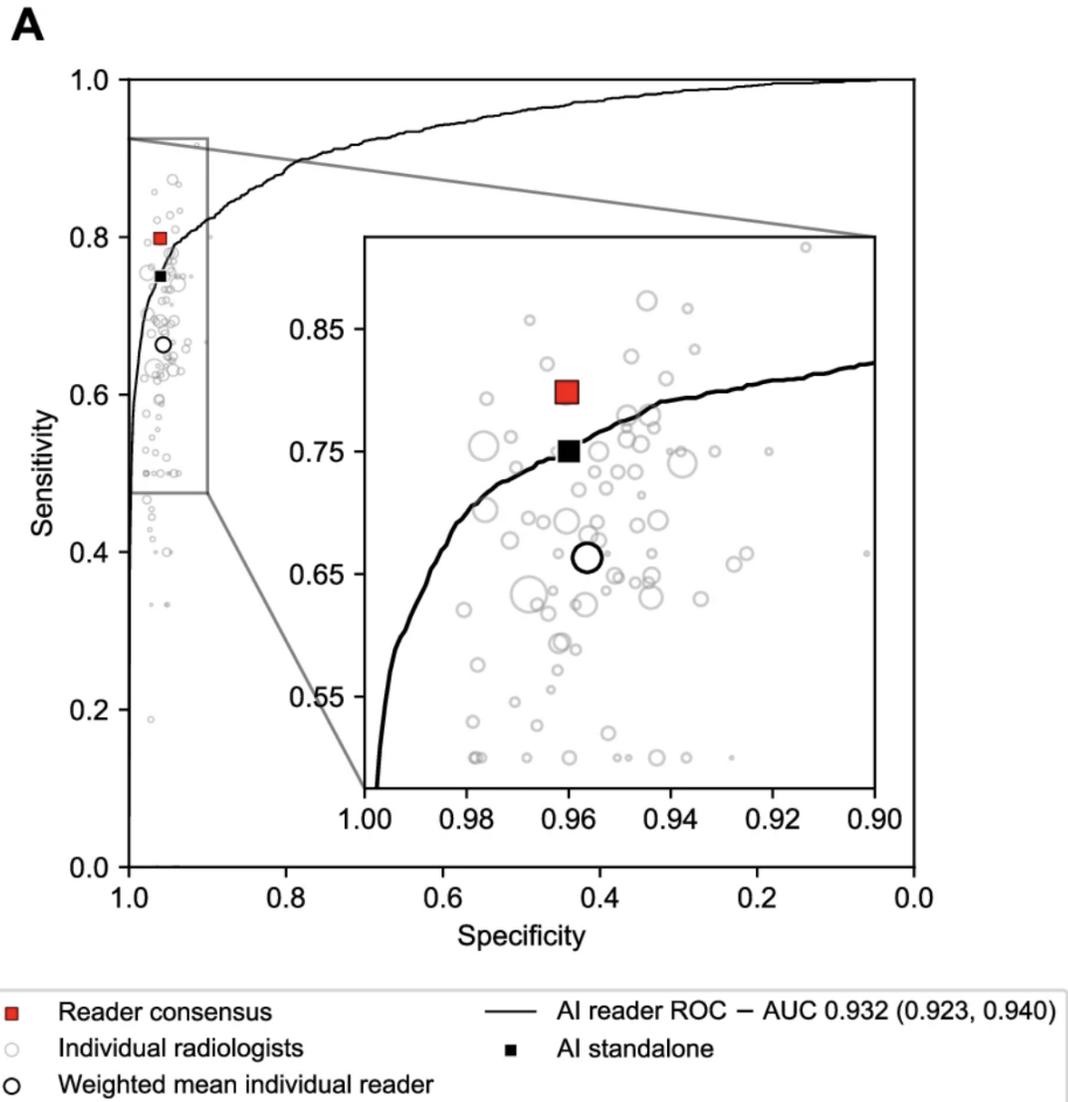
DATA ELEMENTS

- Includes over 200 data elements for every screening episode (e.g. family history, HRT, age, histopathology, reader outcomes)
- Radiologist labelled image annotations localise lesions

STRONG GROUND TRUTHS

- Surgical histopathology confirms cancer outcome
- Interval history confirms all clear outcome

