



Experience with introduction of AI in Breast Cancer Screening in Capital Region of Denmark

**E-Sundhedsobservatoriets Konference, Odense
2025**

Ilse Vejborg,

Chief Physician and Head of Capital Mammography Screening Programme, Capital Region, Denmark

Chairman Danish Society of Breast Imaging, Chairman Danish Quality Database for Mammography Screening

Retrospective Simulation Studies

A collaboration between two Institutes at University of Copenhagen (Computerscience and Public Health), a professor and founder of the AI from Radboud University, NL and Capital Mammography Screening Programme



Two retrospective simulation studies based on

- Results of Double blind readings by experienced full time breast radiologist of 114.421 consecutive womens screening exams versus AI
- Sampling period January 2014 - December 2015. 2 year follow up.
- 791 screen detected cancers, 327 interval cancers and 2107 false positives

Preliminary simulation study:

AI only (no radiologist readings) with a sensitivity matched to experienced breast radiologists sensitivity

- 100% work load reduction
- Lower specificity than the radiologist (94.9% versus 98.1%)
- Signifikant rise in FP: 276,5% rise - 5825 women compared to 2107

"An Artificial-Intelligence-based Mammography Screening Protocol for Breast Cancer: Outcome and Radiologist Workload". Radiology 2022.

Retrospective simulation studies



Main simulation study:

- AI[★] only reader on the lowest risk group (<5 on a risk score on a scale from 1-10)
- Double blind readings by experienced breast radiologists (risk score ≥ 5 - 9,989)
- Direct recall of women with a risk score on ≥ 9.989

Results

- Sensitivity: AI 69.7% versus breast radiologist 70.8%
- Specificity: AI 98.6% versus breast radiologist 98.1%
- Numbers of false positive reduced with 25%

★Transpara version 1.7.0

"An Artificial-Intelligence-based Mammography Screening Protocol for Breast Cancer: Outcome and Radiologist Workload". Radiology 2022.

