Photo by Changhua City

AI-Empowered Precision Breast Cancer Screening

Technological Advances · Clinical Impact · Taiwan Implementation

17th IACCS International Asian Conference on Cancer and Chronic Disease Screening

Cancer and Chronic Disease Screening Aligned with the "4P" Health Model — Predictive · Preventive · Personalized · Participatory

Alt: 17th IACCS International Asian Cancer and Chronic Disease Screening Network Conference

Date: June 8 (Mon), 8:20–16:40 | Venue: 2F, Floramos Hotel, Tianzhong Township, Changhua County, Taiwan (彰化馥御花園酒店 2 樓)

13:05 – 13:25

AI-Empowered Precision Breast Cancer Screening

Chen-Pin Chou, MD · Director of Radiology, Kaohsiung Veterans General Hospital

Lecture Roadmap · **15 Minutes · 5 Core Issues**

15 min

1	2	3	4	5	Q&A
Why AI Now?	**AI Detection**	**Safety & Recall**	**Interval Cancers**	**Taiwan 2030**	**5 Key Questions**
The not satisfied need	*What works, what doesn't*	*Does AI increase false positives?*	*Where AI can and can't help*	*Precision screening roadmap*	*Most asked issues*
2 min	3 min	2 min	3 min	3 min	5 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Lecture Roadmap · **15 Minutes · 5 Core Issues**

15 min

1

Why AI Now?

The not satisfied need

2 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Why do accidents still happen

Why Accidents Still Happen Despite Driver Licenses

Human limitations, environment, and system factors — and how AI autopilot can help

Why accidents still happen

1	Distraction (phone use)	Taking your eyes off the road even for a few seconds can lead to crashes.
2	Fatigue / drowsiness	Drowsy driving slows reaction time and can cause you to doze off.
3	Speeding	Higher speeds reduce your time to react and increase the severity of crashes.
4	Alcohol impairment	Alcohol affects judgment, slows reactions, and impairs coordination.
5	Poor judgment / human error	Risky choices, overconfidence, or momentary lapses can lead to mistakes.
6	Bad weather	Rain, fog, snow, or strong winds reduce visibility and tire traction.
7	Poor road conditions	Potholes, loose gravel, uneven surfaces, or poor lighting increase risk.
8	Vehicle / mechanical problems	Brake failure, worn tires, or engine issues can lead to loss of control.
9	Unexpected actions by others	Pedestrians crossing suddenly or other drivers' sudden stops, turns, or lane changes can cause crashes.

A license proves basic qualification, not perfect driving.

Benefits of AI Autopilot

1. **Continuous monitoring** — Watches lanes, vehicles, and road conditions in real time.
2. **Faster hazard detection** — Alerts or reacts faster to sudden risks.
3. **Collision avoidance support** — Assists braking, lane keeping, and safe distance control.
4. **Reduced human error** — Helps compensate for distraction, fatigue, and delayed reactions.
5. **Smoother driving** — Improves consistency, stability, and traffic flow.
6. **Safer decision support** — Provides driver assistance, but does not replace responsibility.

Licensed Driver (Skill, experience & judgment) + **AI Autopilot (Assistance)** (Sensing, analysis & decision support) = **Safer Journeys Together**

Best safety comes from combining trained drivers, responsible behavior, and intelligent assistance.

AI autopilot can reduce risk, but safe driving still requires human attention and judgment.

Issue 1 · Why AI Now? — Demand, Workforce, and Technology Converge

BACKGROUND · 2 min

~2 min

Workforce shortage	*Reading volume outpaces radiologist supply globally*
Dense breast masking	*Sensitivity drops to ~50–60% in BI-RADS C/D (vs ~80% overall)*
Taiwan: single reading	*No double-check safety net — AI is the only second reader*
Interval cancer burden	*~22% of screening cancers present as interval cancers*
Risk stratification gap	*Same biennial protocol regardless of individual risk*

KEY ANSWER A confluence of workforce shortage, dense breast masking, single-read practice in Taiwan, and a persistent interval cancer problem make AI integration clinically urgent NOW.

Breast Mammography Screening: Age Range Expanded Twice

Policy	Start from	Target Reduction in Cancer Mortality Rate	Mammography Screening Age
Reduce Cancer Mortality Rate by 10% within 4 Years	**2008**	10% (within 4 years)	**50 y/o → 45**
Reduce Cancer Mortality Rate by 33% within 6 Years	**2024**	33% (within 6 years)	**45 y/o → 40** **69 y/o → 74**

行政院衛生署國民健康局委託辦理

乳癌篩檢服務量能規劃研究

Capacity Planning Research for Breast Cancer Screening Services

期末成果報告

執 行 機 構 ：中華民國放射線醫學會

計畫主持人：周春平

共同主持人：李中一

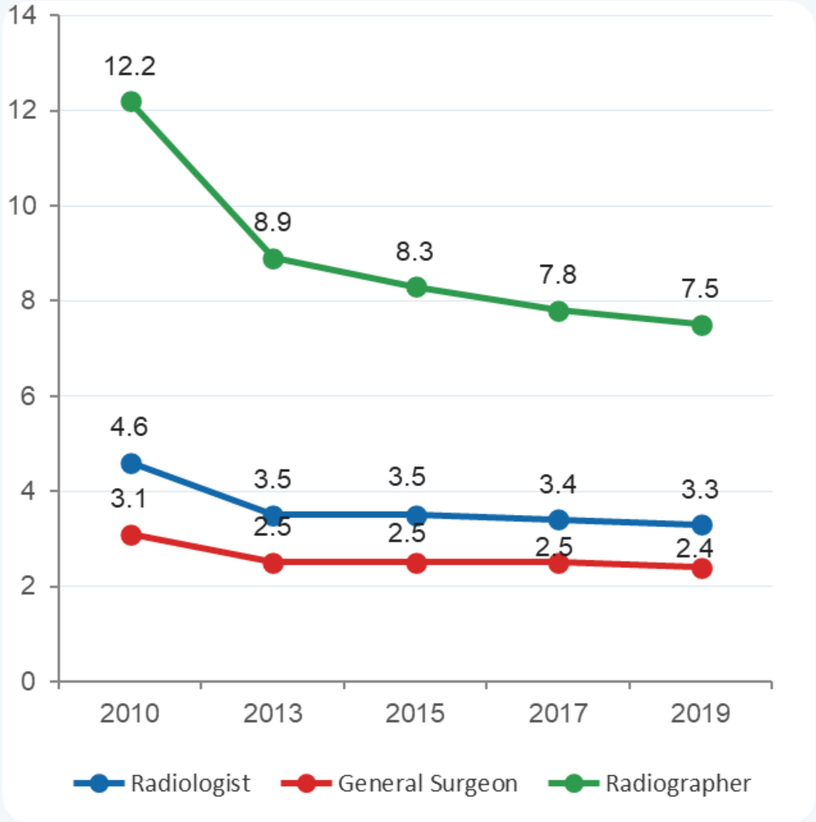
協同主持人：潘慧本、許居誠、黃錦輝、林惠賢、嚴明芳

Execution Period: April 29, 2010 – April 28, 2011

執行期間：99 年 04 月 29 日至 100 年 04 月 28 日

Biennial screening · Screening rate: 40% (2013) → 50% (2019)

Healthcare Personnel per 10,000 Screened Women



Radiologists' Recommendations if Screening Volume Doubles at Their Institution

Recommendation	n	%	Proportion (sorted high → low)
For Hospital Management			
Increase payment for mammography	183	73.8%	73.8%
Increase certified mammography technologists	139	56.1%	56.1%
Increase number of radiologists	136	54.8%	54.8%
Encourage physicians to specialize in mammography	114	46.0%	46.0%
Paid for continuing education & training	109	44.0%	44.0%
Increase staffing for breast ultrasound	108	43.6%	43.6%
Purchase new mammography equipment	93	37.5%	37.5%
Increase number of breast surgeons	65	26.2%	26.2%
Upgrade breast ultrasound equipment	65	26.2%	26.2%
Improve reporting and image viewing systems	63	25.4%	25.4%
Radiologists assist with localization and biopsy	60	24.2%	24.2%
Replace outdated mammography equipment	47	19.0%	19.0%
Improve surgical biopsy and operating room facilities	25	10.1%	10.1%
Other	11	4.4%	4.4%

Note.— Multiple responses allowed; percentages do not sum to 100%. Items sorted by descending percentage. Respondents are radiologists performing mammography at institutions in Taiwan's national breast cancer screening program.

Radiologists' Recommendations if Screening Volume Doubles at Their Institution

Recommendation	n	%	Proportion (sorted high → low)
For Government Health Authority			
Plan workforce expansion in proportion to mammography volume growth	194	78.2%	78.2%
Clearly announce national mammography screening policy to attract more physicians	135	54.4%	54.4%
Provide project-based funding for recall and biopsy procedures	129	52.0%	52.0%
Accurately plan referral volumes to participating screening hospitals	126	50.8%	50.8%
Reward or recognize institutions based on screening volume	106	42.7%	42.7%
Reward or recognize institutions based on breast cancers detected	79	31.9%	31.9%
Regularly publish institutional and physician screening performance data publicly	58	23.4%	23.4%
Strengthen breast diagnostics training in PGY programs (physicians & technologists)	48	19.4%	19.4%
Other	19	7.7%	7.7%

Note.— Multiple responses allowed; percentages do not sum to 100%. Items sorted by descending percentage. Respondents are radiologists performing mammography at institutions in Taiwan's national breast cancer screening program.

2010

iPhone era: iPhone 4

- Release year: 2010
- Chip: Apple A4
- Storage: 16GB / 32GB
- Battery life: Up to 7 hours talk time

A DECADE AND A HALF OF INNOVATION

From a communication smartphone to an AI-powered personal computing device.

	2010		2026
SCREEN	3.5-inch Retina display 960 x 640	>	6.9-inch Super Retina XDR ProMotion 120Hz
CAMERA	5MP Single camera 720p video	>	48MP+48MP+12MP Pro camera system 8K video, Spatial video
POWER	Apple A4 chip 1-core CPU	>	Apple A19 Pro chip 6-core CPU
AI	No AI Basic features Rule-based	>	Apple Intelligence On-device AI Personal & proactive

Built for today. Ready for tomorrow.
Smarter. More personal. More capable.

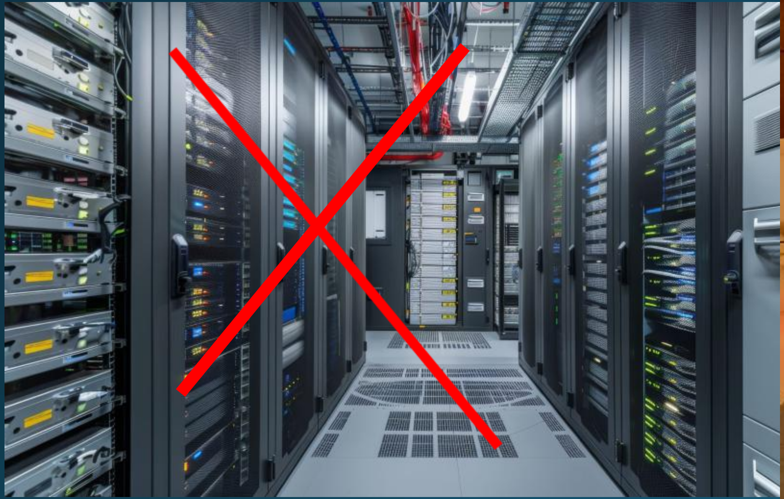
2026

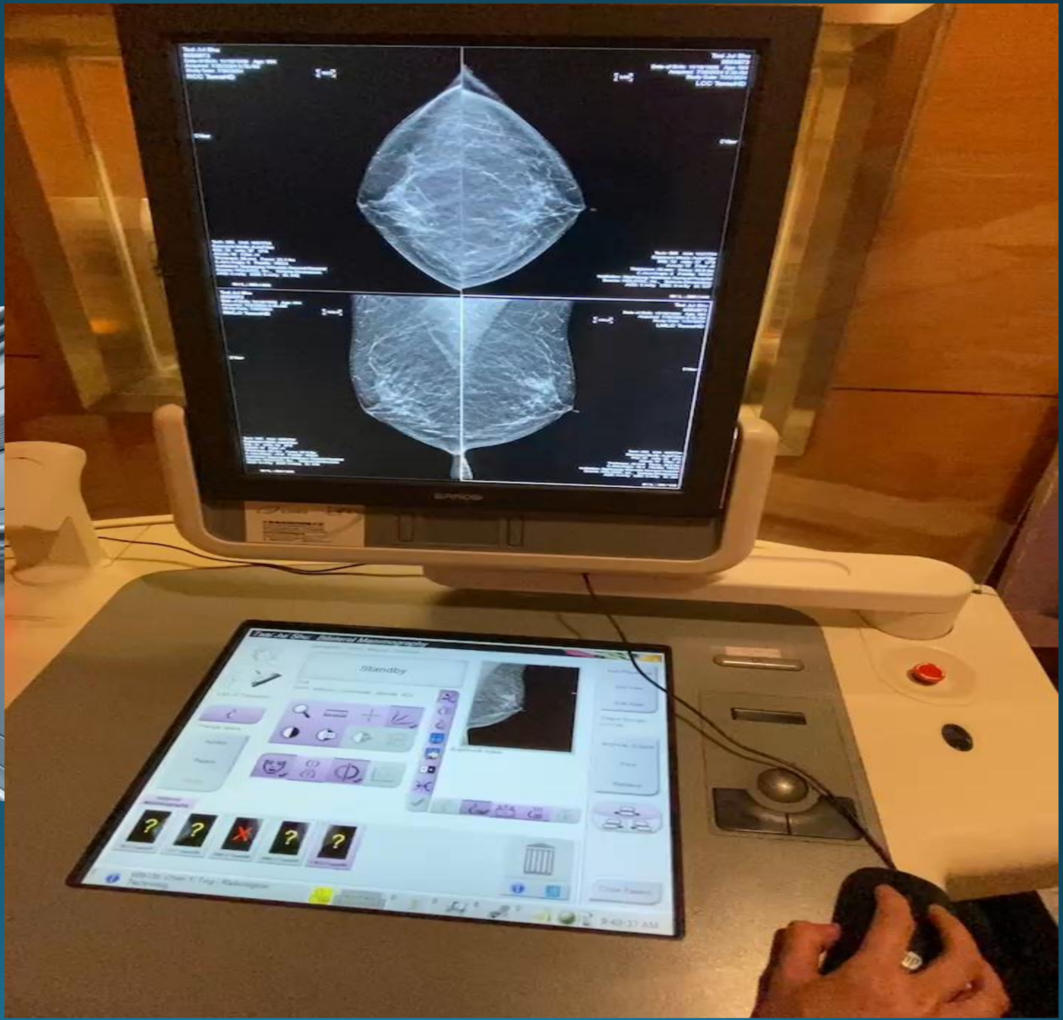
AI-powered iPhone era

- Release year: 2026
- Chip: Apple A19 Pro
- Storage: 256GB / 512GB / 1TB / 2TB
- Battery life: Up to 36 hours video playback

