Clinically Meaningful AI Detection of Interval Cancers

Background and context

Artificial intelligence (AI) is becoming a more regular part of breast cancer screening. Yet as adoption grows, an important question remains: is there robust evidence that these tools improve long-term outcomes or reduce mortality?

Large randomized trials to prove a mortality benefit are costly and take decades. In the meantime, interval cancers—symptomatic cancers that appear after a negative mammogram—serve as a valuable surrogate. Often more aggressive and harder to detect, earlier detection by AI may signal that AI can improve long-term outcomes.

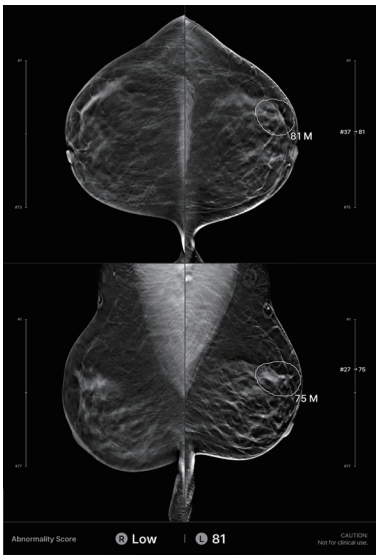
The Lunit INSIGHT Breast Suite enhances 2D/3D mammography reading to improve accuracy, reduce recalls, and support radiologist confidence. Here's a look at the clinical evidence behind it.

Latest clinical research with Lunit INSIGHT DBT

A recent study in Radiology by Bahl et al[1] evaluated Lunit INSIGHT DBT for its ability to detect and correctly localize interval cancers on screening exams. Unlike previous studies reporting only exam-level AI scores, this work validated performance at the lesion level, confirming whether the AI flagged the exact site of cancer.

This distinction matters. Exam-level scoring can make it seem as though AI "caught" a cancer even when highlighting a benign area, overstating its value. By focusing on lesion-specific accuracy, Bahl et al addressed this limitation and offered clearer insight into the practical meanings of AI findings.

In an accompanying commentary[2], Dr. Lee and Dr. Milch called this rigor a critical strength, noting that lesion-level validation is essential for understanding the true clinical relevance of AI in screening.



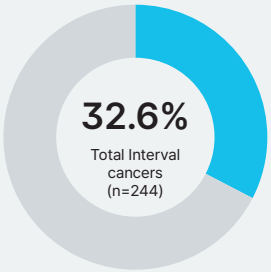
▶ Example of an interval cancer retrospectively detected by the artificial intelligence (AI) algorithm. A 41-year-old woman presented for a screening digital breast tomosynthesis examination; its findings were interpreted as negative. Ten months later, the patient presented with a lump in the left breast and was subsequently diagnosed with grade 3 invasive ductal carcinoma. At retrospective evaluation of the initial screening mammogram, the AI algorithm marked a suspicious lesion (white outline) in the left breast, with high scores of 81 on the craniocaudal view and 75 on the mediolateral oblique view. This area of architectural distortion corresponds to the cancer that was subsequently diagnosed.

Lunit AI correctly localized nearly one-third of interval cancers and has the potential to decrease the interval cancer rate.

Lunit AI could reduce interval cancers by 32.6%.

Interval cancers detected by Lunit AI (73/244)

- Interval invasive cancers detected by AI were significantly larger at the time of surgery (mean size, 37 vs 22 mm; P < .001) and more frequently lymph node–positive (41.3% vs 22.8%; P = .01)
- These results suggest that AI preferentially detects more clinically advanced tumors that, if caught earlier, may have remained more localized and potentially lower stage.



"Our research found Lunit AI could reduce interval breast cancers by more than one-third. Reducing the interval cancer rate of screening tomosynthesis will ultimately lead to improved patient outcomes."

Manisha Bahl, MD, MPH, FSBI
Massachusetts General Hospital

Clinically meaningful research with Lunit INSIGHT MMG

Multiple large studies have tested how AI scores can be used in screening—either to select exams for additional imaging or to compare AI performance with radiologists.
Key methods and results:

Dembrower et al (Lancet Digital Health)[3]

This study tested whether AI scores could be used to select women with negative mammograms for more sensitive screening, promoting early detection of cancers that would otherwise be diagnosed as interval cancers or as next-round screen-detected cancers.