Retrospektivt simulationsstudie



Mammografi
Computert
University

AI we have used is

- Dobbel
kons
- 791 s
inter

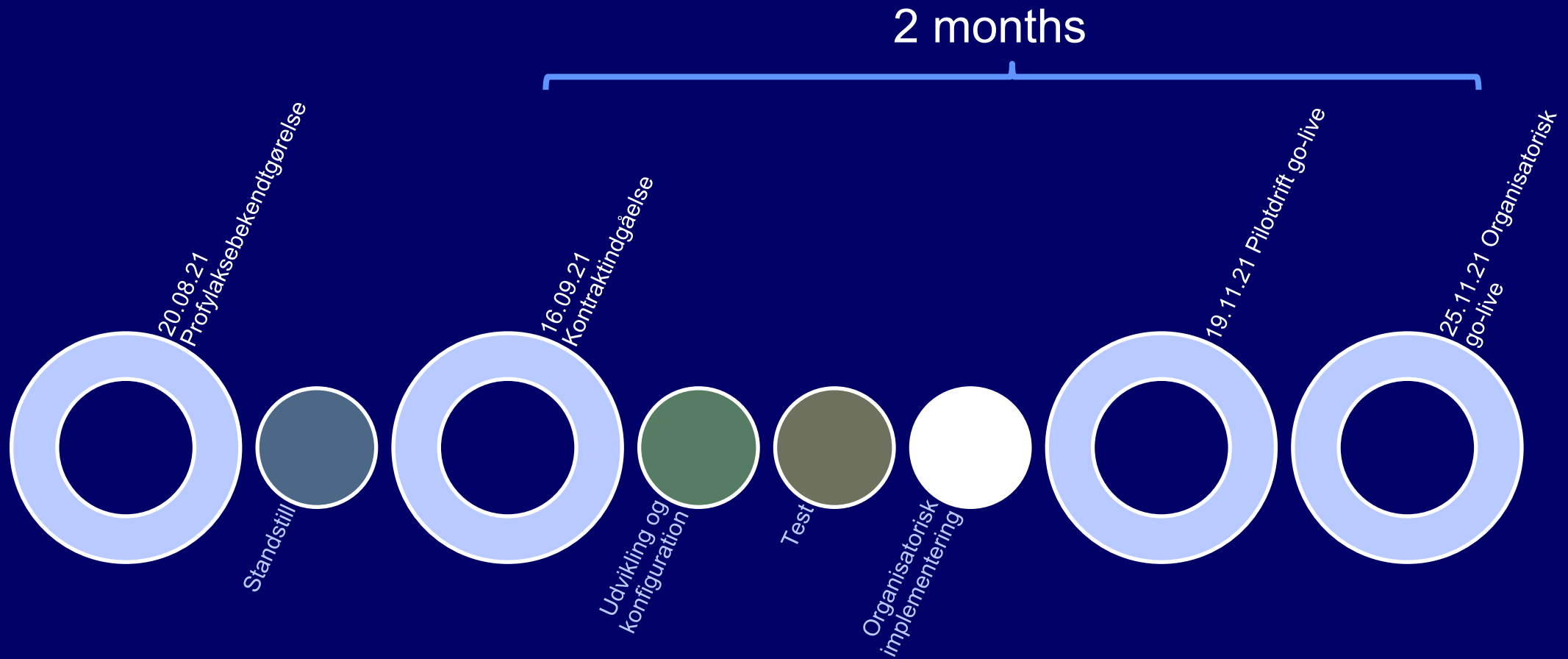
Resultat

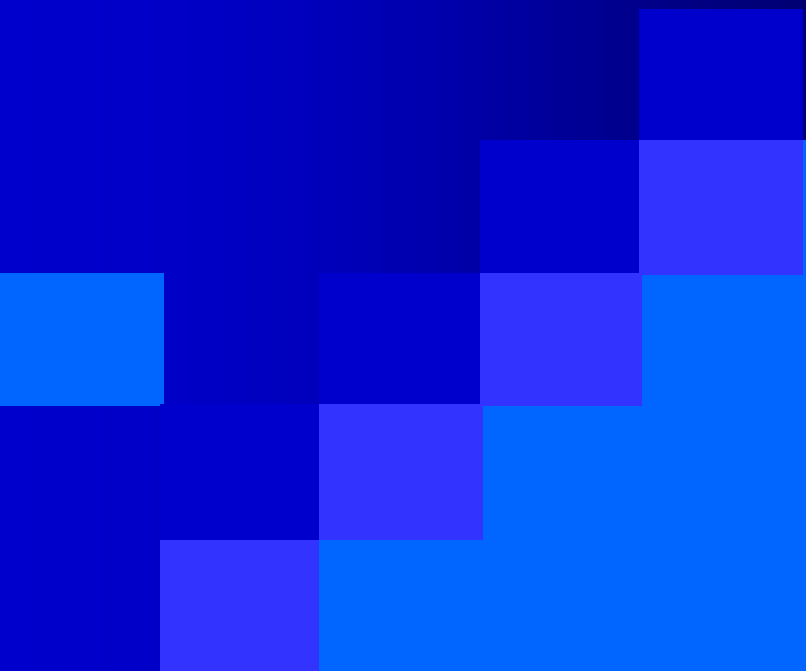
- Sens
- Spec

- Transpara version 1.7.0
- CE and FDA approved
- Trained on > 1 mio. processed mammograms
- from different vendors in Europe and USA
- Data from our regional screening programme was raw data
- ..and not included in the training (no bias!)

tut for
lboud

Procurement and implementation completed in a compressed process of 3 months – in a strong collaboration between CIMT, Human Bytes / Transpara and clinical staff from the Breast Cancer Screening Program in RegionH





