INPLASY

INPLASY202630080

doi: 10.37766/inplasy2026.3.0080

Received: 22 March 2026

Published: 22 March 2026

Corresponding author:
Ting Wei Wang

eltonwang1@gmail.com

Author Affiliation:
Taipei Veteran General Hospital.

AI-integrated versus standard mammography screening pathways: a systematic review and network meta-analysis of diagnostic accuracy

Wang, TW; Yang, S; Wang, YC; Chen, YC; Tsai, YF; Tu, YK; Chiou, HJ; Tseng, LM; Lai, YC; Wu, YT.

ADMINISTRATIVE INFORMATION

Support - NR.

Review Stage at time of this submission - Completed but not published.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202630080

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 22 March 2026 and was last updated on 22 March 2026.

INTRODUCTION

Review question / Objective This review aims to compare the diagnostic accuracy of clinically deployed breast cancer screening interpretation pathways using digital mammography and, in extended analyses, digital breast tomosynthesis. Using a Bayesian network meta-analysis of diagnostic test accuracy, we will compare single human reading, double human reading, stand-alone artificial intelligence (AI), AI as an independent reader, AI reader assistance, and AI triage strategies for the detection of breast cancer in screening populations.

Primary question: among women undergoing breast cancer screening, how do different human and AI-integrated interpretation pathways compare in sensitivity and specificity for cancer detection?

Condition being studied Breast cancer detected in screening mammography programmes. The review focuses on asymptomatic screening populations undergoing breast imaging for early cancer detection, primarily using digital mammography and, in broader analyses, digital breast tomosynthesis.

METHODS

Participant or population Women undergoing routine breast cancer screening in real-world or programme-based screening settings. The target population includes asymptomatic screening participants, typically middle-aged and older women eligible for population or opportunistic screening programmes. Studies restricted to symptomatic diagnostic populations will be excluded.

Intervention Artificial intelligence-enabled breast screening interpretation pathways, including:stand-alone AI,AI as an independent reader,AI reader assistance,AI triage,and other clinically defined hybrid human–AI pathways that produce a binary screening decision relevant to programme implementation.

Comparator Human-reader pathways, including:single human reading,double human reading,and other eligible screening pathways contributing to a connected diagnostic network.Comparators may be direct within-study comparators or indirect comparators through the network structure.

Study designs to be included We will include prospective and retrospective diagnostic accuracy studies conducted in breast cancer screening settings, including:head-to-head comparative studies,workflow simulation studies based on real screening cohorts,paired-reader or paired-pathway studies,randomised or non-randomised comparative screening studies,and additional eligible studies that contribute indirect evidence to a connected diagnostic network.

Eligibility criteria Inclusion criteria:Studies of breast cancer screening pathways in asymptomatic screening populations.Studies evaluating at least one eligible human, stand-alone AI, or human–AI workflow.Studies reporting sufficient information to derive or reconstruct pathway-level 2 × 2 data (TP, FP, TN, FN) at the examination or patient level.Studies using digital mammography, and, in extended analyses, digital breast tomosynthesis.Studies with an acceptable reference standard incorporating pathology for positive examinations and follow-up or registry-based ascertainment for negative examinations when available.Exclusion criteria:Symptomatic or diagnostic-only cohorts.Case reports, narrative reviews, editorials, conference abstracts without sufficient data, and protocols.Studies reporting only lesion-, breast-, image-, patch-, or pixel-level outcomes without examination-level or patient-level pathway data.Studies of AI tools not linked to an interpretable screening decision pathway.Duplicate publications of the same dataset when no additional usable data are provided.We will include prospective and retrospective diagnostic accuracy studies conducted in breast cancer screening settings, including:head-to-head comparative studies,workflow simulation studies based on real screening cohorts,paired-reader or paired-pathway studies,randomised or non-randomised comparative screening studies,and additional eligible studies that contribute indirect evidence to a connected diagnostic network.

Information sources Electronic databases: PubMed, Embase, Web of Science.
Other sources: manual checking of reference lists of included studies and relevant reviews.

Main outcome(s) Primary outcomes are:

Sensitivity
Specificity

These will be synthesised at the workflow-node level using Bayesian network meta-analysis of diagnostic test accuracy. We will also translate pooled estimates into absolute expected outcomes per 10,000 screened examinations at an assumed prevalence of 0.8%, including true positives, false negatives, true negatives, and false positives.

Quality assessment / Risk of bias analysis Risk of bias will be assessed independently by two reviewers using QUADAS-2/QUADAS-3-style domain assessment, and QUADAS-C for comparative diagnostic accuracy studies where appropriate. Domains will include patient selection, index test, reference standard, and flow and timing. Vendor involvement, AI-processability exclusions, calibrated or post-hoc thresholds, and incomplete verification of negative examinations will be explicitly recorded. Disagreements will be resolved by discussion or third-reviewer adjudication.

GRADE-DTA will be used to assess certainty of evidence for the main network-level outcomes.

Strategy of data synthesis We will conduct a Bayesian hierarchical bivariate network meta-analysis of diagnostic test accuracy. Screening pathways will be mapped to prespecified nodes. The primary network will include:

single human,
double human,
stand-alone AI,
AI as an independent reader.

A broader secondary network will additionally include:

AI reader assistance,
AI triage.

True-positive counts will be modelled conditional on diseased examinations and true-negative counts conditional on non-diseased examinations. Sensitivity and specificity will be jointly modelled on the logit scale using correlated random effects. Posterior summaries will be reported as posterior means with 95% credible intervals. Heterogeneity will be summarised using tau and tau-squared. We will also generate HSROC displays, clinical absolute-effect tables, and exploratory ranking

summaries. Formal transitivity assessment will be emphasised in the primary network, and consistency will be assessed using node-splitting and loop inconsistency checks where appropriate.

Subgroup analysis Prespecified subgroup or moderator analyses may include:

modality (DM vs DBT),
study design (prospective vs retrospective),
geographic region,
screening programme type,
vendor involvement,
estimated versus directly observed 2 × 2 data,
and other transitivity-related effect modifiers.

These analyses will be exploratory and interpreted cautiously.

Sensitivity analysis Prespecified sensitivity analyses will include:

exclusion of estimated or derived rows,
exclusion of vendor-authored studies,
exclusion of studies at high risk of bias,
restriction to DM-only studies,
restriction to population-screening studies,
restriction to cohorts with interval-cancer or longer follow-up ascertainment where applicable,
exclusion of overlap-flagged cohorts,
and prior-sensitivity analyses in the Bayesian model.

Leave-one-cohort-out analyses may also be conducted.

Country(ies) involved Taiwan.

Keywords breast cancer screening; mammography; digital mammography; digital breast tomosynthesis; artificial intelligence; deep learning; diagnostic test accuracy; network meta-analysis; workflow; radiologist.

Contributions of each author
Author 1 - Ting Wei Wang.
Author 2 - Sheng Yang.
Author 3 - Yun-Chu Wang.
Author 4 - Yu-Chuan Chen.
Author 5 - Yi-Fang Tsai.
Author 6 - Yu-Kang Tu.
Author 7 - Hong-Jen Chiou.
Author 8 - Ling-Ming Tseng.
Author 9 - Yi-Chen Lai.
Author 10 - Yu-Te Wu.