The Daffodil Centre

THE UNIVERSITY OF MELBOURNE

Mammographic density measurement methods: how well do they identify population breast screeners according to breast cancer risk?

Velentzis LS[1,2], Hughes S[1], Carle C[1], Luo Q[1], Evans J[3], Mann GB[4,5], Nickson C[1,2]

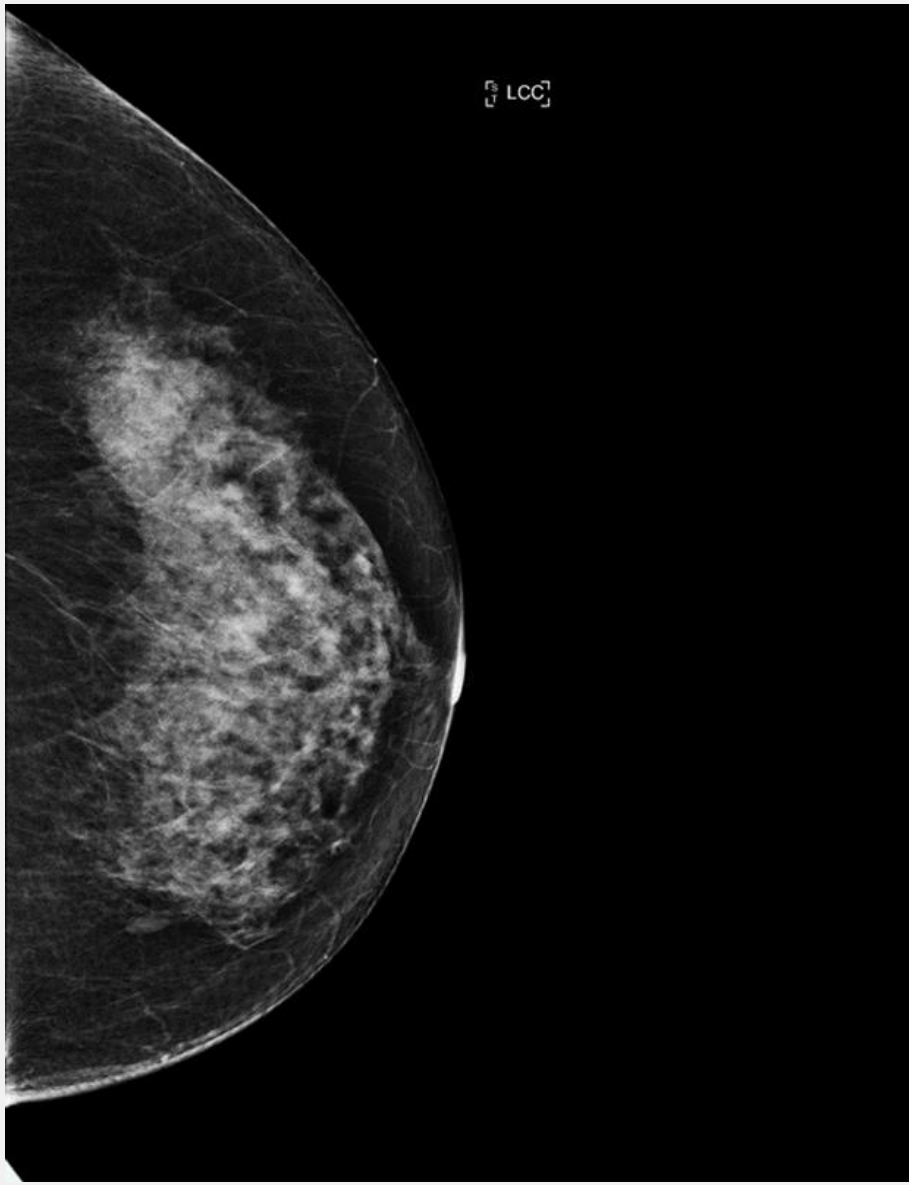
BreastScreen Australia Conference,
Canberra, 15th March 2024