Implementation of AI in Capital Mammography Screening Programme in Denmark

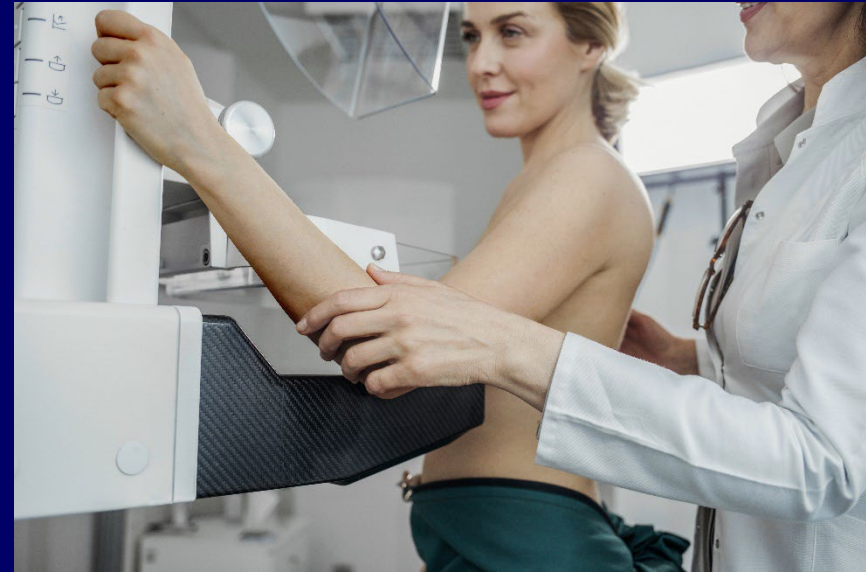
Main goal has been to reduce radiologist
workload keeping quality indicators stable

Screening mammography

- 2 standardized views: CC + MLO
- No clinical examination or UL

Time consumption

- *6-10 minutes in the examination room at the screening clinics (radiographers)*
- **<1-3 min. x 2/ exam centralized double blind readings (two radiologists)**



Screening mammography

- 2 standardized views
- No clinical examination

Hard competition considering workload reduction but..

Target group in DK ≥ 735.000 Q aged 50-69 år; 220.000 Q in RegionH

*Time consumption
6-10 minutes in the
Centralized double*

Extended offer to breast cancer treated women aged 70-79 years; 8150 Q in Capital Region



...s are working)



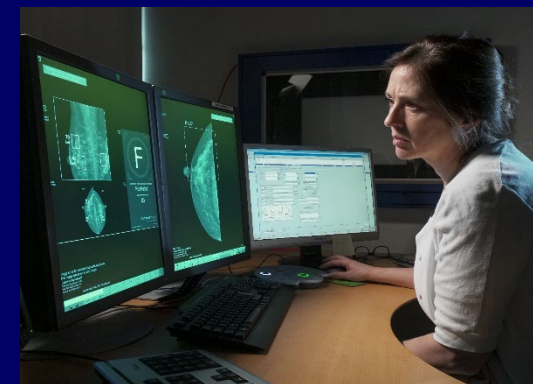
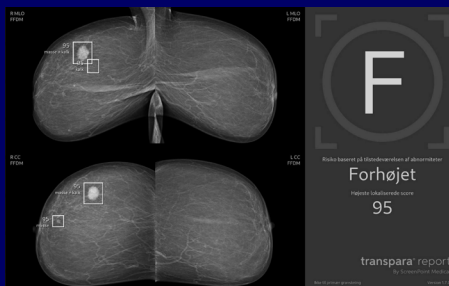
Mammograms analyzed by Transpara AI

Local regional score

Selection of highest regional score

Stratification into **risk categories on a scale from 1-100**

Shown in PACS
(in the end of the exam)



