# The ROSA PROJECT

‘Roadmap for Optimising Screening in Australia – Breast’, investigating risk-based breast cancer screening.

## Chapter 5. Implementation (abridged)

20 March 2023, abridged 1 May 2024

The Daffodil Centre

A partnership between

Cancer Council | The University of Sydney

Produced by the Daffodil Centre for Cancer Council Australia.

The Daffodil Centre

153 Dowling St, Woolloomooloo, NSW 2011

PO Box 572 Kings Cross NSW 1340

info@daffodilcentre.org

daffodilcentre.org



The project 'Roadmap for Optimising Screening in Australia – Breast' was funded by the Australian Government Department of Health and Aged Care (formerly the Department of Health). The project lead was Associate Professor Carolyn Nickson, with contributions from an extensive team of project personnel and various advisory and reference groups as described herein.



**Suggested citation:** The Daffodil Centre (2024). The ROSA Project: Roadmap for Optimising Screening in Australia – Breast. Chapter 5: Implementation (Abridged). 20 March 2023, abridged 1 May 2024. Produced by the Daffodil Centre on behalf of Cancer Council Australia.

The Daffodil Centre acknowledges the Traditional Custodians of Country throughout NSW and recognises the continuing connection to lands, waters, and communities. We pay our respect to Aboriginal and Torres Strait Islander cultures and to Elders past, present, and emerging.

## CONTENTS

# 1 Executive summary

## 1.1 Background

Translating risk-based breast cancer screening evidence into practice in Australia would be a significant undertaking. For any widespread approach to risk-based screening and surveillance in Australia, there are several potential factors likely to affect implementation that operate at many levels including the inner contexts (e.g., health system organisation resources, culture) and outer contexts (e.g., within and between jurisdictional policy, funding), characteristics of individuals (e.g., attitudes, skills), the intervention itself (i.e., risk-based breast screening and surveillance), and the process (e.g., implementation plan). Careful planning that is informed by evidence and stakeholder input can help prepare the Australian health system for any changes required in ways that meet individual and system needs.

## 1.2 Contracted activities

The ROSA project undertook a range of activities to gain insights about implementing risk-based screening in Australia. The topics covered in this chapter and the general approach/methods usedare outlined in Table 1.

*Table 1. Chapter sections and their related ROSA project activities.*

| Chapter section/s | Approach/methods |
|---|---|
| 2. Workforce and organisational readiness (from page 5) | Online surveys of both BreastScreen personnel and health service providers outside BreastScreen providing breast cancer surveillance services about readiness for change in relation to risk-based screening. |
| 3. Trials of risk-based screening (from page 16) | An overview and critical appraisal of trials of risk-based screening. |
| 4. Trialling risk-based breast screening in Australia (from 49) | An analysis of potential trials in the Australian setting. |

## 1.3 Summary of findings

Drawing from the detailed analyses and results described throughout this chapter, the project generated an itemised set of Key findings which were extensively reviewed by the ROSA Expert Advisory Group over May to July 2022, accompanied by summaries of the evidence outlined here. We present the final set of key findings in Appendix 5.1 (page 63).

In summary, our survey of health service providers within and outside the BreastScreen Australia program indicates that a self-selected but diverse group of personnel hold a diverse range of views on potential implementation of risk-based breast screening in Australia. Overall, they are generally supportive, on the proviso that any change in service design and program delivery is based on conclusive evidence, is adequately resourced and not expected to disadvantage subpopulation groups such as women living outside metropolitan areas. Overall, they report that their

organisations are somewhat ready for such change. The study also identified several potential barriers to implementation that warrant attention, including staff capacity and availability, coordination of guidelines and advice between health services, and education of personnel and health service users.

We found six randomised trials currently underway assessing the benefits and/or potential harms of risk-based breast screening. Two assess the effect of supplemental screening for women with dense or extremely dense breasts on screening program outcomes, and four assess risk-based screening that includes reduced screening for some very low risk groups necessitating an assessment of non-inferiority. The trials are collectively assessing a wide variety of interventions, including various screening technologies, screening intervals, age ranges and methods for assessing and categorising risk. Some of these differences may be due to the trial settings, but the variation between studies also highlights the complexity of the evidence related to risk-based breast screening presently, and the range of perspectives about how to best interpret the available evidence.

Our analysis of potential translation of these trials to the Australian setting identifies potential barriers and enablers relating to age ranges, risk assessment, screening intervals and screening tests. These findings, combined with our analysis of the impact of the COVID pandemic on consideration of risk-based breast cancer screening in Australia and stakeholder perspectives on this topic, inform and recommend a strategy for working towards the development of a large-scale trial of risk-based breast cancer screening in Australia. As outlined in this chapter, this would commence with a trial within BSA of routine risk assessment and advice incorporating breast density, to be followed by the design of staged trial protocols for women aged 40-49 and 50-74 as indicated from collection, review and analysis of available evidence.

## 1.4 Glossary of terms

A glossary of selected terms used in this chapter is shown below.

| | |
|---|---|
| **ABUS** | Automated breast ultrasound |
| **BD** | Breast Density. Describes the extent (amount and distribution) of radiopaque tissue in the breast. This is usually perceived through mammography and quantified as either the proportion or area of the breast that is dichotomously dense, or classified through categories such as the BI-RADS breast density categories that combine quantitative and qualitative aspects of the breast density. |
| **BI-RADS** | The American College of Radiology Breast Imaging Reporting & Data System, which includes a framework for categorising breast density through visual assessment. |
| **BSA** | BreastScreen Australia |
| **BSV** | BreastScreen Victoria |
| **BSAPMG** | BreastScreen Australia Program Management Group |
| **CEM** | Contrast-Enhanced Mammography |
| **DBT** | Digital breast tomosynthesis |

| | |
|---|---|
| **EBPAS-36** | Evidence-based Practice Attitude Scale-36 - used to measure individual attitudes towards evidence-based practices |
| **Interval cancer** | Cancer diagnosed following a negative screening episode, within a defined period of the screen (usually 12 or 24 months |
| **Intention-to-treat analysis (ITT)** | A method for analyzing results in a prospective randomized study where all participants who are randomized are included in the statistical analysis and analyzed according to the group they were originally assigned, regardless of what treatment (if any) they received |
| **Overdiagnosis** | Cancers detected by screening that would not have otherwise been found in a woman's lifetime |
| **MD** | Mammographic Density. Another term for breast density, confined to breast density assessed from mammograms. |
| **MRI** | Magnetic Resonance Imaging. |
| **Non-inferiority trials** | Trials assessing whether an intervention is no worse than the comparator (usually current practices). This includes, for example, trials or trial arms assessing less intensive breast screening for lower-risk groups. |
| **ORCA** | Organizational Readiness to Change Assessment. A study instrument used to measure organisational readiness to implement evidence-based practices. |
| **Risk-based surveillance** | Breast cancer surveillance services provided outside BSA through such as primary care, high risk clinics, family cancer centres and specialist breast clinics. This includes breast imaging directed at asymptomatic women on the basis of their breast cancer risk. |
| **Superiority trials** | Trials assessing whether an intervention is better than the comparator (usually current practices). This includes, for example, trials or trial arms assessing more intensive breast screening for higher-risk groups. |
| **Type I error** | Falsely rejecting a null hypothesis that is actually true. For example, finding a difference between interventions on outcomes when there is no difference. |
| **Type II error** | Failing to reject a null hypothesis that is actually false. For example, finding no difference between interventions on outcomes when there is, in truth, a difference. |
| **US** | Ultrasound. |

# 2 Workforce and organisational readiness

## 2.1 Authors

Dr Sabine Deij, Elijah Tyedmers, A/Prof Carolyn Nickson, Dr Andrea Smith, Gabriella Tiernan, Amanda Tattam, A/Prof Natalie Taylor

## 2.2 Background

One aim of the ROSA project is to consider the translation of risk-based screening protocols currently being trialled internationally within the Australian setting. This requires an understanding of Australian health services both within and outside of the BreastScreen Australia (BSA) program, and how these health services might support or be impacted by, the introduction of risk-based breast cancer screening.

The ROSA project began work on this topic in 2018 that is summarised in two reports finalised in 2019, namely:

(i) 'An environmental scan of clinical services involved in or impacted by risk-based screening or surveillance of women without breast cancer symptoms'; and
(ii) 'Stakeholder perspectives on risk-based screening and highlighted many challenges for the trial or implementation of more systematic risk-based screening services in Australia'.

*Abridgement note: These unpublished reports were included as appendices in the full report but have been removed from the abridged report.*

These reports described a range of clinical guidelines, policies and practices in place both within and outside the BreastScreen program for how to assess breast cancer risk and manage screening and surveillance of asymptomatic women at higher-than-average risk (as defined through various guidelines). This analysis highlighted that Australian women can receive different risk advice and management depending on where they live and who they see. The reports also found that stakeholders hold a variety of views about how to best consider options for risk-based screening.

Implementation of risk-based breast screening would involve development of new guidelines, technologies or clinical pathways that would require adaptation from healthcare organisations and changes in behaviour by healthcare professionals, both within and outside the BreastScreen program. With a multidisciplinary workforce located in over 750 screening sites across the country [1], such changes would require substantial 'buy-in' and engagement from many organisations and personnel working across the sector. It is timely to explore current perspectives across that workforce about how ready they and their organisations would be to implement risk-based breast cancer screening.

### 2.2.1 BreastScreen Australia (BSA)

The BSA program is provided by state and territory services and jointly funded by Commonwealth and state and territory budgets. Service delivery models vary between state and territory programs, so that risk-based breast screening may require different implementation strategies. For example, workforce considerations may differ depending on how personnel are engaged by BSA programs and contextual differences such as competing demands for personnel from other health services. Having an appropriate and willing BSA workforce is critical to the successful operation of BSA [2] and the introduction of personalised breast cancer screening [3].

### 2.2.2 Other health services

Outside BSA, many health services such as primary care, high risk clinics, family cancer centres and specialist breast clinics provide risk-based surveillance for breast cancer, drawing on Medicare and hospital-based funding, often also requiring out-of-pocket payments by women involved. These services would be impacted by, and need to make changes to support, more risk-based screening provided as delivered by BSA. These services may also need to change their own guidelines and data management in relation to risk-based surveillance to help ensure consistent and equitable breast cancer screening and surveillance services across the Australian community.

For any provision of risk-based screening by BSA, primary care is likely to play a particularly important role to support women as they receive risk advice and risk-based management, and potentially as an interface between BSA and specialist services for women at particularly high risk. The role of primary care in providing risk assessment to support BSA risk-based screening is less clear(see Chapter 4, section reporting  ROSA Clinical and Health Economics Modelling). Validated tools to assess workforce and organisational readiness

Implementation science provides a theoretical base to help plan for successful implementation of health service trials or evaluations [4,5], offering validated analytic tools that provide a structured means of highlighting elements important to successful implementation such as the strength and providence of the evidence base, the importance of context, and the need for active and deliberate action to bring about change.

Several tools focus on the preliminary stages of the implementation process and provide evidence for the degree of workforce and organisational support implementing new practices. These include:

1. The Evidence-based Practice Attitude Scale-36 (EBPAS-36) [6] instrument, which is a 36-item questionnaire (shortened from its 50-item predecessor (EBPAS-50)) that can be used to measure individual attitudes towards evidence-based practices.
2. The Organizational Readiness to Change Assessment (ORCA) [7] instrument, which measures organisational readiness to implement evidence-based practices. This tool was developed from the Promoting Action on Research Implementation in Health Service (PARIHS) framework.

## 2.3 Aims

1. To understand whether Australian health services personnel within and outside BSA are likely to support the introduction of risk-based breast cancer screening.

2. To understand whether Australian health services personnel within and outside BSA think their organisations are ready for the introduction of risk-based breast cancer screening.

## 2.4 Research question

Are Australian health services personnel working in screening and surveillance likely to support the introduction of risk-based breast cancer screening, and do they think their organisations are ready?

# 2.5 Methods

## 2.5.1 Study populations

The study populations comprised personnel from within BSA and health services personnel outside BSA (Table 2). Some individuals are likely to belong to both populations.

*Table 2. Survey study populations defined according to their roles within or outside the BreastScreen Australia (BSA) program.*

| Survey Population | Eligible participants |
|---|---|
| BSA personnel | Any employee of any BreastScreen state or territory program |
| Health service providers outside BSA | Health service providers involved in screening and risk-based surveillance for breast cancer outside of BSA |

## 2.5.2 Survey design

A separate survey was developed for each population. The survey for the BSA group was developed with feedback from the ROSA BreastScreen Reference Group and the ROSA Expert Advisory Group chair Paul Vardon.

The surveys used questions adapted from two validated implementation science instruments, namely:

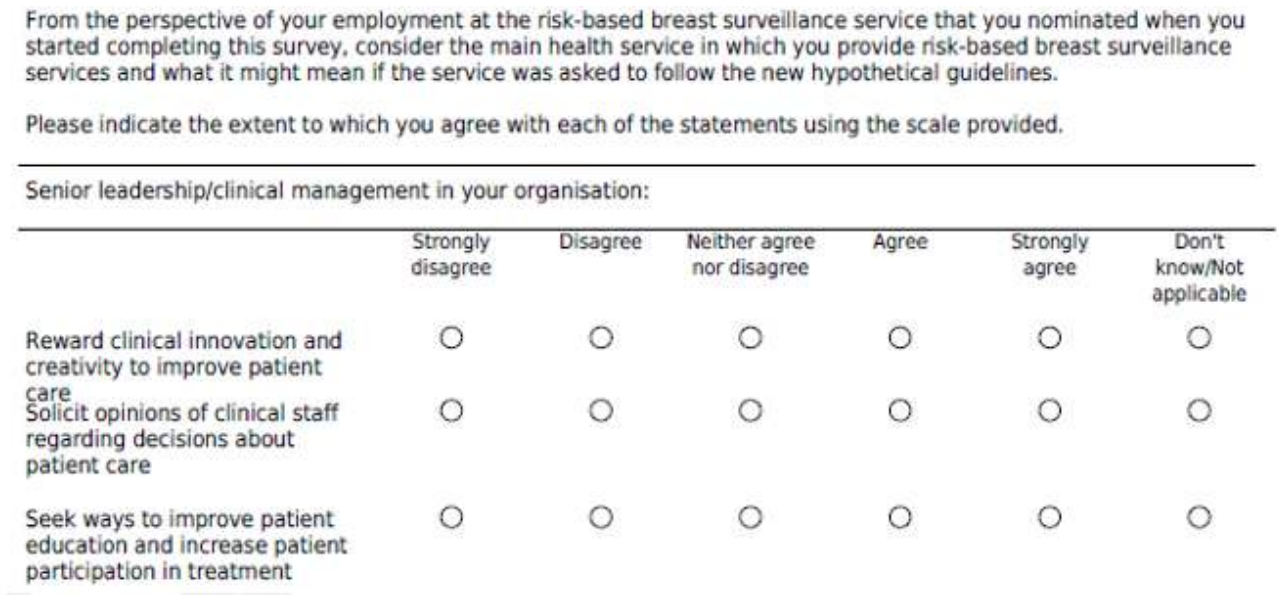
- The 'Organizational Readiness to Change Assessment' (ORCA) instrument.
- The Evidence-based Practice Attitude Scale (EBPAS-36) instrument.

### Organisational readiness for change

The ORCA instrument can be used to assess the willingness of an organisation to accept change, by helping to identify where support might be needed within an organisation during a change process. This is captured through five dimensions:

- 'Leadership culture', which measures respondents' perception of senior leadership's/management's ability to promote a culture that rewards innovation and input from staff towards improving patient care and outcomes.
- 'Staff culture', which measures respondents' perception of staff members' sense of personal responsibility, their cooperation and their willingness towards improving patient care and outcomes.
- 'Leadership', which measures respondents' perception of the individuals in leadership roles that they interact with when carrying out their work, i.e., their supervisors.
- 'Measurement', which covers perceptions of how well an organisation and its leadership motivates its aims and supports staff to understand what they should be doing and giving feedback on their performance within their role.
- 'Opinion leaders', which assesses respondents' perceptions on the role influential people within the organisation play to influence the change processes.

In both surveys, the change that respondents were asked to consider was the implementation of new hypothetical evidence-based guidelines for risk-based cancer screening[1]. An example component of this survey is shown in Figure 1.



From the perspective of your employment at the risk-based breast surveillance service that you nominated when you started completing this survey, consider the main health service in which you provide risk-based breast surveillance services and what it might mean if the service was asked to follow the new hypothetical guidelines.

Please indicate the extent to which you agree with each of the statements using the scale provided.

Senior leadership/clinical management in your organisation:

| | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree | Don't know/Not applicable |
|---|---|---|---|---|---|---|
| Reward clinical innovation and creativity to improve patient care | ○ | ○ | ○ | ○ | ○ | ○ |
| Solicit opinions of clinical staff regarding decisions about patient care | ○ | ○ | ○ | ○ | ○ | ○ |
| Seek ways to improve patient education and increase patient participation in treatment | ○ | ○ | ○ | ○ | ○ | ○ |

*Figure 1. Sample component of survey questions on organisational readiness for change in relation to risk-based breast cancer screening.*

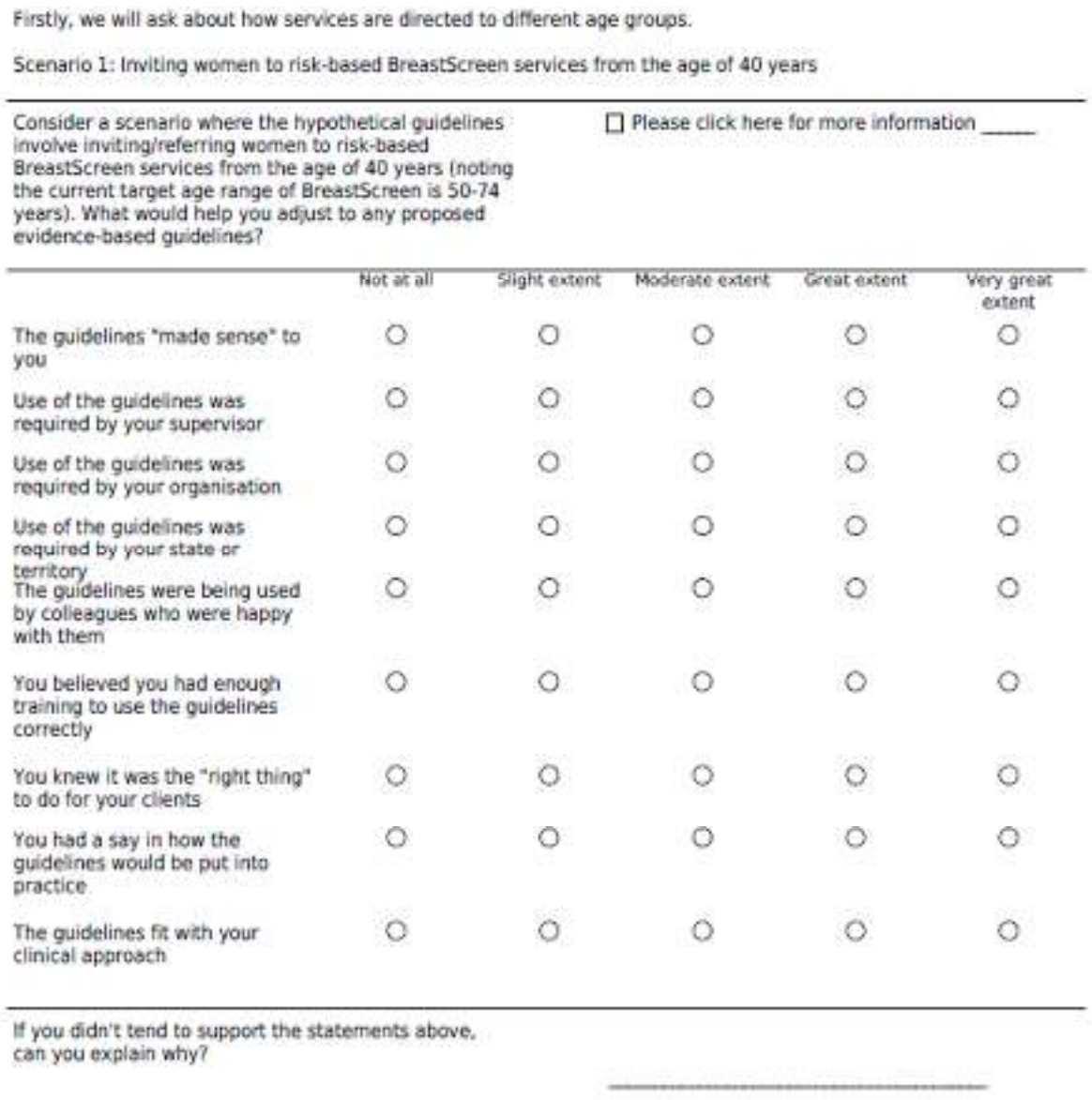
## Individual attitudes to adopting new guidelines

The EBPAS-36 instrument provides a brief and pragmatic measure of attitudes to the use and adoption of evidence-based practice along several dimensions. The instrument questions were adapted for the ROSA survey to gauge attitudes towards implementation of guidelines in general for nine dimensions:

- Openness – openness to new practices.
- Divergence - the perceived divergence of one's usual practice with research-based/academically developed interventions.
- Limitations – the limitations of evidence-based practices.
- Monitoring – negative perceptions of monitoring.
- Balance - the perceived balance between clinical skills and science as important in health service provision.
- Burden - the time and administrative burden with learning evidence-based practices.
- Job security – as related to expertise in evidence-based practices.
- Organisational support - perceived organisational support.
- Feedback - positive perceptions of receiving feedback.