1. The Daffodil Centre, University of Sydney
2. Melbourne School of Population and Global Health, University of Melbourne
3. Monash BreastScreen, Monash Cancer Centre, Moorabbin Hospital, Melbourne
4. Breast Service, The Royal Women's and Royal Melbourne Hospital, Melbourne
5. Department of Surgery, University of Melbourne, Melbourne

Cancer Council | THE UNIVERSITY OF SYDNEY

Mammographic density (MD)

- Reflects variation in breast tissue:
  fat appears dark, connective and epithelial tissues appear light on mammograms
- Independent risk factor for breast cancer
- Reduces mammographic screening sensitivity
- Women with higher MD have:
  - ➡ higher rates of interval cancers
  - ➡ higher false positive rates



How is MD assessed?

- There are various methods
- Subjective visual assessment, e.g. BI-RADS
- Semi-automated methods, e.g. Cumulus
- Automated methods, e.g. Volpara
- No recommendation for MD standardisation

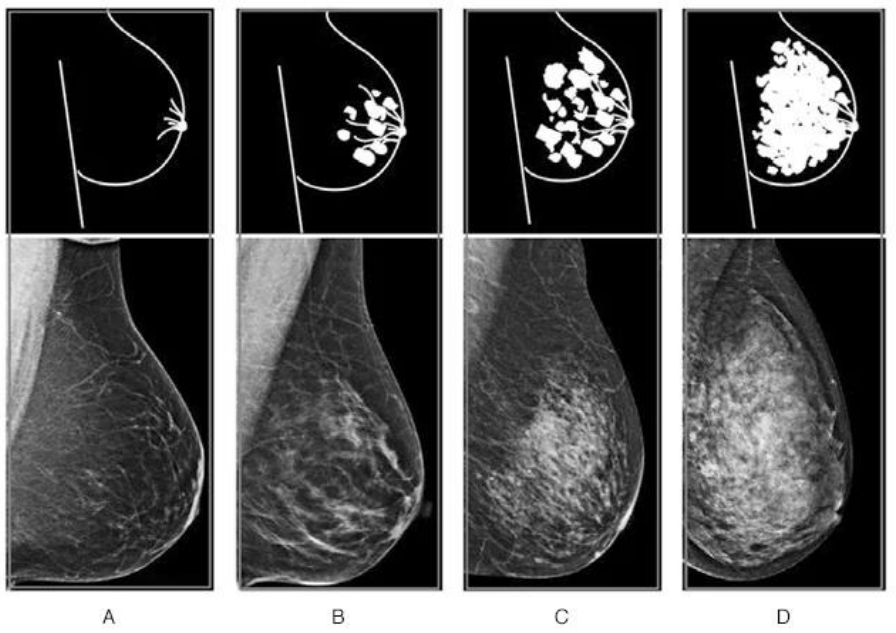
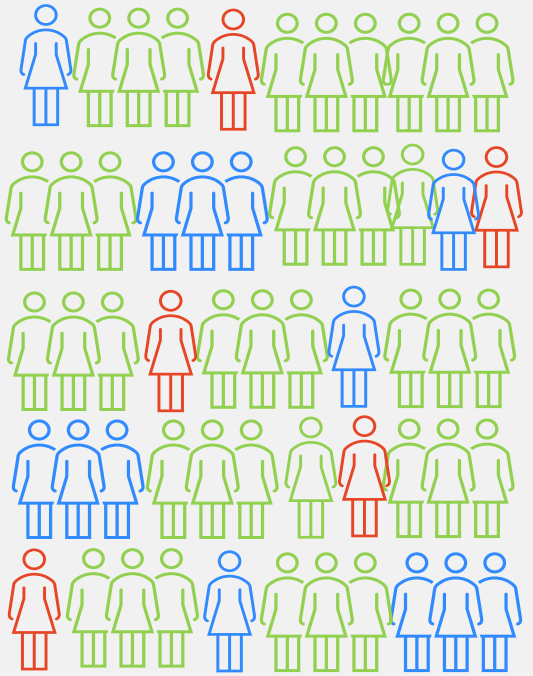


Image: Mayo Foundation
American College of Radiology's Breast Imaging Reporting and Data System

Is there a role for MD in risk-based screening?