Relation between scores

Capital Region:
Score 78 = recall rate på 2,5%

3th of May 2022 AI first reader
of whole low risk group

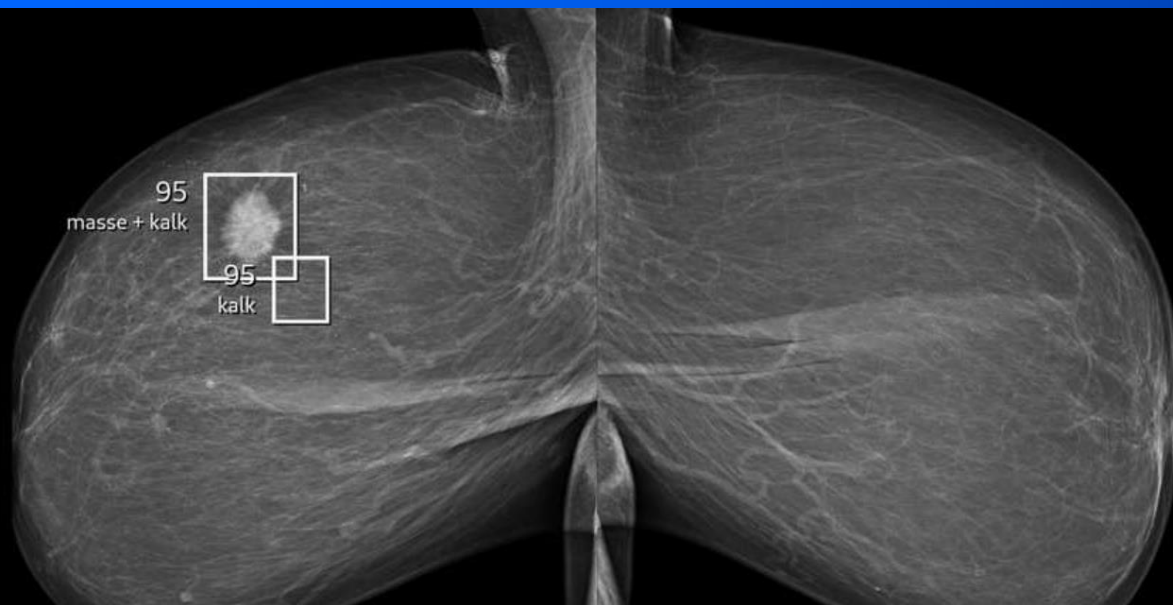
18th of November 2021



Højeste lokaliserede score ved undersøgelse	Risiko baseret på tilstedeværelsen af abnormiteter	Transpara Undersøgelsesresultat
≥ 75	Forhøjet - 1 ud af 10 undersøgelser påviser kræft ved screening* - Svarer til en tilbagekaldelsesrate på 4%*	10 Tjek lokaliseret score
61 - 74	Middel Samlet frekvens for kræft i dette interval svarer til screeningen af befolkningen (6/1000)	9
50 - 60		8
43 - 49		7
39 - 42	Lav - > 99,9% normale test* - Fund vist med markør ≥ 36	6
36 - 38		≤ 5
≤ 35		