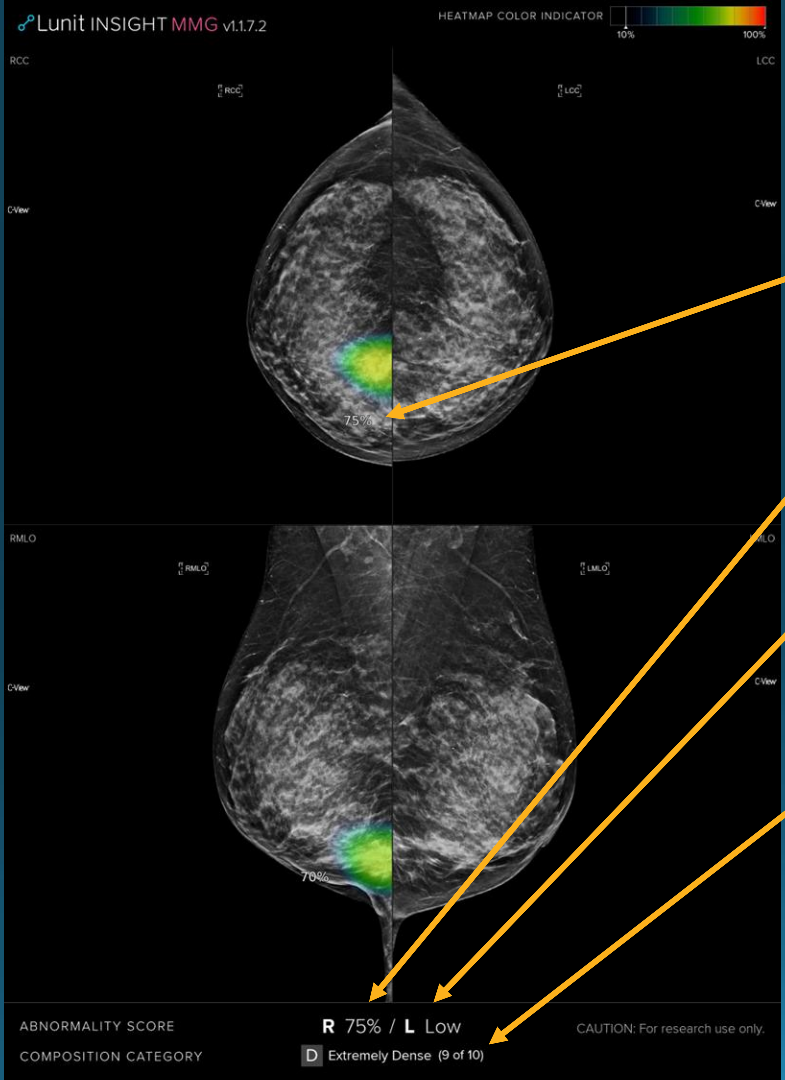
2010: Connect with the world → **2026:** AI that understands you, works for you, and grows with you

AI build in mammography console





Loaded image: DM/ syntheses images from 3D



Lesion score

Unilateral breast score

Case score (max value)

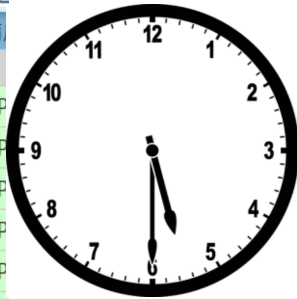
Breast density

Lunit AI triage

AI score

共65 筆查詢結果 1 /3 排序:檢查日期 每頁 25 筆 ☐ 篩選

序	報告	影像	AI Score	病歷號	申請序號	病人姓名	掛號別	病
檢查日期:3天內								
☑ 1		☐	34.09	1942[illegible]	1[illegible]520	林[illegible]	門診	OP
☑ 2		☐	2.44	1943[illegible]	189763	陳[illegible]	門診	OP
☑ 3		☐	13.95	8158[illegible]	190129	曾[illegible]	門診	OP
☑ 4		☐	2.31	1523[illegible]	190333	莊[illegible]	門診	OP
☑ 5		☐	2.01	1811[illegible]	191513	黃[illegible]	門診	OP
☑ 6		☐	0.3	1948[illegible]	191577	黃[illegible]	門診	OP
☑ 7		☐	0.36	4139[illegible]	191839	陳[illegible]	門診	OP
☑ 8		☐	2.72	4308[illegible]	192953	萬[illegible]	門診	OP
☑ 9		☐	0.11	1647[illegible]	193096	程[illegible]	門診	OP
☑ 10		☐	3.1	2491[illegible]	195139	王[illegible]	門診	OP
☑ 11		☐	4.77	1481[illegible]	195425	潘[illegible]	門診	OP
☑ 12		☐	1.22	1726[illegible]	197029	蔡[illegible]	門診	OP
☑ 13		☐	0.45	1176[illegible]	202708	潘[illegible]	門診	OP
☑ 14		☐	2.83	5760[illegible]	202977	傅[illegible]	門診	OP
☑ 15		☐	0.84	5693[illegible]	203039	蔡[illegible]	門診	OP
☑ 16		☐	1.31	1774[illegible]	203680	楊[illegible]	門診	OP
☑ 17		☐	1.09	1629[illegible]	204191	黃[illegible]	門診	OP
☑ 18		☐	1.35	1017[illegible]	204821	巫[illegible]	門診	OP
☑ 19		☐	3.23	1610[illegible]	205443	林[illegible]	門診	OP
☑ 20		☐	0.34	[illegible]	[illegible]	[illegible]	門診	OP



AI score

共65 筆查詢結果 1 /3 排序:AI Score 每頁 25 筆 ☐ 篩選

序	報告	影像	AI Score	病歷號	申請序號	病人姓名	掛號別	病
☑ 1		☐	45.59	1948[illegible]	87	黃[illegible]	門診	OP
☑ 2		☐	34.09	1942[illegible]	20	林[illegible]	門診	OP
☑ 3		☐	17.37	3064[illegible]	18	劉[illegible]	門診	OP
☑ 4		☐	16.0	1948[illegible]	45	蔡[illegible]	門診	OP
☑ 5		☐	13.95	8158[illegible]	29	曾[illegible]	門診	OP
☑ 6		☐	12.82	1922[illegible]	00	游[illegible]	門診	OP
☑ 7		☐	10.27	5258[illegible]	64	邱[illegible]	門診	OP
☑ 8		☐	9.84	1292[illegible]	23	林[illegible]	門診	OP
☑ 9		☐	9.63	1010[illegible]	68	程[illegible]	門診	OP
☑ 10		☐	6.3	1947[illegible]	21	史[illegible]	門診	OP
☑ 11		☐	5.26	4836[illegible]	08	鍾[illegible]	門診	OP
☑ 12		☐	4.77	1481[illegible]	25	潘[illegible]	門診	OP
☑ 13		☐	4.14	1693[illegible]	78	林[illegible]	門診	OP
☑ 14		☐	3.31	1487[illegible]	89	沈[illegible]	門診	OP
☑ 15		☐	3.23	1610[illegible]	43	林[illegible]	門診	OP
☑ 16		☐	3.1	2491[illegible]	39	王[illegible]	門診	OP
☑ 17		☐	2.94	1948[illegible]	88	廖[illegible]	門診	OP
☑ 18		☐	2.84	1596[illegible]	36	張[illegible]	門診	OP
☑ 19		☐	2.83	5760[illegible]	77	傅[illegible]	門診	OP
☑ 20		☐	2.81	1133[illegible]	25	李[illegible]	門診	OP
☑ 21		☐	2.73	1947[illegible]	53	李[illegible]	門診	OP

genius AI™
DETECTION

Patient Manager | User Preferences | About

Patient List | Sessions | Log

Reset Columns | Resend | Notices | Cancel Editing | Create Session | Review

Update Patient List | Merge Patients | Cancel Import | Import... | Suspend And Review

Patient List

Study Date	Name	Patient ID	Reading Priority	CAD Complexity	CAD
2021-08-24	[redacted]	[redacted]	Normal	No findings	Medium
2021-08-24			Normal	No findings	Low
2021-08-24			Normal	Single finding	Low
2021-08-24			Normal	Single finding	Low
2021-08-24			High	Multiple findings	Medium
2021-08-24			Normal	No findings	Low
2021-08-24			Normal	No findings	Low
2021-08-24			Normal	Single finding	Medium
2021-08-24			Normal	Multiple findings	Medium
2021-08-24					+

Lecture Roadmap · **15 Minutes · 5 Core Issues**

15 min

2

AI Detection

What works, what doesn't

3 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2024) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

AI Detection — What 3 Prospective Trials Tell Us

DETECTION · 3 min

~3 min

Trial	Design	n	Detection Δ	Recall Δ	Workload	Reference
PRAIM DE (Germany)	Real-world	463,094	+17.6% ✓*	No change ✓	43% ↓(if triage)	Eisemann, Nat Med 2025
MASAI SE (Sweden)	RCT	106,000	+29% ✓	No change ✓	44.2% ↓	Lång, Lancet Oncol 2023
AI-STREAM KR (Korea)	Prospective	24,500	+13.8% ✓	No change ✓	NR	Chang et al. 2024

+13–29%
Detection increase range
3 prospective trials

~44%
Workload reduction (MASAI RCT)
Lång, Lancet Oncol 2023

minimal
Recall rate change in 3 trials
PRAIM · MASAI · AI-STREAM

** 95% CI: +5.7% to +30.8% | All three trials: no significant recall rate increase — detection gains achieved without increased false positives*

KEY ANSWER Prospective trials across 3 countries consistently show 13–29% higher cancer detection with unchanged recall rates. The evidence is now sufficient for clinical implementation.

2011 Oslo Breast Tomosynthesis Trial

Key Performance Metrics: 2D Mammography vs. 2D+3D (DBT)

Skaane et al. Radiology 2013

Metric	2D	2D+3D	Δ
Detection rate (per 1,000)	6.1	**8.0**	+27%
False positive rate (per 1,000)	61.1	**53.1**	−15%
Positive predictive value	29.1%	**28.5%**	−2%
Reading time	45 s	**91 s**	+100%

+27%
Cancer detection rate
8.0 vs 6.1 per 1,000 screened

−15%
False positive rate
53.1 vs 61.1 per 1,000 — fewer callbacks

+100%
Reading time doubles
91 s vs 45 s — workflow impact

Improved Detection

27% more cancers detected — adding 3D tomosynthesis substantially increases the clinical yield of screening.

- Detection: 8.0 vs 6.1 per 1,000 (2D+3D vs 2D)
- Benefit consistent across all density categories
- Incremental cancers largely early-stage, invasive

Reduced False Positives

15% fewer false positives mean fewer unnecessary callbacks, biopsies, and patient anxiety.

- False positive rate: 53.1 vs 61.1 per 1,000
- PPV marginally unchanged (28.5% vs 29.1%)
- Net benefit: more detection, fewer harms

Workflow Considerations

Reading time doubles to 91 s — a key challenge for implementation planning.

- Full DBT reading: 91 s vs 45 s for 2D alone
- AI triage can reduce reads by 40–50% (PRAIM 2025)
- Staffing models must account for volume impact

Source: Skaane P et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology. 2013;267(1):47–56.