- Current breast screening programs:
  - ➡ based on target age
- A more risk-based approach could use MD to adjust screening protocols for women in particularly
  - ➡ high-risk groups
  - ➡ low-risk groups
- To evaluate the role of MD in risk-based screening, important to understand how different assessment methods compare in screening performance outcomes in population screening



Can MD assessment methods stratify women participating in breast cancer screening?

- Systematic review of studies

  i) to determine how MD assessment methods perform in stratifying women according to screening outcomes in **different** screening settings

  ii) to compare how different methods identify risk groups in the **same** screening population

- Part of the Roadmap to Optimising Screening in Australia (ROSA-Breast) project

Methods

Outcome

- Primary: pooled estimates from included studies of the interval invasive cancer rate difference between the two highest MD categories and the two lowest MD categories
- Secondary: trends in screening outcome rates according to increasing MD categories

Eligibility criteria

- Studies in populations screened with DM* reporting ≥ 1 outcomes for all categories of a method
- Outcomes: interval cancer rates, screening program sensitivity (invasive cancers), false positive rates, screening program specificity, missed cancers (apparent on retrospective review but showing minimal signs)

Searches

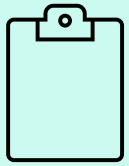
- Medline, Embase, Cochrane Database of Systematic Reviews, International Health Technology Assessment databases
- Jan 2008 – April 2023

Data synthesis

- Each outcome of interest was plotted by (i) MD categories reported by studies and (ii) MD category midpoint percentiles (standardise comparisons)
- Trends in observed outcomes according to MD were calculated
- Pooled estimates (e.g. interval cancer rate difference) were generated using random-effects modelling