An example of the EBPAS-36 questions as implemented in the survey is shown in Figure 2.

[1] The following is an example of the scenario used in the BreastScreen Survey only: Consider a general scenario where your state or territory BreastScreen program introduces risk-based screening which involves providing more intensive imaging (e.g. MRI, digital breast tomosynthesis, contrast enhanced mammography or supplemental ultrasound) to some women according to their risk, where that risk assessment is based on questionnaire information and a breast density assessment.



Firstly, we will ask about how services are directed to different age groups.

Scenario 1: Inviting women to risk-based BreastScreen services from the age of 40 years

Consider a scenario where the hypothetical guidelines involve inviting/referring women to risk-based BreastScreen services from the age of 40 years (noting the current target age range of BreastScreen is 50-74 years). What would help you adjust to any proposed evidence-based guidelines?

☐ Please click here for more information _____

| | Not at all | Slight extent | Moderate extent | Great extent | Very great extent |
|---|---|---|---|---|---|
| The guidelines "made sense" to you | ○ | ○ | ○ | ○ | ○ |
| Use of the guidelines was required by your supervisor | ○ | ○ | ○ | ○ | ○ |
| Use of the guidelines was required by your organisation | ○ | ○ | ○ | ○ | ○ |
| Use of the guidelines was required by your state or territory | ○ | ○ | ○ | ○ | ○ |
| The guidelines were being used by colleagues who were happy with them | ○ | ○ | ○ | ○ | ○ |
| You believed you had enough training to use the guidelines correctly | ○ | ○ | ○ | ○ | ○ |
| You knew it was the "right thing" to do for your clients | ○ | ○ | ○ | ○ | ○ |
| You had a say in how the guidelines would be put into practice | ○ | ○ | ○ | ○ | ○ |
| The guidelines fit with your clinical approach | ○ | ○ | ○ | ○ | ○ |

If you didn't tend to support the statements above, can you explain why?

_____

*Figure 2. An example of the EBPAS-36 questions from the ROSA survey (Section 4 Scenario 1) below.*

These questions were used for a range of hypothetical evidence-based guidelines (hereon described as ‘scenarios’). For each scenario, attitudes were captured across three dimensions:

- Requirements – likelihood of adoption if the guidelines were required by their organisation, state/territory or supervisor.
- Appeal – likelihood of adoption if the guidelines made sense, there was sufficient training and colleagues were happy using them.
- Fit – likelihood of adoption if the guidelines were seen as the ‘right thing’ to do, they fitted with the respondent’s clinical approach and they had a say on how they were implemented.

To cover a wide range of scenarios while minimising the impost on respondents, participants were randomly assigned to consider different subsets of scenarios.

The set of scenarios presented to BSA respondents is shown in Table 3. All participants were invited to respond to Scenario 1, and then they were randomly assigned to one of two groups, namely:

- Group 1, to consider various scenarios based on risk assessment (Scenarios 2-4), or

- Group 2, to consider various scenarios based on imaging modalities for higher-risk women (Scenarios 5-7).

*Table 3. Summary of scenarios presented to BSA respondents.*

| Scenario | Description | Group 1 | Group 2 |
|---|---|---|---|
| Scenario 1 | Inviting women to risk-based BreastScreen services from the age of 40 years | ● | ● |
| Scenario 2 | More intensive imaging based on **questionnaire-based risk assessment** | ● | |
| Scenario 3 | More intensive imaging based on **genetic test results** | ● | |
| Scenario 4 | More intensive imaging based on **mammographic breast density** | ● | |
| Scenario 5 | Higher risk women screened using **digital breast tomosynthesis** | | ● |
| Scenario 6 | Higher risk women screened using **supplemental ultrasound** | | ● |
| Scenario 7 | Higher risk women screened using **MRI** | | ● |

The set of scenarios presented to personnel outside BSA is shown in Table 4. All participants were invited to respond to two common scenarios (1 and 2), and then randomly assigned to one of two groups, namely:

- Group 1, to consider various scenarios based on risk assessment (Scenarios 3-5), or
- Group 2, to consider various scenarios based on imaging modalities for higher-risk women (Scenarios 6-8).

*Table 4. Summary of scenarios presented to health service providers outside BSA.*

| Scenario | Description | Group 1 | Group 2 |
|---|---|---|---|
| Scenario 1 | Inviting women to risk-based BreastScreen services from the age of 40 years | ● | ● |
| Scenario 2 | Risk-based guidelines for women under 40 years | ● | ● |
| Scenario 3 | More intensive imaging based on **questionnaire-based risk assessment** | ● | |
| Scenario 4 | More intensive imaging based on **genetic test results** | ● | |
| Scenario 5 | More intensive imaging based on **mammographic breast density** | ● | |
| Scenario 6 | Higher risk women screened using **digital breast tomosynthesis** | | ● |
| Scenario 7 | Higher risk women screened using **supplemental ultrasound** | | ● |
| Scenario 8 | Higher risk women screened using **MRI** | | ● |

More detail about each scenario was provided as shown in Table 5 below.

*Table 5. Full descriptions for each scenario about attitudes to change.*

| |
|---|
| **Scenario 1: Inviting women to risk-based BreastScreen services from the age of 40 years** |
| Consider a scenario where the hypothetical guidelines involve inviting/referring women to risk-based BreastScreen services from the age of 40 years (noting the current target age range of BreastScreen is 50-74 years). What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 2: Guidelines for women under 40 years (only in HSP survey)** |
| This time consider a scenario where the hypothetical guidelines involve a standardised risk assessment for women at age 30 years, through your health service(s) outside of BreastScreen. What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 3: More intensive imaging based on questionnaire-based risk assessment** |
| Consider a scenario where the hypothetical guidelines involve providing more intensive imaging (e.g. MRI, digital breast tomosynthesis or supplemental ultrasound) to some women according to a questionnaire-based risk assessment, using a validated risk assessment tool. This could include questions about e.g. menopause, hormone therapy use, reproductive history and detailed family history of breast and ovarian cancer. What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 4: Providing digital breast tomosynthesis for screening in higher risk women** |
| This time consider a scenario where the hypothetical guidelines involve providing digital breast tomosynthesis to higher-risk women, assuming the cost of these services would be covered by BreastScreen or subsidised by Medicare (depending on the setting). What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 5: More intensive imaging for clients/patients based on genetic test results** |
| Consider a scenario where the hypothetical guidelines involve providing more intensive imaging (e.g. MRI, digital breast tomosynthesis or supplemental ultrasound) to some women based on the results of a genetic test for high-risk mutations and polygenic risk scores, assuming the cost of these services would be covered by BreastScreen or subsidised by Medicare (depending on the setting). What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 6: Providing supplemental ultrasound to higher risk women** |
| This time consider a scenario where the hypothetical guidelines involve providing a supplemental ultrasound to screening higher-risk women. What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 7: More intensive imaging based on mammographic breast density** |
| This time consider a scenario where the hypothetical guidelines involve providing more intensive imaging (e.g. MRI, digital breast tomosynthesis or supplemental ultrasound) to some women based on their mammographic breast density, assuming the cost of these services would be covered by BreastScreen or subsidised by Medicare (depending on the setting). What would help you adjust to any proposed evidence-based guidelines? |
| **Scenario 8: Providing MRI to higher-risk women** |
| This time consider a scenario where the hypothetical guidelines involve providing MRI for screening of higher-risk women, assuming the cost of these services would be covered by BreastScreen or subsidised by Medicare (depending on the setting). What would help you adjust to any proposed evidence-based guidelines? |

### Benefits and challenges

Respondents were asked several open-ended questions about the potential benefits and challenges in their role as well as for their organisation if a risk-based approach to screening and surveillance were introduced in Australia, as follows:

- *In your view, what might be the key benefits to potentially introducing guidelines for risk-based breast screening and surveillance such as those described in the different scenarios?*
- *In your view, what would be the key challenges in your role at BreastScreen if risk-based approaches to breast cancer screening and surveillance were introduced in Australia?*
- *In your view, what would be the key challenges for BreastScreen if risk-based breast cancer screening and surveillance were introduced in Australia?*

### COVID impacts and general feedback

Respondents were asked how the current COVID-19 pandemic influenced their answers and invited to provide any additional thoughts they wanted to share with the research team and general feedback about the survey. The specific questions were:

- *Does the current COVID-19 pandemic change your answers, and if so, in what way?*
- *Are there any other things you think we should know?*
- *Other comments/feedback?*
- *Did you have any difficulties answering any of the questions in this survey? (No/Yes)*
    - *If 'Yes', please comment [optional]*

## 2.5.3 Recruitment

For BSA personnel, participants were recruited via a survey recruitment flyer and, following consultation with the BSA Program Management Group, an email to each BSA state and territory Program Manager inviting them to opt-in to distribute the survey (and asking them to nominate a key contact in their jurisdiction to distribute the survey details and survey link to staff). All state and territory program managers agreed to distribute the survey and were sent a link to the survey via the Cancer Council Australia website[2]. The survey was also promoted via social media, including through the Cancer Council Australia and University of Sydney Daffodil Centre LinkedIn accounts.[3,4.]

For health service providers outside BSA, we identified services of interest including (but were not limited to) primary health care settings (e.g., GPs), high-risk clinics (e.g., family cancer centres), breast cancer specialists, and genetic testing, imaging and pathology services. Participants were recruited by promoting the study through our existing connections to services established over the course of the ROSA project. Promotional material included study details, an email address for the research team so that potential participants could request further information or speak to a team member, and a direct web link to the online survey. In addition to the mechanisms of promotion on

[2] www.cancer.org.au/about-us/policy-and-advocacy/early-detection-policy/breast-cancer-screening/optimising-early-detection].
[3] www.linkedin.com/posts/cancer-council-aus_do-you-work-for-breastscreen-or-a-related-activity-6927859448082268160-Nrrd?utm_source=linkedin_share&utm_medium=member_desktop_web
[4] www.linkedin.com/posts/the-daffodil-centre_do-you-work-for-breastscreen-or-a-related-activity-6927474316389347328-B9An?utm_source=linkedin_share&utm_medium=member_desktop_web

social media described above, the survey was also promoted through the ROSA Expert Advisory Group and Co-opted Experts panel as well as through communication with peak professional bodies, such as the Royal Australian and New Zealand College of Radiologists, Australian Society for Medical Imaging and Radiation Therapy, Australian Society for Breast Disease, Australian Society of Breast Physicians, Breast Cancer Network Australia, Breast Surgeons of Australia and New Zealand, Cancer Nurses Society of Australia and the Royal Australian College of General Practitioners.

Potential participants who were eligible for both surveys were encouraged to complete the survey twice, for each of their roles.

A draw for one of five $100 gift cards (for each survey) was used as an incentive for participants to complete the survey before the deadline on 9 May 2022. Respondents were able to participate in this incentive by providing their details in a separate form, used solely for this purpose.

### 2.5.4 Data collection

Surveys were hosted online, on the University of Sydney REDCap platform[5,6]. This REDCap platform is hosted on secure and encrypted University-licensed servers within NSW and meets University standards for security, data ownership and privacy. REDCap accounts require approval from the University of Sydney's Information Communications and Technology service department, and all accounts are password-protected. Only approved researchers working on this study had access to the study REDCap platform.

Participants' information was stored securely in a specific study folder in the University of Sydney Research Data Store, with access permitted only by staff authorised as per the approved University of Sydney HREC in January 2022 (ID # 2021/843). For participants who chose to participate in an interview, to receive a copy of the final public report and/or to go into the draw for a gift voucher, participant contact details were stored separately to their survey data. Survey data was not re-identifiable as the survey was completed anonymously.

The data collection period for responses included in this report was 6 April 2022 to 30 June 2022.

### 2.5.5 Data analysis and reporting

Survey data were included in the analysis if score-based questions were at least 80% complete.

Data were analysed separately according to study population (BSA personnel versus other health service providers) due to differences in the recruitment, methods, questions and expected study population for each survey.

#### Quantitative data

Quantitative data was analysed using summary scores of Likert-scale responses for each set of questions (as specified by the study instruments from which they were drawn), also reporting

[5] PA Harris, R Taylor, R Thielke, J Payne, N Gonzalez, JG. Conde, Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support, J Biomed Inform. 2009 Apr;42(2):377-81.
[6] PA Harris, R Taylor, BL Minor, V Elliott, M Fernandez, L O'Neal, L McLeod, G Delacqua, F Delacqua, J Kirby, SN Duda, REDCap Consortium, The REDCap consortium: Building an international community of software partners, J Biomed Inform. 2019 May 9 [doi: 10.1016/j.jbi.2019.103208]

summary statistics (mean, range, median, inter-quartile range) for specific questions. Results were plotted to help identify patterns.

#### Qualitative data

Qualitative data were collated and reviewed in detail with a view to identifying more comprehensive and illustrative responses. We also conducted an initial thematic coding of these data using an inductive thematic qualitative method, where similar quotes were grouped by themes based on salience and frequency [8]. For this report we have selected illustrative examples of qualitative feedback against these initial themes; the coding and themes are being further reviewed and refined for a peer-reviewed manuscript.

### 2.5.6 Participant follow-up

Upon completion of the surveys, participants were asked if they would like to be contacted for an optional interview, would like to receive a summary of the results of this survey and/or would like to participate in a draw for one of five additional $100 gift cards (for each survey).

Participants who indicated that they would like to receive a summary of the results of this survey will receive this summary from the researchers (under the governance of the project ethics approval, and subject to approval by the Australian Government Department of Health and Aged Care (the project funder)).

The draws for the $100 gift cards were conducted on 23 May 2022. Winning participants were notified and sent their electronic gift card via email on 26 May 2022.

### 2.5.7 Ethics approval

Ethics approval was provided by the University of Sydney in January 2022 (ID # 2021/843).

*Abridgement note: Survey results and additional detail on methods will become available through a publication in process. Please also refer to key findings (section 5.1, from page 63).).).*

## 2.6 Conclusion

This study indicates that a self-selected but diverse group of Australian health services personnel working in breast cancer screening and surveillance hold a range of views on potential implementation of risk-based breast screening. However, they are generally supportive, on the proviso that a) any change in service design and program delivery is based on conclusive evidence; and b) services would be adequately resourced to implement and would not disadvantage population groups such as women living outside metropolitan areas. Overall, they report that their organisations are somewhat ready for such change. The study also identifies several potential barriers to implementation that warrant attention, including staff capacity and availability, coordination of guidelines and advice between health services, and education of personnel and health service users.

# 2.7 References

1. Commonwealth of Australia Department of Health and Aged Care. About the BreastScreen Australia Program 2022 [Available from: https://www.health.gov.au/initiatives-and-programs/breastscreen-australia-program/about-the-breastscreen-australia-program#:~:text=the%20screening%20process.-,Where%20you%20can%20have%20a%20breast%20screen,can%20get%20a%20breast%20screen. Last accessed 20 March 2023.
2. Australian Government Department of Health. Population Based Screening Framework. Updated August 2018. [Available from: https://www.health.gov.au/sites/default/files/documents/2019/09/population-based-screening-framework_0.pdf]. Last accessed 20 March 2023.
3. Keogh L, Steel E, Weideman P, Butow P, Collins IM, Emery J, et al. Consumer and clinician perspectives on personalising breast cancer prevention information. The Breast. 2019;43:39-47.
4. Nilsen P. Making sense of implementation theories, models and frameworks. Implementation Science. 2015;10(1):53.
5. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging Research and Practice: Models for Dissemination and Implementation Research. American Journal of Preventive Medicine. 2012;43(3):337-50.
6. Rye M, Torres EM, Friborg O, Skre I, Aarons GA. The Evidence-based Practice Attitude Scale-36 (EBPAS-36): a brief and pragmatic measure of attitudes to evidence-based practice validated in US and Norwegian samples. Implementation Science. 2017;12(1):44.
7. Helfrich CD, Li Y-F, Sharp ND, Sales AE. Organizational readiness to change assessment (ORCA): Development of an instrument based on the Promoting Action on Research in Health Services (PARIHS) framework. Implementation Science. 2009;4(1):38.
8. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.
9. Commonwealth of Australia Department of Health. BreastScreen Australia Workforce Demand Report. 2020. (Provided to the ROSA project in confidence by the Australian Government Department of Health)
10. Moran S, Warren-Forward H. The Australian BreastScreen workforce: a snapshot. Radiographer. 2012;59(1):26-30.
11. BreastScreen Australia Radiographer and Mammographer Workforce Demand Working Group. BreastScreen Australia Radiographer and Mammographer Workforce Demand Report. 2020 2 November.

# 3 Trials of risk-based screening

## 3.1 Authors

Dr Denise Campbell, Chelsea Carle, Suzanne Hughes, Dr Susan Yuill, Sam Egger, Prof Dianne O'Connell, Dr Louiza Velentzis & A/Prof Carolyn Nickson.