Table : Diagnostic Performance — AI System (2D) Alone vs. Radiologists (2D+3D)

Screening cohort: 2,190 women | Breast cancers identified: 34 | Follow-up: 2 years

	AI Alone (2D)		Radiologist (2D+3D)		
Metric	**Events / Total**	**%**	**Events / Total**	**%**	**P-value**
Sensitivity	24/34	**70.6%**	27/34	**79.4%**	0.125
Specificity	1897/2156	**88.0%**	2066/2156	**95.8%**	**<0.001*****
Accuracy	1924/2190	**87.7%**	2039/2190	**95.6%**	**<0.001*****
PPV	24/283	**8.5%**	27/116	**23.3%**	**0.004****
NPV	1897/1907	**99.5%**	2066/2073	**99.7%**	0.899
Recall Rate	283/2190	**12.9%** **(ACR <12%)**	116/2190	**5.3%**	**<0.001*****

***P < 0.01 ***P < 0.001 PPV: Positive Predictive Value NPV: Negative Predictive Value AI: Artificial Intelligence*

Comparative Study of Artificial Intelligence-based System Alone in Synthesized Mammography Versus Radiologist's Interpretation of Digital Breast Tomosynthesis in Screening Women

Hsin-Yi Li[1], Yun Lin[1], Yu-Ting Hong[1], Chen-Pin Chou[1,2]*

[1]Department of Radiology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan, [2]Department of Medical Laboratory Science and Biotechnology, Fooyin University, Kaohsiung, Taiwan

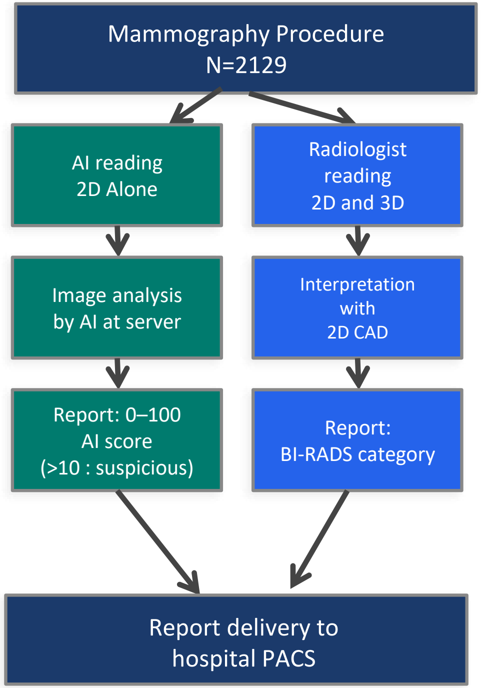
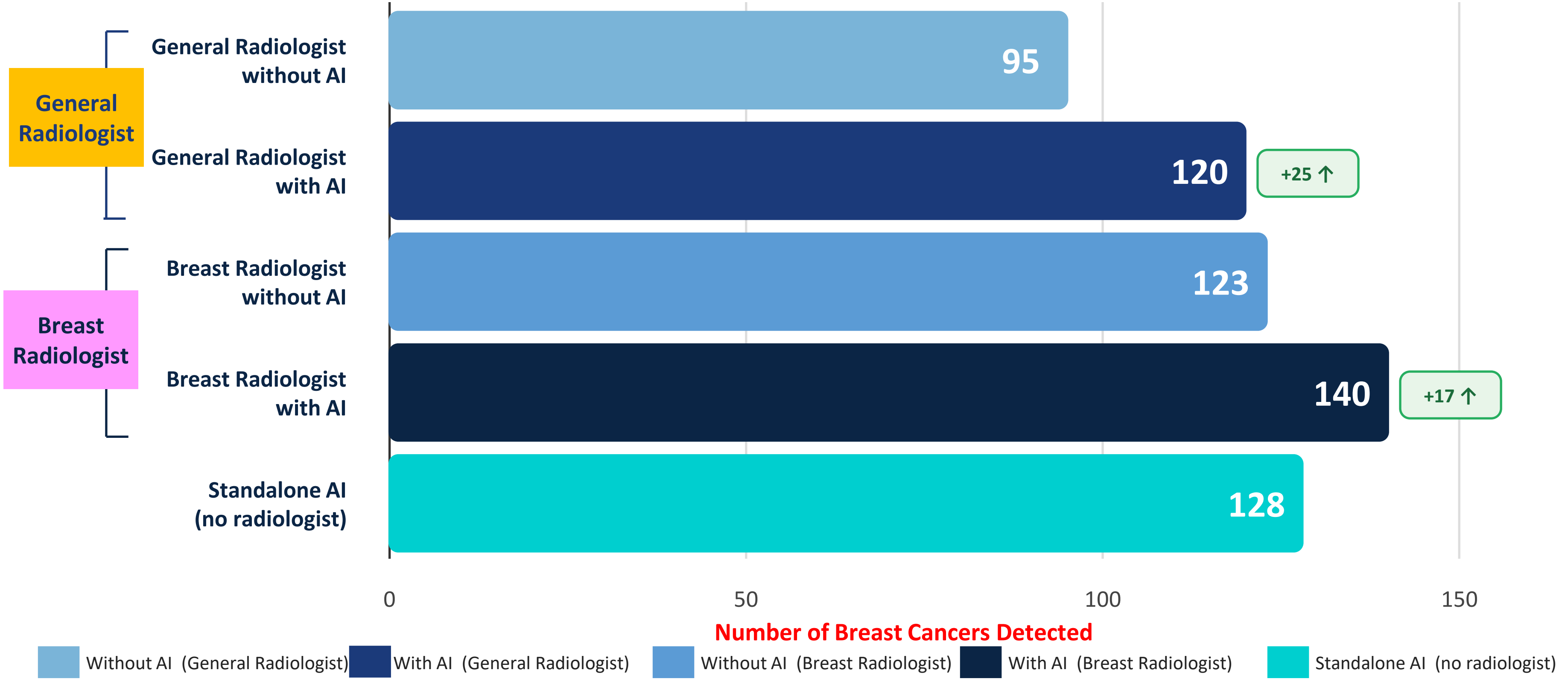


Figure 2. Study workflow — parallel AI and radiologist reading pathways

Breast Cancer Detection by Reading Strategy

AI-STREAM Study · Prospective multicenter cohort · Screen-detected biopsy-proven cancers



Key finding: AI boosted detection for both groups — Breast Radiologist +17 (123→140) and General Radiologist +25 (95→120). Standalone AI detected 128 cancers, between both radiologist groups.

* Breast Radiologist: experts in breast imaging with >10 years of experience   † General Radiologist: radiologists not specializing in breast imaging   **Source:** *AI-STREAM, Nature Communications 2025*

Lecture Roadmap · **15 Minutes · 5 Core Issues**

15 min

3

Safety & Recall

Does AI increase false positives?

2 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Why AI Does NOT Increase Recall Rate — Doing More Good Than Harm

The core concern: Could AI increase false-positive recalls — causing unnecessary biopsies, patient anxiety, and cost?
Prospective evidence says: No.

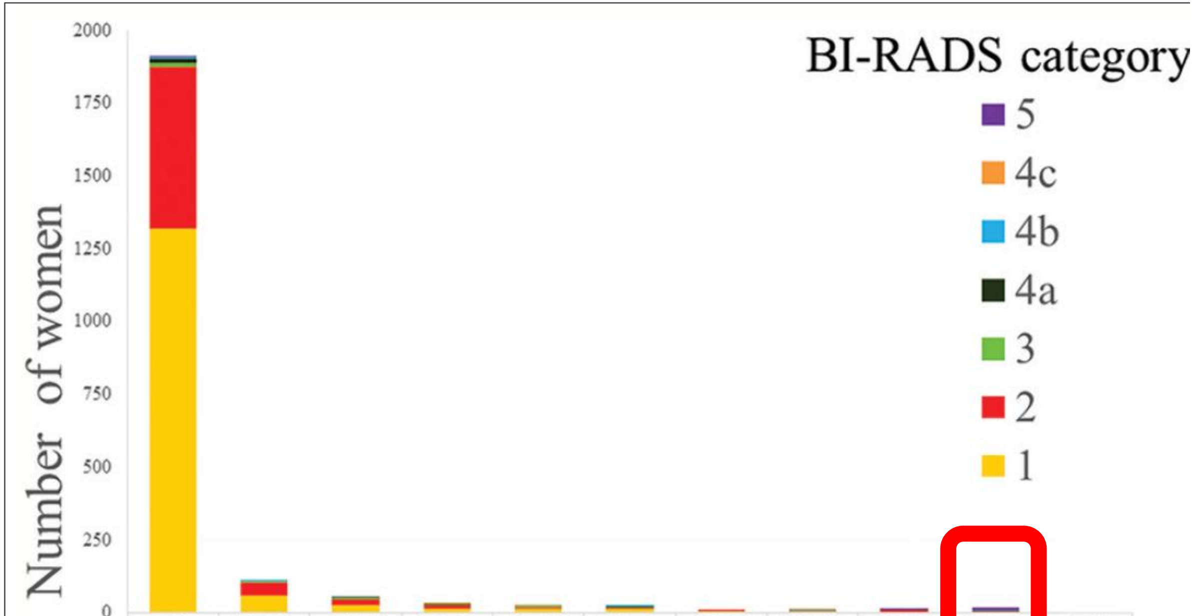
PRAIM DE	MASAI SE	AI-STREAM KR
Eisemann et al. *Nature Medicine 2025*	*Lång et al.* *Lancet Oncology 2023*	*Chang et al. 2024* *(Asian single-read cohort)*
n = 463,094	n = 106,000 (RCT)	n = 24,500
Recall (AI-assisted): **3.7%**	Recall (AI-triage): **2.7%**	Recall (AI-aided): **No change**
Recall (standard): **3.8%**	Recall (standard DR): **2.8%**	Recall (standard): **Baseline**
✓ **No significant difference** — Detection +17.6% with same recall	✓ **No significant difference** — Detection +29% Workload −44%	✓ **No significant difference** — Detection +13.8% Single-read setting
Safety net activates only AFTER both radiologists score normal — AI never independently triggers recall	*AI triage to single-read for low-risk — radiologist still makes all recall decisions*	***Most relevant to Taiwan: AI-CAD in single-reader program — no recall rate increase in Asian population***

Why no harm: AI is advisory, not autonomous. The radiologist retains all recall decisions. AI flags alert only when the human already scored normal — adding a second check, not bypassing clinical judgment.

Li, et al.: AI-based system alone for screening mammography (VGH-KS)

Select >90 (Lunit) for safety net
Only 1.3% (3/2190) additional recall

ACR suggested recall <12%



	<10	≥10-20	≥20-30	≥30-40	≥40-50	≥50-60	≥60-70	≥70-80	≥80-90	≥90-100	total
5	1	0	0	0	0	0	0	0	1	6	8
4c	0	0	0	1	1	0	0	1	0	1	4
4b	2	1	0	0	1	2	0	3	0	3	12
4a	19	0	1	3	1	3	0	0	1	3	31
3	12	5	4	1	1	1	0	0	0	0	24
2	555	44	21	10	4	6	4	3	7	2	656
1	1318	59	26	15	15	12	5	3	1	1	1455

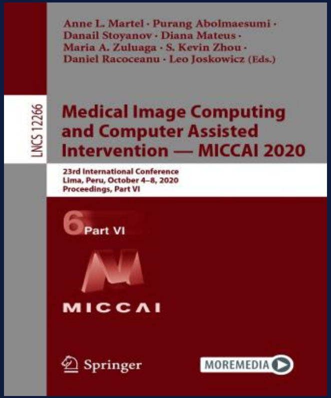
AI-based score

a

AI as a Radiologist's Safety Net

KEY CONCLUSIONS

Chorev et al.
MICCAI 2020
LNCS 12266, pp. 220–229



10.7%
Missed Cancers Detected (DM only)
10.68 alerts / 1,000 cases

80.3%
Sensitivity with AI (DM only)
vs. 77.9% baseline (p<0.001)

95.3%
Specificity with AI (DM only)
vs. 96.4% baseline (p<0.001)

Bootstrapped to real-world prevalence; 2,638 held-out exams enriched with 90 missed cancers (IBM Research / Assuta / Maccabi)

Multimodal Setting (DM + US)

8.51 alerts/1,000 cases → detects 11.7% of missed cancers
Sensitivity: 92.3% vs. 91.3% (p<0.001)
Specificity: 90.8% vs. 91.7% (p<0.001)
AI contribution remains significant even with US

Post-hoc Visibility (Expert Review)

All 90 FN cases reviewed by expert breast radiologist (>20 yrs)
AI identified 11/90 missed cancers
· 8 of 11 were visible post-hoc on index DM
· 3 of 11 were not visible post-hoc
AI caught proportionally more visible cancers (p = 0.018)

Reader Study — 5 Radiologists

120 exams; 10 originally missed cancers (FN); AI at 99% specificity
AI detected 4 additional missed cancers across 3 readers:
· Reader 2: +1 TP · Reader 3: +2 TP · Reader 5: +1 TP
Readers 1 & 4: no change in TP or FP with AI

False Alerts — Important Distinction

Held-out set (real-world): 10.05 false alerts/1,000 (DM only); 8.24 false alerts/1,000 (DM+US) — these are true false positives
Reader study only: AI added NO new false alarms to any reader
⚠ 'Zero false alerts' applies to reader study, not real-world data

Conclusion: AI safety net significantly improves sensitivity in both DM-only and multimodal settings. False alerts exist in real-world data; the reader study showed no added false alarms — a controlled but encouraging finding.

