UTILIZATION OF AN AI DIAGNOSTIC SOFTWARE IN A MAMMOGRAPHY SCREENING PROGRAM SHOWS POTENTIAL FOR HIGHER SCREENING EFFICIENCY AND EFFECTIVENESS

(1) University of St.Gallen, School of Medicine, Chair of Health Economics, Policy and Management, St.Gallen, Switzerland, (2) Krebsliga Ostschweiz, St.Gallen, Switzerland

Marcel Blum (1,2)*, Alexander Geissler (1), David Kuklinski (1), Justus Vogel (1), Jonas Subelack (1), Alena Eichenberger (2), Rudolf Morant (2)

* Marcel.Blum@krebsliga-ostschweiz.ch

AIM & OBJECTIVES

- Artificial intelligence (AI) has increasingly been tested in Mammography Screening Programs (MSP) since the mid-2010. [1-5]
  - May increase effectiveness, i.e., sensitivity and specificity
  - May increase efficiency, i.e., reduce radiologists' workload
- The optimal utilization and thresholds, i.e., when a mammography should be further discussed and/or investigated are not determined or backed up yet by scientific evidence, guidelines or recommendations.

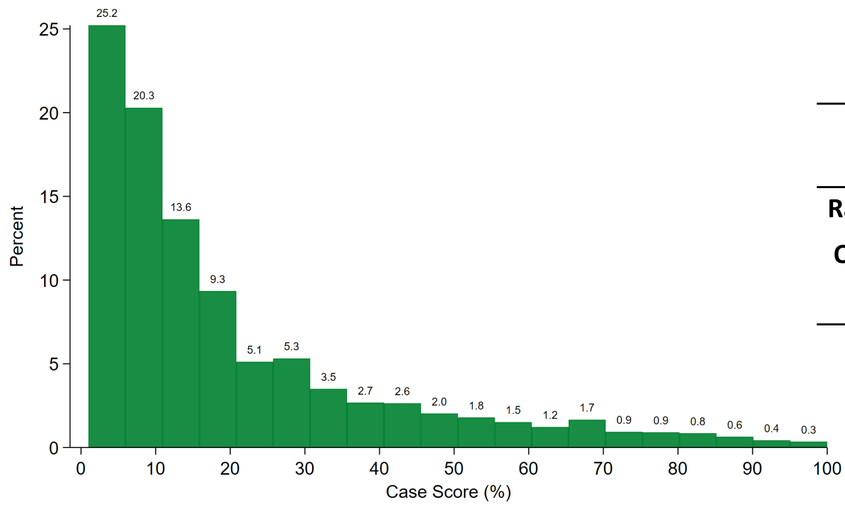
PATIENTS & METHODS

- We used data from 7,546 mammographies done in the MSP "donna" in the cantons of Bern and Solothurn in the ongoing pilot phase between May 2023 and July 2024, where an AI score by ProfoundAI® (pAI) was available.
- pAI reviews each mammography with a case score, which reflects its certainty that a mammography contains a cancer case.
- In this pilot phase, all mammographies with a case score of 55% or higher were discussed in a consensus conference (CC), regardless of the reviews by two radiologists.
- We used explorative, descriptive statistics to show how mammographies were reviewed by the two radiologists and pAI.
- We also simulated two scenarios how AI could be implemented within the MSP. In the first scenario, AI would substitute one of the two radiologists, whereas in the second scenario, AI would preselect the mammographies for double reading.