## 3.2 Background

### 3.2.1 Rationale

The BreastScreen Australia (BSA) program offers free biennial mammography to women between the ages of 40 and 74 years, with annual screening offered to women with a personal history of breast cancer or breast disease or a family history of breast or ovarian cancer (policies differ between state and territory services). In addition, through clinical services outside BreastScreen, women who are at very high risk of developing breast cancer are invited to begin screening at a younger age, are screened more frequently, and are offered supplemental screening such as magnetic resonance imaging (MRI).

As outlined in Chapter 4, for all but very high-risk women, the current approach may not serve all women well. Women with a higher risk of developing breast cancer (particularly those with higher breast density) may be under-screened, as indicated by higher rates of interval cancers and women with a very low risk of developing breast cancer (particularly those with very low breast density) may be over-screened, as indicated by very low rates of interval cancers and relatively higher rates of overdiagnosis. These women may benefit from a screening program in which a woman's risk of breast cancer determines the intensity and/or modality by which she is screened.

The ROSA scoping reviews undertaken in August 2019 identified a number of trials currently underway designed to assess the benefits and harms of a variety of risk-based screening protocols. An objective assessment of the quality of these trials will be a necessary step in determining which, if any, of these risk-based screening protocols would best meet Australian needs.

This report work presents a scoping-level appraisal of the quality of randomised trials which compare risk-based screening with a single screening protocol. Initial appraisals were undertaken in a May 2020 and updated in September 2021. This report presents the updated consolidated results of these appraisals and includes additional assessments to address some of the limitations of the May 2020 appraisals.

### 3.2.2 Aims

To identify, summarise and critically appraise the published or documented details of registered ongoing randomised controlled trials of risk-based breast cancer screening.

## 3.3 Methods

### 3.3.1 Trial selection criteria

Trial selection criteria are shown in Table 6.

*Table 6. Selection criteria for registered ongoing risk-based screening trials*

| | ***Inclusion criteria*** | ***Exclusion criteria*** |
|---|---|---|
| **Population** | Asymptomatic women aged ≥ 40 years undergoing breast cancer screening including age subgroup of women eligible for screening e.g., women aged 40-49 or 44-50 years<br>Breast cancer screening participants identified by screening program as being at higher risk of breast cancer | Study population restricted to:<br>• Women undergoing breast imaging as follow-up for breast cancer or DCIS, or for breast abnormalities, or suspect mammogram<br>• Selected populations enriched for cancer<br>Study population include women aged < 40 years and mean or median age of population < 50 years<br>High risk women not identified as part of a screening program<br>Screening program does not include assessment of any risk |
| **Intervention** | Risk-based screening program i.e. breast cancer risk is assessed and screening protocol (i.e. frequency, modality, initiation and/or cessation) is determined by risk | |
| **Comparator** | Standard non-risk–based screening<br>No additional screening for higher risk group | |
| **Outcome** | Screen-detected invasive breast cancer and/or DCIS<br>Recalls<br>False positives<br>Positive predictive values<br>Interval cancers<br>Screening program sensitivity<br>Negative predictive values<br>Screening program specificity<br>Cost-effectiveness | Predicted program sensitivity<br>Diagnostic sensitivity based on cross-sectional data |
| **Trial design** | Randomised controlled trial | Cohort studies<br>Case-control studies |
| **Trial status** | Ongoing or only interim results published | Results for primary outcome published |
| **Publication type** | Journal article, website, clinical trial registry record | |
| **Language** | English | |

## 3.3.2 Trial registry searches

Registered, ongoing clinical trials of risk-based breast cancer screening were identified by searching registry websites.
The following clinical trial registries were searched:

- Clinicaltrials.gov using the terms:
  "breast cancer", "tomosynthesis"
  "breast cancer", "screen" and "MRI"
  "breast cancer", "screen" and "ultrasound"
  "breast cancer", "screen", "mammogram" and "contrast-enhanced"
  "breast cancer", "screen" – interventional - phase 2 or Phase 3;
- Australian New Zealand Clinical Trials Registry (ANZCTR) using the terms: "tomosynthesis",
  "breast cancer" and "MRI"
  "breast cancer" and "ultrasound"
  "breast cancer" and "contrast-enhanced"
  "breast cancer" and "screen" – interventional - randomised;
- International Clinical Trials Registry Platform (ICTRP) using the terms:

"tomosynthesis" and "breast'
"ultrasound" and "breast"
"MRI" and "breast"
"contrast-enhanced mammography" and "breast"
"screen" and "breast".

These searches were most recently performed in March 2021.

### 3.3.3 Protocol and publication searches

To gather additional details to assist in the description and appraisal of the quality of included trials, registry records of included trials were checked for listed interim publications or protocols and the internet searched for protocols and interim publications using the trial registration number, acronym and/or title. If no protocols or interim publications were found, further information was sought from the trial principal investigators.

### 3.3.4 Data extraction

Prespecified study details describing the trial population, intervention, comparator, outcomes, trial status, trial methodologies and designs relevant to quality appraisal (described below) were extracted.

### 3.3.5 Quality appraisal

Based on the information collected, the quality of the ongoing trials was appraised using five pre-specified criteria under three domains:

| ***Domain*** | ***Criteria*** |
|---|---|
| Risk of bias | A risk of bias assessment using a risk-of-bias tool |
| Power calculations | Whether statistical power calculations had been undertaken |
| | Assessment of the power calculations |
| Statistical analysis plans | Whether a statistical analysis plan was publicly available |
| | Assessment of the reporting of planned statistical analyses |

All quality appraisals were based on the primary outcomes specified in the registered trial details and were undertaken by at least two independent assessors.

### 3.3.6 Risk of bias assessment

Risk of bias assessments are important because they identify issues that could skew the results. It is important to measure the risk of bias related to a study because there are factors that can systematically affect the observations and conclusions of a study and cause them to be different from the truth (Higgins 2011). Studies affected by bias can be inaccurate and lead to an over- or under-estimation of the true effect of an intervention. This can, in turn, lead to inappropriate clinical recommendations, wasted resources, and result in harm to consumers (NHMRC 2019).

The risk of bias was assessed using version 2 of the Cochrane risk-of-bias tool for randomised trials (Higgins 2019); specifically, the effect of allocation to risk-based screening on the primary outcomes was assessed using this tool. Version 2 of the tool assesses **five sources of bias:**

1. bias due to the randomisation process,
2. bias due to deviations from the intended interventions,
3. bias due to missing outcome data,
4. bias due to measurement of the outcome, and
5. bias due to selection of the reported results.

Each source of bias is assessed using a number of signalling questions and these are shown in Table 13 in the Appendix.

This tool was designed to assess the risk of bias for trials with reported results and used in its entirety to assess risk-based screening trials with reported interim results, but not studies without published results as missing outcome data and details of the actual analyses are only available once results are available. To assess the risk of bias for ongoing trials with no published interim results we used an abridged and modified version of the Cochrane tool. This modified tool assessed only biases due to the randomisation process, deviations from the intended interventions and measurement of the outcome. It assessed the risk of bias due to each of these sources of bias as "provisionally low", "provisionally some concerns", "provisionally high" or "insufficient information". In addition, the assessment of biases due to the randomisation process and deviations from the intended interventions were modified as some of the criteria considered in the original tool, such as baseline differences between intervention groups, are dependent on results being available.

Deviations from the intended intervention are a major potential source of bias in screening trials, as these trials are vulnerable to contamination of the intervention or the comparator groups depending on whether the intervention is perceived as advantageous or disadvantageous, if the participants are not blinded. Measurement bias is also an important potential source of bias because the results could be skewed if outcomes in both groups are not ascertained in the same way or the assessors are not blinded. Four trials planned non-inferiority analyses. These analyses aim to show that outcomes for risk-based screening are not unacceptably worse than those for current screening programs. Such analyses are particularly vulnerable to type I errors, i.e. failing to find that risk-based breast screening is unacceptably worse when it actually is (Piaggio 2012; Mo 2020). For risk-based breast screening trials the effects of specific deviations will depend on the outcome. For example, for the outcome of more advanced stage of cancer at diagnosis, deviations likely to dilute differences between arms in non-inferiority trials could include non-adherence, cross-overs and dropouts, such as where participants in the control arm at higher risk of breast cancer undergo supplementary screening outside of the screening program, participants in the intervention arm at lowest risk of breast cancer undergo current rather than reduced intensity screening or where the proportion of women that drop out is greater in the control arm.

We are interested in the effect of the intervention of risk-based breast cancer screening in the real-world context of population-based screening, rather the effect of adherence to the intervention. Intention-to-treat analysis[7] is considered the appropriate analysis for determining the effect of the intervention when assessing superiority (i.e. in superiority trials or trial arms) as it minimises type I

[7] Intention-to-treat analysis is a method for analyzing results in a prospective randomized study where all participants who are randomized are included in the statistical analysis and analyzed according to the group they were originally assigned, regardless of what treatment (if any) they received. (McCoy CE. West J Emerg Med. 2017 Oct;18(6):1075-1078).

errors due to deviations in these trials (i.e., finding a difference when there is no difference). In contrast, an intention-to-treat analysis will increase the risk of type I errors in non-inferiority analyses (Schumi 2011). There is no clear guidance as to which analysis is appropriate for determining non-inferiority. The Cochrane risk of bias tool endorses intention-to-treat analysis, the appropriate analysis when assessing superiority but does not consider the assessment of non-inferiority. The alternative to an intention-to-treat analysis is a per protocol analysis with adjustment for confounding if deviations are associated with prognostic factors. The consensus appears to be to ideally minimise protocol deviations and, if this is not achieved, to present both intention-to-treat and per protocol analyses adjusted for confounders if the deviations are associated with prognostic factors (Schumi 2011; Mo 2020; Piaggio 2012).

The Cochrane tool assesses the risk of bias of trials in which individuals are randomised. One of the ongoing trials is a cluster-randomised trial in which groups of individuals referred to as clusters, are randomised. These trials are subject to additional sources of bias, in particular recruitment bias. To assess these sources of bias we included additional signalling questions designed to address the specific considerations for cluster-randomised controlled trials described by Eldridge et al 2021.

Each source of bias and its signalling questions for the modified tool are shown in Table 13 in the Appendix. Assessments using the modified tools were considered provisional as they were made in the absence of some details and may change with the publication of results. This modified risk of bias tool has been independently reviewed by another methodologist (SY).

### 3.3.7 Statistical power calculations and statistical analysis plan

Statistical power calculations are important because they ensure that there are sufficient participants in the trial to detect statistically significant differences in superiority trials or unacceptable differences in non-inferiority trials between trial arms.

Power calculations indicate, with reference to the pre-specified analysis plans, whether the investigators running the trial have calculated the minimum number of patients needed for the study (i.e. in the intervention and comparator groups) to be able to detect differences between patient groups when they truly exist. Assessments of available power calculation details for ongoing risk-based were undertaken by 2 reviewers using a checklist comprised of items developed by Charles et al., 2009 (Charles 2009) and additional items pertaining to non-inferiority trials (Piaggio 2012). Examples of the type of item in the checklist include: ‘Was statistical power specified to detect a clinically important difference?’; ‘Were expected values for the control group specified?’; and ‘Is the expected size of difference in outcomes due to the intervention specified?’.

A statistical analysis plan describes the planned analysis/es for a clinical trial. Publicly available pre-specified statistical analysis plans are important because they reduce the risk of investigators using multiple analyses to obtain the most favourable results. To achieve this, the important details of the plan which take into account of the study’s aims and outcomes, need to be reported and be publicly available prior to the publication of results. The sources of the details of planned statistical analyses were assessed with planned analyses considered publicly available if they had been published or were available on the internet, and where interim results had been published, prior to the publication of interim results. Assessments of the reporting of the details of planned statistical analyses were undertaken by two reviewers using relevant items from a checklist designed for this purpose published by Gamble et al., 2017 (Gamble 2017).

### 3.3.8 Impact of COVID-19 pandemic

Information was sought in early March 2021 from principal investigators of trials started before 2020 as to whether the running of the trial had been impacted by the COVID-19 pandemic. Investigators were asked to outline what changes had occurred if their trial was impacted.

## 3.4 Results

### 3.4.1 Search results

Six relevant trials were identified, the details of which were available on the ClinicalTrials.gov registry (https://clinicaltrials.gov). The trial names are:

- My Personalized Breast Screening (MyPeBS)
- Women Informed to Screen Depending on Measures of Risk (WISDOM)
- Tailored Screening for Breast Cancer in Premenopausal Women (TBST)
- Breast Cancer Screening with MRI in Women Aged 50-75 Years with Extremely Dense Breast Tissue: the DENSE Trial
- Breast Screening – Risk Adaptive Imaging for Density (BRAID)
- What is the Best Interval to Screen Women 45-49 and 70-74 for Breast Cancer? (MISS)

A reference table for these trials is shown in Table 7, with additional detail shown in Table 8 (page 30) and Table 9 (page 31).

*Table 7. The six population level trials of risk-based screening included in the ROSA quality appraisal of international trials.*

| ***Acronym and age range*** | ***Location*** | ***Trial period*** | ***Risk groups*** | ***Comparator*** | ***Intervention*** | |
|---|---|---|---|---|---|---|
| | | | | | ***Intervals*** | ***Supplemental screening tests*** |
| **MyPeBS** (40-70) | France, Italy, UK, Belgium and Israel | 2019 - 2025 | BCSC/T-C scores (4 groups) | Various (Annual/biennial/ triennial screening, with mammography/DBT± supplemental US) | 1-4 years | US/ABUS, MRI |
| **WISDOM** (40-74) | USA | 2016 - 2020 | BCSC score (4 groups) | Annual mammography | 1-2 years None <50y | MRI |
| **TBST** (45-50) | Italy | 2013 - 2022 | BI-RADS 1-2 vs 3-4 | Annual mammography | 2 years for BI-RADS 1-2 | N/A |
| **DENSE** (50-75) | Netherlands | 2011 - 2019 | Extremely dense (Volpara D) | Biennial mammography | No change | MRI |
| **BRAID** (50-70) | UK | 2019 - 2026 | BI-RADS C-D | Triennial mammography | 18 months | Abbreviated MRI, ABUS, CEM |
| **MISS** (45-49) | Italy | 2020 - 2026 | BI-RADS A-C versus D. | Uncertain (most likely annual tomosynthesis) | 2 years for BI-RADS A-C | N/A |

ABUS = automated breast ultrasound; BCSC = Breast Cancer Surveillance Consortium; BI-RADS = Breast Imaging and Reporting Data; CEM = contrast-enhanced mammography; MRI = magnetic resonance imaging; NA = not applicable; T-C = Tyrer Cuzick; US = ultrasound

### 3.4.2 Available protocols and results

The ClinicalTrials.gov trial registry provides information about study design, population, outcomes, interventions and comparators, however registry records provide few if any details about the randomization process or the planned statistical methods. Therefore, a more detailed protocol was

sought for each of the trials. Published protocols in peer-reviewed journals were found for the DENSE, WISDOM and TBST trials; the MyPeBS trial protocol was found on a trial website and an unpublished protocol for the BRAID trial was obtained from clinical investigators in April 2020. Interim results had been published for one trial, the DENSE trial (Table 8). We were unable to obtain a more detailed protocol for the MISS trial. The primary contact person for the trial was asked for a copy of the study protocol in March and April 2021 but none was provided. A general search of the internet using the name of the trial and those of the three lead investigators did not identify a protocol for this study.

### 3.4.3 Trial aims

The primary aims of the identified ongoing trials can be broadly grouped according to whether they assess the non-inferiority of risk-based screening for all women, or the superiority of supplemental screening only for specific groups of higher risk women.

The MyPeBS, WISDOM, TBST and MISS trials assess whether risk-based screening in which screening is reduced for some women is non-inferior to current programs in which most women are screened in the same way. Specifically:

- Two trials, namely the MyPeBS and WISDOM trials, aim to determine if personalised screening based on a 5-year estimated risk of breast cancer is non-inferior to standard country-specific age-based screening practices with respect to the rates or proportion of stage 2B or more advanced cancers.
- Two trials, namely the TBST and MISS trials, aim to assess the impact, on the incidence of interval cancers (TBST) or more advanced cancers (TBST and MISS), of biennial rather than annual screening for premenopausal women with lower breast density (BI-RADS categories 1 and 2 in the TBST trial and BI-RADS categories A to C in the MISS trial).

The WISDOM trial also aims to determine whether biopsy rates are lower with personalised screening.

The DENSE and BRAID trials aim to assess whether supplemental screening within screening programs for women with denser breasts improves outcomes. Specifically:

- The DENSE trial aims to assess the effectiveness of offering MRI in addition to mammography to women with extremely dense breasts (>75% mammographic density). The investigators are comparing the interval cancer rate for biennial screening with and without supplemental MRI for those with extremely dense breasts.
- The BRAID trial is a 4-arm trial investigating whether breast cancer detection rates will improve when women with dense breasts (BI-RADS C or D) are offered supplemental imaging. Women in the control arm undergo mammographic screening every 3 years, and women in the three intervention arms undergo mammographic screening every 3 years and receive additional mammographic screening at 18 months and supplemental imaging at baseline and at 18 months using one of three imaging modalities, either abbreviated-MRI, automated whole breast ultrasound or contrast-enhanced spectral mammography. The trial will assess the effect of providing both supplemental screening and more frequent screening for women with dense breasts compared to standard triennial mammographic screening.

## 3.4.4 Impact of COVID-19 pandemic

Most of the trials are embedded in population-based screening programs (i.e. individuals are recruited into studies when they attend as part of a population-based screening program). Three studies (MyPeBS, TBST, BRAID) responded to the ROSA team's email requests for information regarding the impact of COVID-19 on trial progress. All three reported a major impact of COVID on their trials. Changes reported included interruption of recruitment, postponement of screening questionnaire completion, allowing some activities to be done remotely, and cancellation of some planned visits which would have been face-to-face and lengthy (i.e. one hour). Each of these trials recruit women from breast cancer population-based screening programs and as these programs paused their services at different times in 2020/2021, recruitment to the trials was paused accordingly. Details of the impact of COVID-19 on each trial, based on responses received, is presented in Table 10.

## 3.4.5 Quality appraisals

The quality appraisals of ongoing risk-based screening trials are summarised for trials with interim results in Table 11. Trials without interim results are summarised in Table 12. Further details including responses to each of the risk of bias assessment signalling questions, assessments of the reporting of each power calculation item and the planned statistical analyses item are shown in the Appendix (Table 13, Table 14 and Table 15).

### Quality appraisals based on protocols with interim results

There was one trial with interim results, the DENSE trial, a superiority trial reporting interval cancer rates.

#### *Risk of bias*

For the DENSE trial, the primary outcome, interval cancer incidence, was assessed and all five sources of bias assessed with the Cochrane risk of bias tool were rated as "low" risk (Table 11 and Table 13).

Importantly, the risk of bias due to deviations from the intended intervention was assessed as low. In this trial, consent was obtained only from those randomised to the intervention arm, in effect blinding those in the control group. As the intervention in this trial was likely to be perceived as advantageous, this strategy minimised the risk of contamination.