Algorithm testing – standalone results

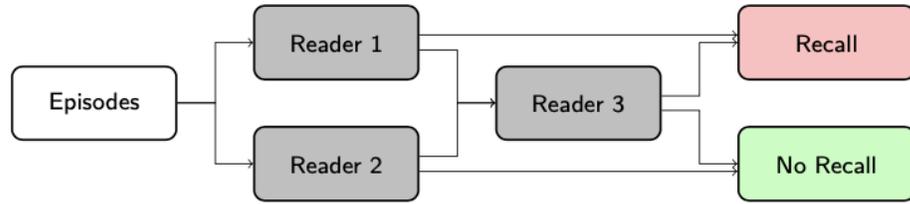


The AI Reader applied retrospectively to our randomly sampled 2016-2019 testing dataset (149,105 episodes) demonstrated:

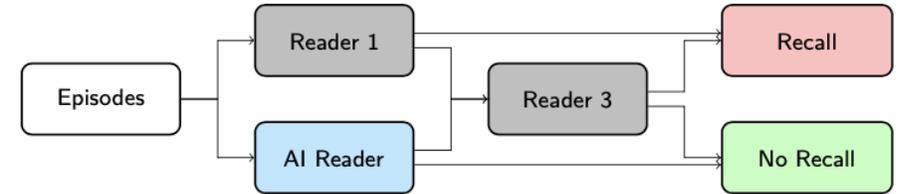
- AUROC = 0.93
- AI reader above weighted mean individual reader (95.6% specificity, 66.7% sensitivity)
- AI reader below current reader consensus system (96.1% specificity, 79.8% sensitivity)