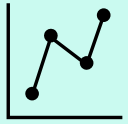
*DM: digital mammography

Results

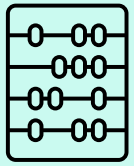


1980 records =>28 articles (26 cohorts) included

5 MD assessment methods:
BI-RADS (n=20), Volpara (n=6), texture resemblance (n=1),
STRATUS (n=1), DenSeeMammo (n=1)



8 studies reported interval cancer rates

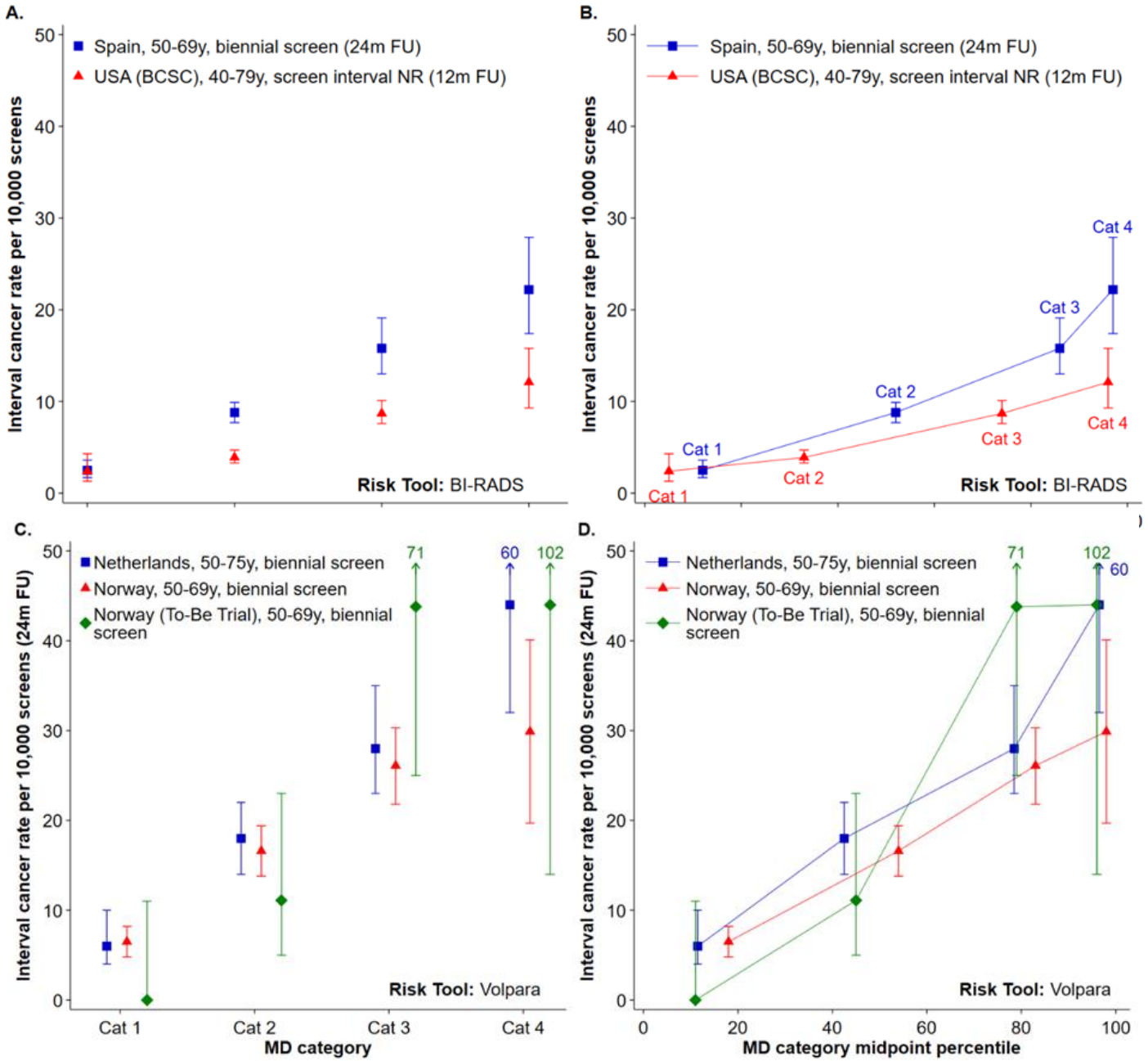


Variation between studies
Setting: organised screening program vs screening in institutions/clinics,
Screening intervals: biennial vs annual/biennial/triennial vs not reported
Age ranges: commencing at 40/50/55y; exiting at 69/74/75y
Screening round: repeat screeners (round 2+) vs first-time screeners