#### *Power calculations*

Power calculations were undertaken for the DENSE trial. This trial was powered to detect a clinically important decrease of 1.95 interval cancers per 1,000 women over a single screening interval which, based on the trialists' assumptions, equates to a 44% reduction from 4.4 to 2.45 interval cancers per 1,000 women over a single screening interval. The results of the assessment of the power calculations are shown in Table 11 and Table 14. The calculations were compliant with 12 (86%) of the 14 items applicable to a superiority trial with one primary outcome; the only relevant items not reported were adaptions for interim analyses and planned sensitivity analyses to examine the impact of assumptions and robustness of the estimated sample size. The study was powered to detect a difference for a single screening event rather than over three screens.

#### *Statistical analysis plans*

A statistical analysis plan for the DENSE trial primary outcome, interval cancer rates, was part of the protocol (Emaus 2015) published prior to the publication of the results (Bakker 2019). It was

reasonably comprehensive with many of the required features adequately described (Table 11and Table 15). The areas that were less comprehensively covered or not reported were;

- adjustment of the significance level due to interim analyses and guidelines for stopping the trial early,
- the measure of the treatment effect to be reported,
- the level of significance and confidence intervals to be used,
- the methods for controlling the Type I error potentially arising as a result of the outcome being measured at multiple time points, and
- methods for handling missing data.

## Quality appraisals based on protocols only

Interim results were not available for the MyPeBS, WISDOM, TBST, BRAID and MISS trials, and therefore we were only able to assess three of the five potential sources of bias; namely biases due to the randomisation process, due to deviations from the intended intervention, and due to measurement.

### *Risk of bias*

All sources of bias were assessed with the modified Cochrane risk of bias tool. The risk of bias was rated either “provisionally low”, “provisionally some concerns” or “insufficient information” for all three sources of bias in the five trials; the risk of bias was not rated at “provisionally high” for any of these sources of bias in any of the trials (Table 12). Further details including responses to each of the risk of bias assessment signalling questions are shown in the Appendix (Table 13).

For most studies we were unable to find sufficient details to provisionally assess the risk of bias due to the randomisation process. The exception was the MyPeBS trial which was rated at “provisionally low” risk of bias. The one cluster-RCT, the BRAID trial, was assessed at “provisionally low” risk of bias due to the cluster-RCT-specific source of bias, the timing of identification or recruitment of participants.

The risk of bias due to deviations from the intended interventions is managed by blinding and using a conservative analysis that minimises falsely finding superiority or non-inferiority (type I errors) due to deviations at the expense of falsely not finding superiority or non-inferiority (type 2 errors). None of the trials were blinded so all were potentially at risk of bias due to deviations from the intended interventions. Some trials did not attempt to manage this source of bias and were rated as “provisionally some concerns” for this source of bias, particularly the TBST which is planning an intention-to-treat analysis of non-inferiority. The MyPeBS trial is unblinded but a per protocol analysis of non-inferiority and an intention to treat sensitivity analysis is planned; adjustment for confounding associated with non-compliance which could reduce the risk of this source of bias, is suggested but was not explicitly stated. Only the WISDOM trial planned to directly address this source of bias. It did so by randomising only those who chose to be randomised to either group and offering women who preferred risk-based screening or current annual screening, thus screening according to their preference. There was insufficient information to assess the risk of bias due to deviations for this trial as the planned analysis populations (intention-to treat of per protocol) were not reported and thus their appropriateness could not be assessed.

As none of the trials described how the outcomes were to be assessed or whether the assessors would be blinded, the risk of measurement bias was rated according to whether the assessors could have been aware of the intervention received and whether their assessment could be influenced by knowledge of the intervention received. For cancer detection and biopsy rates, this was considered

possible and the risk of measurement bias was, therefore, rated “provisionally of some concern”. For interval cancer and cancer stage outcomes, this was considered highly unlikely and the risk of measurement bias was considered “provisionally low”. The exception to this was stage outcomes in trials where the intervention included a comprehensive risk assessment that identified women at high risk of breast cancer (WISDOM and MyPeBS) which could impact the assessment of cancer staging for these women. For these trials, in the absence of any information, it was considered that there was insufficient information to make a provisional assessment of bias due to the measurement of stage outcomes.

***Power calculations***

There was variation in the reporting of power calculations by trials. This is to be expected to some degree due to the different designs and interventions being employed. Overall, details of power calculations for at least one primary outcome were included in published or provided protocols for 4 of the 5 trials with unpublished results, the WISDOM, TBST, MyPeBS and BRAID trials.

We contacted the trial coordinator for the MISS trial to ascertain whether a protocol was publicly available, however, the response provided was that the protocol had not yet been published. We therefore only had the details presented as part of the NCT registration on the ClinicalTrials.gov website, which were insufficient to enable the assessment of power calculations for this trial.

Of the two trials that planned two primary outcomes, the WISDOM and TBST trials, only the WISDOM trial reported power calculations for both outcomes; hence the power calculations for five trial-outcome combinations were assessed. The checklist contained 17 items for assessment. The results of the assessments are shown in Table 12 and Table 14.

Compliance with checklist items ranged from 50% to 67% across the five trial-outcome combinations (excluding items that were non-applicable to a trial outcome). Items with 100% compliance were specification of hypothesis to be tested (superiority, non-inferiority or equivalence), non-inferiority margin for all non-inferiority trials, statistical power to detect a difference and assumptions for control group, and the appropriateness of calculations for endpoints of interest. None of the trial protocols reported adjusting for potential loss-to-follow-up and/or non-compliance in their calculations or the sources of information for assumptions for the effect of the intervention, and neither of the studies with planned interim analyses (BRAID and TBST) reported adapting the alpha level (an indication of potential type I error) for interim analyses.

Only the WISDOM trial provided all the information required to check the calculations and reported planned sensitivity analyses. Of the three trials assessing the non-inferiority of risk-based screening, only the TBST trial provided a rationale for the reported non-inferiority margin, the maximum difference in outcome between risk-based and standard screening programs at which a risk-based screening program would not be considered inferior to a standard screening program. All three non-inferiority studies reported high non-inferiority margins which raised concerns as to their clinical acceptability. In the absence of a rationale for the non-inferiority margins of increases in stage IIB or higher cancers of 53% per year in the WISDOM trial, and 25% over 4 years in the MyPeBS trial, these margins are unlikely to be clinically acceptable. They might be acceptable if based on a widely accepted benchmark which is appropriate for the population of interest. The TBST trial reported a non-inferiority margin of a 70% increase in interval cancers for women aged 45-50 years based on the rate of interval cancers considered acceptable by the European Community Guidelines for women aged 50-69 years. Whether this margin could be considered clinically acceptable will depend on whether the benchmark used is considered appropriate for this population.

***Planned statistical analysis***

At least some details of planned statistical analyses are publicly available for each of the trials however a publicly available comprehensive statistical analysis plan was found only for the MyPeBS trial. A statistical analysis plan for the BRAID trial is not publicly available but was provided by the trial investigators. Based on the details available, either published or provided, compliance with the checklist items applicable ranged from 67% to 74% for the two trials with statistical analysis plans and 32% to 47% for the other three trials (Table 12 and Table 15). All trials reported trial design, number of primary outcomes and the timing of final analysis and outcome assessments.

The areas that were less comprehensively covered or not reported were: details of the randomisation method (not provided for 3 trials); statistical methods for analysis (not provided for 2 trials); measure of treatment effect to be reported (not provided for 3 trials); the level of significance to be used (not reported for 4 trials); whether confidence intervals would be reported was missing for 4 trials; adjustment of the significance level due to interim analyses and guidelines for stopping the trial early (not provided for the 3 trials reporting interim analyses). For the three trials with outcomes measured at multiple timepoints, the method for controlling the Type I error was not described.

Other areas for which reporting was poor included testing the assumptions required for the proposed statistical methods, and methods for handling missing data.

The four non-inferiority trials were labelled as such, however the type of trial was not stated, but was implied, for the two superiority trials. Of the four trials undertaking non-inferiority analyses only one reported planned adjusted analyses.

## 3.5 Discussion

We found six randomised trials currently underway assessing the benefits and/or potential harms of risk-based breast cancer screening. Two assess the effect of supplemental screening for women with dense or extremely dense breasts on screening program outcomes, and four assess risk-based screening that includes reduced screening for some very low risk groups necessitating an assessment of non-inferiority.

The trials are collectively assessing a wide variety of interventions, including various screening technologies, screening intervals, age ranges and methods for assessing and categorising risk. Some of these differences may be due to the trial settings, but the variation between studies also highlights the complexity of the evidence related to risk-based breast screening, and the range of perspectives about how to best collect and interpret the available evidence.

The MyPeBS trial program is being implemented across a range of settings each with their own comparator (standard) screening protocols, providing an example of a single trial program that can enable multiple protocols to be assessed within a unified framework.

The ultimate aim of the interventions being trialled is to improve mortality rates and/or screening performance, or to reduce screening intensity for some without compromising mortality rates, however the assessment of the impact of screening changes on mortality rates requires lengthy follow-up that is not feasible. A diagnosis of more advanced disease is considered an acceptable surrogate for breast cancer mortality in the MyPEBS trial and “safety” of the program in the WISDOM trial and is a primary outcome for the four non-inferiority trials. Since interval cancer rates are cancers detected between screens, unlike the outcome of more advanced tumour stage at

diagnosis, these will be increased by additional imaging outside of screening programs and include early-stage disease, and thus are an imperfect surrogate for mortality in these circumstances.

### 3.5.1 Insights from our quality appraisal

Based on the information available, none of the trials were considered to have a high or provisionally high risk of bias for any of the sources of bias assessed and all except the MISS trial reported power calculations. Some remaining sources of bias may be reduced as additional information is produced by the trials.

Assessment of bias due to deviations from intended interventions highlighted the difficulties faced by investigators designing trials of risk-based breast cancer screening. Deviations from the intended interventions are a major potential source of bias for screening trials. In the absence of blinding there are several deviation scenarios that could distort results with the type of distortion dependent on the outcome. The potential impacts are particularly complex for screening performance outcomes and non-inferiority trials.

Non-inferiority assessments are highly susceptible to this source of bias, and when it occurs its management is more difficult as there is no clear single analytic approach that minimises the risk of incorrectly finding non-inferiority. At least three of the trials sought to manage this source of bias. The DENSE and WISDOM trials are designed to minimise deviations from the interventions and the MyPeBS trial may be planning analyses for non-inferiority that are considered most likely to minimise this bias, a per protocol analysis with compliance defined and possibly adjustment for potential confounders, and an intention to treat sensitivity analysis. For the five trials without any results, if the trial investigators monitor for, document and subsequently report at the end of the intervention period that there were minimal deviations as a result of the trial context in both groups, then these studies will likely be at low risk of bias due to deviations from their intended intervention.

None of the trials provided details as to how the primary outcomes were or will be ascertained. In the absence of any details, the risk of outcome measurement bias was dependent on assumptions as to whether the assessment of a specific outcome could be influenced by knowledge of the intervention received. If details become available as to how these outcomes were ascertained, for example whether those detecting cancer were blinded as to the intervention received or whether those staging cancers were likely to be influenced by knowledge of the intervention received, these ratings could change.

Power calculations were available for five of the six trials, with at least 50% compliance. The underlying assumptions were reported however the bases for these assumptions were less frequently reported. Assessment of the power calculations highlighted the additional challenges intrinsic to non-inferiority trials. To provide clinically meaningful evidence these trials need to specify a clinically acceptable non-inferiority margin – the maximum difference in outcome between risk-based and standard screening programs at which a risk-based screening program would not be considered inferior to standard screening program – and this margin needs to be based on clear justifications. Non-inferiority margins in all four trials were relatively high, raising concerns as to whether they would be clinically acceptable in the absence of any justification. A justification was provided in one instance, however, its applicability may be questionable. Publication of the full details of the rationale of the choice of the margin of non-inferiority, such as a clinically acceptable and applicable benchmark, may allay some or all of these concerns.

All trials provided some details of their planned statistical analyses but publicly available comprehensive statistical plans that reduce the risk of multiple analyses being used to obtain the

most favourable results were available for only the DENSE and MyPeBS trials. Trials with statistical analysis plans reported 65% and 74% of pre-specified planned statistical analyses items and remainder reported 32-47% of these items. Items reported routinely on trial registry records such as trial design and primary outcomes were available for all trials whereas none of the trials reported how missing data would be handled. The provision of publicly available comprehensive statistical analysis plans for the BRAID, WISDOM, MISS and TBST trials before the publication of their results will minimise the risk of selecting and analysis that provides the most favourable results.

In the absence of blinding there are several deviation scenarios potentially distorting results in ways that depend on trial outcomes. The potential impacts are particularly complex for screening performance outcomes and non-inferiority trials. Given this complexity, ideally trials should be designed to manage deviations, plan to monitor for deviations and adjust for any resulting confounding.

### 3.5.2 Limitations of the information available

Due to the nature of this review, the quality appraisal was primarily limited to published or provided protocols and trial registrations, and thus a full assessment of sources of bias was not possible for 5 of the 6 trials. Our appraisals identified a number of limitations as to the details available for sources of bias that could be assessed, power calculations and statistical analysis plans. There was often insufficient information to assess the potential bias due to randomisation. There were no details as to how outcomes would be measured and consequently assessment of the potential risk of outcome measurement was based on assumptions and knowledge of Australian clinical pathways, which may be quite different to clinical pathways in the countries where these trials are taking place. The absence of important details justifying power calculations and of the statistical analysis plan raise uncertainties about the clinical significance and analysis of prospective results. Provision of these details in statistical analysis plans would clarify these issues.

### 3.5.3 Limitations of these scoping level appraisals

This appraisal of the quality of the identified ongoing risk-based trials has limitations. Firstly, the risk bias assessments for five of the six trials are only provisional; they do not cover all sources of bias and the assessment of bias due the measurement of outcomes is based on several assumptions. As more information becomes available these sources of bias will need to be reassessed. Secondly, no attempt has been made to provide an overall quality score for each trial. The development of an overall quality score requires careful consideration of the relative importance of each assessed component prior to undertaking the final assessments. This was not considered feasible as only the risk of bias assessment tool provided an overall rating; the tools used to assess the reporting of planned statistical analyses and power calculations did not provide any ratings and attempting to develop an overall score would likely be very subjective. Furthermore, we assessed the completeness of reporting but not the quality of the planned statistical analyses as the trials are very different in terms of hypotheses, analyses required and outcomes, rendering comparisons of the quality of planned statistical methods inappropriate.

### 3.5.4 Implications for future trials of risk-based screening

The general principle of trials is to avoid deviations from the intended intervention through the trial design. Some trials assess only whether more intensive screening is superior to current practice.

For these superiority trials, an intention-to-treat analysis is considered most appropriate as it is a conservative approach that minimises falsely finding superiority (a type I error[8]).

Trials of risk-based breast screening protocols that include a screening component that is less intensive than the screening currently offered (non-inferiority trials) are required to show, for ethical reasons, that mortality outcomes (or surrogates thereof) for the intervention are not inferior to those for current screening practice. This is challenging in practice because non-inferiority methods and analyses are highly sensitive to protocol deviations. Even random deviations can impact findings from non-inferiority trial. For non-inferiority trials, there is no universal agreement about whether intention-to-treat or per protocol is the best approach. However, it is reasonable to minimise protocol deviations and if this is not achieved, to present both (i) per protocol analyses with monitoring of, and where necessary adjustment for the influence of deviations on the finding, and (ii) intention-to-treat analyses (such as planned by the MyPEBS trial).

Some principles that can help minimise bias in trials of risk-based breast cancer screening include the following:

1. Minimise deviations through the trial design e.g. as done by DENSE and WISDOM to reduce falsely finding the intervention superior or non-inferior (type I errors) and also falsely finding the intervention non-superior or unacceptably inferior (type II errors[9]). Monitor deviations from the intended intervention.
2. Accept that there will be residual deviations.
3. Plan the appropriate analysis; where:
    a. Superiority trials require intention-to-treat analyses to avoid falsely finding a difference where there is none.
    b. Non inferiority trials require both per protocol analyses adjusting for confounding as identified through monitoring deviations from the intended intervention and intention-to-treat analyses.
4. For non-inferiority trials, justify *a priori* the margins for non-inferiority (such as the benchmark for interval cancer rates as specified for the TBST trial).

These principles highlight that effective trial design for risk-based breast screening is a significant undertaking that requires expertise in trial design.

[8] A type I error involves falsely rejecting a null hypothesis that is actually true. For example, finding a difference between interventions on outcomes when there is no difference.
[9] A type II error involves failing to reject a null hypothesis that is actually false. For example, finding no difference between interventions on outcomes when there is, in truth, a difference.

# 3.6 Tables

*Table 8. Sources of the details of included trials.*

| ***Trial name (acronym) and ID*** | ***Registry details*** | ***Trial contact or lead investigator*** | ***Protocol identified (source or publication)*** | ***Start date*** | ***Planned completion date*** | ***Status*** | ***Publication of interim results (publication)*** | ***Planned statistical analysis (source)*** |
|---|---|---|---|---|---|---|---|---|
| Randomized, Comparison of Risk-Stratified versus Standard Breast Cancer Screening in European Women Aged 40-70 (**MyPEBS**)<br>NCT03672331 | https://clinicaltrials.gov/ct2/show/NCT03672331 | Dr S Delaloge | Yes (trial website) | 2019 | Dec 2025 | Recruiting | No | Yes<br>(protocol) |
| Women Informed to Screen Depending on Measures of Risk (Wisdom Study) (**WISDOM**)<br>NCT02620852 | https://clinicaltrials.gov/ct2/show/NCT02620852 | A Fiscalini | Yes (Esserman 2017) | 2016 | Dec 2020 | Recruiting | No | Yes<br>(Esserman 2017; Eklund 2018) |
| Tailored Screening for Breast Cancer in Premenopausal Women (**TBST**)<br>NCT02619123 | https://clinicaltrials.gov/ct2/show/NCT02619123 | Dr P Mantellini | Yes (Paci 2013) | 2013 | Jan 2022 | Recruiting | No | Yes<br>(Paci 2013) |
| What is the Best Interval to Screen Women 45-49 and 70-74 for Breast Cancer? (**MISS**)<br>NCT04590560 | https://clinicaltrials.gov/ct2/show/NCT04590560 | Dr F Falcini | No | 2020 | Feb 2026 | Recruiting | No | Yes (trial registration record) |
| Breast Cancer Screening with MRI in Women Aged 50-75 Years with Extremely Dense Breast Tissue: the DENSE Trial (**DENSE**)<br>NCT01315015 | https://clinicaltrials.gov/ct2/show/NCT01315015 | Dr C van Gils | Yes (Emaus 2015) | 2011 | Dec 2019 | Active not recruiting | Yes (Bakker 2019) | Yes<br>(Emaus 2015) |
| Breast Screening - Risk Adaptive Imaging for Density (**BRAID**)<br>NCT04097366 | https://clinicaltrials.gov/ct2/show/NCT04097366 | Dr F Gilbert | Yes (personal communication from lead investigator, Dr F Gilbert) | 2019 | Oct 2026 | Recruiting | No | Yes<br>(unpublished protocol) |