Algorithm testing – screening pathway integration scenarios

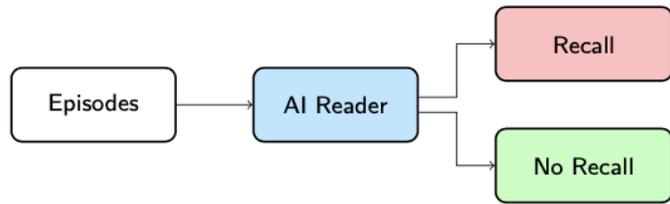
A Standard of care



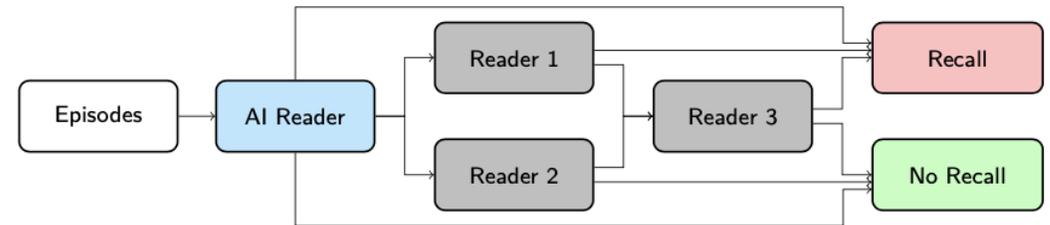
D AI reader-replacement



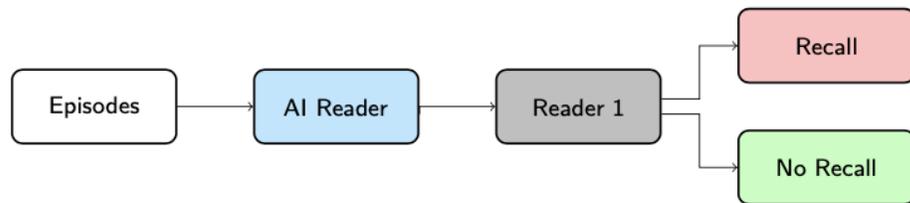
B AI standalone



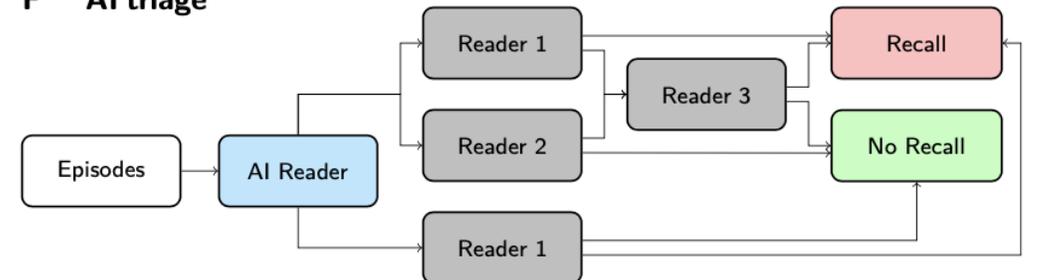
E AI band-pass



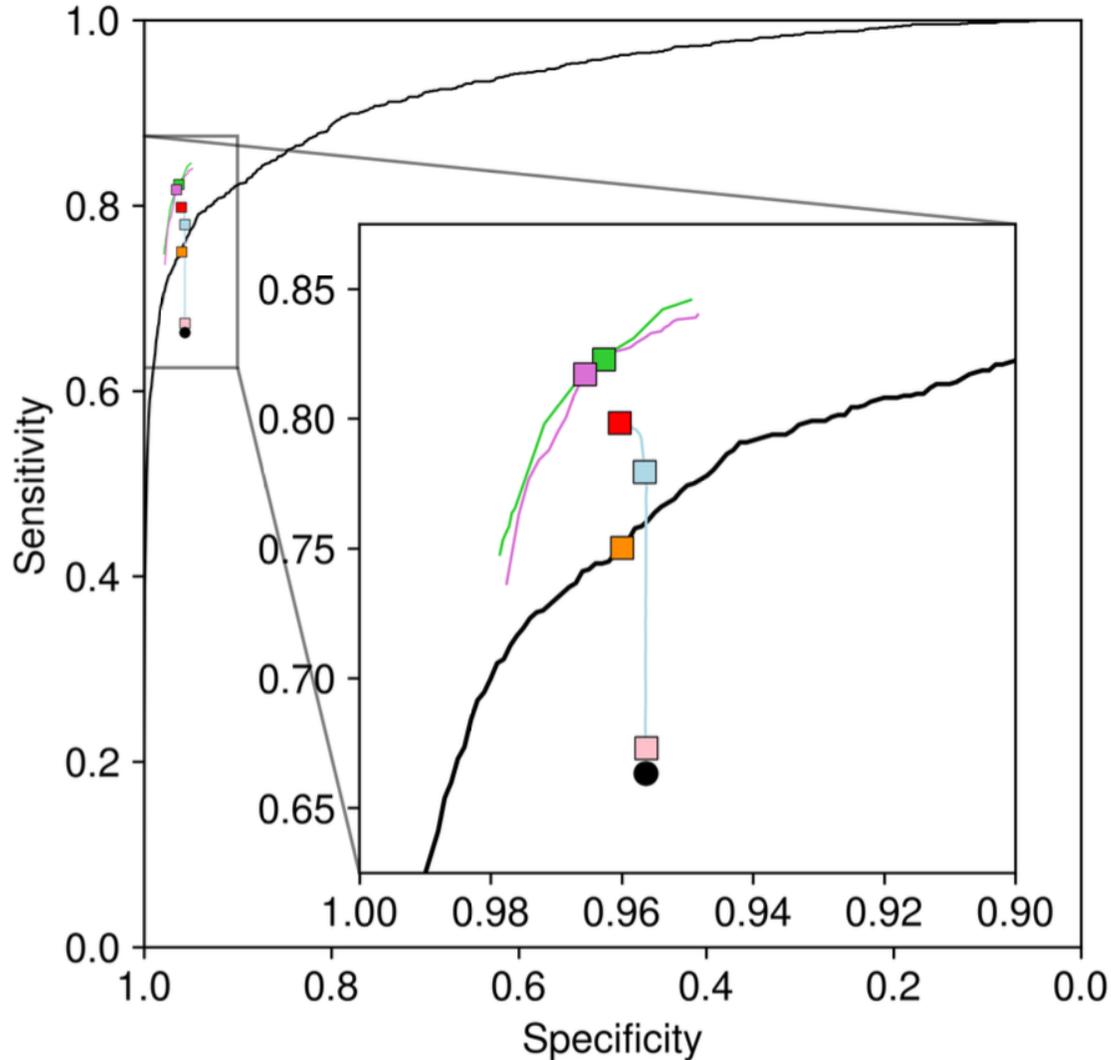
C AI single-reader



F AI triage



Algorithm testing – screening pathway integration results

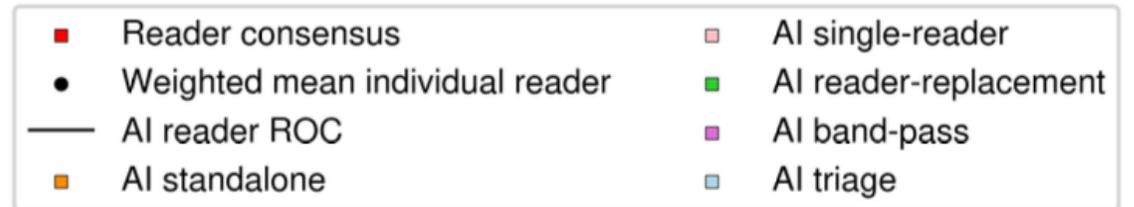


Both AI reader-replacement and AI band-pass improved performance over the human reader consensus

