
AI Integration in Breast Screening Across Diverse Groups



Artificial Intelligence (AI) has shown promise in improving the efficiency and accuracy of breast cancer screening. While existing studies highlight population-level benefits, evidence regarding its performance across specific patient subgroups remains limited. The ARIES study—a large-scale, multicentre retrospective analysis of over 306,000 screening cases in the UK—addresses this gap. It assesses AI-integrated workflows against standard double reading across geographic regions, age groups, breast density categories and ethnic backgrounds. The findings inform the safety, effectiveness and operational implications of deploying AI in real-world screening contexts.

Subgroup Safety and Clinical Effectiveness

The study evaluated two AI-assisted workflows: supporting independent reader (sIR) and double reader triage (DRT). Both approaches use AI as an independent second reader alongside a human counterpart. Clinical outcomes such as cancer detection rate (CDR) and positive predictive value (PPV) were assessed across centres and subgroups. The AI workflows passed non-inferiority tests in all clinical metrics and subgroups, demonstrating their safety relative to standard human double reading.

Must Read: [Explainable AI in Mammographic Breast Cancer Screening](#)

Although the absolute differences in CDR were small, there was a consistent trend of slightly reduced detection in AI workflows, ranging from -0.24 to -0.08 per 1,000 cases. However, these reductions were within acceptable bounds given the study's retrospective nature and the AI workflows' design for operational, rather than detection, enhancement. Crucially, AI performance mirrored the subgroup disparities observed in human readings, such as higher detection rates in older women and those with denser breast tissue, without amplifying those disparities. Moreover, PPV improved across all demographics, suggesting fewer false positives and reduced patient anxiety.

Operational Efficiency and Workload Reduction

A significant advantage of integrating AI in screening workflows was the reduction in operational workload. Both sIR and DRT configurations demonstrated substantial savings: workload was reduced by 42.5% and 39.6% respectively. These reductions were consistent across all participating centres, indicating the generalisability of the benefit. Furthermore, AI integration led to a decrease in recall rates, enhancing downstream efficiency by reducing unnecessary assessments.

Among the two AI workflows, sIR produced slightly greater workload savings, while DRT achieved better recall reduction and improved PPV. These variations reflect the intended design of each workflow, with sIR optimised for operational support and DRT for balancing efficiency with clinical outcomes. The combined effect of these reductions contributes to the long-term sustainability of breast screening programmes, especially amidst increasing screening volumes and workforce pressures.

Potential for Enhanced Cancer Detection

While retrospective studies cannot directly measure improvements in cancer detection, the ARIES study explored AI's standalone performance in identifying interval cancers (ICs)—cases missed during initial screenings but diagnosed later. The AI system flagged 41.2% of ICs, with the highest flag rate observed in patients with dense breast tissue, a known challenge in mammography. This result indicates the potential of AI to enhance early detection when used as an additional reader (XR), complementing the sIR or DRT configurations.

Extrapolating from previous prospective evaluations, the study estimated that integrating AI as an XR could increase the overall CDR by 1.2 per 1,000 cases. Such an increase would offset the marginal CDR reductions observed in the operationally focused workflows and suggests a hybrid deployment strategy may yield both safety and performance benefits. This finding underscores the importance of flexible AI roles in screening,

tailored to clinical goals and resource constraints.

The ARIES study provides robust, stratified evidence supporting the integration of AI in breast screening. Across diverse populations and regions, AI workflows achieved non-inferior or superior performance on clinical metrics while delivering meaningful reductions in workload and recall rates. Importantly, the AI system maintained performance equity across subgroups, demonstrating its potential for safe deployment. Additionally, the high IC flag rate offers a compelling case for using AI as an additional reader to enhance cancer detection. Together, these results support the scalable, equitable and sustainable use of AI in national breast screening programmes.

Source: [BMJ Health & Care Informatics](#)

Image Credit: [iStock](#)

Published on : Fri, 20 Jun 2025