>70%

R MLO
FFDM

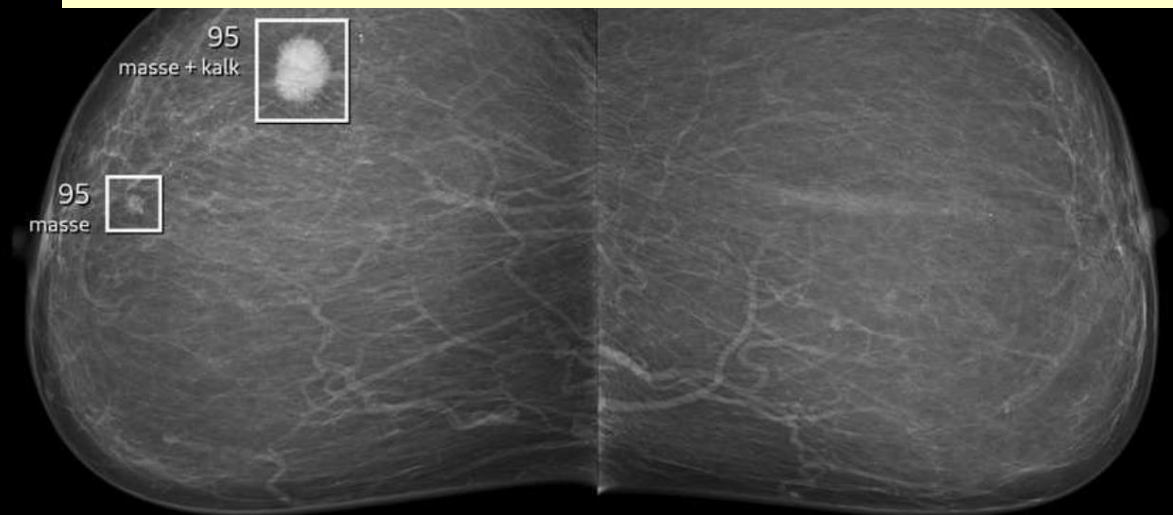


L MLO
FFDM



**Highest regional score decides
the final risk score**

R CC
FFDM



L CC
FFDM

Risiko baseret på tilstedeværelsen af abnormiteter

Forhøjet

Højeste lokaliserede score

95

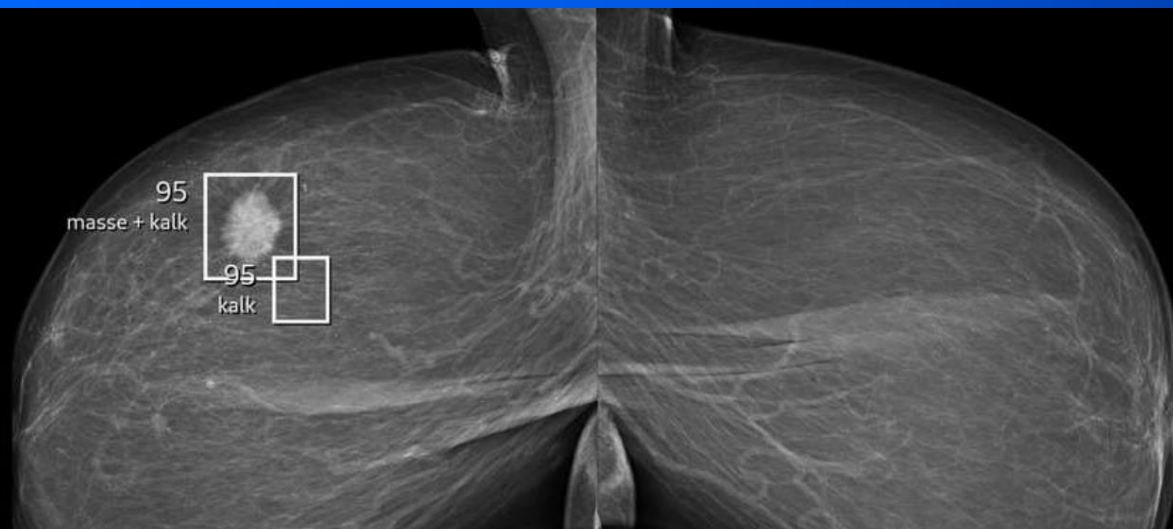
transpara® report

By ScreenPoint Medical

Ikke til primær granskning

Version 1.7.0

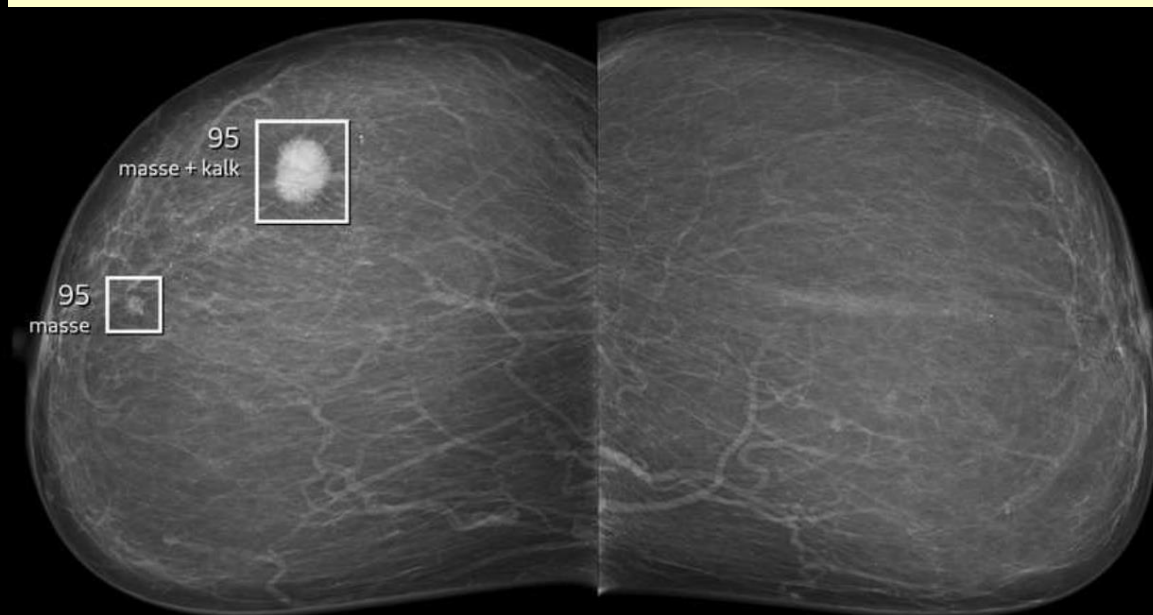
R MLO
FFDM



L MLO
FFDM

AI has no previous exams to compare with- but the radiologists have them!

R CC
FFDM



L CC
FFDM



Risiko baseret på tilstedeværelsen af abnormiteter

Forhøjet

Højeste lokaliserede score

95

transpara® report

By ScreenPoint Medical

Ikke til primær granskning

Version 1.7.0

Workflow in Capital Region DK

AI+Single or double reading?

Women with **low risk score**

from 3/5 2022 all with score ≤ 42 (<36 from 18/11 2021-3/5 2022)



AI (first reader) + one breast radiologist (second reader)

Consensus list in case of disagreement

Always a radiologist who decide!

Women with **intermediate or high risk score**



Double blind readings as usual by two breast radiologists (with AI assistance)

(no direct recall)

National Performance Indicators

(Danish Quality Database for Mammography Screening)

Performance Indicator (Number)	Invitation round						
	First 2008- 2009/2010	Second 2010- 2011/12	Third 2012- 2013/14	Fourth 2014- 2015/16	Fifth 2016-2018	Sixth 2018-2020	7th 2020- 2023
2 a. Participation (%invited)	76%	82%	84%	83%	83%	84%	83%
b. Coverage (% target)	75%	75%	77%	76%	79%	79%	79%
4. Recall rate	3%	2,7%	2,7%	2,5%	2,4%	2,4%	2,4%