- AI reader-replacement (96.3% specificity, 82.3% sensitivity) GREEN
- AI band-pass (96.6%, 81.7%) PURPLE
- Current reader consensus system (96.1% specificity, 79.8% sensitivity) RED

Human reading workload was significantly reduced

- AI reader-replacement - by 48%
- AI band-pass – by 81%



RCT with ethics approval commencing this year

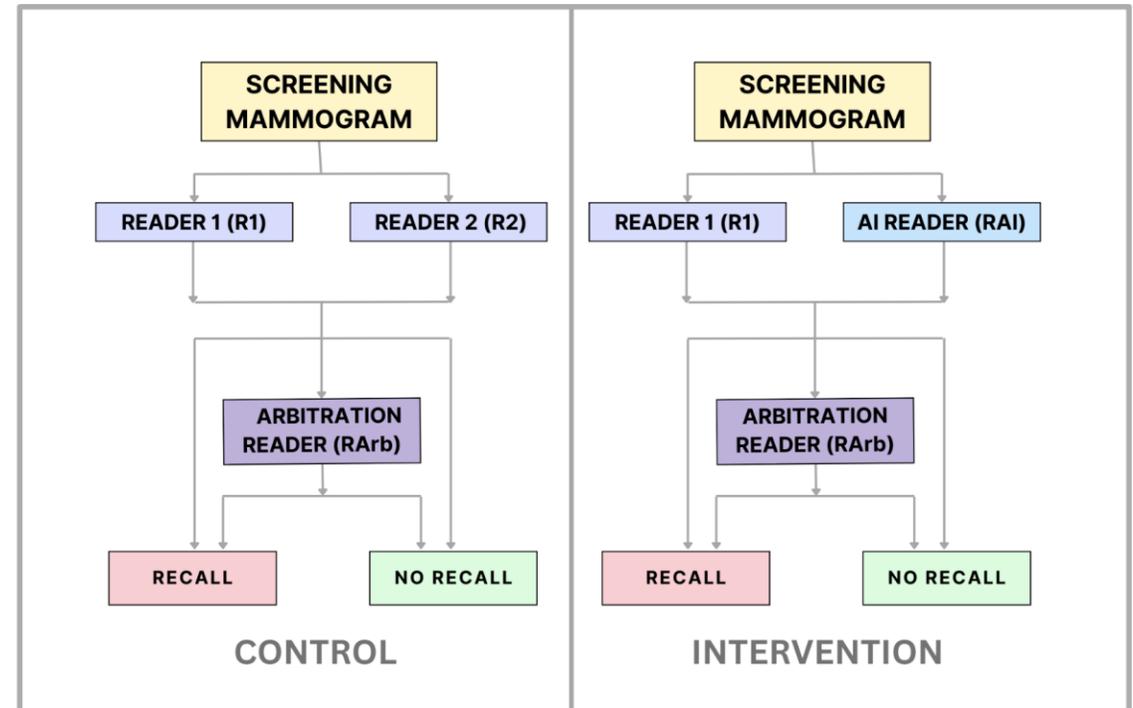
CLINICAL STUDY PROTOCOL

A Multi-State, Single-Blinded, Randomised Controlled Trial to Assess if the Implementation of a Deep Learning Classification Model Improves Mammogram Reading in Breast Cancer Population Screening