Lecture Roadmap · 15 Minutes · 5 Core Issues

15 min

4

Interval Cancers

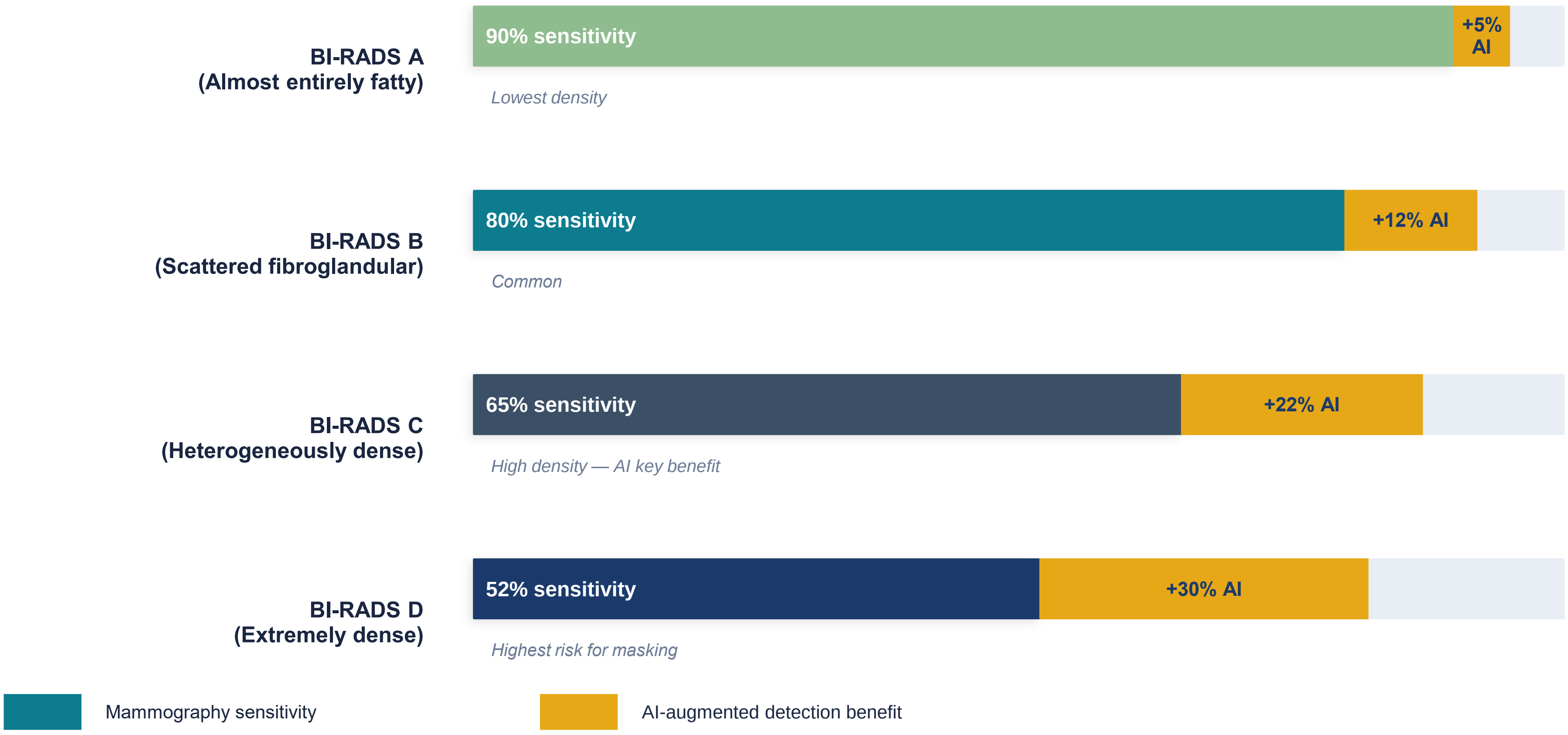
Where AI can and can't help

3 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Dense Breasts & AI Risk Prediction — Nuanced Evidence

Q4: Breast Density, Sensitivity Drop, and AI Benefit



Note: Values illustrative based on published ranges. McKinney et al. Nature 2020; EUSOBI 2022; Eisemann et al. 2025

Interval Cancers (IC) — Where AI Can and Cannot Help

INTERVAL CANCERS · 3 min

~3 min

~20–30%	~19.3%*	~50%
Missed Cancers (Perception Error) *Visible on prior mammogram but not detected*	**Minimal-Sign Interval Cancers** *AI score 10 in retrospective review*	**True Interval Cancers (Occult)** *No mammographic visible signs*
▼	▼	▼
AI localises lesion → Safety net alert	AI scores highly after read → Safety net activates	AI CANNOT detect → Ultrasound or MRI required

** Lång 2021 (Eur Radiol): 83/429 IC cases AI score 10 correctly localised · Up to 47% relative IC reduction possible (Larsen 2022, Scenario 2: IC rate 1.7→0.9/1,000)*

11.2%

IC reduction at 4% recall

At a 4% recall threshold, AI safety-net protocols could reduce interval cancer rates by 11.2%.

1 in 3

ICs scored AI risk = 10

143 of 429 interval cancers received the highest AI risk score of 10 at prior screening.

23%

of poor-outcome ICs identified

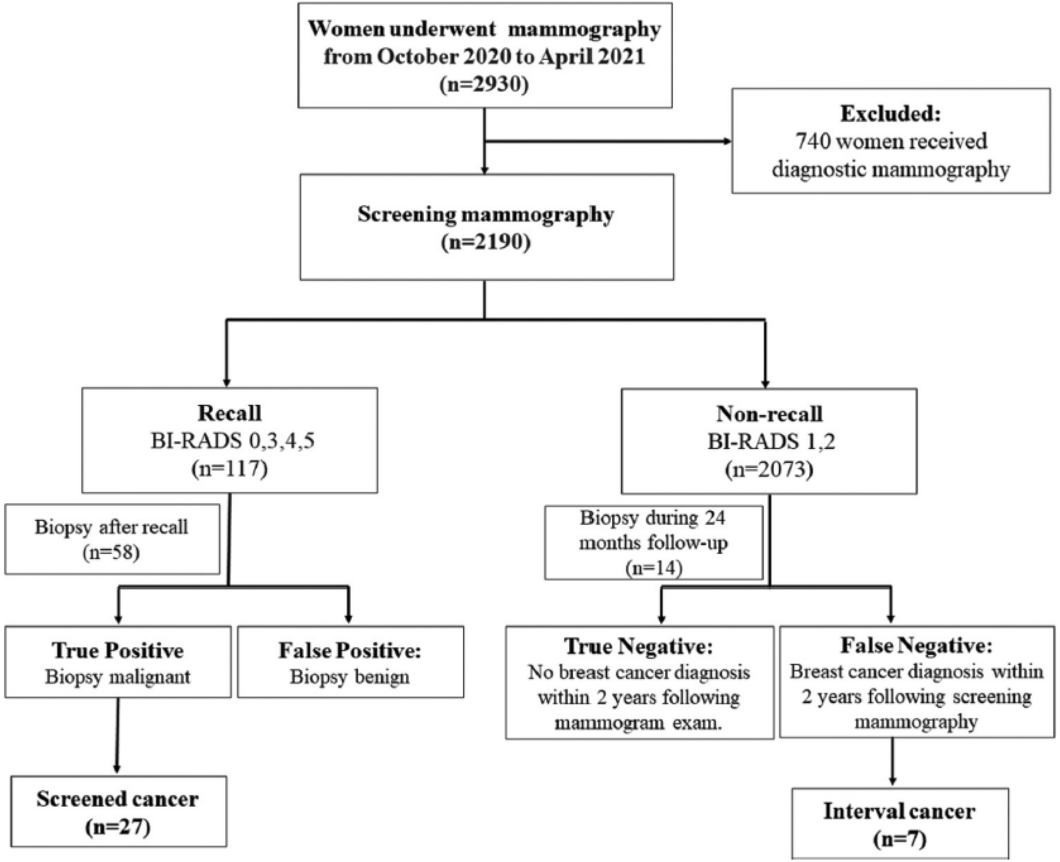
8 of 35 interval cancers with grave outcome (death or stage IV) were correctly localised and flagged as high risk by AI.

KEY ANSWER AI can address ~50% of interval cancers (perception errors + minimal-sign). The remaining ~50% are truly occult — AI cannot replace supplemental MRI/ultrasound for high-risk women.

Comparative Study of Artificial Intelligence-based System Alone in Synthesized Mammography Versus Radiologist's Interpretation of Digital Breast Tomosynthesis in Screening Women

Hsin-Yi Li[1], Yun Lin[1], Yu-Ting Hong[1], Chen-Pin Chou[1,2*]

Radiologist IC: 7/2,190 (3.2/1,000 screens) on BDT screening



Lunit AI Triage — Screening Mammography (N = 2,190)

Oct 2020 – Apr 2021 · VGHKS · AI threshold ≥10% = recall · AI sensitivity 70.6%

Lunit Insight MMG v1.1.2.0

All Scans · N = 2,190

87%

AI Score < 10% — Low Risk
n = 1,907 women (87%)

13%

AI Score ≥ 10% — High Risk
n = 283 women (13% recall rate)

AI alone detected 2/7 ICs (28.6%)

97%

Missed by Both

AI(−) Radiologist(−)

n = 1,843 · BI-RADS 1–2

Benign: 12

Malignant: 5 ⚠

IDC 3 · DCIS 1 · Mucinous 1

Interval cancers (missed by AI & radiologist)

5

3%

Caught by Radiologist

AI(−) Radiologist(+)

n = 64 · BI-RADS 0,3,4,5

Benign: 19

Malignant: 5

DCIS 3 · IDC 2

Radiologist rescued what AI missed

5

VS

AI Thresh

81%

Caught by AI

AI(+) Radiologist(−)

n = 230 · BI-RADS 1–2

Benign: 2

Malignant: 2

IDC 1 · DCIS 1

AI rescued what radiologist missed

2

19%

Caught by Both

AI(+) Radiologist(+)

n = 53 · BI-RADS 0,3,4,5

Benign: 12

Malignant: 22 ★

DCIS 6 · IDC 12 · ILC 2 · Others 2

Highest yield group (AI + radiologist agree)

22

Total cancers found: **5** Missed (both) + **5** Radiologist only + **2** AI only + **22** Both agree = **34** cancers

Li HY, Lin Y, Hong YT, Chou CP. J Radiol Sci 2024;49:59–65 · Lunit Insight MMG v1.1.2.0 · Threshold ≥10% · ⚠ = interval cancers missed by both AI and radiologist · ★ = highest-yield group

Comparative Study of Artificial Intelligence-based System Alone in Synthesized Mammography Versus Radiologist's Interpretation of Digital Breast Tomosynthesis in Screening Women

Hsin-Yi Li[1], Yun Lin[1], Yu-Ting Hong[1], Chen-Pin Chou[1,2*]

[1]Department of Radiology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan, [2]Department of Medical Laboratory Science and Biotechnology, Fooyin University, Kaohsiung, Taiwan

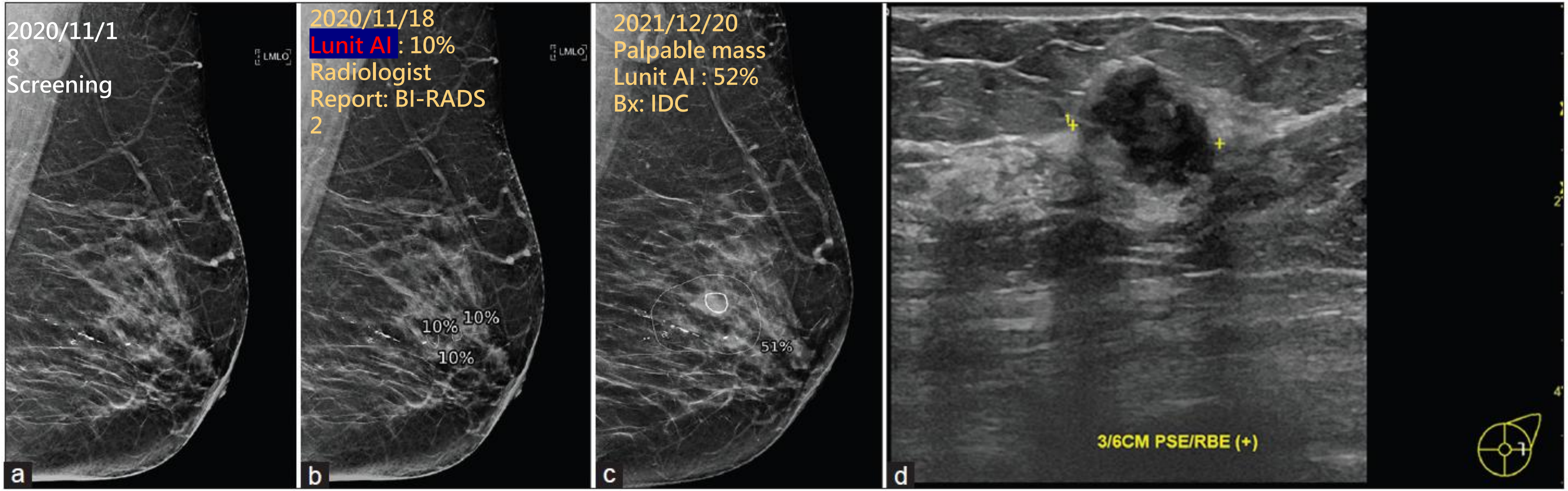


Figure 4: A 50-year-old woman with interval breast cancer. (a) A synthesized mammography (SM) mediolateral oblique view of the left breast assigned as breast imaging reporting and data system 2 by the radiologist at a screening examination. (b) The artificial intelligence (AI) system identified an asymmetry in the upper hemisphere of the left breast, reporting an AI-based score of 10. (c) A diagnostic SM in the subsequent year, prompted by a palpable lump, shows an irregular mass in a similar region previously pinpointed by the AI system, leading to a significantly heightened AI score of 51. (d) A targeted breast ultrasound confirmed an irregular mass

AI (Synthesized Mammography) vs. Radiologist (DBT)

Li HY et al. · J Radiol Sci 2024;49:59–65 · Kaohsiung Veterans General Hospital, Taiwan

J Radiol Sci 2024
n = 2,190 women
Retrospective · Taiwan

Study Design

Retrospective · Single institution · Taiwan
Oct 2020 – Apr 2021
n = 2,190 screening women
Age mean 56.2 yrs · 77% age ≥50
Breast density 66% Category C (dense)
24-month follow-up for interval cancers

Two Reading Scenarios

AI Alone
Lunit Insight MMG
SM images only
Score ≥10% = recall

VS

Radiologist
DBT + SM + R2 CAD
5 rads · avg 12 yrs exp
BI-RADS 0/3/4/5 = recall

Cancer Cases (n=34)