*Table 9. Overview of ongoing risk-based screening randomised controlled trials (RCTs).*

| ***Trial name and ID*** | ***Population*** | ***Intervention*** | ***Comparator*** | ***Outcomes*** |
|---|---|---|---|---|
| ***Trials comparing risk-based screening with standard non-risk-based screening*** | | | | |
| Randomized, Comparison of Risk-Stratified versus Standard Breast Cancer Screening In European Women Aged 40-70<br>(**MyPeBS**)<br><br>NCT03672331<br><br>France, Italy, UK, Belgium and Israel | Women aged 40-70 years affiliated to a social security or national healthcare system<br>With no prior DCIS or breast cancer, atypical breast lesion, lobular carcinoma in situ or chest wall irradiation or known or suspected very high-risk germline mutation | Personalised risk-based screening protocol for 4 years, according to estimated 5-year risk of breast cancer.<br>Risk determined using algorithm incorporating BCSC score and Tyrer-Cuzick score for women with more than one first degree relative with breast or ovarian cancer. Both scores will be modified to incorporate genotyping results and will be adjusted for country-specific breast cancer incidence.<br><br>Risk stratified screening protocols are as follows:<br>Low risk (<1% 5-year risk): Quadrennial mammogram for all women (i.e at study entry and end)<br>Average risk (1-<1.67% 5-year risk): Biennial mammogram for all women + ultrasound or ABUS for women with “high” breast density<br>High risk (1.67-<6% 5-year risk): Annual mammogram for all women + ultrasound or ABUS for women with “high” breast density<br>Very high risk (≥6% 5-year risk): Annual mammogram + MRI for all women<br><br>Supplemental tomosynthesis and/or ultrasound will be performed in this arm according to standard screening guidelines in each participating country (i.e. per comparator) | Mammogram with or without supplemental imaging according to guidelines in each participating country for 4 years:<br><br>Belgium (Brussels, Leuven): Biennial mammogram +/- tomosynthesis for women aged 50-69 years<br><br>Italy (4-6 regions): Biennial mammogram for all women aged 50-69 years, and up to 74 years in some regions. Annual mammogram for women aged 45-49 in some regions<br><br>UK (Cambridge, Manchester, Leeds): Triennial mammogram for women aged 50-73 years<br><br>Israel (national): Biennial mammogram for women aged 50-74 years +/- tomosynthesis +/- ultrasound per radiologist<br><br>France (national): Biennial mammogram for women aged 50-74 years + ultrasound in all women with dense breasts | ***Primary outcome***<br>*4 years follow-up (end of intervention)*<br>Stage 2 or higher breast cancer incidence **– non-inferiority**<br><br>***Secondary outcomes***<br>*4 years follow-up (end of intervention)*<br>Stage 2 or higher breast cancer incidence – superiority<br>False positive rate<br>Benign biopsy rate<br>Anxiety<br>Quality of life<br>Cost-effectiveness<br>Stage specific breast cancer and DCIS incidence<br>Overdiagnosis rate<br>Interval cancer rate<br><br>*10 years and 15 years follow-up*<br>Cumulative incidence of all breast cancer and stage 2 or higher breast cancer<br>Breast cancer-specific survival |

| ***Trial name and ID*** | ***Population*** | ***Intervention*** | ***Comparator*** | ***Outcomes*** |
|---|---|---|---|---|
| | | | *Guidelines for screening mammography at time of protocol writing* | |
| Women Informed to Screen Depending on Measures of Risk (**WISDOM**)<br><br>NCT02620852<br><br>USA | Women aged 40-74 years<br>With no prior DCIS or breast cancer | Personalised risk-based screening protocol for 5 years, according to estimated 5-year risk of breast cancer.<br>Risk determined using the BCSC model, genetic testing for rare high/moderate-penetrance mutations in nine genes and polygenic risk score for 96 lower-risk common genetic variants with known association to breast cancer.<br><br>Risk stratified screening protocols are as follows:<br>Lowest risk (aged 40-49 with <1.3% 5-year risk): No screening until age 50<br>Average risk (aged 50-74; or aged 40-49 with ≥1.3% 5-year risk): Biennial mammogram (if individual does not meet elevated or highest risk criteria)<br>Elevated risk (aged 40-49 with BI-RADS 4, or ≥0.75% 5-year risk of ER-breast cancer based on age and ethnicity; or women in top 2.5th percentile of risk by 1-year age category; or ATM, PALB2 or CHEK2 mutation carrier without a positive family history* of breast cancer): Annual mammogram (if individual does not meet highest risk criteria)<br>Highest risk (BRCA1/2, TP53, PTEN, STK11, CDH1 mutation carrier; or ATM, PALB2, or CHEK2 mutation carrier with positive family history of breast cancer; or ≥ 6% 5-year risk; or had mantle radiation when aged 10-30): Annual mammogram + MRI<br><br>*Family history: first degree relative with breast cancer, two second-degree relatives with breast cancer, or one second-degree relative diagnosed prior to age 45 | Annual mammogram | ***Primary outcome***<br>*5 years follow-up*<br>Proportion of cancers stage IIB or higher **– non-inferiority**<br>Biopsy rate<br><br>***Secondary outcomes***<br>*5 years follow-up*<br>Stage IIB or higher breast cancer rate<br>Interval cancer rate<br>Systemic therapy rate<br>Mammogram recall rate<br>Breast biopsy rate<br>DCIS rate<br>Chemoprevention uptake rate<br>Participant preference – risk-based vs annual screening (in self-assigned cohort)<br>Participant adherence to assigned screening schedule<br>Breast cancer anxiety (PROMIS anxiety scale)<br>Decisional regret (Decision Regret Scale)<br>Ultra-low risk cancer rate |
| Tailored Screening for Breast Cancer in Premenopausal Women (**TBST**)<br><br>NCT02619123 | Premenopausal women aged 44-45 years resident in screening centre catchment area invited to attend for mammographic screening<br>With no prior DCIS or breast cancer, family not at high risk for breast cancer and no diagnosis of other cancer in last 5 years | Risk-based screening for women aged 45-50 years according to breast density (BI-RADS classification).<br><br>Risk stratified screening protocols are as follows:<br>Low risk (low breast density; BI-RADS 1-2 on baseline mammogram): Biennial mammogram until aged > 50 years | Annual invitation to mammography for women aged 45-49 years<br>After the age of 50 years, all women will continue to be screened in the usual service screening programme | By arm and breast density group:<br>***Primary outcomes***<br>*3 years and 6 years follow-up*<br>Cumulative incidence of interval cancer **– non-inferiority** |

| ***Trial name and ID*** | ***Population*** | ***Intervention*** | ***Comparator*** | ***Outcomes*** |
|---|---|---|---|---|
| Italy | | High risk (high breast density; BI-RADS 3-4 on baseline mammogram): Annual mammogram<br><br>After the age of 50 years, all women will continue to be screened in the usual service screening programme | (In Italy biennial mammogram for all women aged 50-69 years, and up to 74 years in some regions. Annual mammogram for women aged 45-49 in some regions) | Cumulative incidence of T2+/node- positive breast cancer **– non-inferiority**<br><br>***Secondary outcomes***<br>*3 years and 6 years follow-up*<br>False positive rates<br>Cumulative incidence of breast cancer<br>*1, 2, 3 ,4, 5 years and 6 years follow-up*<br>Mammography screening attendance |
| What is the Best Interval to Screen Women 45-49 and 70-74** for Breast Cancer?<br>(**MISS**)<br><br>NCT04590560<br><br>Italy | Women aged 45-49 years resident in four locations in Italy<br><br>With no prior DCIS or breast cancer, no familial risk for breast cancer and no concurrent participation in another clinical trial on breast cancer screening | Biennial tomosynthesis<br><br>OR<br><br>Risk-based screening for women aged 45-49 years according to breast density (BI-RADS classification):<br>Low risk (breast density; BI-RADS category A-C): Biennial tomosynthesis until aged 50 years<br>High risk (breast density; BI-RADS category D): Annual tomosynthesis | Unclear - Annual tomosynthesis? (aim is to compare screening intervals not screening modalities) | **Primary outcome**<br>*6 years follow-up*<br>Cumulative incidence of cancers stage II or higher **– non-inferiority**<br><br>**Secondary outcomes**<br>*6 years follow-up*<br>Participation rate within 3 months of invitation<br>Proportion of women allocated biennial screen who have a screen performed prior to next 2-year screen<br>Breast cancer detection rate<br>Overall recall rate<br>Recall rate involving an invasive procedure<br>Interval breast cancer rate<br>Cumulative breast cancer incidence<br>Resource expenditure<br>Prevalence of dense breast in the target population |
| ***Trials comparing different or additional screening modalities with standard screening for higher risk groups*** | | | | |

| ***Trial name and ID*** | ***Population*** | ***Intervention*** | ***Comparator*** | ***Outcomes*** |
|---|---|---|---|---|
| Breast Cancer Screening With MRI in Women Aged 50-75 Years With Extremely Dense Breast Tissue<br>(**DENSE**)<br><br>NCT01315015<br><br>Netherlands | Asymptomatic women aged 50-75 years participating in population-based screening program<br>With **extremely dense breasts** (Volpara grade 4/D) and a negative mammogram | Biennial MRI + mammogram for 4 years (3 screening rounds) | Biennial mammogram for 4 years (3 screening rounds) | ***Primary outcome***<br>*6 years follow-up*<br>Incidence of interval cancer<br><br>***Secondary outcomes***<br>*6 years follow-up*<br>Tumour size, stage, grade, histology and molecular subtype<br>Mortality rate (MISCAN program)<br>Cost-effectiveness (MISCAN program)<br>Quality of life (MRI group)<br><br>*4 years follow-up*<br>MRI screen-detected cancer<br>MRI referral rate<br>PPV (MRI group)<br>Number of biopsies per MRI referral |

| ***Trial name and ID*** | ***Population*** | ***Intervention*** | ***Comparator*** | ***Outcomes*** |
|---|---|---|---|---|
| Breast Screening - Risk Adaptive Imaging for Density (**BRAID**)<br><br>NCT04097366<br><br>**Cluster**-RCT<br><br>UK | Women aged 50-70 years undergoing triennial population-based screening (NHS-BSP)<br>With **dense breasts** (BI-RADS C with high chance of masking or D) on baseline (current) mammogram (negative or positive)<br>With no known BRCA mutation or < 50% risk of being a BRCA carrier | Mammogram + abbreviated-MRI at baseline and 18 months; mammogram at 3 years<br>or<br>Mammogram + ABUS at baseline and 18 months; mammogram at 3 years<br>or<br>Mammogram + contrast-enhanced spectral mammogram at baseline; contrast-enhanced spectral mammogram only at 18 months; mammogram at 3 years | Triennial mammogram | ***Primary outcome***<br>*3 years follow-up*<br>Cancer detection rates<br><br>***Secondary outcomes***<br>*3.5 years follow-up*<br>Stage II or higher cancer incidence<br>Cancer detection rate<br>Interval cancer rate<br>Recall rate<br>Sensitivity of supplemental imaging<br>Specificity of supplemental imaging<br><br>*0.5 year and 1.75 years follow-up*<br>Cancer detection rate<br>Recall rate<br>Sensitivity of supplemental imaging<br>Specificity of supplemental imaging<br><br>*1 year follow-up*<br>Cost-effectiveness of each modality |

ABUS = automated breast ultrasound; BCSC = Breast Cancer Surveillance Consortium; BI-RADS = Breast Imaging and Reporting Data; DCIS = ductal carcinoma in-situ; HRT = hormone replacement therapy; MRI = magnetic resonance imaging; NA = not applicable; NHS-BSP = National Health Service – Breast Screen Programme

** details for women aged 70-74 not included because this part of the study is not an RCT

*Table 10. Reported impact of COVID-19 on trials in response to requests by the ROSA team in March 2021 for information.*

| ***Trial name and ID*** | ***Reported Impacts*** |
|---|---|
| Randomized, Comparison of Risk-Stratified versus Standard Breast Cancer Screening i-n European Women Aged 40-70 (**MyPeBS**) | Major impact on trial. Accrual interrupted for 4 months. Have had to adapt the trial structure, postpone some of the questionnaires, allow some activities to be done remotely etc. The initial inclusion visit is meant to take 1 hour but this is not happening due to COVID-19. |
| Women Informed to Screen Depending on Measures of Risk (**WISDOM**) | No response to our queries |
| Tailored Screening for Breast Cancer in Premenopausal Women (**TBST**) | The pandemic affected breast cancer screening heavily during March-April-May 2020. However, participation rate in the trial does not seem to be affected. |
| What is the Best Interval to Screen Women 45 to 49 and 70 to 74 for Breast Cancer? (**MISS**) | No response to our queries |
| Breast Cancer Screening With MRI in Women Aged 50 to 75 Years With Extremely Dense Breast Tissue (**DENSE**) | No response to our queries |
| Breast Screening - Risk Adaptive Imaging for Density (**BRAID**) | The trial stopped at the end of March 2020 as they paused the breast screening program from where they recruit women for the trial. The trial restarted in mid-July when breast screening recommenced. Screening is reduced to 60% so that is affecting recruitment to the trial. Screening and the trial paused again in January 2021 and restarted in mid-February 2021. |

MRI = magnetic resonance imaging

*Table 11. Quality appraisals of ongoing risk-based screening trials **with interim results.***

| Trial | Outcome assessed | Risk of bias due to | | | | | Power calculations | | Details of planned statistical analysis | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Randomisation process | Deviations from the intended interventions | Missing outcome data | Outcome measurement | Selection of the reported result | Details available | % required detail reported | Publicly available | % required detail reported |
| **DENSE** | Interval cancers | Low | Low | Low | Low | Low | Yes | 86 | Yes | 65 |

*Table 12. Quality appraisals of ongoing risk-based screening trials with no interim results – based on protocols only.*

| Trial | Outcome/s assessed | Risk of bias due to | | | | Power calculations | | Details of planned statistical analysis | |
|---|---|---|---|---|---|---|---|---|---|
| | | Randomisation process | Timing of identification or recruitment of participants in a cluster-RCT | Deviations from the intended interventions | Outcome measurement | Details available | % required detail reported | Publicly available | % required detail reported |
| **BRAID** | Cancer detection rate | Insufficient information | Provisionally low | Provisionally some concerns | Provisionally some concerns | Yes | 50 | Some details | 67 |
| **WISDOM** | Stage IIB or higher<br><br>Biopsy rate | Insufficient information to assess | Not applicable | Insufficient information | Insufficient information<br><br>Provisionally some concerns | Yes | 50<br><br>57 | Some details | 32<br><br>32 |
| **MyPeBS** | Stage II or higher | Provisionally low | Not applicable | Provisionally some concerns | Insufficient information | Yes | 67 | Yes | 74 |
| **TBST** | Interval cancer<br>T2+ and/or node-positive cancers | Insufficient information | Not applicable | Provisionally some concerns | Provisionally low<br><br>Provisionally low | Yes | 50 | Some details | 38 |
| **MISS** | Stage II or higher | Insufficient information | Not applicable | Insufficient information | Provisionally low | No | Not applicable | Some details | 47 |

# 3.7 References

1. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. New England Journal of Medicine. 2019;381(22):2091-102.

2. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. BMJ 2009;338:b1732.

3. Eklund M, Broglio K, Yau C, Connor JT, Stover Fiscalini A, Esserman LJ. The WISDOM Personalized Breast Cancer Screening Trial: Simulation Study to Assess Potential Bias and Analytic Approaches. JNCI Cancer Spectrum. 2019;2(4):pky067.

4. Eldridge S, Campbell MK, Campbell MJ, Drahota AK, Giraudeau B, Reeves BC et al. Revised Cochrane risk of bias tool for randomsied trials (RoB 2). Additional considerations for cluster-randomized trials (RoB 2 CRT). Available at https://drive.google.com/file/d/1yDQtDkrp68_8kJilUdbongK99sx7RFI-/view (accessed February 2022).

5. Emaus MJ, Bakker MF, Peeters PH, Loo CE, Mann RM, de Jong MD, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. Radiology. 2015;277(2):527-37.

6. Esserman LJ, Anton-Culver H, Borowsky A, Brain S, Cink T, Crawford B, et al. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. NPJ Breast Cancer. 2017;3(1):34.

7. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017;318(23):2337-2343.

8. Higgins J, Altman D. Chapter 8: Assessing risk of bias in included studies. In: Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011) ed 2011.

9. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.0: Cochrane; 2019 (updated July 2019). Available from: www.training.cochrane.org/handbook.

10. Mo Y, Lim C, Watson JA, White NJ, Cooper BS. Non-adherence in non-inferiority trials: pitfalls and recommendations. BMJ 2020;360:m2215.

11. MyPeBS. Protocol number: UC-0109/18052017 Version number 1.2 27 July 2018 Available from: https://cancer-rose.fr/my-pebs/wp-content/uploads/2019/12/MyPEBS-Protocol-.V1.2-du-27.07.18-.pdf (accessed February 2022).

12. NHMRC. Guidelines for Guidelines: Assessing risk of bias. 2019 [updated 29 August 2019. Available from: https://nhmrc.gov.au/guidelinesforguidelines/develop/assessing-risk-bias.

13. Paci E, Mantellini P, Giorgi Rossi P, Falini P, Puliti D. [Tailored Breast Screening Trial (TBST)]. Epidemiologia e Prevenzione. 2013;37(4-5):317-27.

14. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, for the CONSORT Group. Reporting of noninferiority and equivalence randomized trials: Extension of the CONSORT 2010 statement. JAMA 2012;308(24):2594-2604.