Local HREC/ERM Project Number: 2024.045/RMH100480

Version: #1.3

Date: 23/02/2024



AI Report Card



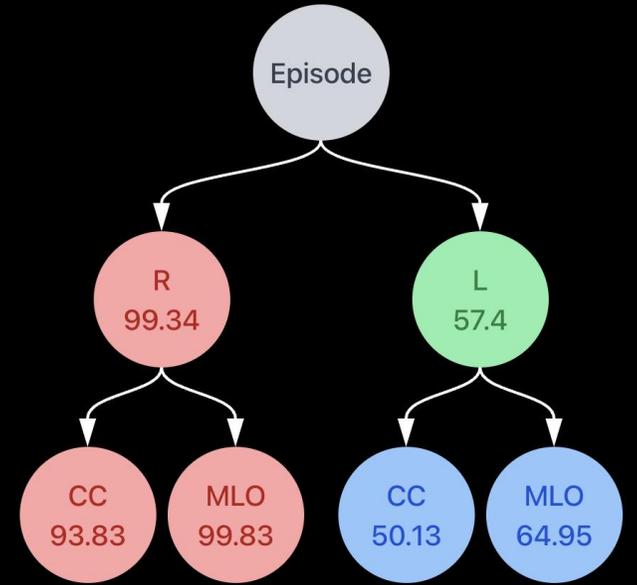
[R-MLO]



[L-MLO]

LATERALITY: R
VIEWPOSITION: MLO

LATERALITY: L
VIEWPOSITION: MLO

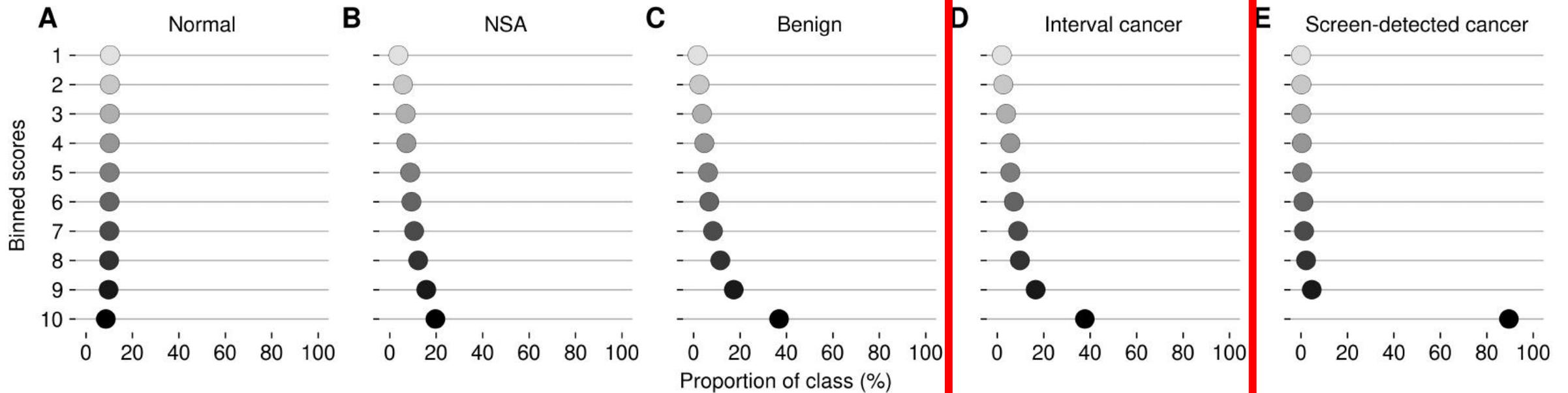


Recall Threshold Score	AI Result (Episode)
95	Recall

Reader Assessment Outcome (Episode)	Result Match (Episode)
Interval cancer	False negative

Result Card (R)	
AI Percentile Score (R)	99.3395
Reader 1 (R)	17 No Recall
Reader 2 (R)	No Recall

What about signal for interval cancer?



