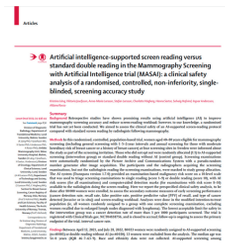
Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study

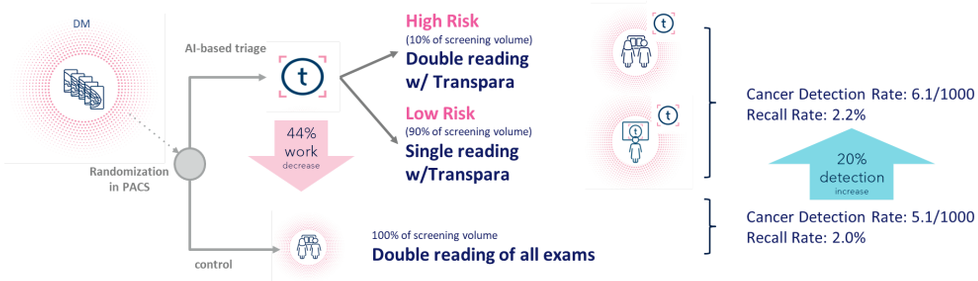
K Lång, V Josefsson, A-M Larsson, S Larsson, C Högberg, H Sartor, S Hofvind, I Andersson, A Rosso



Study design

In this randomised, controlled, population-based trial, women aged 40–80 years eligible for mammography screening (including general screening with 1·5–2-year intervals and annual screening for those with moderate hereditary risk of breast cancer or a history of breast cancer) at four screening sites in Sweden were informed about the study as part of the screening invitation. Those who did not opt out were randomly allocated (1:1) to AI-supported screening (intervention group) or standard double reading without AI (control group). Screening examinations were automatically randomised by the Picture Archive and Communications System with a pseudo-random number generator after image acquisition. The participants and the radiographers acquiring the screening examinations, but not the radiologists reading the screening examinations, were masked to study group allocation. The AI system (Transpara version 1.7.0) provided an examination-based malignancy risk score on a 10-level scale that was used to triage screening examinations to single reading (score 1–9) or double reading (score 10), with AI risk scores (for all examinations) and computer-aided detection marks (for examinations with risk score 8–10) available to the radiologists doing the screen reading. Here we report the prespecified clinical safety analysis, to be done after 80 000 women were enrolled, to assess the secondary outcome measures of early screening performance (cancer detection rate, recall rate, false positive rate, positive predictive value [PPV] of recall, and type of cancer detected [invasive or in situ]) and screen-reading workload. Analyses were done in the modified intention-to-treat population (ie, all women randomly assigned to a group with one complete screening examination, excluding women recalled due to enlarged lymph nodes diagnosed with lymphoma). The lowest acceptable limit for safety in the intervention group was a cancer detection rate of more than 3 per 1000 participants screened. The trial is registered with ClinicalTrials.gov, NCT04838756, and is closed to accrual; follow-up is ongoing to assess the primary endpoint of the trial, interval cancer rate

Results



PLUS: no change in arbitration volume.

Conclusion

AI-supported mammography screening resulted in a similar cancer detection rate compared with standard double reading, with a substantially lower screen-reading workload, indicating that the use of AI in mammography screening is safe. The trial was thus not halted and the primary endpoint of interval cancer rate will be assessed in 100 000 enrolled participants after 2-years of follow up.

SPM-SMR-001-126 Rev A