Screen-detected: 27 (79.4%)
Interval cancers: 7 (20.6%)

Invasive cancer: 67.6%
IDC: 52.9% · ILC: 5.9%
Mean tumor size: 15.8 mm

Metric	AI Alone	95% CI	Radiologist	95% CI	P value
Sensitivity	70.6%	52.5–84.9	79.4%	62.1–91.3	0.125
Specificity	88.0%	86.5–89.3	95.8%	94.9–96.6	★ <0.001
Accuracy	87.7%	86.3–89.1	95.6%	94.6–96.4	★ <0.001
PPV	8.5%	5.5–12.4	23.3%	15.8–31.8	★ 0.004
NPV	99.5%	99.0–99.8	99.7%	99.3–99.9	0.899
Recall Rate	12.9%	11.5–14.3	5.3%	4.3–6.2	★ <0.001

Key Takeaways

Similar Sensitivity
AI 70.6% ≈ Rad 79.4%
(non-inferior, P = 0.125)

Lower Specificity & PPV
Spec 88% vs 96%
PPV 8.5% vs 23.3% (P <0.001)

Higher Recall Rate
AI 12.9% vs Rad 5.3%
2.4× more false positives

Best Use Case
AI as triage / pre-assessment
before radiologist DBT review

★ significant · AI: Lunit Insight MMG v1.1.2.0 · DBT: Hologic Dimensions · SM: C-view (Hologic) · BI-RADS 5th ed. · IRB: VGHKS 3.3013.2118 · DOI: 10.4103/jradiolsci.JRADIOLSCI-D-23-00027

genius AI™
DETECTION

Genius AI DBT Triage — Screening Mammography (N=1,780)

July 2021 – June 2022 · AI threshold ≥2% = recall · Total cancers: 18

AI threshold
≥2% = recall

Women underwent mammography
July 2021 – June 2022 (n=2,679)

Excluded (n=899)
Symptomatic breast disease
Diagnostic mammograms for abnormal screening

50% | 50%

Screening Mammography
(n = 1,780)

50% | 50%

No-Recall — Genius AI < 2%
n = 892 (50%)

Recall — Genius AI ≥ 2%
n = 888 (50% recall rate)

98% | 2% | 91% | 9%

AI alone detected 3/6 ICs (50%)

Missed by Both

AI(−) Radiologist(−)

No-Recall · BI-RADS 1–2

n = 876 (98%)

Benign: 5

Cancer: 3 ⚠

Missed by both

3

Caught by Radiologist

AI(−) Radiologist(+)

No-Recall · BI-RADS 0,3,4,5

n = 16 (2%)

Benign: 4

Cancer: 0

Radiologist found what AI missed

0

vs
AI
2%

Caught by AI

AI(+) Radiologist(−)

Recall · BI-RADS 1–2

n = 808 (91%)

Benign: 18

Cancer: 3 ⚠

AI found what radiologist missed

3

Caught by Both

AI(+) Radiologist(+)

Recall · BI-RADS 0,3,4,5

n = 80 (9%)

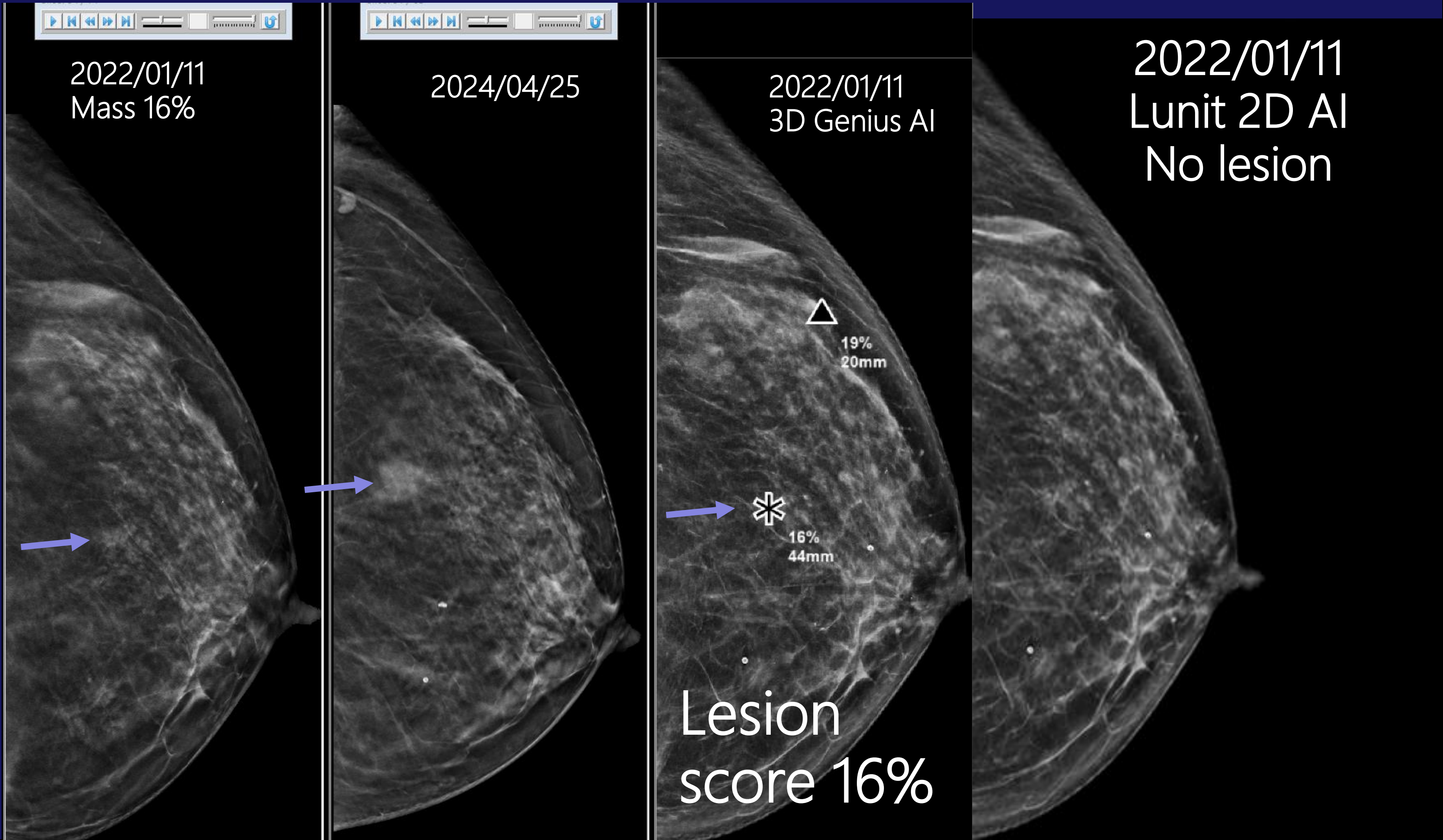
Benign: 22

Cancer: 12 ★

Highest yield (both agree)

12

Total cancers found: 3 Missed (both) + 0 Radiologist only + 3 AI only + 12 Both agree = 18



2022/01/11
Mass 16%
2024/04/25
2022/01/11
3D Genius AI
19%
20mm
16%
44mm
Lesion
score 16%
2022/01/11
Lunit 2D AI
No lesion

AI Triage First — Fast-Track Normal Exams, Focus Radiologists on High-Risk Cases

Low-Risk Lane
(Normal scans — fast track)

High-Risk Lane
(Suspicious — specialist reads)

Breast Cancer Risk Prediction

Artificial intelligence models substantially outperform classical clinical risk scoring

Tyrer-Cuzick

Classical clinical model

History-based scoring only — family history, hormonal, and reproductive factors. No imaging input; marks the ceiling of clinical-only risk assessment.

Mirai

Deep learning — 5-year mammography risk

Predicts 5-year risk from mammography images alone. Validated across US, Sweden, and Taiwan; extracts imaging biomarkers invisible to human readers.

Gjesvik 2024 (Lunit Insight)

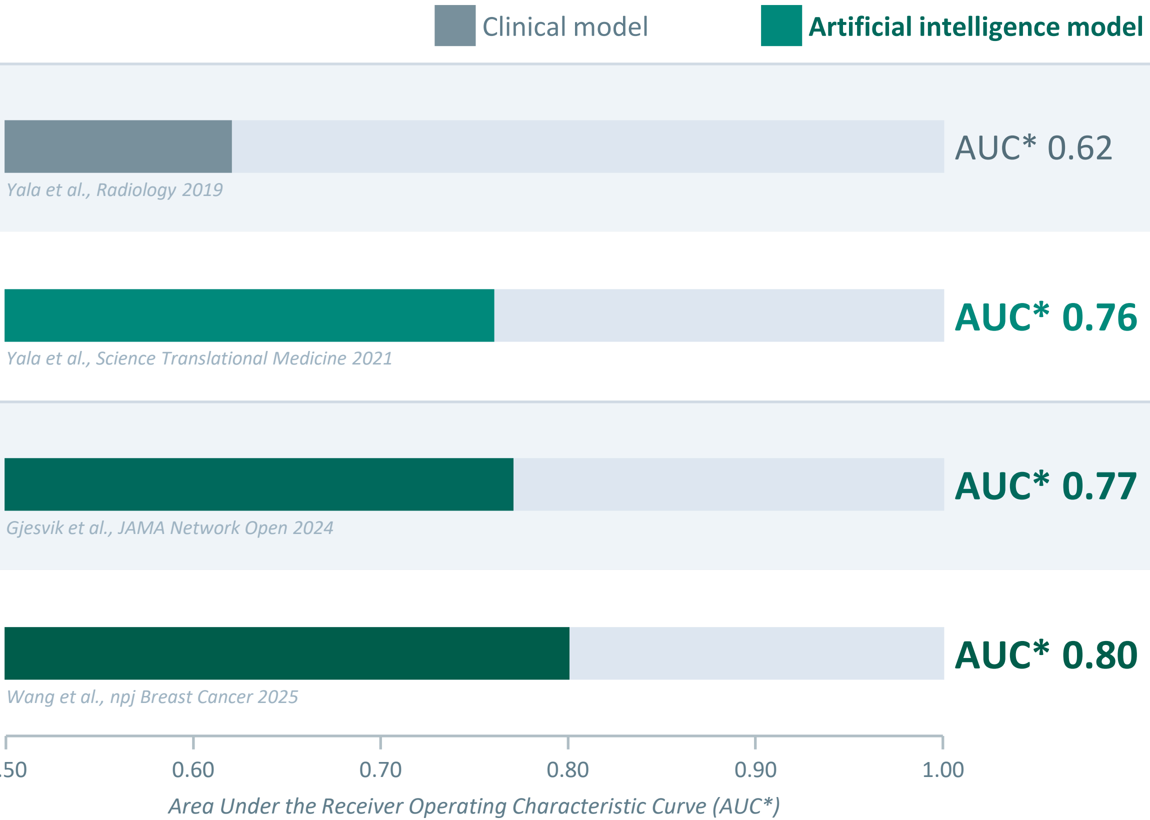
Artificial intelligence — BreastScreen Norway (3 rounds)

Change in AI score across 3 consecutive screening rounds predicts interval cancer (AUC 0.77, round 3). Detects subclinical cancer 4-6 years before diagnosis.

MTP-BCR (Wang 2025)

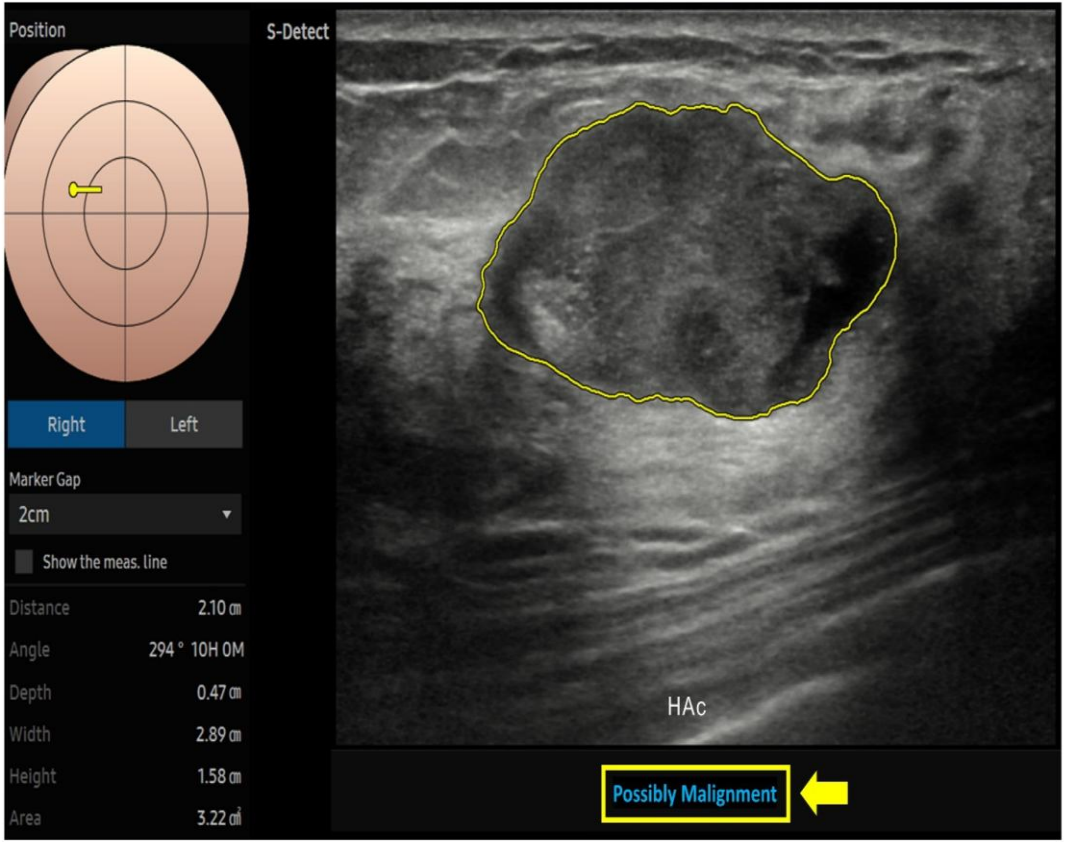
Deep learning — multi-timepoint longitudinal analysis

Integrates up to **5 prior mammography exams + clinical risk factors**. **Achieves 10-year AUC 0.80** (95% CI 0.78-0.82); outperforms all single-timepoint models.