Results

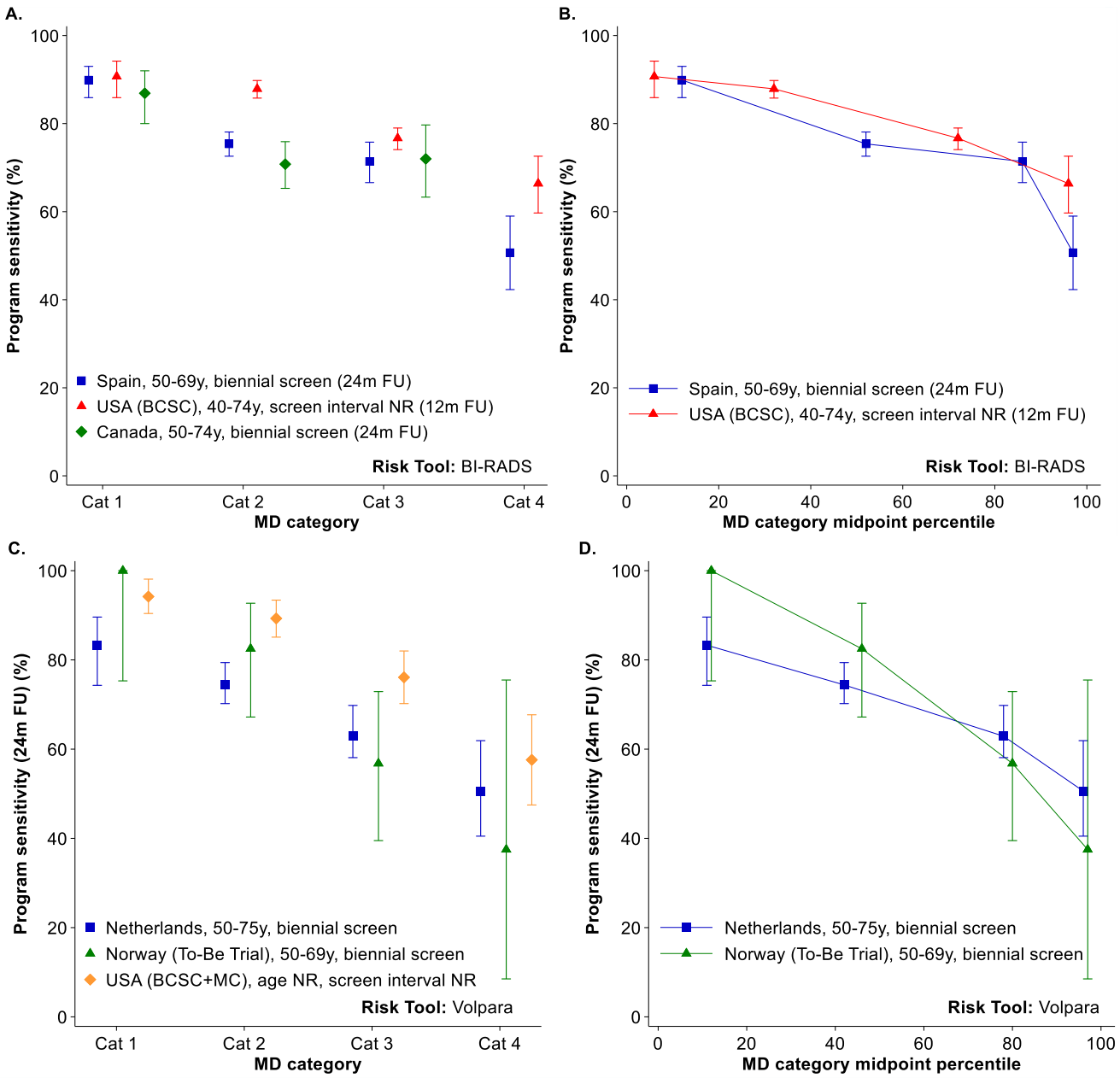
i) How MD measurement methods perform in stratifying women according to screening outcomes in different screening settings

Results – interval invasive cancers



- Graphs
  A,B: BI-RADS studies
  C,D: Volpara studies
- Significant trend of increasing interval cancer rates with increasing MD category ($p<0.001$ for all studies)

Results – screening program sensitivity



- Graphs
  A,B: BI-RADS studies (n=3)
  C,D: Volpara studies (n=3)

