

<https://doi.org/10.29289/259453942025V36S11087>

1087 – Mammography AI-based breast cancer risk prediction versus traditional clinical risk calculators: a systematic review and comparative evidence synthesis

Frederico Nassif Gomes¹, Talita Mariane Franco¹, Julia Cristina Ferreira¹, Maria de Lourdes Nassif¹, Jan Pawel Andrade Pachnicki¹

¹Pontifícia Universidade Católica do Paraná – Curitiba (PR), Brazil.

Introduction: Conventional breast cancer risk calculators using subjective clinical criteria have shown decreasing levels of generalizability across different contexts. Concurrently, risk models that utilize artificial intelligence (AI) for analysis of mammographic screening examinations are on the rise, primarily for identifying changes that lie outside of human vision.

Objective: This study evaluated the performance of AI-based image screening techniques for breast cancer risk quantification compared with that of clinical calculators to determine their role in population screening.

Methods: A systematic review without meta-analysis was conducted under PRISMA 2020 guidelines. A preliminary search in four databases between December 2025 and February 2026 identified 1,130 studies, of which 175 were screened, and 26 were included in the quantitative and qualitative analyses. Three reviewers screened the studies, and issues were resolved by consensus. Articles evaluating mammographic AI models for predicting breast cancer risk *versus* traditional calculators (e.g., Gail, BCSC, or Tyrer-Cuzick) were included. Articles that used other imaging methods or other comparators, as well as non-original studies, were excluded. Studies were mostly published between 2022 and 2024, had retrospective design, and were concentrated in the United States/Europe, with only one study in Brazil/Latin America. The area under the receiver operating characteristic curve (AUC) and C-index were the key reported metrics.

Results: In the reviewed studies, AI models were superior to conventional calculators (AUC 0.713 *vs.* 0.595, 95% confidence interval [CI]), as indicated in short- and medium-term risk prediction settings. The hybrid model was superior to the imaging-only model (AUC 0.745 *vs.* 0.715, 95%CI), while the static model proved inferior to the longitudinal model in risk prediction (AUC 0.705 *vs.* 0.738, 95%CI).

Conclusion: This review points out the relative benefit of the AI models over clinically based calculators, especially hybrid and longitudinal models, possibly in the context of risk-based screening. Nevertheless, methodological heterogeneity and the lack of Latin American studies limit the generalizability of the findings to other ethnic and racial groups, illustrating the need for prospective studies in diverse settings.

Ethics: Public data – Literature review.

Keywords: breast cancer; mammography; artificial intelligence; risk stratification; screening.