RESULTS

Pilot Phase Evaluation

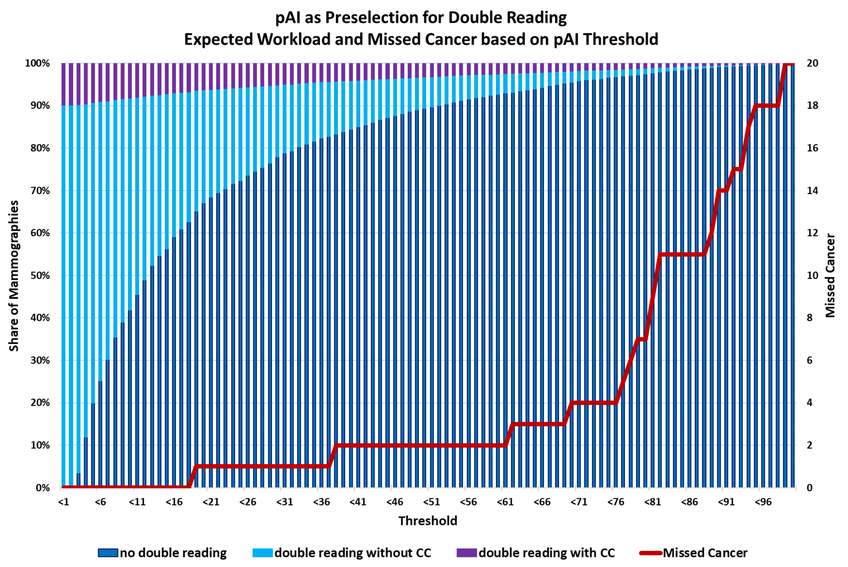
- Half of all mammographies received a case score of less than 12% and the mean case score was 19.9%. The share of mammographies that were assigned with a case score of 55% or higher was 8.45%.
- Only 17.8% of all mammographies discussed in a CC were suspiciously reviewed by the radiologists and had a pAI case score ≥ 55%, but these had the highest investigation rate as well as the highest sensitivity.
- Of the 22 detected cancer cases, 19 were marked by the radiologists and pAI, two by radiologists only, and one by pAI only.



Responsible for the Concensus-Conference (CC)			% Investigation per CC	Cancer case per Investigation
Radiologists & pAI	17.8%	→	69.0%	29.7%
Only Radiologists	44.7%	→	45.9%	1.3%
Only pAI	37.5%	→	7.2%	7.7%

Utilization Simulation

- Substituting one of the two radiologists would have reduced the overall workload by around 43%, whereas using pAI as preselection would have reduced the workload by almost 60%.
- In both simulations no mammography with a confirmed cancer case would have been missed.



CONCLUSIONS & DISCUSSION

- The results indicate that the utilization of pAI as a third reader led to a higher workload, but also to a higher sensitivity, i.e., higher breast cancer detection rate.
- A higher breast density was not associated with a systematically higher case score, emphasizing the AI's ability to detect abnormalities in more dense breasts. [6]
- Results indicate that pAI could be used as a second reader or as a preselection method without loss of effectiveness but with an increase of efficiency, i.e., a workload reduction of radiologists. [1,2]
- All findings will be validated with a larger retrospective sample of more than 27,000 mammographies in the MSP "donna" between 2022 and 2023.

REFERENCES

1. Dembrower, K., Wåhlin, E., Liu, Y., Salim, M., Smith, K., Lindholm, P., Eklund, M., Strand, F. (2020). 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, 2(9), e468-e474
2. Marinovich, M. L., Wylie, E., Lotter, W., Lund, H., Waddell, A., Madeley, C., Pereira, G., Houssami, N. (2023). Artificial intelligence (AI) for breast cancer screening: BreastScreen population-based cohort study of cancer detection. EBioMedicine, 90.
3. Lång, K., Dustler, M., Dahlblom, V., Åkesson, A., Andersson, I., & Zackrisson, S. (2021). Identifying normal mammograms in a large screening population using artificial intelligence. European radiology, 31, 1687-1692.
4. Dembrower, K., Crippa, A., Colón, E., Eklund, M., & Strand, F. (2023). Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. The Lancet Digital Health, 5(10), e703-e711.
5. Pacilè, S., Lopez, J., Chone, P., Bertinotti, T., Grouin, J. M., & Fillard, P. (2020). Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. Radiology: Artificial Intelligence, 2(6), e190208.
6. Bergan, M. B., Larsen, M., Moshina, N., Bartsch, H., Koch, H. W., Aase, H. S., ... & Hofvind, S. (2024). AI performance by mammographic density in a retrospective cohort study of 99,489 participants in BreastScreen Norway. European Radiology, 1-11.