*** AUC (Area Under the Curve)**: measures how well a model separates women who will develop breast cancer from those who will not. AUC = 1.0: perfect prediction; AUC = 0.5: no better than chance.

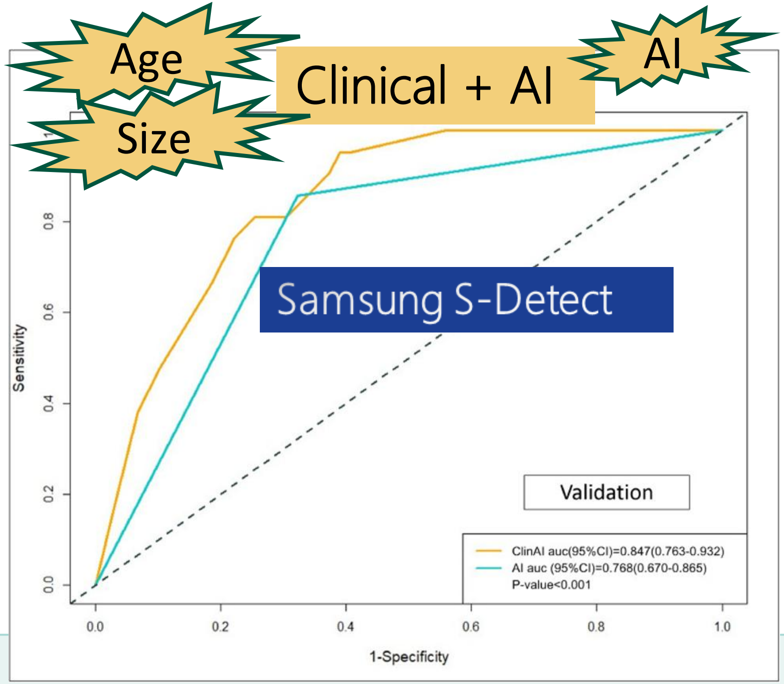
AI Risk Prediction — Earlier Detection, Better Discrimination

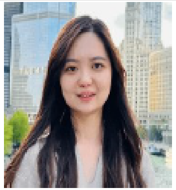


Ultrasound Med Biol. IF: 2.4 Q2 2024 Sep;50(9):1372-1380.
doi: 10.1016/j.ultrasmedbio.2024.05.012. Epub 2024 Jun 19.

Enhancing Breast Cancer Diagnosis: A Nomogram Model Integrating AI Ultrasound and Clinical Factors

Zi-Han Yu [1], Yu-Ting Hong [2], Chen-Pin Chou [3]



Time	Instructor		
14:10-14:30	Bayesian Network and Deep Learning Approaches for Personalized Multi-State Disease Natural History Modeling	 **Ting-Yu Lin** **林庭瑀博士** National Taiwan University 臺灣大學	
14:30-15:00	AI-Empowered Digital Twin Frameworks for Precision Screening		
15:00-15:20	**Break** *Changhua Christine Hospital Museum Tour Departure (by Invitation)*		
15:20-15:50	Reinforcement Learning–Enabled Digital Twins for Risk-Adaptive Screening Strategies with LLM	 **Ting-Yu Lin** **林庭瑀博士** National Taiwan University 臺灣大學	 **Ming-Fang Yen** **嚴明芳教授** Taipei Medical University 臺北醫學大學
15:50-16:10	Digital Twin AI Demonstration for Precision Screening Applications		
16:10-16:20	**Closing Remark**	 **Chen-Yang Hsu** **許辰陽理事長** Taiwan Association of Medical Screening 台灣篩檢學會	

Trial Leadership & Risk Protocols

PROTOCOL & RESULTS MATRIX

Dr. Laura J. Esserman, MD, MBA
PRINCIPAL INVESTIGATOR · UC SAN FRANCISCO
Co-Investigators: Dr. Alexander Borowsky (Pathology Lead, UC Davis) · Dr. Allison Stover Fiscalini (Clinical Operations Director)

Dr. Suzette Delaloge, MD, PhD
GLOBAL COORDINATOR · GUSTAVE ROUSSY, FRANCE
Co-Investigators: Dr. Corinne Balleyguier (Lead Radiologist) · Dr. Sandrine de Montgolfier (Ethics & Policy Director)

Clinical Metric	WISDOM Trial (United States)	MyPeBS Trial (Europe)
High-Risk Ratio *Distribution & Triage*	~32% HIGH RISK ~32% classified as High Risk Includes "Elevated" and "Highest" genomic & clinical categories (Breast Cancer Surveillance Consortium v2 + Polygenic Risk Score).	~35% HIGH RISK ~35% classified as High Risk Includes "High" and "Very High" risk brackets calculated via Marmorisk/TC + Polygenic Risk Score 313.
Scan Protocols *Risk-Based Suggestions*	Very High Risk: Alternating Mammogram & MRI every 6 months Average/Low Risk: Biennial (2-year) mammography starting at age 50	Very High Risk: Annual Mammogram + Annual MRI from age 30 Low Risk: Extended 4-year mammography intervals
Safety Endpoints *Stage IIB+ Incidence*	NON-INFERIORITY MET · JAMA 2025 **30.0 cases** per 100,000 person-years (risk-stratified arm) vs. **48.0 (standard arm)** No advanced cancers detected in highest-risk tier.	PRIMARY COMPLETION ESTIMATED 2027 Evaluating non-inferiority rates Joint baseline profiles match WISDOM curves, suggesting comparable safety modeling.
Secondary Yields *Biopsies & Genetics*	Biopsies: No significant difference (943 vs 1,029, p=0.1) Genetics: **30% of pathogenic variant carriers lacked any family history**	Biopsies: Projecting lower benign biopsy recalls in low-risk group Genetics: Joint analysis shows identical genomic triage accuracy

15 min

5

Asia/Taiwan 2030

Precision screening roadmap

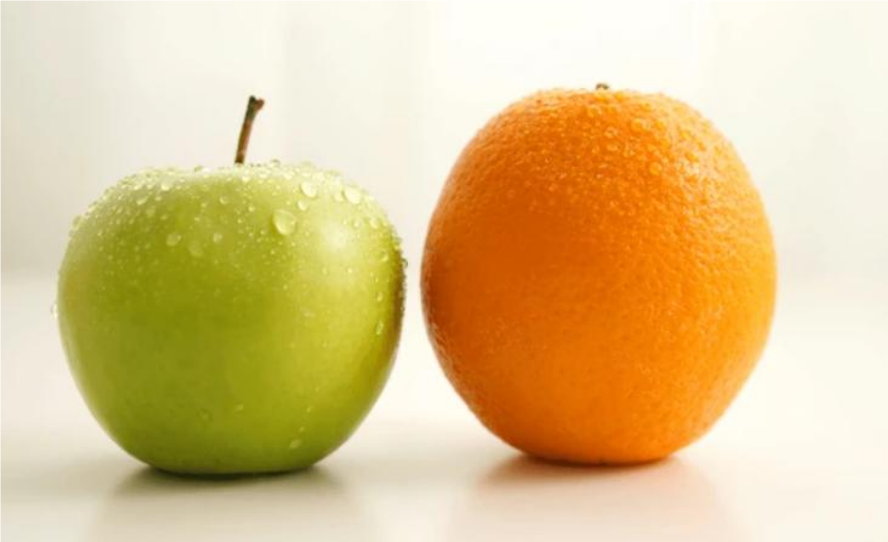
3 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Why real-world data

Designed study
Selected cases

Real-world
Un-selected cases



Study Design Influences AI Performance Metrics

Study Type	What It Means	What's the Problem	Good for Real-Word?
Selected Cases	Only includes obvious cancer or clearly normal cases	Makes AI look better than it really is (AUC too high)	✗
Pair Testing	AI and radiologists read the same cases	More fair, but still not real clinical flow	Somewhat
Simulated Test Sets	Test set is picked or mixed, not real screening order	Unreal mix of cases, not like real practice	✗
Real-World Test Set	Real screening order all types of cases included	Most realistic and reliable test	Yes – Best choice

Issue 5 · Asia/Taiwan Implementation — Two-Phase Roadmap to future

IMPLEMENTATION · 3 min

~3 min

Phase 1 · Retrospective Validation 6–12 months

1. ≥50,000 exams from national programme (multiple HPA centers)
2. AI processing: Transpara / Vara / Lunit — stratify by density & age
3. Link cancer registry; identify interval cancers
4. Baseline report: AI score distribution, sensitivity, specificity

Phase 2 · Prospective Safety Net Implementation 12–36 months

1. ≥5 HPA sites, IRB approval, PRAIM safety net model
2. Radiologist reads first → AI flags ~1.5% highest-risk exams after
3. Interim analysis: target ≥15% IC reduction, no recall increase
4. Cost-Effectiveness analysis → national rollout recommendation to Health Promotion Administration

Primary Endpoints:

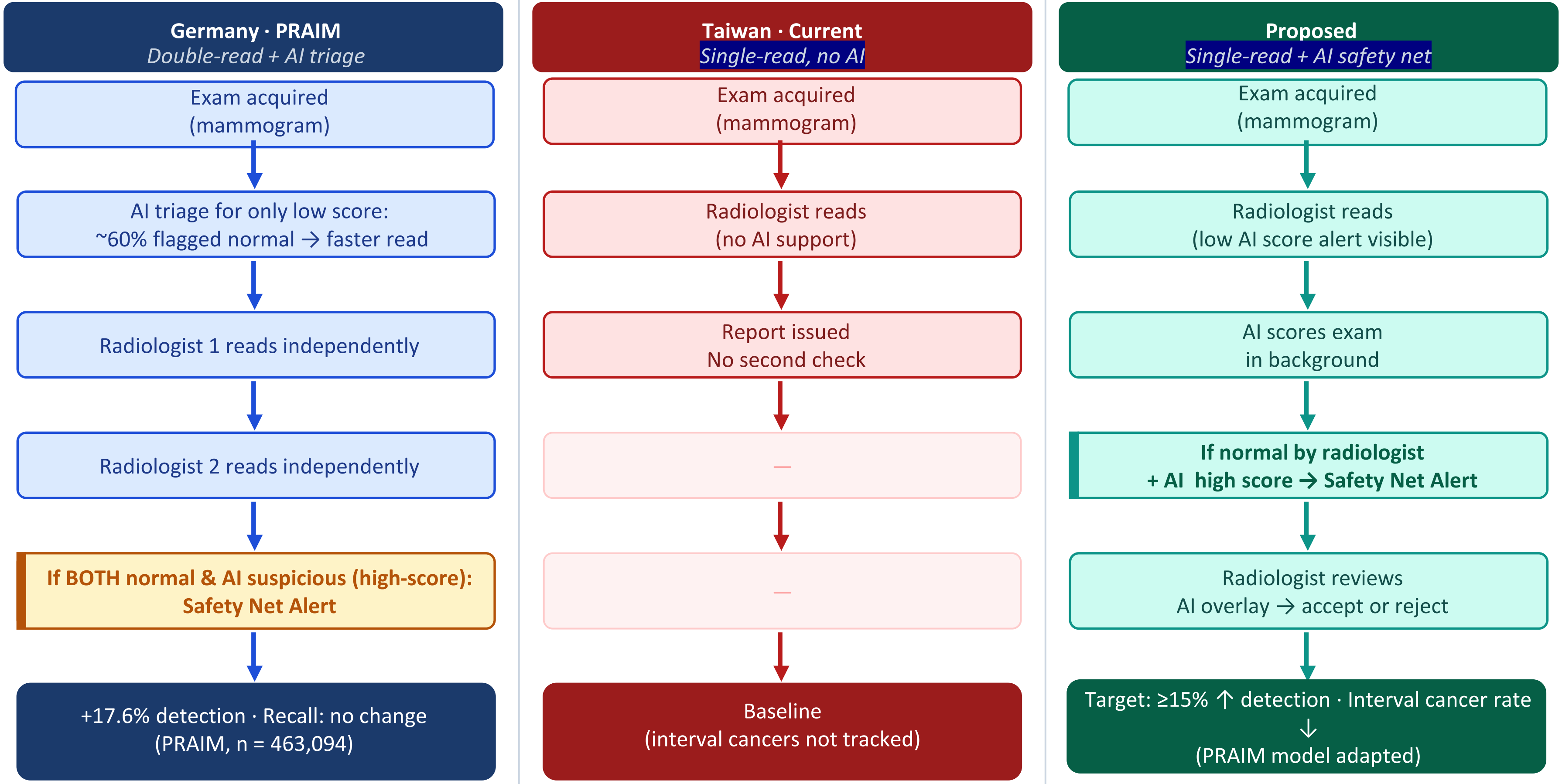
- Detection: ≥15% SDC↑
- Interval Cancers: ≥15% IC↓
- Recall Rate: No change
- Economics: Cost/effect accept.