False-positive rate	2.0%	2.1%	2.1%	1.9%	1,8%	1,8%	1,7%
Detection rate (IC+DCIS)	0.93%	0.62%	0.67%	0.61%	0.62%	0,61%	0,66%
5. Interval cancer rate (Interval IC / Interval IC+ screen detected < 12 / 12-24 months after)	NA	NA	12% 21%	11% 19%	11% 20%	12,4% 21%	12,4% 20,6%
6. Invasive % (IC / IC+DCIS)	87%	86%	86%	86%	87%	85%	83,8%
7. Lymph node neg %	70%	75%	78%	81%	76%	78%	76%
8. Small tumor ≤1cm %	37%	39%	37%	37%	37%	37%	37%
9. Benign : malign operation ratio	1:6	1:7	1:8	1:9	1:10,5	1:10	1:10
10.BCS % (BCS / BCS+ mastectomy)	80%	81%	83%	No longer in use	Not in use	Not in use	

National Performance Indicators

(Danish Quality Database for Mammography Screening)

Performance Indicator (Number)	Invitation round					
	First 2008- 2009/2010	Second 2010-2011/12	Third 2012- 2013/14	Fourth 2014- 2015/16	Fifth 2016-2018	Sixth 2018-2020
2 a. Participation (%invited)	76%	82%	84%	83%	83%	84%
b. Coverage (% target)	Even a small increase in recall rate decreases the benefit! 1 diagnostic mammography (incl. clinical examination, UI and evt. needle biopsy) matches = 30-50 single readings					90%
4. Recall rate						94%
False-positive rate						8%
Detection rate (IC+DCIS)						61%
5. Interval cancer rate (Interval IC / Interval IC+ screen detected < 12 / 12-24 months a)						1%
6. Invasive % (IC / IC+DCIS)						5%
7. Lymph node neg %	70%	75%	78%	81%	76%	77%
8. Small tumor ≤1cm %	37%	39%	37%	37%	37%	37%
9. Benign : malign operation ratio	1:6	1:7	1:8	1:9	1:10,5	1:10
10.BCS % (BCS / BCS+ mastectomy)	80%	81%	83%	No longer in use	Not in use	Not in use