TRUE NEGATIVES

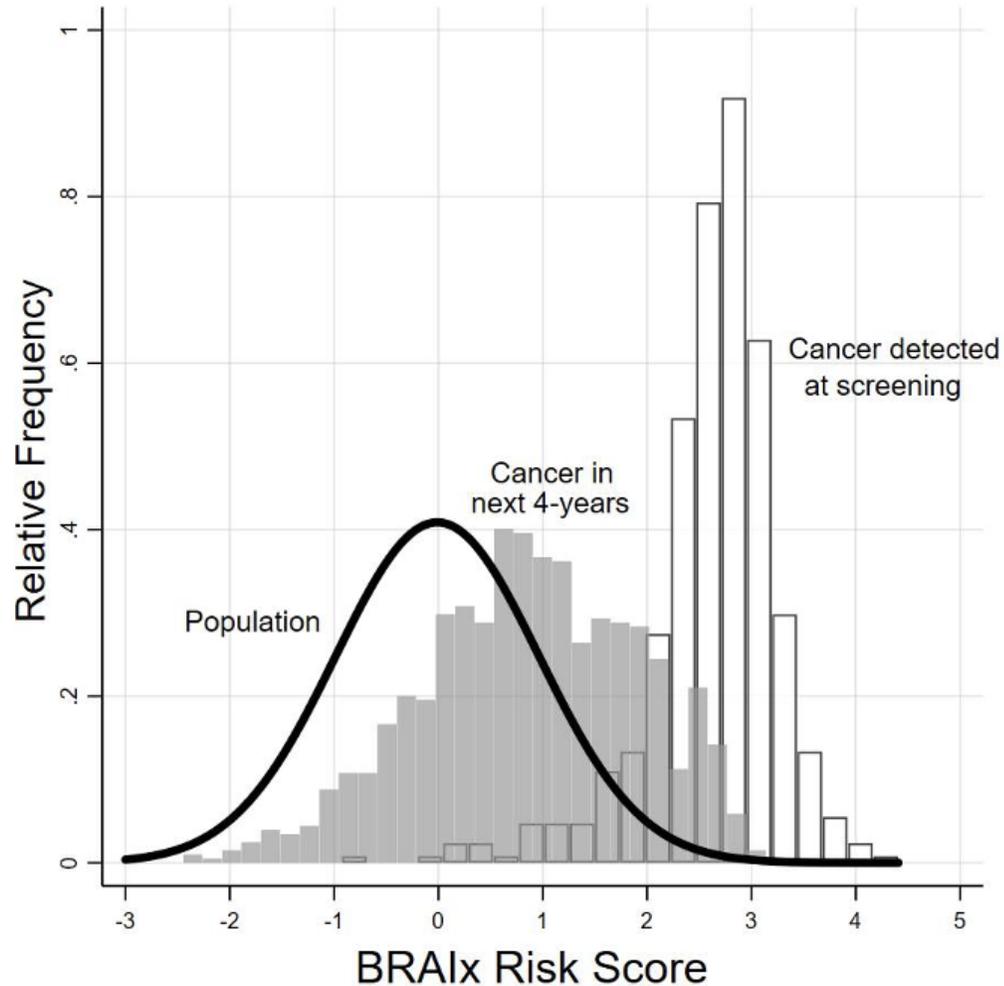
FALSE POSITIVES

FALSE NEGATIVES

TRUE POSITIVES

40% of interval cancers in top decile AI reader score

Use of AI also highlighting mammogram as a risk biomarker



Source: Hopper, John, Nguyen, Tuong Linh, Elliott, Michael S, BRAIx consortia, and Frazer, Helen, Braix Risk Score: An Automated Mammogram-Based Biomarker for Breast Cancer Created by Applying Artificial Intelligence. Available at SSRN: <https://ssrn.com/abstract=4764786> or <http://dx.doi.org/10.2139/ssrn.4764786>

Current breast cancer risk models (e.g. Tyrer-Cuzick) used to guide screening have only moderate discrimination

Mammography contains informative indicators of risk not captured by questionnaires, and not discernible by radiologists or simple volume-of-density measurements

AI models are now demonstrating improved risk prediction and are more discriminatory by themselves and in conjunction with traditional risk models.

These models have the potential to replace conventional risk prediction models.