15. Schumi J & Wittes JT. Through the looking glass: understanding non-inferiority. Trials 2011;12:106.

# 3.8 Appendix

*Table 13. Summary of risk of bias responses and assessments, power calculations availability and public availability planned statistical analyses for the ongoing clinical trials.*

| ***Bias domain and signaling question*** | ***DENSE*** | ***BRAID*** | ***WISDOM*** | ***MyPeBS*** | ***TBST*** | ***MISS*** |
|---|---|---|---|---|---|---|
| **Study design** | Parallel | Cluster | Parallel | Parallel | Parallel | Parallel |
| **Outcomes being assessed** | Interval cancers | Cancer detection rate | Stage IIB or higher Biopsy rate | Stage II or higher | Interval cancers T2+ and/or node-positive | Stage II or higher |
| **Framework** | Superiority | Superiority | Non-inferiority for stage IIB or higher | Non-inferiority | Non-inferiority | Non-inferiority |
| **1a. Bias arising from the randomisation process** | | | | | | |
| 1a.1 Was the allocation sequence random? | Y | NI | Y | PY | NI | NI |
| 1a.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions? | PY | NI | NI | PY | PY | NI |
| 1a.3 Did baseline differences between intervention groups suggest a problem with the randomisation process? | N | Not assessable - Can only be assessed if results are available | | | | |
| **Risk-of-bias judgment** | **Low** | **Insufficient information** | **Insufficient information** | **Provisionally low** | **Insufficient information** | **Insufficient information** |
| **1b. Bias arising from the timing of identification or recruitment of participants in a cluster-randomised trial** | | | | | | |
| 1b.1 Were all the individual participants identified and recruited (if appropriate) before randomization of clusters? | **NA** | **N** | **NA** | **NA** | **NA** | **NA** |
| 1b.2 If N/PN/NI to 1b.1: Is it likely that selection of individual participants was affected by knowledge of the intervention assigned to the cluster? | | **PN** | | | | |
| 1b.3 Were there baseline imbalances that suggest differential identification or recruitment of individual participants between intervention groups? | | Not assessable - Can only be assessed if results are available | | | | |
| **Risk-of-bias judgment** | | **Provisionally low** | | | | |
| **2. Bias due to deviations from intended interventions** | | | | | | |
| 2.1 Were participants aware of their assigned intervention during the trial? | Y | Y | Y | PY | Y | Y |
| *2.1.1 Are deviations likely from the assigned intervention in the intervention group because the participants know they are in a trial?* | Applicable only to trials without results | PN | PN | PY | PY | PY |

| ***Bias domain and signaling question*** | ***DENSE*** | ***BRAID*** | ***WISDOM*** | ***MyPeBS*** | ***TBST*** | ***MISS*** |
|---|---|---|---|---|---|---|
| *2.1.2 Are deviations likely from the assigned intervention in the control group because the participants know they are in a trial?* | | PY | PN | PY | PN | PN |
| 2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial? | PY | PY | PY | PY | PY | PY |
| 2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context? | PN | Not assessable - Can only be assessed if results are available | | | | |
| 2.4 If Y/PY/NI to 2.3: Were these deviations likely to have affected the outcome? | NA | | | | | |
| 2.5 If Y/PY to 2.4: Were these deviations from intended intervention balanced between groups? | NA | | | | | |
| 2.6 Was an appropriate analysis used/ *is an appropriate analysis planned* to estimate the effect of assignment to intervention? | Y | Y | NI | PY | N – interval cancers<br>NI -T2+/N-positive cancers | NI |
| 2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants appropriately ? | NA | NA | NI | NA | PY – interval cancers<br>NI -T2+/N-positive cancers | NI |
| **Risk-of-bias judgment** | **Low** | **Provisionally some concerns** | **Insufficient information** | **Provisionally some concerns** | **Provisionally some concerns** | **Insufficient information** |
| **3. Bias due to missing outcome data** | | | | | | |
| 3.1 Were data for this outcome available for all, or nearly all, participants randomised? | Y | Not assessable - Can only be assessed if results are available | | | | |
| 3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data? | NA | | | | | |
| 3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value? | NA | | | | | |
| 3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value? | NA | | | | | |
| Risk-of-bias judgment | **Low** | | | | | |
| **4. Bias in measurement of the outcome** | | | | | | |
| **Primary outcome (1)** | ***Interval cancers*** | ***Cancer detection rate*** | ***Stage IIB or higher*** | ***Stage II or higher*** | ***Interval cancers*** | ***Stage II or higher*** |
| 4.1 Was the method/ *is the planned method* of measuring the outcome inappropriate? | N | PN | NI | NI | PN | NI |
| 4.2 Could measurement or ascertainment of the outcome have differed/ *differ between intervention groups?* | PN | PN | NI | NI | PN | PN |

| ***Bias domain and signaling question*** | ***DENSE*** | ***BRAID*** | ***WISDOM*** | ***MyPeBS*** | ***TBST*** | ***MISS*** |
|---|---|---|---|---|---|---|
| 4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors/ *would outcome assessors* be aware of the intervention received by study participants? | PY | PY | PY | PY | PY | PY |
| 4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been/ *be* influenced by knowledge of the intervention received? | PN | PY | NI | NI | PN | PN |
| 4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was/ *will be* influenced by knowledge of the intervention received? | NA | PN | PN | PN | NA | NA |
| **Risk-of-bias judgment** | **Low** | **Provisionally some concerns** | **Insufficient information** | **Insufficient information** | **Provisionally low** | **Provisionally low** |
| **Primary outcome (2)** | | | ***Biopsy rate*** | | ***T2+ and/or node-positive cancers*** | |
| 4.1 Was the method/ *is the planned method* of measuring the outcome inappropriate? | | | NI | | NI | |
| 4.2 Could measurement or ascertainment of the outcome have differed/ *differ* between intervention groups? | | | NI | | PN | |
| 4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors/ *would outcome assessors be* aware of the intervention received by study participants? | | | PY | | PY | |
| 4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been/ *be* influenced by knowledge of the intervention received? | | | PY | | PN | |
| 4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was/ will be influenced by knowledge of the intervention received? | | | NI | | NA | |
| **Risk-of-bias judgment** | | | **Provisionally some concerns** | | **Provisionally low** | |
| **5. Bias in selection of the reported result** | | | | | | |
| 5.1 Were the data that produced the result analysed in accordance with a prespecified analysis plan that was finalized before unblinded outcome data were available for analysis? | PY | Not assessable - Can only be assessed if results are available | | | | |
| 5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements within the outcome domain? | N | | | | | |
| 5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data? | N | | | | | |
| **Risk-of-bias judgment** | **Low** | | | | | |

| *Bias domain and signaling question* | *DENSE* | *BRAID* | *WISDOM* | *MyPeBS* | *TBST* | *MISS* |
|---|---|---|---|---|---|---|
| **6. Extra study reporting quality questions** | | | | | | |
| E1. Were power calculations reported for the primary outcome? | Y | Y | Y for both primary outcomes | Y | Y for interval cancers<br>N for T2+/N-positive cancers | Not found |
| E2. Is a statistical analysis plan publicly available (including on website or trial registry) | Y - part of published protocol | Some details included in trial registration record | Some but not all details published | Y - part of publicly available protocol | Some details published | Some details included in trial registration record |

TBST = Tailored Screening for Breast Cancer in Premenopausal Women; WISDOM = Women Informed to Screen Depending on Measures of Risk; MyPeBS = My Personalized Breast Screening; BRAID = Breast Screening – Risk Adaptive Imaging for Density; DENSE = Breast Cancer Screening With MRI in Women Aged 50-75 years With Extremely Dense Breast Tissue: the DENSE Trial; MISS = What is the Best Interval to Screen Women 45-49 and 70-74 for Breast Cancer Y = yes; PY = probably yes; PN = probably no; N = no; NA = not applicable; NI = no information.

*Table 14. Summary of assessments for each of the pre-specified power calculation items for the ongoing clinical trials.*

| *Parameter* | *Description* | *DENSE* | *BRAID* | *WISDOM* | *WISDOM 2* | *MyPEBS* | *TBST* |
|---|---|---|---|---|---|---|---|
| Primary outcome/s | | Interval cancer rate at 6 years follow-up | Cancer detection rate at 3-year follow-up | Proportion of cancers diagnosed at stage IIB or higher at 5-year follow-up | Biopsy rate at 5-year follow-up | Incidence of stage 2 or higher cancer at 4-year follow-up | Cumulative incidence of interval cancers and of T2+/node-positive cancers at 3- and 6-year follow-up |
| Calculations for all primary outcomes listed in the protocol | If more than one outcome is specified, was sample size/ power or detectable difference estimated for each? | NA - single primary outcome | NA - single primary outcome | Yes | Yes | NA - single primary outcome | No - only calculated for interval cancers not T2+/node-positive cancers |
| Type of trial | Is the aim to demonstrate intervention superiority, non-inferiority or equivalence? | Superiority | Superiority | Non-inferiority | Superiority | Non-inferiority | Non-inferiority |
| If non-inferiority trial: | | | | | | | |
| *Margin of non-inferiority specified?* | If non-inferiority trial, is the margin of non-inferiority provided (i.e. Smallest detectable difference that should not be missed)? | NA | NA | Yes - 50 stage IIB or higher cancers per 100,000 women in one year | NA | Yes - 120 stage IIB or higher cancers per 100,000 women over 4 years | Yes - 21 interval cancers per 10,000 women over 6 years |
| *Rationale for choice?* | Is the rationale for the choice of the margin of non-inferiority provided? | NA | NA | No | NA | No | Yes - based on level of interval cancers in women 50-69 years considered acceptable by the European Community Guidelines |

| ***Parameter*** | ***Description*** | ***DENSE*** | ***BRAID*** | ***WISDOM*** | ***WISDOM 2*** | ***MyPEBS*** | ***TBST*** |
|---|---|---|---|---|---|---|---|
| Appropriateness of calculation | Does the method for calculating sample size/power or detectable difference match the outcomes/endpoints of interest | Yes | Yes | Yes | Yes | Yes | Yes |
| Type I error (alpha) | Alpha level specified? | 0.05 | No | No | No | 0.025 | 0.05 |
| *One- or two-tailed?* | One- or two-tailed test specified? | One-tailed | No | No | No | One-tailed | No |
| *Adapted for interim analyses or multiple comparisons?* | If interim analyses or multiple comparisons are planned, was this allowed for? | No - powered to detect difference after a single screen – two interim analyses planned with primary outcome 2 years after third screen | No | NA – no interim analyses reported | NA – no interim analyses reported | NA – no interim analyses reported | No |
| Power (1-beta) | Was statistical power specified to detect a clinically important difference? | 80% | 87% | 90% | 90% | 80% | 90% |
| Assumptions for control group | Expected values (e.g. proportion, mean) for the control group specified? | Yes | Yes | Yes | Yes | Yes | Yes |
| *Source of information* | Are the values based on results from a previous trial, pilot study, observational data, results of systematic review or other source? | Yes | Yes | No | No | Yes | Yes |
| Assumptions for intervention effect | Is the expected size of difference in outcomes due to the intervention specified? | Yes | Yes | Yes | Yes | Yes | No |
| *Source of information* | Is the assumed intervention effect based on that for another similar intervention or trial, observational data or results of meta-analysis, or other source? | Yes | No | No | No | No | NA |

| ***Parameter*** | ***Description*** | ***DENSE*** | ***BRAID*** | ***WISDOM*** | ***WISDOM 2*** | ***MyPEBS*** | ***TBST*** |
|---|---|---|---|---|---|---|---|
| *Magnitude of effect to be detected* | Is the magnitude of the treatment effect realistic and large enough to be considered to be clinically important (superiority) /the non-inferiority margin small enough to be clinically acceptable (non-inferiority)? | Yes - powered to detect a decrease of 1.95 interval cancers per 1,000 women over a single screening interval which based on the trialists' assumption equates to a 44% reduction from 4.4 to 2.45 interval cancers per 1,000 women over a single screening interval | Yes - powered to detect an increase of 10 cancers detected per 1,000 women which based on triallists' assumptions equates to a 100% increase from 10 to 20 cancers per 1,000 women if there is a focus on avoiding overtreatment | No - a non-inferiority margin of an increase of 50 stage IIB or higher cancers per 100,000 women **in one year** which based on the trialists' assumptions equates to a **53% increase** from 95 to 145 stage IIB or higher cancers per 100,000 women per year would be of concern | Yes-powered to detect a decrease of 1.1 biopsies per 100 women which based on trialists' assumptions equates to a 5% decrease from 22 to 20.9 biopsies per 100 women (time frame not reported - assume over a 5-year period) is of modest clinical importance | No - a non-inferiority margin of an increase of 120 stage IIB or higher cancers per 100,000 women over 4 years which based on trialists' assumptions equates to a **25% increase** from 480 to 600 stage IIB or higher cancers per 100,000 women **over 4 years** would be of concern | No - a non-inferiority margin of an increase of 21 interval cancers per 10,000 women over 6 years which based on the trialists' assumptions equates to a **70% increase** from 30 to 51 interval cancers per 10,000 women **over 6 years** would be of concern - based on level of interval cancers in women 50-69 years considered acceptable by the European Community Guidelines |
| Allowance for loss-to-follow-up and/or non-compliance? | Is the required sample size/expected power/detectable difference adjusted for potential loss to follow-up/non-compliance? | Yes | No | No | No | Yes | No |
| Completeness of reporting | Is all information provided to allow replication/ checking of calculations? | Yes | No | No | Yes | No | No |
| Sensitivity analysis | Were sensitivity analyses done or planned to examine the impact of assumptions and robustness of the estimated sample size/power or detectable difference? | No | No also do not consider clustering of observations where all participants attending a clinic are randomised to the same arm | Yes calculate power of 83% for a lower non-inferiority margin of 0.035% ie increase of 35 stage IIB or higher cancers per 100,000 women in one year which equates to a **37% increase** from 95 to 130 stage IIB or higher cancers per 100,000 women per year | No | No | No |

NA = not applicable; NR = not reported

*Table 15. Summary of assessments for each of the pre-specified planned statistical analyses items for the ongoing clinical trials.*

| *Item* | *Description* | *DENSE* | *BRAID* | *WISDOM* | *MyPEBS* | *TBST* | *MISS* |
|---|---|---|---|---|---|---|---|
| **Resources available** | | Emaus 2015 and NCT registration | Trial protocol provided by investigators and NCT registration | Esserman 2017, Eklund 2018 and NCT registration | Publicly available trial protocol and NCT registration | Paci 2013 and NCT registration | NCT registration |
| **Trial design** | Brief description of trial design including type of trial) and allocation ratio | Parallel group RCT Allocation ratio 1 (intervention):4 (control) | Parallel group cluster RCT Allocation ratio 1:1:1:1 | Parallel group RCT Allocation ratio 1:1 | Parallel group RCT Allocation ratio 1:1 | Parallel group RCT Allocation ratio 1:1 | Parallel group RCT Allocation ratio 1:1:1 |
| **Randomisation** | Randomization details, e.g., whether any minimization or stratification occurred | Computer-generated random schedule in permuted blocks of random block size stratified by hospital and regional screening organisation | Randomisation by whole screening clinic stratified by study centre | NR | Randomization using permuted block lists with a random block size stratified by country, age and prior mammogram. | NR | NR |
| **Number of primary outcomes** | | 1 | 1 | 2 | 1 | 2 | 1 |
| **Sample size calculation** | Full sample size calculation or reference to sample size calculation in protocol | Reported for primary outcome | Reported for primary outcome | Reported for both primary outcomes | Reported for primary outcome | Reported for one of two primary outcomes | NR |
| **Framework** | Superiority, equivalence, or noninferiority hypothesis testing framework | Implied superiority | Implied superiority | Non-inferiority for stage IIB+ and implied superiority for biopsy rate | Non-inferiority.<br>If found to be non-inferior will then test superiority hypothesis as secondary outcome | Non-inferiority for both primary outcomes (interval cancer and T2+/node -positive cancer) | Non-inferiority |
| **Statistical interim analysis and stopping guidance** | Information on **interim** analyses including time points | Two interim analyses 2 years after first and second screening rounds ie at 2 years and 4 years follow-up | Interim analysis after first round of supplemental imaging completed | NR | No planned interim analyses | Interim analysis at 3 years average follow-up | No planned interim analyses |
| | Any planned adjustment of the significance level due to interim analysis | NR | NR | Not applicable | Not applicable | NR | Not applicable |
| | Details of guidelines for stopping the trial early | NR | NR | Not applicable | Not applicable | NR | Not applicable |
| **Timing of final analysis** | Timing of final analysis, e.g., all outcomes analysed collectively or timing stratified by planned length of follow-up | Assume final analyses at 6 years follow-up.<br>No planned longer-term outcomes or analyses reported | Analysis of primary outcome after 3 years follow-up with risk analyses at 6 years follow-up | Analysis at 5 years follow-up.<br>No planned longer-term outcomes or analyses reported | Analysis of primary outcome after 4 years follow-up with later analyses at 10 years and 15 years follow-up | Analysis at 6 years average follow-up.<br>No planned longer-term outcomes or analyses reported | Analysis at 6 years follow-up.<br>No planned longer-term outcomes or analyses reported |

| *Item* | *Description* | *DENSE* | *BRAID* | *WISDOM* | *MyPEBS* | *TBST* | *MISS* |
|---|---|---|---|---|---|---|---|
| **Timing of outcome assessments** | Time points at which the outcomes are measured including visit "windows" | Primary outcome will be measured 2 years after each of 3 screening rounds. | After initial screen and subsequent program screen with 6-month windows | At 5 years only | At 4, 10 and 15 years | Outcomes will be measured after each screen | At 6 years only |
| **Confidence intervals and *P* values** | Level of statistical significance - alpha | NR | NR | NR | One-sided alpha of 0.025 | NR | NR |
| | Description and rationale for any adjustment for multiplicity and, if so, detailing how the type 1 error is to be controlled | NR | Network meta-analysis methods to take account of the multiple treatments | NR | NR | NR | NR |
| | Confidence intervals to be reported | NR | NR | NR | NR | NR | 95% |
| **Analysis populations** | Definition of analysis populations, e.g., intention to treat, per protocol, complete case, safety | Intention to treat | Intention to treat | NR | Per protocol for non-inferiority (compliance defined) with intention to treat as sensitivity analysis | Intention to treat for interval cancers. NR for T2+/N- positive cancers | NR |
| **Analysis methods** | What analysis method will be used | Chi squared test or Fisher's exact test (if numbers are low) | Logistic regression using network meta-analysis methods to take account of the multiple treatments, and different treatment allocations by centre<br>Unclear if analysis properly accounts for cluster design | NR | Kaplan-Meier with log rank test, and Cox proportional cause-specific hazards regression | NR | Two proportion Z-test |
| | How will the treatment effects be presented | NR | OR (assumed) | NR | HR | NR | NR |
| | Any adjustment for covariates (non-inferiority trials)– | Not applicable | Not applicable | NR | Hazard ratios adjusted for the stratification factors. Multivariate model will also be constructed using relevant risk factors of breast cancer on the different time-to-event endpoints | NR | NR |

| *Item* | *Description* | *DENSE* | *BRAID* | *WISDOM* | *MyPEBS* | *TBST* | *MISS* |
|---|---|---|---|---|---|---|---|
| | Methods used for assumptions to be checked for statistical methods | Assumed expected cell sizes will be checked for Chi squared test as report will use Fisher's exact test if numbers are low | Not applicable | NR | NR | NR | Not applicable |
| | Details of alternative methods to be used if distributional assumptions do not hold, e.g., normality, proportional hazards, etc. | Will use Fisher's exact test if numbers are low | Not applicable | NR | NR | NR | Not applicable |
| **Missing data** | Reporting and assumptions/statistical methods to handle missing data (e.g., multiple imputation) | NR | NR | NR | NR | NR | NR |
| **Additional analyses** | Details of any additional statistical analyses required | Planned additional complier-average causal effect analysis | NR | Planned sensitivity analysis which excludes cancers detected at first screen. | Planned sensitivity analyses: (1) as if all participants had complied using causal inference methods; (2) excluding all prevalent cases | NR | NR |

NR = not reported

# 4 Trialling risk-based breast screening in Australia

## 4.1 Authors

A/Prof Carolyn Nickson, Paul Grogan

## 4.2 Background

Throughout the ROSA project, stakeholders have expressed a wide range of views about whether Australia would require a localised clinical trial or trials of risk-based based breast cancer screening, the alternatives being to either (i) continue with the current approach to screening until evidence from international trials demonstrates the value of a specific approach to risk-based screening, or (ii) pilot some form of risk-based breast screening expected to suit the Australian health setting and population and monitor outcomes in relation to implementation within a program setting. International trials are assessing a range of screening protocols, with varying potential for translation to the Australian setting.

The ROSA project has generated insights that inform whether a trial of risk-based breast screening is required in Australia, and how such a trial might best be designed and implemented.

## 4.3 Aims

To provide an overview of how various international trial protocols could translate to the Australian setting and propose a potential design and framework for an Australian trial program.

## 4.4 Methods

We produced above a high-level framework of potential barriers and enablers to implementing screening protocols being assessed through international trials. This outline draws on various ROSA activities including stakeholder feedback provided through the ROSA surveys of personnel within and outside the BSA program (Section 2, page 5). We then provide an assessment whether Australia requires its own trial of risk-based screening, a general design for that trial, and a framework for how this might best be implemented.

