Policy1-Breast

Model description

Version 1.0, September 2022

1 Model overview

Policy1-Breast is a continuous-time, stochastic, multiple-cohort population micro-simulation model of breast cancer screening, diagnosis and treatment in Australia. It is designed to evaluate costs, benefits and harms of breast cancer screening and other interventions in the Australian population.

The Policy1-Breast model is part of a larger Policy1 modelling platform using a world-leading modelling approach and including information about multiple aspects of cancer and risk behaviours, which can be drawn on to help specify different areas of the whole model.

Each individual woman in the simulation is assigned a range of attributes including age, lifetime breast cancer risk, breast cancer natural history up to clinical diagnosis, and life-course breast density.
Breast cancer screening is then modelled as an overlay, for current and alternative BreastScreen Australia screening protocols, to evaluate the likely clinical benefits and harms of various approaches to population breast cancer screening. BreastScreen Australia currently offers free biennial two-view screening mammography, actively inviting women 50-74 years of age (extended from 69 years in 2013). Women 40-49 years can attend BreastScreen for free screening, if they request it, but are not targeted by the program.

Structurally, Policy1-Breast comprises a set of sub-models (Figure 1). The population to be simulated and the calendar period are specified for each model run. Women enter the simulation at age 40 years and exit the simulation either by dying or reaching the end of the calendar period modelled (Figure 2). Simulated time generally advances in one-year steps however multiple events can happen within one calendar year (e.g. screening episode and death) on a weekly basis (programmed via an 'event scheduler').

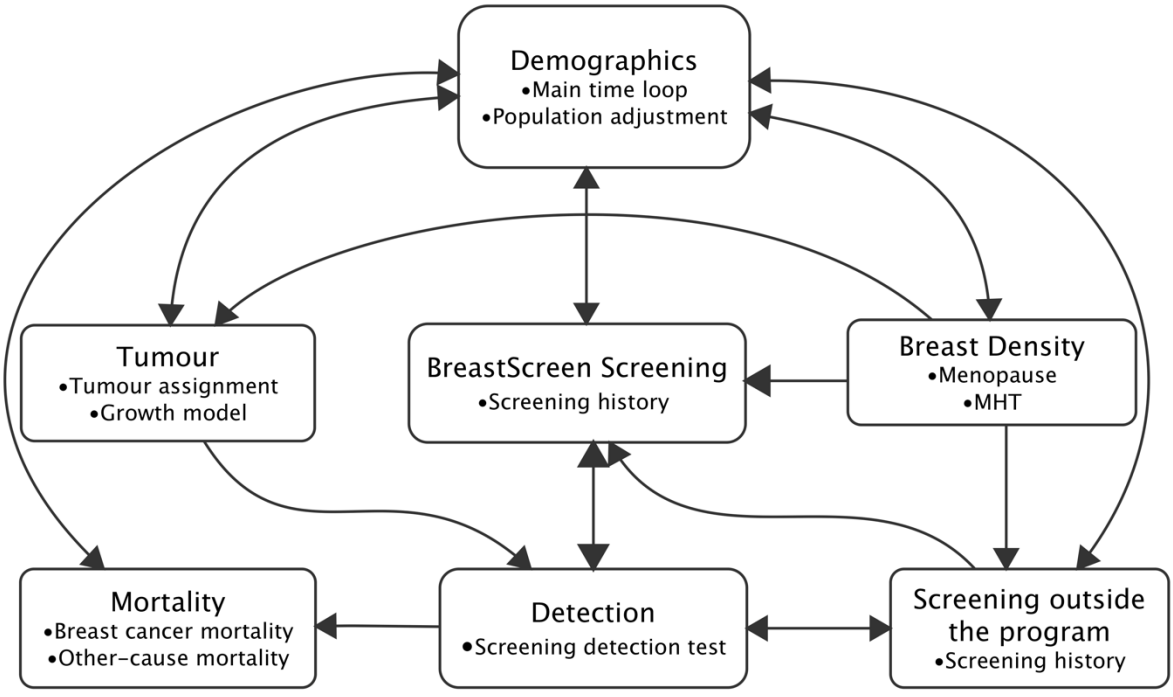

Figure 1: Policy1-Breast schematic diagram demonstrating the relationship between sub-models. 'Screening outside the program' describes opportunistic and risk-based screening which is not part of the national screening program.

Policy1-Breast is based on detailed observed data (key data sources are shown in Table 3, and model parameters are described in *Table 4*). To replicate observed distributions of input variables, where possible the parameter values are assigned by randomly sampling from observed individual or aggregate data (rather than modelling from summary statistics of distributions).

