Abstract

There are currently no single disruptors to breast cancer screening akin to the impact of human papillomavirus testing and vaccination on cervical cancer screening. However, there is a groundswell of interest to review the BreastScreen Australia program to consider more risk-based screening protocols and to establish whether to routinely inform women about their breast density. We propose a framework for a considered, evidence-based review.

Population-level effectiveness of breast cancer screening is ultimately measured through its impact on breast cancer mortality, and this has been realised in Australia. Effectiveness can also be measured through treatment intensity, estimated overdiagnosis, false-positive screens and health economics measures. Key levers to improve such population-level outcomes include screening participation, screening test sensitivity and specificity, risk assessment and screening protocols.

We propose that the review of the program should fall under an evidence-based, consensus-guided framework comprising four complementary elements: improved evidence on current program performance for population risk subgroups; regularly updated evidence on key levers for change; clinical trials and population simulation modelling working in tandem; and consensus-based decision making about the degree of improvement required to justify change.

Informing women about their breast density is feasible and would be valued by some BreastScreen clients to help understand the accuracy of their screening test. However, without agreed protocols for screening women with dense breasts, increases in supplemental screening as observed in other settings would, in Australia, shift screening costs to clients and Medicare. This would reduce equity of access to population screening, and maintaining BreastScreen’s usual standard of monitoring and quality management (such as screen-detected and interval cancer diagnoses, and imaging and biopsy rates) would require data linkage between BreastScreen and other services.
Introduction

BreastScreen Australia was phased in from 1991 and fully implemented in 1995, and now offers biennial mammographic screening targeted to women aged 50–74 years (available from age 40). Its primary purpose is to reduce breast cancer mortality and this has been realised in Australia with a 41–52% reduction in mortality for screening participants, and a 21% reduction in population-level breast cancer mortality. BreastScreen’s design is largely unchanged since its inception. The screening test has changed from film to digital mammography and, more recently, some BreastScreen services have incorporated digital breast tomosynthesis (DBT) imaging in women recalled for assessment after a positive mammographic screen. The target age range was extended from 69 to 74 years in 2013, although some jurisdictions had previously re-invited women up to 74 years. Management of women with a strong family history of breast cancer or a history of breast disease has varied between jurisdictions. BreastScreen Western Australia reports breast density information as part of the screening result.

While there are no single disruptors to breast cancer screening akin to the impact of human papillomavirus testing and vaccination on cervical cancer screening, there is a groundswell of interest to review the current program as a result of factors such as improved breast cancer risk assessment tools, changes in breast imaging outside the screening program and a shift towards more personalised medicine. There is community interest in informing women about the harms and benefits of screening and, more specifically, in providing advice about breast density to indicate the sensitivity of the screening test and help estimate breast cancer risk.

In this context, we propose a framework for a considered, evidence-based review of the current national screening program.

How do we best assess the effectiveness of breast cancer screening?

The effectiveness of breast cancer screening is ultimately measured through its impact on breast cancer mortality. However, there are some well-established secondary benefits and harms that should be considered. Early detection of breast cancer reduces the intensity of treatment, with short- and long-term sequela, and this benefit will only increase as breast cancer treatment is further targeted to tumour subtypes and as clinicians aim for more conservative management. As such, information on tumour size, stage, hormone receptor and human epidermal growth factor receptor 2 (HER2) status at diagnosis help indicate the benefit of breast cancer screening in the contemporary health setting, with suitable consideration of overdiagnosis (detection of cancers that would not otherwise become clinically detected).

It follows that where access to treatment is below national standards, such as for remote communities and Indigenous populations, the effectiveness of screening could be improved through improved treatment access. Although treatment access is outside the remit of the BreastScreen program, it is an important public health consideration.

In terms of harms, overdiagnosed cancers confer harm to women through their diagnosis and treatment, while also reducing the cost-effectiveness of screening. Overdiagnosed cancers cannot be distinguished from other cancers at detection, so rates of overdiagnosis are uncertain, most likely in the order of 11% of all diagnoses in the population and 19% of screen-detected cancers. Until overdiagnosed cancers can be differentiated at diagnosis, their burden may be reduced through reduced treatment intensity via selective use of adjuvant therapies after surgical excision, and possible monitoring rather than treatment of low-grade ductal carcinoma in situ.

False-positive screening episodes are common (approximately 10% of first-round and 3% of subsequent-round screens) and a significant impost to BreastScreen participants and BreastScreen assessment services, deterring some women from future participation. Interval cancers (cancers diagnosed following a negative screen and before the next scheduled screen) comprise a mix of cancers missed at screening and new, fast-growing cancers that were not perceptible on the mammographic screen; interval cancers occurring in the first year after screening are more likely to be missed, and are an important indicator of screening effectiveness.