## 4.5 Results

### 4.5.1 Current international trials

As described under Chapter 4 (Risk-based screening protocols), the ROSA project produced an overview and quality appraisal of current international trials of risk-based breast screening. The six trials are summarised in Table 16 (page 50).

*Table 16. The six population level trials of risk-based screening included in the ROSA quality appraisal of international trials. Trial protocols are described at https://clinicaltrials.gov/ct2/show/[NCT#].*

| Acronym and age range | Full name and trial reference | Location | Trial period | Risk groups | Comparator | Intervention | |
|---|---|---|---|---|---|---|---|
| | | | | | | Intervals | Screening tests |
| **MyPeBS** (40-70) | Randomized, Comparison of Risk-Stratified versus Standard Breast Cancer Screening in European Women Aged 40-70 (MyPEBS), NCT03672331 | France, Italy, UK, Belgium and Israel | 2019 - 2025 | BCSC/T-C scores (4 groups) | Various (Annual/biennial/triennial screening, with mammography/DBT± supp US) | 1-4 years | Supp US/ABUS, supp MRI |
| **WISDOM** (40-74) | Women Informed to Screen Depending on Measures of Risk (Wisdom Study) (WISDOM), NCT02620852 | US (California) | 2016 - 2020 | BCSC (4 groups) | Annual mammography | 1-2 years None <50y | Supp MRI |
| **TBST** (45-50) | Tailored Screening for Breast Cancer in Premenopausal Women (TBST), NCT02619123 | Italy | 2013 - 2022 | BI-RADS 1/2 vs 3/4 | Annual mammography | 1-2 years | No |
| **DENSE** (50-75) | Breast Cancer Screening with MRI in Women Aged 50-75 Years with Extremely Dense Breast Tissue: the DENSE Trial (DENSE), NCT01315015 | Netherlands | 2011 - 2019 | Extremely dense (Volpara D) | Biennial mammography | No change | Supp MRI |
| **BRAID** (50-70) | Breast Screening – Risk Adaptive Imaging for Density (BRAID), NCT04097366 | UK | 2019 - 2026 | BI-RADS C/D | Triennial mammography | 18 months | Abbreviated MRI, ABUS, CEM |
| **MISS** (45-49 and 70-74) | What is the Best Interval to Screen Women 45-49 and 70-74 for Breast Cancer? (MISS), NCT04590560 | Italy | 2020 - 2026 | BI-RADS A-C versus D. | Uncertain (most likely annual tomosynthesis) | 2 years for BI-RADS A-C | N/A |

.

That critical appraisal concluded that no current international trials were considered to have a high or provisionally high risk of bias for any of the sources of bias assessed. We noted that some remaining sources of bias may be reduced as additional information is produced by these trials.

There are numerous differences between these international trials (in addition to their settings, populations and timing), including:

- age ranges of participants,
- approaches to risk assessment and stratification,
- screening intervals, and
- screening tests.

The variability of trial designs and settings means that detailed extrapolation and analysis of their findings is required in relation to the Australian setting. As described in Chapter 1, the ROSA project has been undertaking a comprehensive supplemental activity of case studies of four BreastScreen jurisdictions (including three service levels within one jurisdiction), with the analysis to be completed in early 2023. While we finalise the results of that analysis, some key considerations based on our existing activities are summarised below. These will be further refined as part of the *case studies* report.

## Age ranges

As indicated in Table 17, the six international trials include various age ranges, and this leads to potential enablers and barriers to their translation to the Australian setting, based on current target and eligible age ranges in the BreastScreen Australia (BSA) program, and stakeholder perspectives about priority age groups for risk-based breast screening.

*Table 17. Age ranges in the various trials of risk-based screening, and how this relates to potential enablers and barriers to implementing these trials in Australia.*

| Age range | Trials | Potential enablers in Australia | Potential barriers in Australia |
|---|---|---|---|
| 40-70/74 | MYPEBS, WISDOM | Corresponds to the current age range eligible for BSA screening | Women aged 40-49 are not currently targeted for BSA screening, so trial recruitment for this age group would require additional preparation by BSA services. |
| 44/45-49/50 | MISS, TBST | Recruitment to the trial would involve less disruption to current BSA services | Women aged 44-49 are not currently targeted for BSA screening, so trial recruitment for this age group would require additional preparation by BSA services. |
| 50-70/75 | DENSE, BRAID | Accords with the current age range targeted for BSA screening. | None identified. |
| 70-74 | MISS | Accords with the current age range targeted for BSA screening. | While there are no practical barriers identified, the level of support among Australian stakeholders for focusing trial efforts on older women within the target age range, and for de-intensifying screening in this age group, is not yet determined. |

## Risk assessment

As indicated in Table 18, the approaches to risk assessment used by the six trials included questionnaires, breast density and polygenic risk score information. Potential barriers related to translating the approaches used include (i) the lack of routine breast density assessment in Australia, (ii) the challenges of collecting detailed questionnaire data, and (iii) potential ethical considerations around assessing polygenic risk scores. Potential enablers include community interest and support for advice about breast density.

*Table 18. Risk assessment and stratification used in the various trials of risk-based screening, and how this relates to potential enablers and barriers to implementing these trials in Australia.*

| Risk assessment and stratification | Trials | Potential enablers in Australia | Potential barriers in Australia |
|---|---|---|---|
| Questionnaire based tools (BCSC/T-C scores) ± polygenic risk scores | (MYPEBS, WISDOM) | Women routinely complete questionnaires at each BreastScreen screening round, to which additional risk-related questions could be added.<br>Resources such as the iPrevent tool have been developed for the Australian population using questionnaire-based risk assessment tools, and could potentially be adapted to help women interpret their breast cancer risk and consider preventative strategies. | There is concern from BSA personnel and senior BSA representatives about the resources that would required to collect risk information and advise women about their risk, and how these would be funded (see Section 2 of this Chapter (Workforce and organisational readiness)).<br>Existing clinical guidelines provide different information about breast cancer risk and the introduction of an additional risk assessment method may add to the risk of women receiving different advice from different health services. |
| BI-RADS breast density | (TBST, BRAID, MISS) | Community and workforce support for breast density assessment and notification.<br>Extensive evidence about how BI-RADS breast density is associated with breast cancer screening outcomes and potential improvements in outcomes with targeted screening technologies (Chapter 2 and Chapter 3). | While BI-RADS visual breast density assessment is routinely used in the United States, Australian radiologists would require training (and QA checks) to reliably provide these assessments, noting that this approach to breast density assessment has limited inter- and intra-reader reliability.<br>BI-RADS assessment would add to work time required of radiologists, noting current concern from BSA personnel and senior BSA representatives about radiologist workforce capacity in the current program (see Section 2 of this Chapter (Workforce and organisational readiness)). |
| Volpara breast density | (DENSE) | We identified a body of evidence about how Volpara breast density is associated with breast cancer screening outcomes and potential improvements in outcomes with targeted screening technologies (Chapter 2 and Chapter 3). | Volpara breast density measurement would require resourcing and changes to hardware and information systems, and some training of various personnel. |

| | | | |
|---|---|---|---|
| | | Volpara has been assessed in some BreastScreen Australia settings. | |

## Screening intervals

The approach to screening intervals used by the six trials is shown in Table 19. As outlined in Table 16, the trials had a different set of screening intervals as a comparator, including annual, biennial and triennial screening. Trials incorporating annual screening can potentially build on existing annual screening policies at BSA. Trials involving screening intervals longer than 2 years may face opposition among Australian stakeholders.

*Table 19. Screening intervals used in the various trials of risk-based screening, and how this relates to potential enablers and barriers to implementing these trials in Australia.*

| Screening intervals | Trials | Potential enablers in Australia | Potential barriers in Australia |
|---|---|---|---|
| Annual | MyPeBS, WISDOM, TBST | Annual screening is currently offered to some BSA participants. | Current BSA policies for annual screening are difficult to evaluate and differ between state and territory programs; more rigorous management and oversight would be required. |
| 2 years | MyPeBS, WISDOM, TBST, MISS | This is the standard screening interval for the current BSA program. | None identified in relation to screening intervals. |
| 4 years | MyPeBS | Some stakeholders have indicated that they would support less frequent screening. | Some stakeholders have indicated that they would be strongly opposed to less frequent screening. |

## Screening tests

Imaging being used in the trials includes supplemental ultrasound, automated breast ultrasound (ABUS), contrast enhanced mammography (CEM), magnetic resonance imaging (MRI), abbreviated MRI, and digital breast tomosynthesis (DBT) (Table 20). The ROSA project *case studies* will provide valuable insights to help refine this analysis.

*Table 20. Breast imaging being used in the various trials of risk-based screening, and how this relates to potential enablers and barriers to implementing these trials in Australia.*

| Screening tests | Trials | Potential enablers in Australia | Potential barriers in Australia |
|---|---|---|---|
| Ultrasound | MyPeBS | Routinely used in BreastScreen assessment services | Staff shortages, additional reading time, additional resources required to manage increased false positives |
| ABUS | MyPeBS | None identified. | As above plus not in widespread use |
| CEM | BRAID | Used in some risk-based surveillance settings | Clinical services and staff required to support safe provision of contrast |
| MRI/ Abbreviated MRI | WISDOM, DENSE | Used in some risk-based surveillance settings | Clinical services and staff required to support safe provision of contrast |
| DBT | MISS | Routinely used in a majority of BreastScreen assessment services | Staff shortages for additional reading time<br>Additional resources required to manage increased false positives |

## 4.5.2 An Australian trial

### Purpose

While various trials of risk-based screening are underway internationally, their findings will not translate directly to the Australian setting due to differences in health service delivery and funding and the profile of the Australian population. Given the complexity of international trials and, on our analysis (as above), the limitations in applying findings to Australian settings, Australia is likely to require a trial designed specifically for and conducted in Australia. A trial tailored for Australian settings would provide a rigorous, independent and accountable framework to develop, test and evaluate:

- integration of different screening technologies in BreastScreen clinical pathways
- routine risk assessment and advice
- information systems and reporting
- staff training programs
- communication tools and acceptability to women
- costs, within a rigorous and accountable framework
- risk-group level performance indicators.

In addition, a trial would have a range of longer-term benefits such as:

- Attracting and developing Australian expertise to support monitoring and evaluation of interventions
- Establishing a framework for ongoing evidence-based improvements.

### Trial design

While the ROSA project findings and recommendations provide guidance on priority protocols to consider for a trial, due to the diversity of evidence about specific breast imaging tools as applied to screening populations and uncertainties in Australian health costs, our analyses do not identify a single risk-based screening protocol that is likely to yield the best balance of benefits, harms and cost-effectiveness.

International triallists have faced the same uncertainty of evidence but progressed nonetheless with selecting feasible trial protocols expected to improve the balance of benefits and harms.

In the Australian setting, we recommend a trial of routine risk assessment and advice incorporating breast density in a BreastScreen setting as a first step towards trialling risk-based screening protocols, with management of women identified as higher risk (in line with either current policies for annual screening and referral to high-risk clinics), and potential referral of women with high breast density and high breast cancer risk to additional or supplemental imaging. As this trial stage would focus on outcomes related to routine risk assessment and advice, to help progress this activity we recommend consideration that additional or supplemental imaging could be provided either (i) through BreastScreen assessment services or (ii) outside the program but requested and managed by BreastScreen, with a view to BreastScreen-led provision in the longer term. These options should be considered in close consultation with the service providers involved and with reference to the *Population-based screening framework*[10] to ensure adherence to the principles for setting up and managing screening programs in Australia.

[10] Clinical Principal Committee: Standing Committee on Screening. Population Based Screening Framework. Commonwealth of Australia as represented by the Department of Health 2018/ ISBN: 978-1-76007-370-1

Concurrently, we recommend further work to design a trial or trials of risk-based screening protocols that are likely to improve the balance of benefits and harms of screening for different risk groups, are feasible in Australia in terms of resourcing and workforce capacity and will not disadvantage populations such as women living in regional or remote areas.

Our recommended high-level trial design would have two distinct intervention protocols, based on age:

1. **For women aged 50-74 (the current target age range for BreastScreen):**
    - Randomisation would most likely occur at the BreastScreen screening site level (i.e. a cluster-randomised trial).
    - Risk assessment and most risk-based screening would be provided by BreastScreen, using well-validated risk assessment tools.
    - Data would be collected and managed primarily by BreastScreen.
    - The primary objective would be non-inferiority, if it can be determined, of the risk-stratified screening strategy in terms of incidence rate of breast cancer of stage 2 and higher (as a surrogate for mortality), compared to standard screening, measured after four years of intervention, with superiority assessed as a secondary outcome.
    - Follow up data would be collected for 15 years from study entry for evaluation of long-term cumulative breast cancer incidence and breast cancer-specific survival.

    This trial protocol would be similar to the MyPeBS trial protocols.

2. **For women aged 40-49 or 45-49 (women currently eligible but not targeted for BreastScreen services):**
    - Randomisation would occur at risk assessment (i.e. not at the individual level), and may be applied at an individual level or clustered by a health service site; this would need to be determined through co-design with relevant services.
    - Risk assessment and risk-based screening may be provided through a combination of BreastScreen and other health services, using well-validated risk assessment tools.
    - Data would be collected and managed through some form of centralised record-keeping system to be determined.
    - The primary objective would be non-inferiority, if it can be determined, of the risk-stratified screening strategy in terms of incidence rate of breast cancer of stage 2 and higher (2+), compared to current screening and risk-based surveillance, measured after four years of intervention, with superiority assessed as a secondary outcome.
    - Follow up data would be collected for 15 years from study entry for evaluation of long-term cumulative breast cancer incidence and breast cancer-specific survival.

This trial protocol would involve routine breast cancer risk assessment for women turning either 40 or 45 years of age, triaging screening and surveillance up to age 49 according to risk.

The protocol targetted to younger women is recommended in the context of women aged 40-49 usually representing approximately 11% of BreastScreen participants, involving over 230,000 screens every two years,[11] yielding 6% of all screen-detected cancers[12] (and 11% of interval cancers[13]). As described in Section 2 of this Chapter (Workforce and organisational readiness),

[11] Australian Institute of Health and Welfare 2021. BreastScreen Australia monitoring report 2021. Cat. no. CAN 140. Canberra: AIHW. (2018-2019 figures, derived from Table S1.1, S4.1, S4.3).
[12] Australian Institute of Health and Welfare 2021. BreastScreen Australia monitoring report 2021. Cat. no. CAN 140. Canberra: AIHW. (2019 figures, derived from S4.1, S4.3).
[13] Australian Institute of Health and Welfare (2022) BreastScreen Australia monitoring report 2022, catalogue number CAN 150, AIHW, Australian Government. (2015-2017 combined figures, derived from Table S6.3 and Table S6.6).

there is considerable stakeholder interest in extending some form of targetted screening to women younger than 50 years. The 19 shortlisted risk-based screening protocols reported in the ROSA clinical and health economics modelled evaluation (Chapter 4) include six scenarios where screening would commence at age 40 and three scenarios where screening would commence at age 45. To help identify priority protocols to consider for a trial, we would recommend further examination of the ROSA project modelled estimates for this age group (noting that outcomes are currently reported for the whole target age range e.g. 40-74 or 45-74).

## An Australian trial program

We recommend that any Australian trial of risk-based breast screening be part of a *trial program* incorporating complementary activities such as implementation studies, evidence reviews and modelled evaluations. A framework for the recommended trial program for Australia is shown in Figure 3, adapted from the European MyPeBS trial (Appendix Figure 5, page 62).

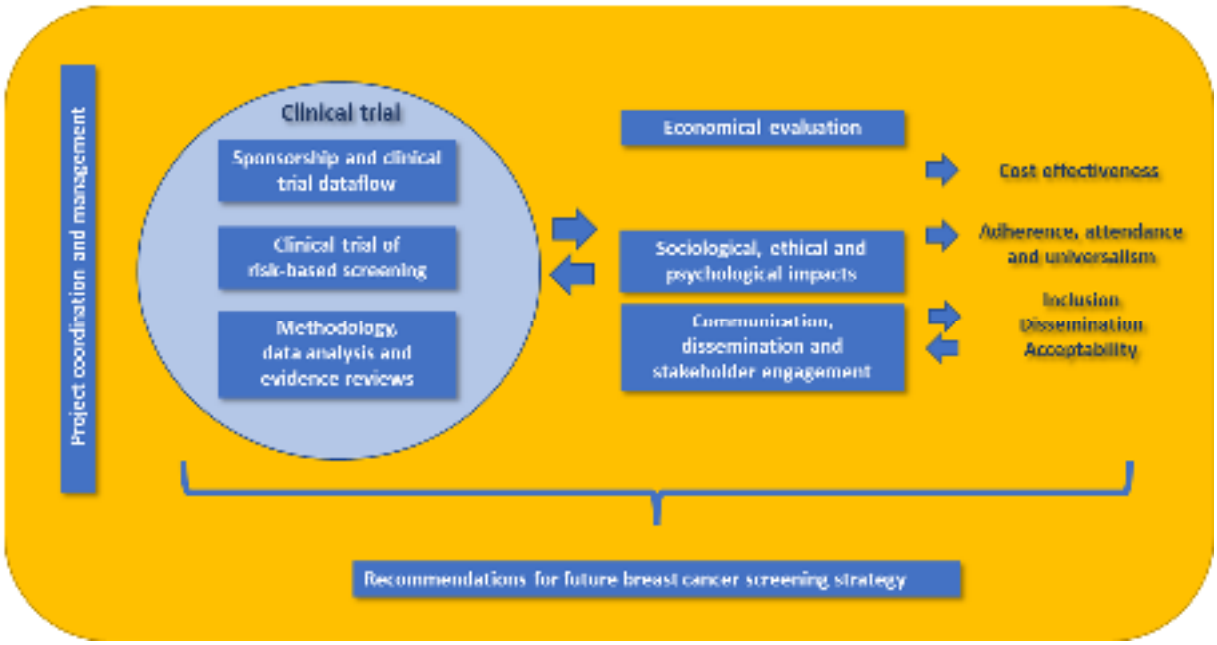



*Figure 3. An overview of the recommended trial program, modified from the MyPeBS trial overview (Appendix, Figure 5).*

The ROSA project recommendations provide guidance on the systems, knowledge and engagement required to support trial implementation, and how to ensure existing disparities in screening uptake are monitored and managed throughout a trial. For any trial, we recommend that the trial design should follow the criteria for assessment of population screening as outlined in the Australian Population Based Screening Framework (2018)[14], to the condition, test, assessment, screening program, and treatment and ongoing management, for each risk group being considered. This will help ensure that any trial would consider the benefits, harms and costs not only for the whole population but for women at all levels of cancer risk.

Thus, the trial program would include a strong focus on BreastScreen services especially for women aged 50-74 years but also incorporate primary care services and high-risk clinics such as family cancer centres and specialist breast clinics. This coordinated approach is critical in Australia, given

[14] https://www.health.gov.au/resources/publications/population-based-screening-framework

the overlap of BreastScreen screening and Medicare subsidised risk-based surveillance, a clear role for primary care in assessing risk for women not yet enrolled in the BreastScreen program, and established specialist high-risk services such as family cancer centres and breast clinics in more complex risk assessment and surveillance.