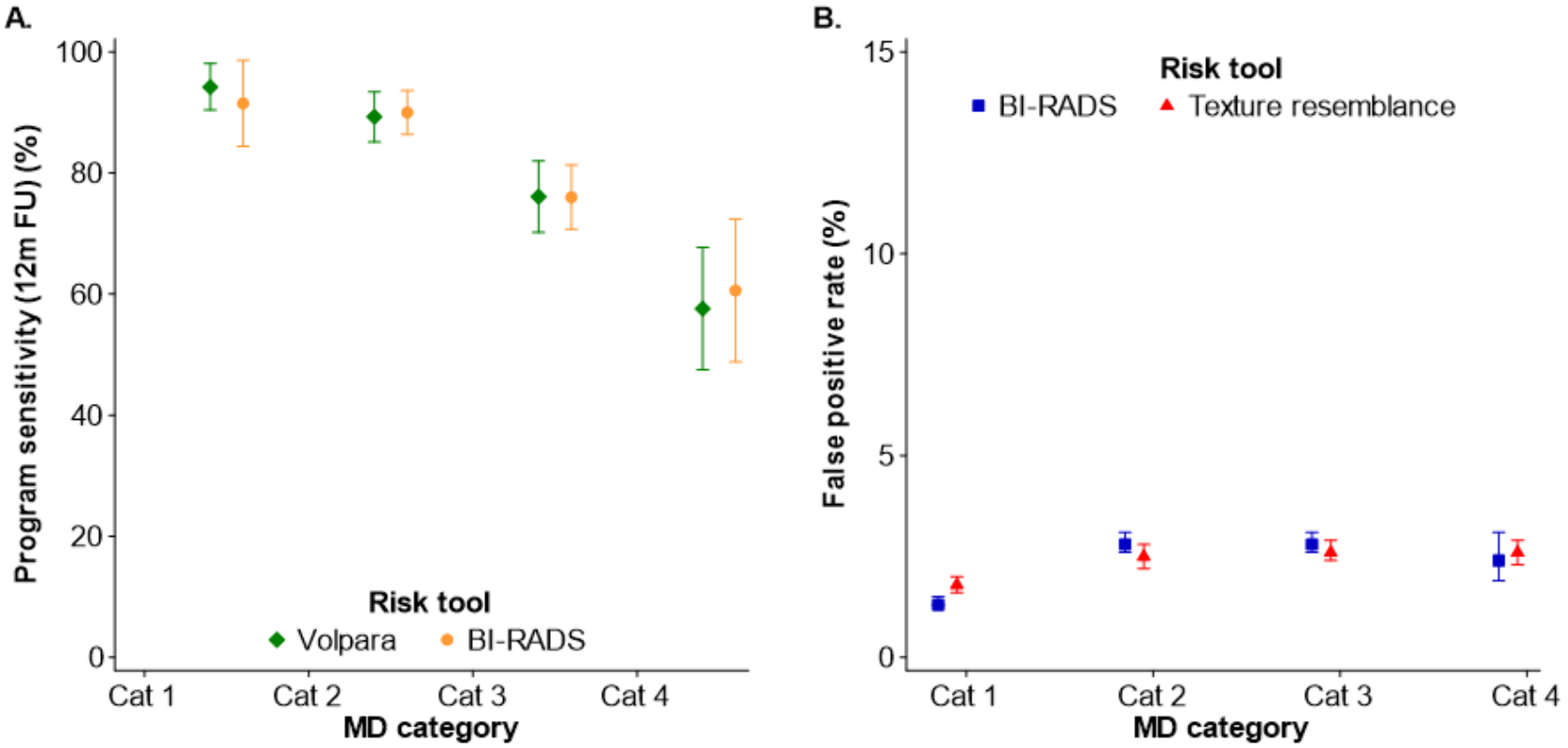
- Significant trend of decreasing program sensitivity with increasing MD category ($p<0.05$ for all studies)

Results

ii) How do different MD measurement methods compare in identifying risk groups in the same screening settings

Results

Only two studies comparing different methods in the same setting
Graph A: program sensitivity in US study (BI-RADS vs Volpara)
Graph B: false positives in European study (BI-RADS vs mammographic texture resemblance
Methods were consistent in their performance -> suggests driver is the setting



Conclusions

- Expected trends of poorer outcomes with increasing MD categories
  - increasing interval cancer rates
  - decreasing program sensitivity
- Most reported MD measurement methods: BI-RADS and Volpara, limited evidence on others
- No study reported clear discrimination of both high- and low-risk groups for interval cancers, for either BI-RADS or Volpara
- Meta-analysis of studies reporting invasive interval cancer rates by MD categories supports the use of BI-RADS and Volpara for directing efforts to reduce interval cancers within screening
- Local validation studies are required before any implementation

Thank you