- Women with the **highest 5% of AI scores** were directed into an enhanced assessment stream (MRI).
- This strategy led to a **potential additional detection of 27% of interval cancers.**

Kühl et al (European Radiology)[4]

This study compared AI to first-reading radiologists using two AI cut-offs matched to reader performance – AI(sens) for **mean sensitivity** and AI(spec) for **mean specificity.** In both scenarios, AI outperformed radiologists.

- **126 interval cancers detected** with AI(sens).
- **117 interval cancers detected** with AI(spec).
- **39 interval cancers detected** by first readers (p<0.0001).

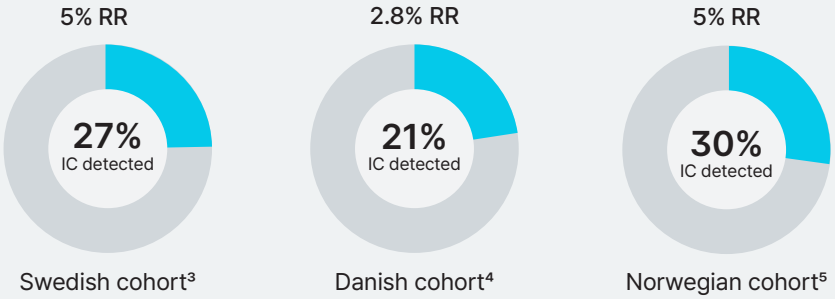
Larsen et al (Radiology AI)[5]

This study explored how different AI thresholds affect the trade-off between false positives and false negatives. To mirror their double-reading program, they set the **top 5% of AI scores** as positive and the remaining **95% as negative.** (Please note: other thresholds may have better fit in other screening programs and need to be considered).

- At this threshold, AI classified **30% of interval cancers** as positive (333 of 1110 cases).
- Authors emphasized that optimal thresholds depend on each screening program and must balance cancer detection with acceptable recall rates.

Lunit INSIGHT MMG demonstrated effectiveness in detecting interval cancers in 10+ peer-reviewed studies. Below is a summary of a few studies.

▶ Standalone IC detection by Lunit, using an operating point chosen to **maintain a low Recall Rate (RR), inferior to 5%.**



Another study by Nanaa et al (Radiology)[6]* focused on AI's ability to both detect and correctly localize interval cancers at a high-specificity setting. At 96% specificity, AI identified 23.5% of interval cancers, with 76.9% correctly localized.

Breakdown of AI performance by screening category (correctly localized cancers by AI):

- 12.6% of true negative interval cancers (initially read as normal/benign).
- 46% of minimal sign interval cancers (uncertain findings at screening).
- 50% of false-negative interval cancers (suspicious findings present but missed).

Key Takeaways

- Interval cancers are aggressive, rapidly growing tumors with poor prognosis.
- Interval cancer rate can serve as a surrogate marker for long-term outcomes; reducing it may lower morbidity and mortality.
- AI tends to catch more advanced interval cancers that, if found earlier, may be lower stage and more localized.
- AI scores can be used to select patients for an enhanced assessment stream (e.g. MRI) for earlier detection.
- AI can correctly detect and localize up to one-third of interval cancers earlier, during screening.
- AI score cut-off/operating point depends on factors such as desired recall rates.

Disclaimer: For full disclaimer, scan the QR code and refer to sections 1 and 2.

*Study was performed with MMG v1.1.2.0 approved under MDD and does not pertain to the current EU MDR-certified version of Lunit INSIGHT MMG.

References

1. Bahl M et al. AI to Reduce the Interval Cancer Rate of Screening Digital Breast Tomosynthesis. Radiology. 2025 Jul;316(1):e241050
2. Lee C et al. Clinically Meaningful AI Detection of Interval Breast Cancer at Digital Breast Tomosynthesis Screening. Radiology 2025; 316(1):e251860
3. Dembrower Ket al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. The Lancet Digital Health, Volume 2, Issue 9, e468 – e474
4. Kühl J et al. Population-wide evaluation of artificial intelligence and radiologist assessment of screening mammograms. Eur Radiol. 2024 Jun;34(6):3935-3946
5. Larsen M et al. Performance of an Artificial Intelligence System for Breast Cancer Detection on Screening Mammograms from BreastScreen Norway. Radiol Artif Intell. 2024 May;6(3):e230375.
6. Nanaa M et al. Accuracy of an Artificial Intelligence System for Interval Breast Cancer Detection at Screening Mammography. Radiology. 2024 Aug;312(2):e232303