KEY ANSWER Start with Phase 1 retrospective validation (6–12 mo) before committing resources. PRAIM safety-net model is most translatable to Taiwan's single-read system. Target: ≥15% detection gain.

Safety Net Workflow

Germany · PRAIM vs Taiwan Current vs Taiwan Proposed

PRAIM · Nat Med 2025



Source: Eisemann N et al. AI-supported mammography screening (PRAIM). Nature Medicine 2025. Taiwan proposed workflow adapted from PRAIM safety-net model

Figure · **Breast Screening in 2030 — Radiologist Role Transformation**

~1 min

Dimension	TODAY (2024–2026)	BREAST SCREENING 2030
Role	Detector & Reporter	**Integrator & Decision Architect**
Workflow	Read → Report → Recall	**(1) AI pre-read → Flag review → Management** **(2) Read → Report →AI Flag → Review (or recall)**
Screening	Age-based, uniform	**AI + genetic risk-stratified, personalized**
Dense Breasts	Limitations acknowledged	**AI-guided DBT / CE-Mammo / MRI triage**
Risk Assessment	BI-RADS only	**AI biomarker + Polygenic Risk Score + temporal change**
Workload	High; rising volume	**AI triage reduces normal reads by ~44%**

Developing Mammography AI in Taiwan

A straightforward analysis of opportunities, risks, and next steps compared to international models

The Chances

Dense Breast Issue
Over 80% of premenopausal and ~55% of postmenopausal East Asian women have dense breast tissue (BI-RADS C/D), making cancers harder to spot on 2D scans. AI acts as a super-powered magnifying glass.

Overworked Doctors
Radiologists are in short supply. AI can instantly flag normal scans, allowing doctors to focus solely on high-risk cases.

The Risks

Strict Data Laws
Taiwan's 2025 NHI Data Act strictly limits how patient scan data can be shared, making training custom AI models highly regulated and slow.

Screening Rates
Despite free biennial mammography for women aged 40–74, Taiwan's screening rate is only 40–50% (Taiwan Breast Cancer Society, 2026). AI cannot help if patients do not get screened.

The Strategy

Focus on 3D Scans
Build AI optimized for modern 3D-DBT imaging and breast ultrasounds, rather than saturated 2D systems.

Self-Pay Market
Bypass complex national healthcare reimbursement models by selling directly to premium, private wellness clinics.

What We Learn from Lunit (Korea)

Global Expansion: Lunit earned 88% of 2024 revenue overseas; Q1 2026 hit 97%, deployed in 4,800+ institutions across 55 countries — global-first strategy.

Fast Approvals: Korea's 2024 Immediate Market Entry system cut approval time from up to 490 days to 80–140 days for AI-based medical devices with international clinical evidence.

Sensitivity: AI Alone vs. Radiologist (DBT)

Source: Li HY et al., J Radiol Sci 2024 · n=2,190 · AI sensitivity (SM alone) vs radiologist sensitivity (DBT+SM)

Ultimate Verdict: Do not just build another 2D scanner AI. Taiwan needs 3D-DBT + ultrasound AI, targeting the 50–60% of eligible women who remain unscreened.

STRATEGIC DIRECTION

AI Autonomy Levels in Mammography Reading (L0 – L5)

Adapted from SAE J3016 · Mapping AI automation to breast cancer screening workflows

Mammography AI Framework

RADIOLOGIST-led reading (L0–L2)

L0 — No AI Assistance
Human only

- Radiologist reads every image alone
- No AI system — no CAD present
- DBT + film at workstation, unaided

e.g. Traditional reading; pre-CAD era

L1 — Passive AI Alert
Alert only

- Radiologist reads; may check AI marks
- Highlights regions of interest (ROI)
- Radiologist decides to look at marks or not

e.g. R2 Image Checker / iCAD 2D CAD

L2 — AI Decision Support
Assist always

- Must review AI score before finalising
- Score + lesion heatmap shown for all cases
- Radiologist adjusts BI-RADS after AI result

e.g. Lunit Insight MMG shown post-read

AI-led reading (L3–L5)

L3 — AI-First Triage
Conditional

- Reviews AI-flagged suspicious cases only
- Triages worklist; fast-tracks normal cases
- Low-score cases skipped; high-score reviewed

e.g. PRAIM safety net — AI sorts worklist

L4 — AI Autonomous Read
Supervised auto

- Reviews recalled (BI-RADS 0/3/4/5) only
- Issues BI-RADS 1/2 reports without radiologist
- Radiologist handles flagged recalls only

e.g. AI standalone reader (shortage areas)

L5 — Fully Autonomous AI
Full auto

- No radiologist required
- AI reads, reports, and manages all recalls
- Full screening pipeline — autonomous

e.g. Experimental / not yet approved (2024)

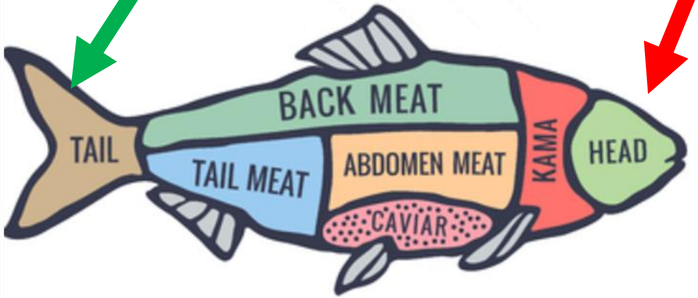
← Radiologist reads all cases · Increasing AI automation · AI reads all cases →

Adapted from SAE J3016 · BI-RADS = ACR 5th ed. · DBT = Digital Breast Tomosynthesis · CAD = Computer-Aided Detection · L5 = experimental only

Triage in Breast Cancer Screening(<2%, >90%)

Triage for less reading time

Safety net as an alarm



Thanks for Your **Attention**

PRAIM · Nat Med 2025

MASAI · Lancet Oncol 2023

AI-STREAM 2024

Larsen · Eur Radiol 2022

Q&A · 5 min

Limitations · What AI Cannot Do — Critical Domains

LIMITATIONS

~1 min

Technical

- Performance varies by vendor, equipment, population
- Training data from Western populations → needs Asian validation
- Performance drift — ongoing post-market validation required

Clinical

- No prospective breast cancer mortality data yet
- Detection benefit ≠ survival benefit until RCT mortality data
- AI optimised for detection, not management decisions

Regulatory

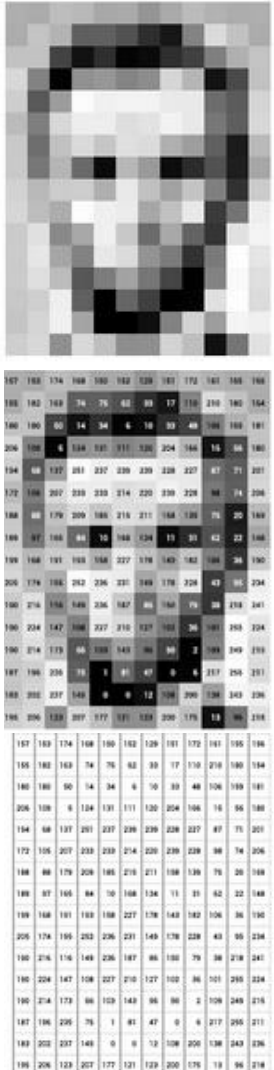
- TFDA AI/ML framework for medical devices still evolving
- FDA/EU AI Act: high-risk classification, mandatory surveillance
- Regulatory pathway must be planned before procurement

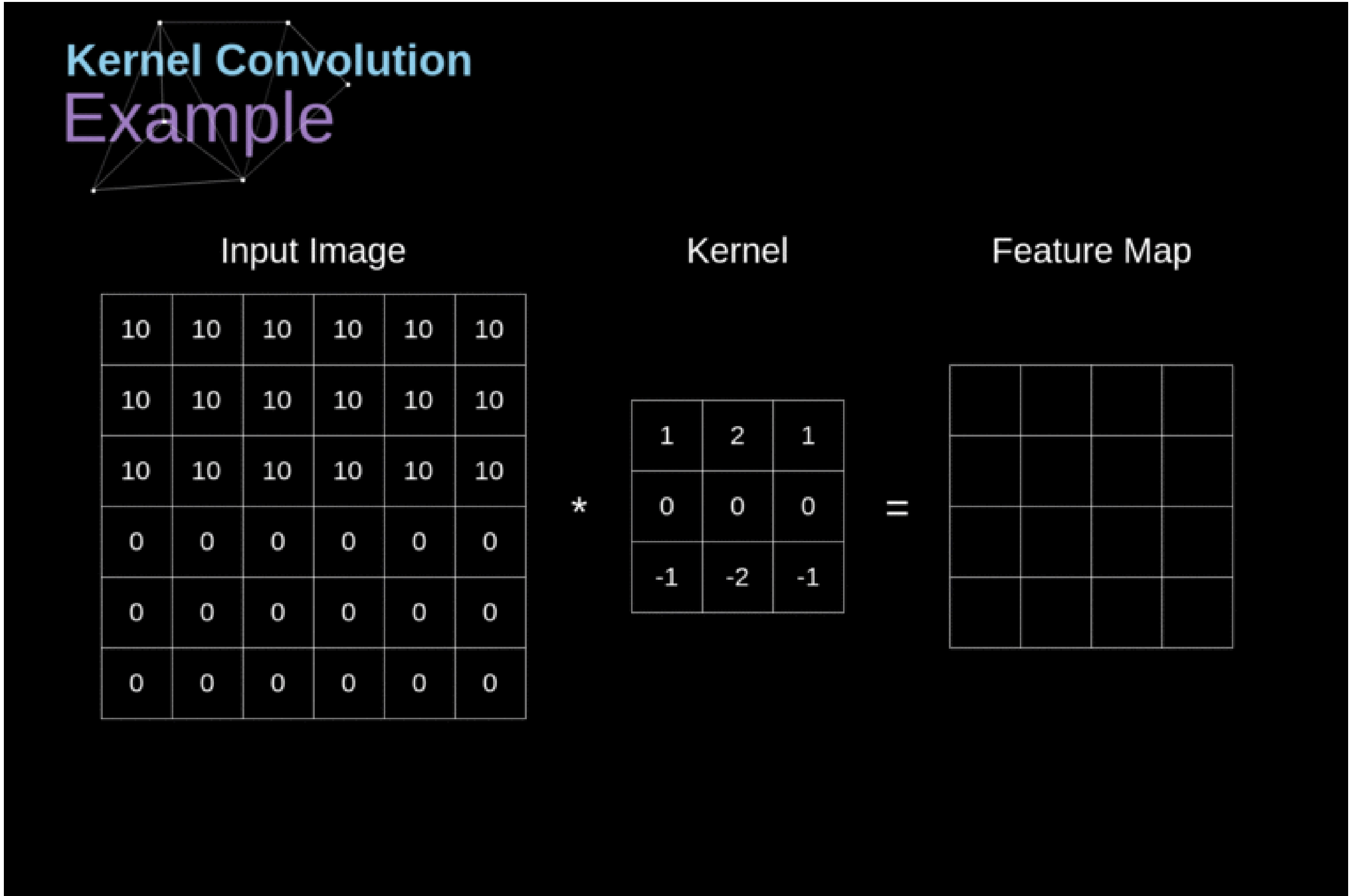
Equity

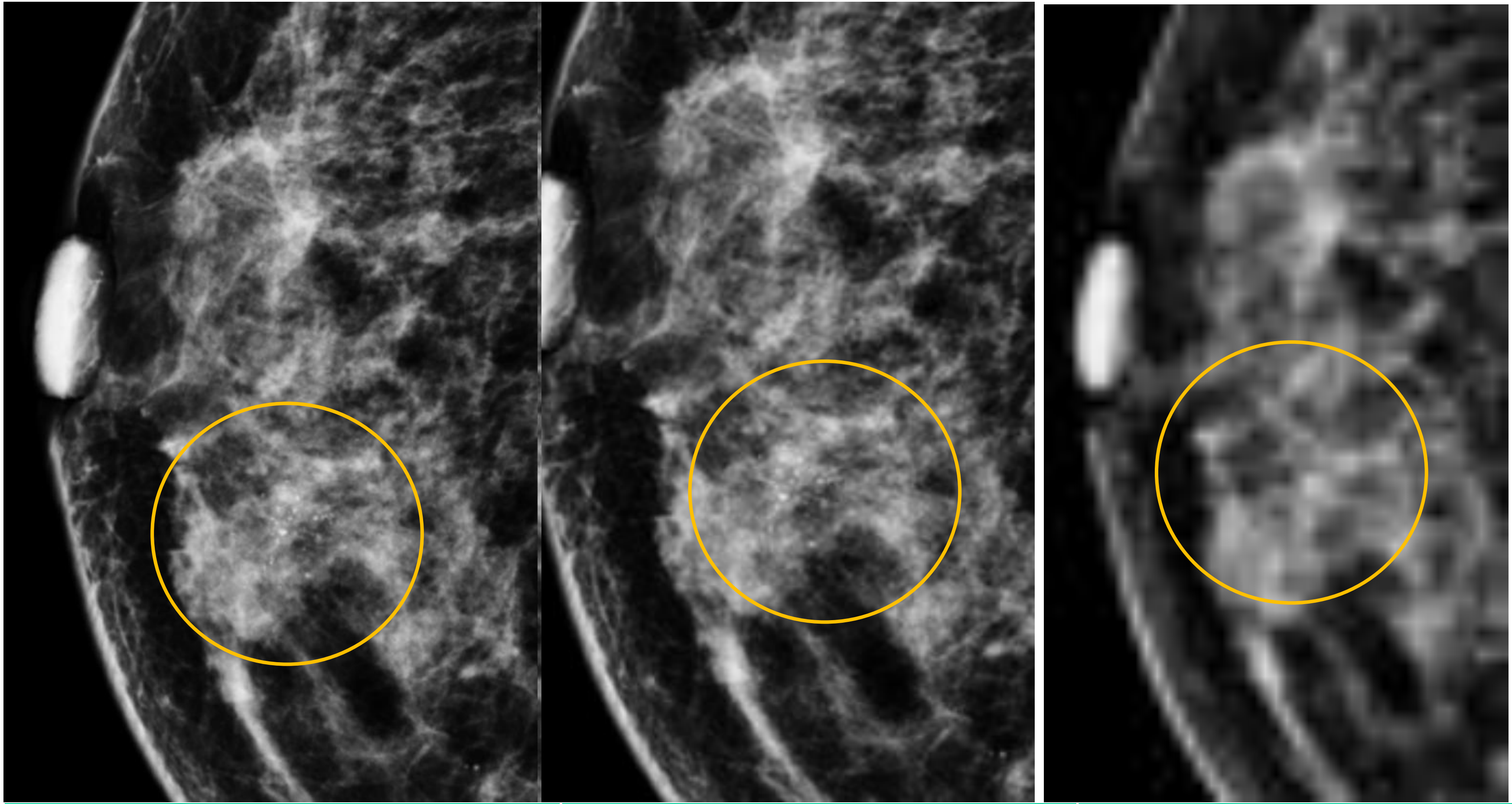
- Limited Asian/Taiwanese data in AI training sets
- ⚠ External validation on Taiwan cohort required before deployment

KEY ANSWER AI requires radiologist oversight, local validation, regulatory approval, and equity checks — it is advisory, not autonomous. Never deploy without Taiwan-specific validation.