## Equitable access

As reported throughout ROSA reports, BreastScreen participation is currently low in some culturally and linguistically diverse (CALD) communities, and among Indigenous women and women living in remote or very remote locations. Women in the BreastScreen target age range can and do access services within and outside BreastScreen, including services prior to commencing BreastScreen participation, concurrently with BreastScreen participation, and as a result of referral from BreastScreen (i.e., to high-risk clinics), and access is known to differ according to location of residence.

Risk-based screening could potentially worsen existing disparities, without careful, evidence-based planning. For example:

- It is possible that in population groups with lower-than-average BreastScreen participation, women may be deterred further by requirements to provide more information to participate in screening.
- The quality of risk assessment may be reduced for women in CALD communities and Indigenous women if relying on self-reported information without ensuring appropriate language and cultural content and support services are readily available.
- It may be challenging to ensure that women living in remote or very remote locations have equal access to more personalised, risk-based breast imaging services. For example, the balance of benefits and harms for screening tests with higher sensitivity but increased false positive recall rates may be worse for women living further away from BreastScreen assessment services.

These considerations should be central to the recommended trial program.

## Safety monitoring

Safety monitoring throughout the trial program would include:

- monitoring the performance of screening protocols to ensure they are performing as expected in each risk group;
- prompt independent review of interval cancer cases and interval cancer rates arising in all trial arms and risk groups; and
- monitoring potential downstream impacts on related health services to avoid harms such as cost-shifts to other services and negative psycho-social or cost impacts for women.

## Leadership and governance

We recommend that the trial program should be led through a collaboration of academic, clinical and policy and consumer organisations. While any large-scale trial would, by necessity, have its own clear governance structure following NHMRC recommendations, we recommend that the broader trial program would also have governance arrangements in place which, as for the ROSA project, would include engagement with key BreastScreen Australia committees and representatives, professional peak bodies, policy advisors and consumers. (Consumer involvement in trials is mandated; there would, however, be a benefit in exceeding mandatory requirements and involving consumer organisations to support the trial.)

The recommended trial program would benefit from access to a centralised BreastScreen governance committee or board with sufficient expertise and independence to advise and support the trial program independent of current program management issues. This would also align with the Australian Population Based Screening Framework (2018) key principle for screening program governance and management to 'Clearly define leadership, advisory and decision-making processes'.

## Stakeholder and consumer perspectives

Commencing with its 2019 desktop review of stakeholder perspectives and findings from a survey of stakeholder groups (Appendix section 5.1, page 63), the ROSA project has highlighted the importance of stakeholder engagement in this domain. Accordingly, the recommended trial program includes a broader scope than the MyPeBS trial in terms of stakeholder engagement.

While BreastScreen Australia and its state and territory programs are essential stakeholders for consideration of risk-based breast screening in Australia, there are other Australian stakeholders to consider, each with their own relationships and connections. This includes people or organisations with a professional connection to breast cancer screening or surveillance and related cancer control services, consumer advocacy groups, and consumer representatives. Different stakeholder groups contribute to the Australian discourse about breast cancer screening on topics such as the benefits and harms of screening, access to screening services, consideration of overdiagnosis and, most recently, breast density assessment as part of routine breast screening service provision.

Understanding consumer perspectives is an active area of research with various recent publications and reports contributing valuable evidence in this domain. This includes the development and pilot testing of an online decision aid for women considering risk-stratified breast screening,[15] analyses of how women would value personalized breast cancer risk assessment in relation to risk-based breast screening[16] and attitudes to different methods of providing risk assessment and advice,[17] and analysis of how women respond to advice that they have higher breast density.[18]

For example, a 2020 report commissioned by the Australian Government Department of Health[19] found that although there are many resources available to assess breast cancer risk, Australian women had a limited understanding of breast cancer risks, and various stakeholders held a mix of views about women's understanding of breast density, family history and age as risk factors. The authors concluded that '*An area consistently identified as challenging by stakeholders was*

[15] Lippey J, Keogh L, Campbell I, Mann GB, Forrest L. Development and pilot testing of an online decision aid for women considering risk-stratified breast screening. J Community Genet. 2022 Feb;13(1):137-141. doi: 10.1007/s12687-021-00571-y. Epub 2022 Jan 21. PMID: 35060087; PMCID: PMC8799785.
[16] Wheeler JCW, Keogh L, Sierra MA, Devereux L, Jones K, IJzerman MJ, Trainer AH. Heterogeneity in how women value risk-stratified breast screening. Genet Med. 2022 Jan;24(1):146-156. doi: 10.1016/j.gim.2021.09.002. Epub 2021 Nov 30. PMID: 34906505.
[17] Sierra MA, Wheeler JCW, Devereux L, Trainer AH, Keogh L. Exploring Implementation of Personal Breast Cancer Risk Assessments. J Pers Med. 2021 Sep 30;11(10):992. doi: 10.3390/jpm11100992. PMID: 34683136; PMCID: PMC8541275.
[18] Darcey E, Hunt EJ, Keogh L, McLean K, Saunders C, Thompson S, Woulfe C, Wylie E, Stone J. Post-mammographic screening behaviour: A survey investigating what women do after being told they have dense breasts. Health Promot J Austr. 2021 Oct;32 Suppl 2:29-39. doi: 10.1002/hpja.396. Epub 2020 Sep 10. PMID: 32754972.
[19] Allen + Clarke, Understanding informed decision making, a literature review, stocktake and stakeholder insights about Australian women's attitudes to participating in population-based breast screening,18 June, 2020 (https://www.health.gov.au/resources/publications/breastscreen-australia-understanding-informed-decision-making-a-literature-review-about-australian-womens-attitudes-to-participating-in-population-based-breast-screening)

*communicating complex clinical information when the science may not provide settled evidence on the direction to take or when there is no clear consensus on what to do. This is problematic for both women who are deciding on whether to participate in breast-screening as well as clinical staff who are providing advice*.'

Any approach to risk-based screening should be done in a way that sustains or improves on usual BreastScreen participation. A 2020 government-funded report on strategies to increase participation in cancer screening programs, including BreastScreen[20] provides valuable insights about how BreastScreen participation can best be increased and made more equitable for the current approach to breast screening (Figure 4). The detailed evidence contained in this report, which includes many insights from BreastScreen personnel, is a valuable reference.

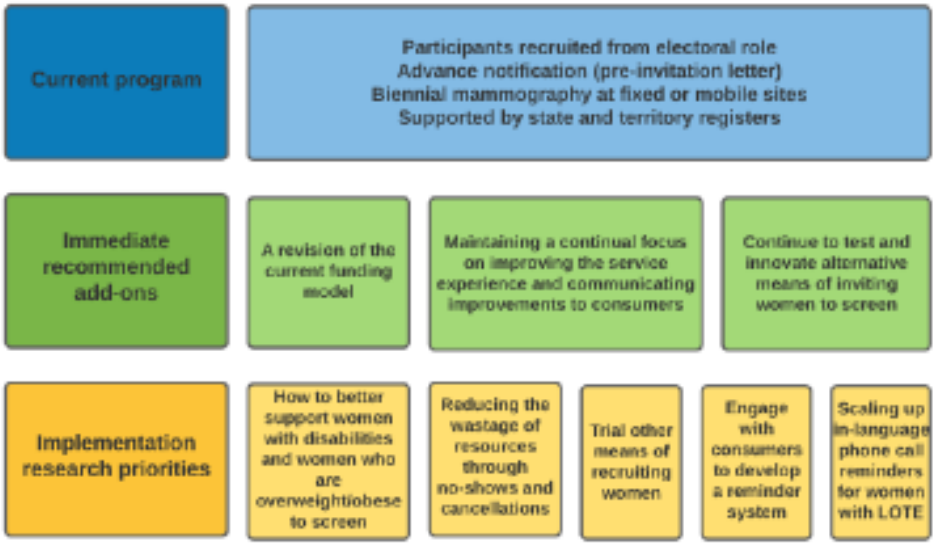



*Figure 4. Key components for consideration in the development of a long-term strategy to improve participation in BreastScreen Australia, as reported in Nightingale et. al. (2020)*

While ROSA project engagement with the Breast Cancer Network of Australia (BCNA) has led to the distribution of information about the project to a wide community of motivated consumers and health professionals, the project has not been funded to undertake activities directly related to consumer readiness for risk-based screening nor to develop strategies for consumer engagement with any implementation of risk-based screening. Given that the majority of Australian women appear to currently support the idea of being told about their breast density,[21] ongoing engagement with stakeholders and consumers is an essential part of any consideration of risk-based breast cancer screening in Australia.

## 4.6 Conclusion

There are potential barriers and enablers to translating any international trials to the Australian setting. On the basis of our analysis, we recommend a strategy for working towards the

[20] Nightingale C, Verbunt E, Creagh N, Brotherton J, English D, Flander L, Jenkins M, Saville M and Kelaher M. Development of a Strategic Approach to Achieve Increased Participation in the Bowel, Breast and Cervical National Cancer Screening Programs. Final Report. 31 July 2020.

[21] Nickel, B., Dolan, H., Carter, S., Houssami, N., Brennan, M., Hersch, J., Verde, A., Vaccaro, L., McCaffery, K. (2022). ?It's about our bodies? we have the right to know this stuff?: A qualitative focus group study on Australian women's perspectives on breast density. Patient Education and Counseling, 105(3), 632-64

development of a large-scale trial of risk-based breast screening in Australia, commencing with a trial of routine risk assessment and advice incorporating breast density, and the design of trial protocols for women aged 40-49 and 50-74 as indicated from our evidence collection, review and analysis reported to date.

While routine risk-based screening would be likely to involve tailored screening tests and possible tailored screening intervals, current methods already used by BreastScreen for annual screening of higher risk women (e.g. data collection and reporting systems and communications to clients and their GPs) may provide a foundation for implementing more detailed risk assessment. Given that GPs and women have access to multiple guidelines about breast cancer risk assessment and risk-based management, this should be considered in the design and implementation of any additional risk assessment and advice by the BreastScreen program. Additionally, there may be downstream impacts of advice for BreastScreen clients about breast cancer risk and risks related to breast density. This includes psycho-social impacts, requests for imaging outside BreastScreen, and potential underestimation of breast cancer risk for women in lower-risk groups.

Any risk-based breast screening should maintain the integrity of the BreastScreen program as specified in the Australian Population Based Screening Framework[22], so that '*screening is an organised, integrated process where all activities along the screening pathway are planned, coordinated, monitored and evaluated through a quality improvement framework*'. This supports continued management of women at different levels of risk by the BreastScreen program, with the exception of a small proportion of very high-risk women suited to substantially more intensive surveillance through family cancer centres.

[22] https://www.health.gov.au/resources/publications/population-based-screening-framework

# 4.7 Appendix

## 4.7.1 MyPeBS

The MyPeBS program is described as a set of 'work packages' (Figure 5), including an external link to the US 'Wisdom' trial. A similar link between these existing trials may be possible for an Australian trial program, including sharing of study instruments and protocols.

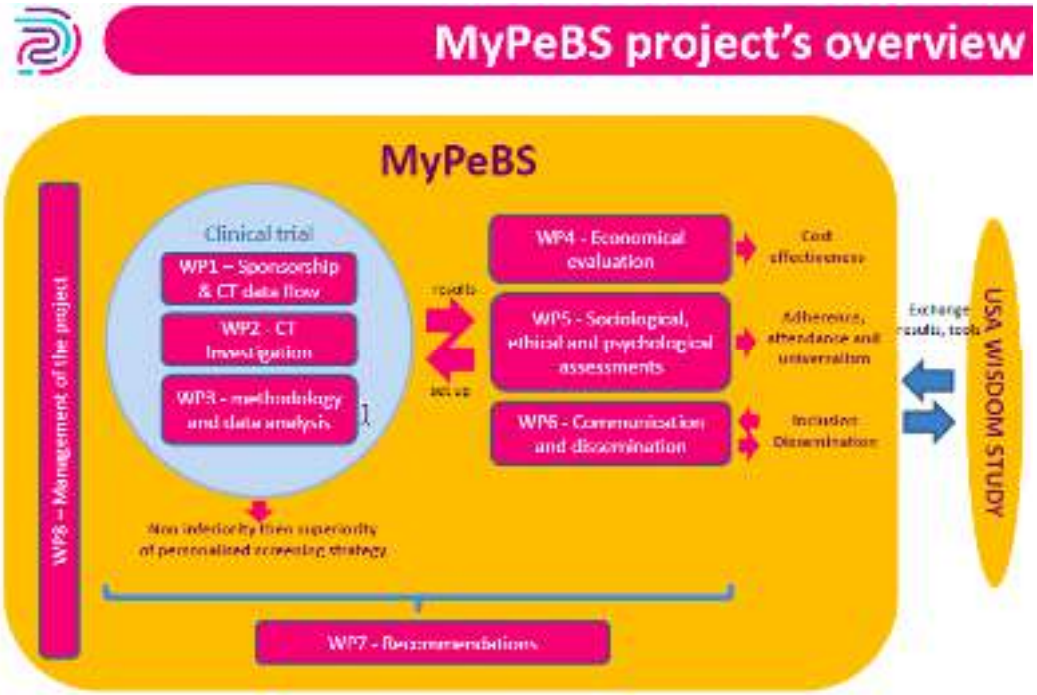



*Figure 5. MyPeBS trial program overview.*

# 5 Appendices

## 5.1 Key findings

Drawing from the detailed analyses and results described throughout the chapter, the project generated a set of key findings which were reviewed by the ROSA Expert Advisory Group over May to July 2022. The final set of key findings is outlined below.

### Q1. Are Australian health services personnel working in screening and surveillance likely to support the introduction of risk-based breast screening, and do they think their organisations are ready?

#### Considerations for implementation

*BreastScreen Australia*

1. On average, BreastScreen personnel (representing a range of experience, roles and state/territory locations) consider BreastScreen to have good readiness for change to more risk-based screening in terms of leadership culture, staff culture, leadership, measurement (how well an organisation and its leadership motivates its aims and supports staff to understand what they should be doing and giving feedback on their performance within their role) and opinion leadership (the role influential people within the organisation play to influence the change processes), with some variation of views among respondents.
2. On average, BreastScreen personnel (representing a range of experience, roles and state/territory locations) have a wide range of attitudes towards adoption of new evidence-based guidelines, with greatest value placed on openness to new practices and the time and administrative burden with learning new evidence-based practices.
3. On average, in response to a range of scenarios involving risk-based breast screening, BreastScreen personnel (representing a range of experience, roles and state/territory locations) indicated a good likelihood of adopting specific risk-based screening guidelines, with mixed views on whether they should have a say in how guidelines should be put into practice.

*Other health services*

1. On average, health services personnel outside BreastScreen (representing a range of experience, roles and state/territory locations) have mixed views about whether their organisations are ready for change to more risk-based screening in terms of leadership culture, staff culture, leadership, measurement (how well an organisation and its leadership motivates its aims and supports staff to understand what they should be doing and giving feedback on their performance within their role) and opinion leadership (the role influential people within the organisation play to influence the change processes).
2. On average, health services personnel outside BreastScreen (representing a range of experience, roles and state/territory locations) have a wide range of attitudes towards adoption of new evidence-based guidelines, with greatest value placed on openness to new practices, feedback, monitoring and the time and administrative burden with learning new evidence-based practices.
3. On average, health services personnel outside BreastScreen (representing a range of experience, roles and state/territory locations) had mixed views about the likelihood of adopting specific risk-based screening guidelines, with greater value placed on the appeal of guidelines (if

the guidelines made sense, there was sufficient training and colleagues were happy using them) and fit (if the guidelines were the 'right thing' to do, they fitted with the respondent's clinical approach and they had a say on how they were implemented) rather than requirements (if the guidelines were required by their organisation, state/territory or supervisor).

### Priority evidence gaps

1. More detailed analyses of ROSA survey data.
2. Qualitative research with health services personnel within and outside BreastScreen (such as follow-up interviews of ROSA online survey respondents).
3. Additional ROSA surveys targeted to health services providers working in remote and rural settings.

## Q2. What are the current registered ongoing randomised controlled trials (RCTs) of risk-based breast cancer screening, and what is the quality of these studies?

### Key evidence

1. Of the six trials outside Australia assessing various protocols for risk-based screening, all trials are methodologically valid.

### Considerations for implementation

1. Most current trials are awaiting primary outcomes.
2. Trials usually assess tumour stage as the primary outcome, as a surrogate for mortality.
3. Four current trials are assessing reduced screening intensity in lower risk groups, which requires a non-inferiority framework.
4. No trial evidence is expected to translate directly to Australia due to differences in health systems.
5. Various trials currently underway involve methods and instruments that are likely to be relevant to a trial in the Australian setting.

### Priority evidence gaps

1. A trial conducted in the Australian setting.

## Q3. How could BreastScreen routine data collection and reporting be enhanced to support risk-based screening?

*These findings refer to an activity described in the joint ROSA/AIHW report [23](unpublished).*

### Considerations for implementation

1. Current BreastScreen data collection and reporting for women aged 50-74 includes some information by factors of interest for risk-based screening, but there are opportunities for this to be enhanced.
2. A suitable change management protocol to support enhanced BreastScreen data collection and reporting would incorporate a clear, evidence-based, reasonably independent and robust governance framework, well-defined decision-making bodies including representatives with

[23] Australian Institute of Health and Welfare & Cancer Council Australia. Enhanced BreastScreen data collection and reporting: An activity under the Roadmap for Optimising Screening in Australia (ROSA) project. Submission date 11 November 2022

operational expertise and advisors with scientific expertise about the items being considered, clear mechanisms for making decisions and careful consideration of national and state-level policies and guidelines, and resources to support the development, implementation and quality assurance of data collection and reporting processes.

## Q4. How does the COVID pandemic impact on consideration of risk-based breast screening?

*These findings refer to a special report provided to the Australian Government Department of Health and Aged Care in 2020 [24] (unpublished).*

### Considerations for implementation

1. BreastScreen adaptations to providing services during the COVID pandemic included various approaches to prioritising which women should be screened first during recovery periods. This may provide insights about implementing more targeted approaches to screening invitations.
2. COVID impacts on observed BreastScreen participation and potential changes in the profile of screened women is expected to impact routinely reported outcomes for the BreastScreen program for some time, and this may impact evaluations of the effectiveness of risk-based screening protocols in the future.

## Q5. What are stakeholder perspectives on risk-based breast screening?

*These findings refer to ROSA Stakeholder Perspectives report provided to the Australian Government Department of Health and Aged Care in 2019 [25](unpublished).*

### Considerations for implementation

1. Stakeholder groups consider mortality benefit, reduced treatment intensity, reduced interval cancers and minimised overdiagnoses to be priority considerations to build consensus on risk-based breast screening.
2. There is a lack of consensus among stakeholders about how breast cancer risk should be assessed, how breast density should be measured and if and how screening should be tailored according to breast cancer risk.
3. Stakeholder interest and advocacy for breast density notification is significant, with a range of views around whether women should be advised about their breast density, and whether breast density advice should be provided without policies and resources in place to provide screening and surveillance services tailored to their breast density.
4. There is increasing effort from commercial interests to promote new technologies to health services and consumers in relation to breast density and risk assessment, including add-ons to mammography machines currently used by BreastScreen services.

[24] Optimising Early Detection of Breast Cancer in Australia (OEDBCA). Interim Report: COVID-19. 09 November 2020 (unpublished).
[25] Optimising Early Detection of Breast Cancer in Australia project report: Stakeholders. 1 November 2019 (unpublished).