Early data on Recall rate

- 6.Screening Round (1.July 2018- 31.September 2020):

2,5%

- 7.Screening Round **Before AI**. (1st October 2020- 17th of November 2021; 63.682 q):

3.09%

Women with a previous breast cancer diagnosis was highly prioritized over the normal screening population; same distribution (4.6% /4.7%) of Q having history of BC operation before and after AI

After AI (18th, November 2021- 31st of December 2022; 79.270 q):

Recall rate **before** increase of threshold: **2.72%**

Recall rate **after** increase of threshold: **2.29%**

In total with AI: **2.46%**

Recall rate for **low risk**: **0.40%** (14 cancers/ 53.438 us= detection rate 0,026%)

Recall rate for **intermediate and high**: **6.65%**

Screening with AI as 1st reader (18th Nov. 2021- 31st December 2022) = 67.41% (53.438 / 79.270 screenings)

Early data on Recall rate

- 6.Screening Round (1.July 2018- 31.September 2020):

2,5%

- 7.Screening Round (1.October 2020- 31.September 2021):

Low risk group

- 14 cancers amongst 215 recalled women
- All cancers were new or lesions changed since last exam
- AI has no previous images to compare with- radiologists do!

*Women with a previous
distribution (4.6% /4.7%)*

ing population; same

After AI (18th, Nov. 2021)

Recall rate **before**

Recall rate **after** in

In total with AI:

2.48%

Recall rate for **low risk**:

0.40% (14 cancers/ 53.438 us= detection rate 0,026%)

Recall rate for **intermediate and high**:

6.65%

Screening with AI as 1st reader (18th Nov. 2021- 31st December 2022) = 67.41% (53.438 / 79.270 screenings)

Work load reduction for radiologists reading

(18/11 2021 – 17/10 2022)

66.9% read by AI as 1st reader

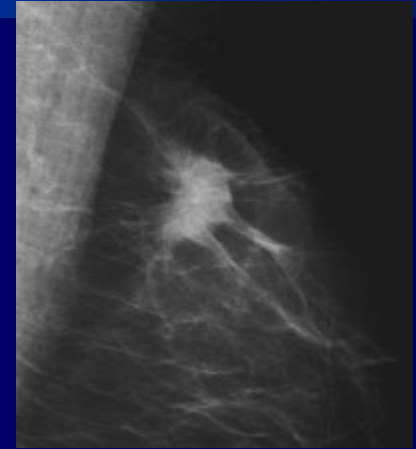
=> 33,5% workload reduction

($\geq 35\%$ after change of level)



Lauritzen AD, Lillholm M, Lynge E, Nielsen M, Karssemeijer N, Vejborg I.
Early Indicators of the Impact of Using AI in Mammography Screening for Breast Cancer
Radiology. 2024 Jun;311(3):e232479. doi: 10.1148/radiol.232479.

Cancer detection rate



Before AI: 60,751 screenings from 1/10-2020 to 17/11-2021
With AI: 56,894 screenings from 18/11-2021 to 17/10-2022
Look ahead: ≥ 180 days.

CDR (before AI) = 0.70%
CDR (with AI) = 0.82% (P < .01)

Cancer Detection (Early Indicators)

Before AI: 60,751 screenings from 1/10-2020 to 17/11-2021

With AI: 56,894 screenings from 18/11-2021 to 17/10-2022

Based on a needle biopsy or pathology following surgery within ≥ 180 days from screening visit positive for IC and/or DCIS.