Finally, health economics measures are essential to help ensure equitable and optimal allocation of health resources, and to account for the positive and negative impacts of screening on quality of life.

Some measures of screening program performance are important for quality control, but are not direct measures of the effectiveness of the program. For example, overall program sensitivity is an important indicator of screening test sensitivity and of outcomes.
experienced by screening participants. However, program sensitivity is not a direct measure of screening program effectiveness due to unclear differentiation between new and missed interval cancers, and because program sensitivity increases with increased overdiaagnoses.\(^{17}\)

**Key levers for improving the performance of breast cancer screening**

Key levers to improve the effectiveness of screening at a population level are screening participation, screening test sensitivity and specificity, and screening protocols. Although some of these factors (e.g. participation) are used as ongoing performance indicators, they are also potential ‘levers’ to improve overall performance of the program.

**Screening participation**

Screening participation improves program performance through increased opportunity for early detection. BreastScreen participation is lower than expected from early screening trials (55% versus 70%) and has been stable for more than 20 years.\(^ {15}\) Although there is some screening outside the program (equivalent to about 3.5% BreastScreen participation)\(^ {4}\), marked changes in screening participation are more likely to arise through engaging the remaining population. Factors influencing screening participation include socio-economic and cultural factors, distance to screening services, personal health and disability, worry or perceived risk about breast cancer, and the quality and extent of information provided to potential program participants and health professionals.\(^ {18}\)

**Screening test sensitivity and specificity**

BreastScreen Australia uses screening protocols associated with increased breast cancer detection, namely two-view mammography and independent double-reading.\(^ {19}\) Improved screening test sensitivity, particularly for women with dense breasts, could potentially reduce rates of interval cancers currently missed by mammography. This may involve alternative or supplemental imaging technologies; however, only DBT and ultrasound have been investigated in prospective trials in population screening, and little is understood about concomitant increases in overdiagnosis.

Improved screening test sensitivity is often accompanied by reduced specificity, leading to increased false-positive outcomes. Ultrasound screening in women with dense breasts and negative screening mammograms will detect additional breast cancers; however, ultrasound also increases false positives, and there is little evidence of reduced interval cancer rates or breast cancer mortality.\(^ {19,20}\) DBT for population screening (combined with, or instead of, mammography) significantly increases breast cancer detection.\(^ {21}\) Recall rates (and hence false positives) vary greatly between settings\(^ {22}\) and longer-term outcomes are not yet available.

In the future, artificial intelligence information systems drawing on related clinical data may help improve test accuracy. However, the evidence on systematic application of artificial intelligence remains formative and unclear. Educational strategies to help optimise radiology practice and promote consistency in service provision therefore present a more immediate opportunity to improve outcomes. For example, the BREAST test assesses BreastScreen radiologists’ ability to correctly identify cases from mixed case-control test sets, with radiologists scored for peer comparison and able to review the accuracy of their own responses. This has been shown to improve screen-readers’ performance on test sets\(^ {22}\) to a degree that may improve overall screening program performance.

**Screening protocols**

Options for risk-based screening protocols include tailoring screening intervals based on estimated risk of an interval cancer. Although this approach may be easiest to implement, earlier detection of some interval cancers is likely to be offset by a significant increase in false positives.\(^ {23}\)

A better balance of benefits and harms is more likely to be achieved through screening protocols that combine tailored screening intervals with individualised use of screening technologies, assigned according to breast cancer risk and/or risk of masking due to breast density. This would require accurate and feasible population-level risk assessment. Some questionnaire-based risk assessment tools are now well validated on large populations from a variety of ethnicities\(^ {24}\), with one tool now validated in Australia.\(^ {25}\) Existing risk assessment tools are being updated to incorporate breast density and genetic information\(^ {26,27}\), although these are yet to be validated.

Implementing risk-based protocols at a population level would require numerous decisions, such as how to stratify women into risk groups (e.g. according to questionnaire, breast density and genetic information), how often risk should be re-assessed over the life course of screening participation, and optimal screening protocols for specific categories of risk.

**A framework for review**

We propose an evidence-based, consensus-guided framework for reviewing breast cancer screening. The proposed framework assesses screening effectiveness in the era of personalised medicine, allows review of multiple factors that may together warrant change, and gives evidence-based consideration of the benefits, harms and costs of various approaches to breast cancer screening.
This holistic approach will generate evidence required for policy makers to better assess any future changes to the program, with time to design appropriate health services. The framework comprises four complementary elements:

1. Improved evidence on current program performance for different risk groups

Current screening performance for different population risk subgroups (such as those that might be used as a basis for more personalised screening protocols) is important for understanding the current program, and as a comparator for monitoring and evaluating any future intervention; however, it could be better understood. Valuable outcome measures include tumour stage, size, hormone receptor and HER2 status at diagnosis, rates of false-positive screens and interval cancers, treatment access, overdiagnosis estimates and health economics measures (such as quality-adjusted life years, years of life gained and cost-effectiveness measures). Current work includes various projects and activities funded by the Australian Government Department of Health and a national data linkage project combining BreastScreen, cancer registry and death records.