Convolutional Neural Network, CNN







1184*1681 | 600*852 | 100*140

Lecture Summary · 5 Take-Home Messages

1. AI NOW because of workforce gaps, dense breast masking, Taiwan single-read risk, and unacceptable interval cancer rates.

2. Prospective evidence is strong: 13–29% detection gain, 44% workload reduction, zero recall rate increase across 3 major trials.

3. AI addresses ~50% of interval cancers — only perception errors and minimal-sign cases. True occult cancers still need MRI/ultrasound.

4. Dense breast AI benefit is clinically plausible (+18.7% PRAIM) but not yet statistically significant. AI risk models (AUC 0.76–0.80) far outperform clinical models (0.61–0.63).

5. Taiwan: start with retrospective validation, then PRAIM safety-net model. Validate on local data before any deployment.

Mammography AI in Asia: A Strategy Guide

For medical directors, hospital leaders and professors · 2025

AI Strategy
Briefing · 2025

Why AI Is Needed

Dense breast tissue hides cancer

More than 80% of Asian women under 50 have thick, dense breasts. Standard 2D X-ray misses up to 30% of tumors in dense breasts. AI finds what the human eye can easily miss. [1,2]

Too few radiologists

The number of breast scans is rising fast but the number of specialist doctors is not keeping up. AI raises junior doctors to near-specialist level (accuracy 0.88 vs 0.90) and cuts reading time. [3]

Breast cancer is increasing

Rates are rising across Asia — Japan (74), Singapore (73), Korea (62) per 100,000 women. Asia has fewer deaths per case than other regions, but still misses many cancers at an early, treatable stage. [4,5]

What Makes This Hard

Every country has different rules

Taiwan, China, Japan and Korea each require separate approvals for AI tools. There is no shared Asian pathway. This makes launching a product slow and expensive across the region. [6]

Most women are not screened

Across Asia, only 1 in 5 women gets a breast scan (18% pooled rate). Japan 47%, Korea 57%, but much of Southeast Asia is below 30%. AI cannot help women who never get scanned. [7]

Cost is a major block

Almost half of radiologists worldwide say budget is the biggest barrier to using AI (49.5%). Many Asian hospitals lack the computer systems and the government payment plans to make it work. [8]

What Leaders Should Do

Use 3D scans, not old 2D scans

AI trained on 3D breast scans scored 0.93 accuracy in Asian women with dense breasts — higher than the average specialist doctor (0.90). Invest in 3D scan AI, not outdated 2D-only tools. [3]

Roll out in stages by country

Japan, Korea, Singapore: add AI alongside doctors in national programs. Taiwan, Malaysia: use AI to sort the reading list. Indonesia, Thailand, Vietnam: start with mobile scan units plus AI. [4,6]

Push for faster approvals

Korea cut AI approval time from 490 days down to 80–140 days. Other Asian countries should follow this model. In a trial, AI also cut the time from scan to biopsy result by 30%. [9,10]

18% of Asian women get breast scans [7]

0.93 AI accuracy in dense breasts [3]

−30% Less time from scan to biopsy [10]

4,800+ Hospitals using Lunit AI (Korea) [11]

Key message: Asian women have denser breasts and fewer specialist doctors than the West — exactly where AI helps most. Use 3D scan AI, roll out in stages, and speed up approvals.

[1] KJR 2025 **[2,3]** Lim WK et al., JMIR Form Res 2025 **[4]** Bray F et al., CA Cancer J Clin 2024 **[5]** Chang JM et al., KJR 2025 **[6]** Taiwan NHI Data Mgmt Act 2025 **[7]** Bhandari AK et al., APJCP 2024 **[8]** ESR Survey 2024 **[9]** Korea MFDS 2024 **[10]** PMC11953103 RCT 2025 **[11]** Lunit IR 2024–26

Lecture Roadmap · 15 Minutes · 5 Core Issues

15 min

Q&A

5 Key Questions

Most asked issues

5 min

All figures from peer-reviewed prospective trials: PRAIM (Nat Med 2025) · MASAI (Lancet Oncol 2023) · AI-STREAM (2024) · Larsen (Eur Radiol 2022)

Limitations · What AI Cannot Do — Critical Domains

LIMITATIONS

~1 min

Technical

- Performance varies by vendor, equipment, population
- Training data from Western populations → needs Asian validation
- Performance drift — ongoing post-market validation required

Clinical

- No prospective breast cancer mortality data yet
- Detection benefit ≠ survival benefit until RCT mortality data
- AI optimised for detection, not management decisions

Regulatory

- TFDA AI/ML framework for medical devices still evolving
- FDA/EU AI Act: high-risk classification, mandatory surveillance
- Regulatory pathway must be planned before procurement

Equity

- Limited Asian/Taiwanese data in AI training sets
- ⚠ External validation on Taiwan cohort required before deployment

KEY ANSWER AI requires radiologist oversight, local validation, regulatory approval, and equity checks — it is advisory, not autonomous. Never deploy without Taiwan-specific validation.

Q&A Q1 of 5 · Does AI increase false-positive recall rates?

SHORT ANSWER: No — consistently across all 3 major trials.

PRAIM (n=463,094)	Recall AI 37.4 vs control 38.3/1,000 — no significant difference (Nat Med 2025)
MASAI RCT (n=106,000)	Recall 2.7% vs 2.8% — non-inferior, detection +29% (Lancet Oncol 2023)
AI-STREAM (n=24,500)	No recall change in Asian single-read cohort — most relevant to Taiwan
Why no harm	AI is advisory only — radiologist retains all recall decisions; safety net activates AFTER human reads normal

BOTTOM LINE Three prospective trials across Germany, Sweden, and Korea consistently show AI improves detection WITHOUT increasing false-positive recall. This addresses the most common clinician objection.

 Can AI replace a second reader in double-reading programmes?

SHORT ANSWER: Yes, with non-inferior performance — MASAI RCT proved this.

MASAI RCT design	AI replaced one radiologist in half the cohort (randomised); radiologist still made final recall decision
Detection outcome	+29% cancer detection vs standard double reading, 44% workload reduction
Non-inferiority	Lancet Oncology 2023: formal non-inferiority endpoint met; full MASAI results Lancet 2026
Taiwan relevance	Taiwan already uses single reading — AI adds a second check, not removes one. Lower implementation bar.

BOTTOM LINE MASAI is the strongest evidence: a randomised trial showing AI-assisted reading is non-inferior to double reading with substantial workload reduction. For single-read systems like Taiwan, AI adds a net safety gain.

 · **Does AI really help in dense breasts — or is this overstated?**

SHORT ANSWER: Clinically plausible but statistical evidence is not yet definitive.

PRAIM Table 4	Dense breast BCDR +18.7% — BUT 95% CI spans −0.5% to +41.4% (crosses zero = not significant)
Non-dense breast	Same trial: non-dense BCDR +16.6%, 95% CI 1.9–33.4% — statistically significant
AI risk models	AUC 0.76–0.80 for 5-year risk prediction vs 0.61–0.63 for clinical models — this IS significant
Asian women	Higher baseline density — dedicated Asian cohort validation is essential before conclusions

BOTTOM LINE AI detection benefit in dense breasts is real but not yet statistically proven in large prospective data. The AI risk stratification benefit IS robust. More dense-breast-specific Asian data needed before clinical claims.

 · **How much can AI realistically reduce interval cancer rates in Taiwan?**

SHORT ANSWER: Up to 47% is theoretically possible — realistically ~15–20% is the evidence-based target.

Larsen 2022 Scenario 2	IC rate 1.7 → 0.9 per 1,000 = 47% relative reduction — but retrospective simulation, not prospective
Achievable fraction	AI addresses ~50% of ICs (perception + minimal-sign); ~50% are occult — AI cannot help these
PRAIM safety net	Only ~1.5% of exams flagged — focused, high-precision intervention
Taiwan target	≥15% IC reduction is used as the primary endpoint in proposed Phase 2 validation study

BOTTOM LINE 47% is a ceiling from simulation data, not a real-world guarantee. A ≥15% IC reduction is the evidence-based realistic target for a Taiwan prospective study. Occult cancers require supplemental imaging, not AI.

genius AI™

DETECTION

Genius AI DBT Triage — Screening Mammography (N=1,780)

July 2021 – June 2022 · AI threshold ≥2% = recall · Total cancers: 18

AI threshold
≥2% = recall

Women underwent mammography
July 2021 – June 2022 (n=2,679)

Excluded (n=899)
Symptomatic breast disease
Diagnostic mammograms for abnormal screening

50%

50%

Screening Mammography
(n = 1,780)

50%

50%

No-Recall — Genius AI < 2%
n = 892 (50%)

Recall — Genius AI ≥ 2%
n = 888 (50% recall rate)

98%

2%

91%

9%

Missed by Both

AI(−) Radiologist(−)

No-Recall · BI-RADS 1–2

n = 876 (98%)

Benign: 5

Cancer: 3

Missed by both

3

Caught by Radiologist

AI(−) Radiologist(+)

No-Recall · BI-RADS 0,3,4,5

n = 16 (2%)

Benign: 4

Cancer: 0

Radiologist found what AI missed

0

vs
AI
2%

Caught by AI

AI(+) Radiologist(−)

Recall · BI-RADS 1–2

n = 808 (91%)

Benign: 18

Cancer: 3

AI found what radiologist missed

3

Caught by Both

AI(+) Radiologist(+)

Recall · BI-RADS 0,3,4,5

n = 80 (9%)

Benign: 22

Cancer: 12 ★

Highest yield (both agree)

12

Total cancers found: 3 Missed (both) + 0 Radiologist only + 3 AI only + 12 Both agree = 18

 What should Taiwan's Health Promotion Administration do FIRST?

SHORT ANSWER: Retrospective validation on Taiwan data before any procurement decision.

Phase 1 priority	≥50,000 exam retrospective study from existing HPA national programme data (6–12 months)
Why validate first	AI trained on Western data — Asian breast density patterns and cancer subtypes differ; bias risk is real
Which AI platform	Transpara, Vara, or Lunit all suitable — must compare on Taiwan data, not rely on Western validation
Phase 2 target	PRAIM safety-net model: radiologist reads first, AI flags ~1.5% after — directly applicable to single-read Taiwan

BOTTOM LINE Do NOT procure AI based on Western trial data alone. Taiwan must validate locally first. Phase 1 retrospective study is low-risk, high-value, and can be started immediately with existing national programme data.

Thank You

"AI will not replace radiologists.

Radiologists using AI will replace those who don't."

Eisemann et al. Nature Medicine 2025 (PRAIM, n=463,094) · Lång et al. Lancet Oncology 2023 (MASAI RCT)

Larsen et al. European Radiology 2022 (IC reduction modelling) · Lång et al. European Radiology 2021 (AI score 10)

McKinney et al. Nature 2020 (AI vs radiologist) · WISDOM trial · MyPeBS trial (risk stratification)

Chen-Pin Chou, MD · Director of Radiology, Kaohsiung Veterans General Hospital

Can AI Help Radiologists Find More Breast Cancers?

PRAIM Study · Nature Medicine 2025 · 461,818 women · Germany

The Problem

- Radiologists read hundreds of mammograms per week
- Double reading required
- Radiologist shortage
- Some cancers still missed

The AI Solution

Normal Triage
56% of exams = quickly confirmed normal

Safety Net
Alerts radiologist if AI thinks a case is suspicious

The Results

+17.6%

more cancers detected

6.7 vs 5.7 per 1,000 women

Recall rate: unchanged ✓
PPV biopsy: 64.5% vs 59.2% ✓

Watch Out!

DCIS ↑ +67.6%
(overdiagnosis risk?)

Observational study
(not randomized)

No survival data yet
2–3 yr follow-up needed

Bottom Line: **AI-supported double reading detects 18% more cancers, keeps recall rates stable, and cuts reading time — ready for routine screening**

463094	12	119	2021–2023	Germany
Women Screened	Sites	Radiologists	Study Period	Country

Eisemann, Bunk et al. · Nat Med 2025;31:917–924 · doi:10.1038/s41591-024-03408-6