Percentage screen detected small invasive cancers ≤ 1 cm

Small cancer rate (before AI) = 36.60%

Small cancer rate (with AI) = 44.93% (P = .02)

Percentage lymph node neg. screen detected Invasive cancers

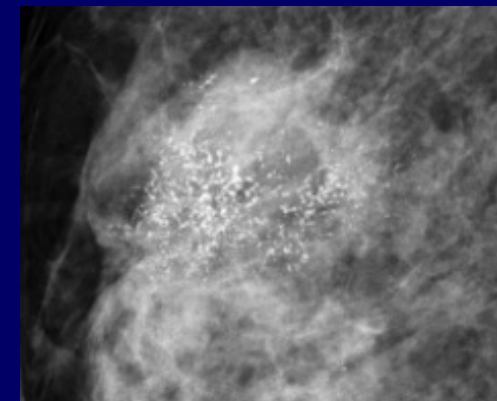
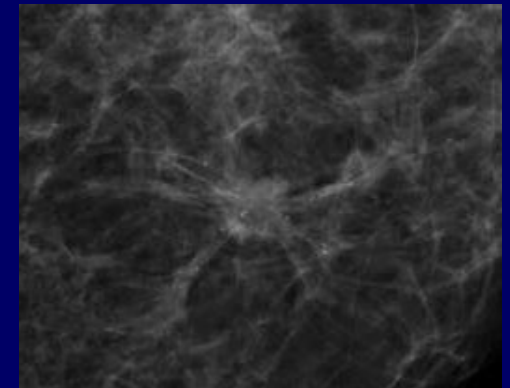
Node negative rate (before AI) = 76.67%

Node negative rate (with AI) = 77.78% (P = .73, NS)

Distribution of screen detected IC versus DCIS

IC / IC + DCIS (before AI) = 84.87%

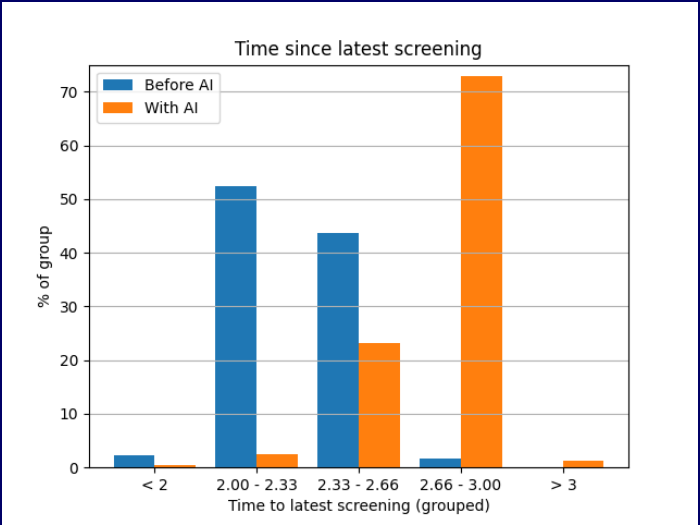
IC / IC + DCIS (with AI) = 79.58% (P = .04)



Population characteristics and possible bias

	BEFORE AI 1/10-2020 to 17/11-2021	WITH AI 18/11-2021 to 17/10-2022
SCREENED WOMEN	60,751	58,246
MEDIAN AGE (IQR)	58 (54, 64)	58 (54, 64)
MEAN BI-RADS DENSITY ($\pm$ STD)	1.8 ($\pm$ 0.8)	1.8 ($\pm$ 0.8)
PREVIOUS BC SURGURY	2,799 (4.6%)	2,736 (4.7%)
AVG. SCREENING INTERVAL	2 year, 121 days	2 years, 256 days

Cancer Detection Rate and Screening Interval before and after AI



Group	CDR, Before AI	CDR, With AI	P-value
2.00 - 2.33 years	0.73% (0.63%, 0.84%)	1.03% (0.45%, 1.61%)	0.25 (ns)
2.33 - 2.66 years	0.64% (0.53%, 0.75%)	0.96% (0.77%, 1.15%)	0.002 (**)
2.66 - 3.00 years	0.66% (0.08%, 1.25%)	0.77% (0.68%, 0.87%)	0.74 (ns)

Conclusion



- Background for implementation: Very promising results in our large retrospective simulations study
- Prospective results:
 - AI is a valuable tool for risk stratification on basis of analysis of the mammograms (> 70 % stratified as low risk)
 - Substantial workload reduction in readings for breast radiologists ($\geq 35\%$)
 - $\geq 20\%$ reduction in recalls
 - Early quality indicators show at least as good results as previously
- Whats next?
 - Data on interval cancers

Thank you for your attention!