2. Evidence about key levers for change

We propose regular updates on evidence about initiatives shown to increase screening participation, mechanisms to improve population-level screening sensitivity and specificity, the feasibility and utility of population-level risk stratification, and the performance of risk-based screening protocols.

3. Clinical trials and population simulation modelling

Population-level trials offer the highest quality evidence, although they are slow to generate long-term outcomes and can only test limited protocols. Several protocols are being trialled internationally (e.g. WISDOM, DENSE and MyPeBS). Population simulation modelling offers estimates for a wide range of screening protocols triaged to a variety of risk groups, with outcomes including survival and health economics measures, assumptions and costs tailored to local settings. This is an active area of work in Australia and internationally. Trials and population modelling work in tandem to generate evidence about how the complex system of breast cancer screening, with its observable and unobservable benefits and harms, might respond to key levers.

4. Consensus-based decision making about the degree of improvement that is required to justify change

Any changes to current screening protocols should be shown to significantly improve on the current program for the whole population without inadvertently reducing its effectiveness for specific subpopulations, with an acceptable balance of costs, benefits and harms. Although modelling studies can generate estimates of net value, such as cost-effectiveness ratios of various strategies and the number needed to screen per cancer death prevented, determining what constitutes sufficient and appropriate evidence to warrant changes to policy and practice requires expert population health and high-level guidance suited to the local health system. Cancer Council Australia is leading a Federal Government–funded project gathering evidence to support consensus-based decision making, aligning with the framework presented here. In addition, information on consumer preferences for risk assessment, advice and management would help maximise the feasibility and uptake of risk-based screening protocols.

On breast density notification

Various advocacy groups are campaigning for BreastScreen participants to be informed about their personal breast density and its possible effect on screening test accuracy. Routine breast density assessment is possible given the range of validated tools available, and such data would be valuable for program monitoring and evaluation. However, evidence on the most effective method for assessing breast density and a pathway for clinical management of women with dense breasts is not yet established.

Supplemental screening for dense breasts increases cancer detection rates and false-positive outcomes, and the balance of benefits and harms is not yet clear. Although work continues in evaluating screening protocols based on breast density and other risk factors, there is a lack of evidence about the relative benefits and harms of breast density notification without offering specified screening protocols. Such practice has increased supplemental screening in other settings; in Australia, this would mean shifting screening costs to Medicare and to clients (introducing inequitable access to population screening). Routine notification has also highlighted nonconsensus and unmet knowledge needs among health service providers and women making decisions about supplemental screening.
These potentially adverse outcomes could be mitigated by carefully planned, monitored and evaluated implementation\textsuperscript{38}, although privately funded supplemental screening and its imposition on women and the health system is likely to remain an issue. Cancer diagnoses and rates of imaging and biopsy in benign cases outside the BreastScreen program would require careful monitoring to capture complete information about the benefits, harms and costs of screening.

There are some concerns that assessment of breast density creates an ethico-legal responsibility to offer women personalised screening protocols to minimise the masking effect of dense tissue. Without clear evidence on optimal protocols for women with dense breasts, this creates pressure to implement overly simplistic risk-based screening protocols without fully understanding the impact of such a change. As described above, evidence is being generated, but this will take time.

**Conclusion**

The proposed framework assesses screening effectiveness in the era of personalised medicine, allows review of multiple factors that may together warrant change, and gives full evidence-based consideration of the benefits, harms and costs of various approaches to breast cancer screening. The framework requires a coordinated approach to generate the evidence required for policymakers, with time to prepare appropriate information and health services to support women, screening programs and health professionals, should changes to screening be warranted.

**Acknowledgements**

NH receives research support through a National Breast Cancer Foundation (Australia), Breast Cancer Research Leadership Fellowship.

**Peer review and provenance**

Externally peer reviewed, commissioned.

**Competing interests**

CN receives salary funding via the project ‘Optimising Early Detection of Breast Cancer in Australia’, funded by the Australian Department of Health and led by Cancer Council Australia. She jointly leads the development of the AutoDensity automated breast density measurement tool, through a collaboration between the University of Melbourne and the CSIRO. PB is the Chief Executive Officer of the start-up DetectEd-X, which provides radiology training and research.

**Author contributions**

CN wrote the first draft and other authors provided comments, and critically reviewed subsequent versions. All authors approved the final version for submission.

**References**