Source: Comparison of Mammography AI Algorithms with a Clinical Risk Model for 5-year Breast Cancer Risk Prediction: An Observational Study. Vignesh A. Arasu, et.al. Radiology 2023 307:5

We now envision AI risk based personalised screening

- **Currently**

- Women 50-74 years of age invited and encouraged to attend BreastScreen – mammograms at 2-year intervals
- Women 40-49 and >75 welcome to attend but not actively recruited
- Very small number of clients with very strong family history screened annually

- **Future**

- Baseline mammogram screen at 40 years of age
- Individual risk score provided to client for discussion with GP
- Personalised screening profile based on risk profile (e.g., annual, 3-year intervals, 3D screening, contrast imaging)
- Watch for *change* in risk score across multiple screening episodes

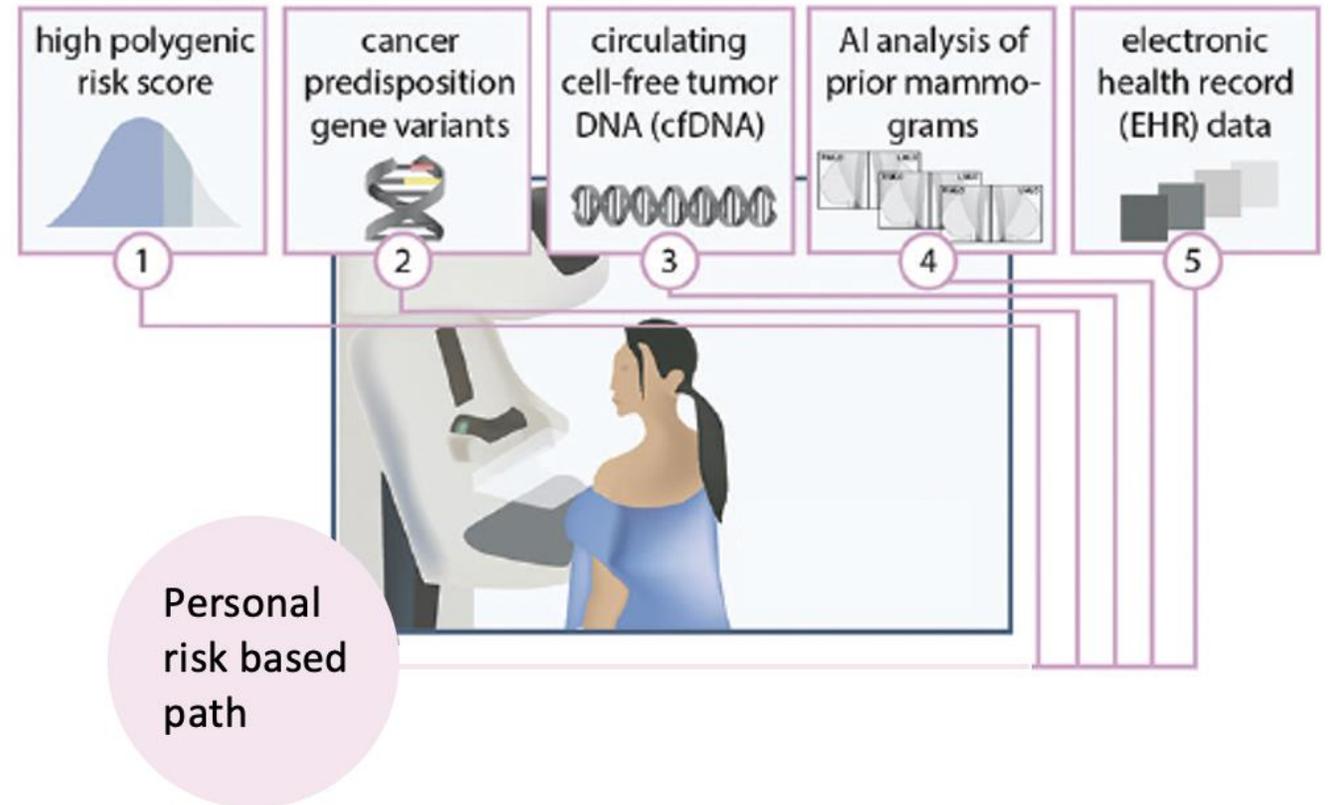
AI will enable personalised screening in the future

Population screening programs have the ingredients required, and can be an exemplar, for a precision health system

Era of double reading of every mammogram in a one-size-fits-all system is coming to a close

AI will enable us to realise the long-held vision of risk stratified, personalised screening

New foundation models will accelerate discovery with multi-modal data



Summary

Breast cancer is a major public health problem

Screening has been effective but there are challenges with the “one size fits all approach”

Developments in technology, particularly with data, and AI, are highlighting the potential for risk-based, personalised screening to save more lives

Breast cancer screening has the potential to be an exemplar for a precision health future