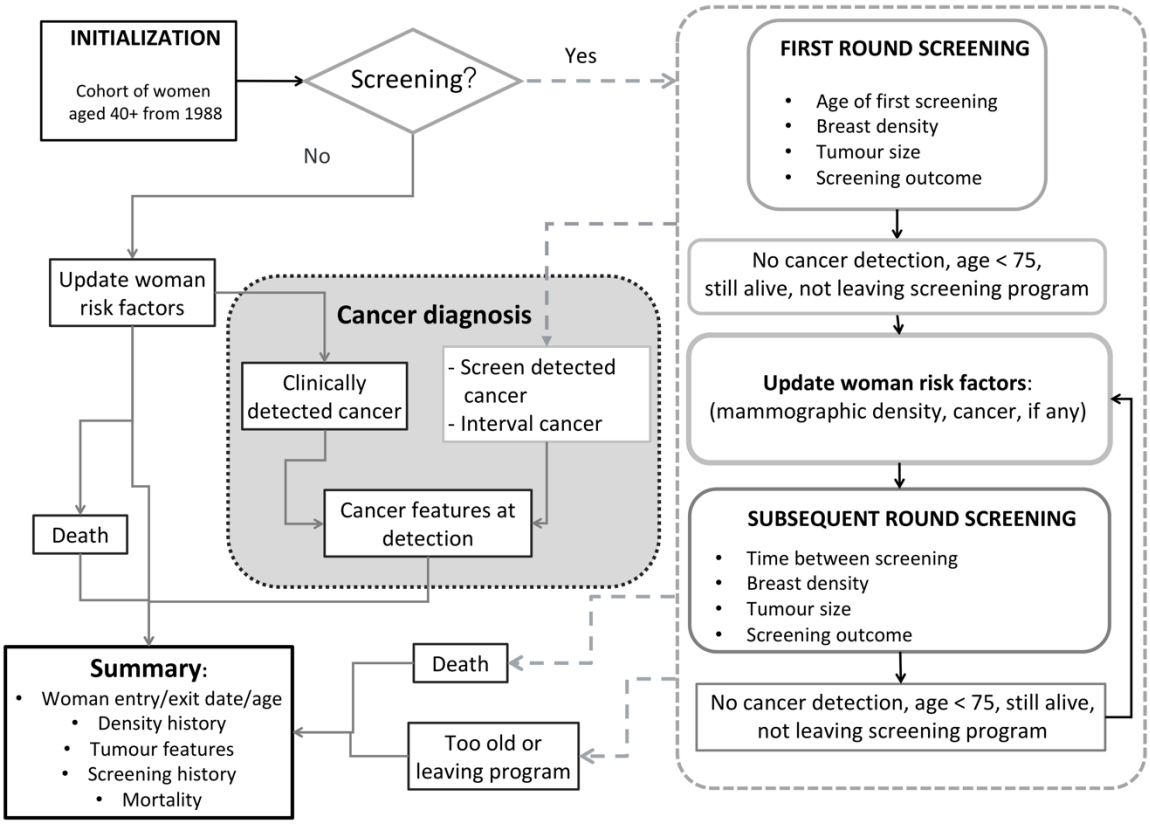

Figure 2: Progression of individual women through the model. The population is initialised for women aged 5+ years starting from 1988 (the baseline cohort), and then progressed through to the age range and calendar period required for each model run, with new individuals joining the simulation to replicate births as well as a 'migration effect' in older age groups to maintain the observed age female population distribution by year.

2 Model specifications

2.1 Demographics

To reproduce the age distribution of Australian women by year, the female estimated resident population by 5-year age group as at 30 June each year is interpolated to age by single year for women aged 40+ for the first year modelled, and women aged 40 for subsequent years. This age-year distribution is then used to specify the age-year distribution of the population modelled, scaled as appropriate to the total number of women simulated. This leads to complete agreement between observed and simulated age-year distributions (Figure 3).

The maximum age of recorded death in Policy1-Breast is 99 years. Simulated women exit the simulation at age 100.

Population projections used ABS estimates which aim to reflect current Australian trends in fertility, mortality and migration.

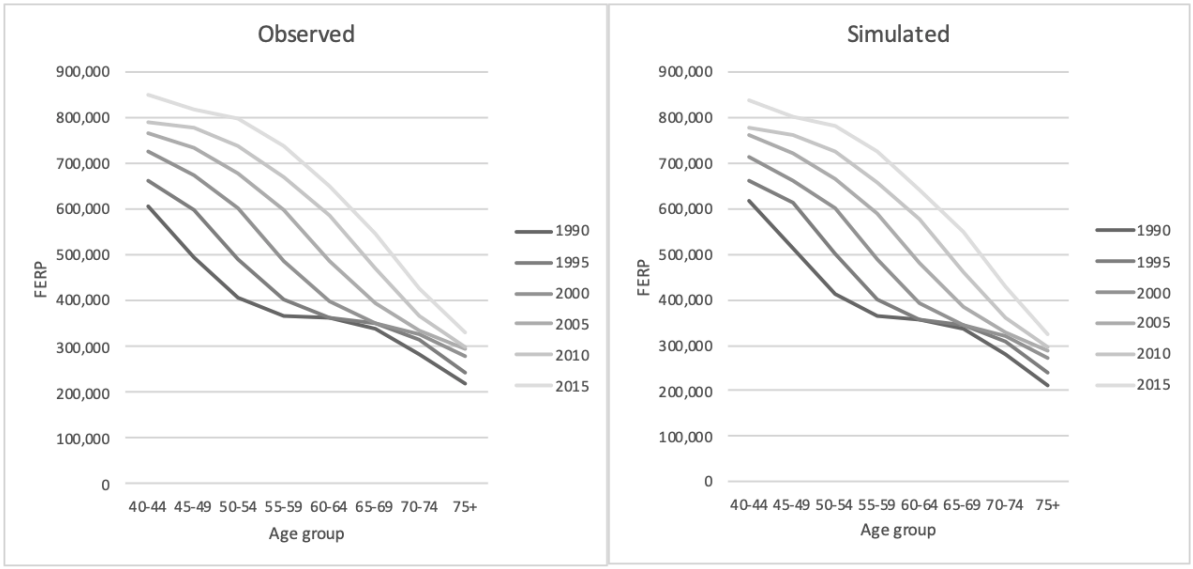

Figure 3: Observed and simulated Female Estimated Resident Population (FERP) by 5-year age group and selected calendar years. Simulated outputs are scaled up to represent national figures.

2.2 Breast density

Breast density is specified through three key variables: an 'upper density path', 'lower density path' and a 'density path locator' (Figure 4). An example is shown where breast density is defined as percent density. The gap between the paths represents the potential range of breast density values for any age. The path locator is a smooth, continuous function so that a woman's percent density shifts smoothly between the upper and lower paths over time.

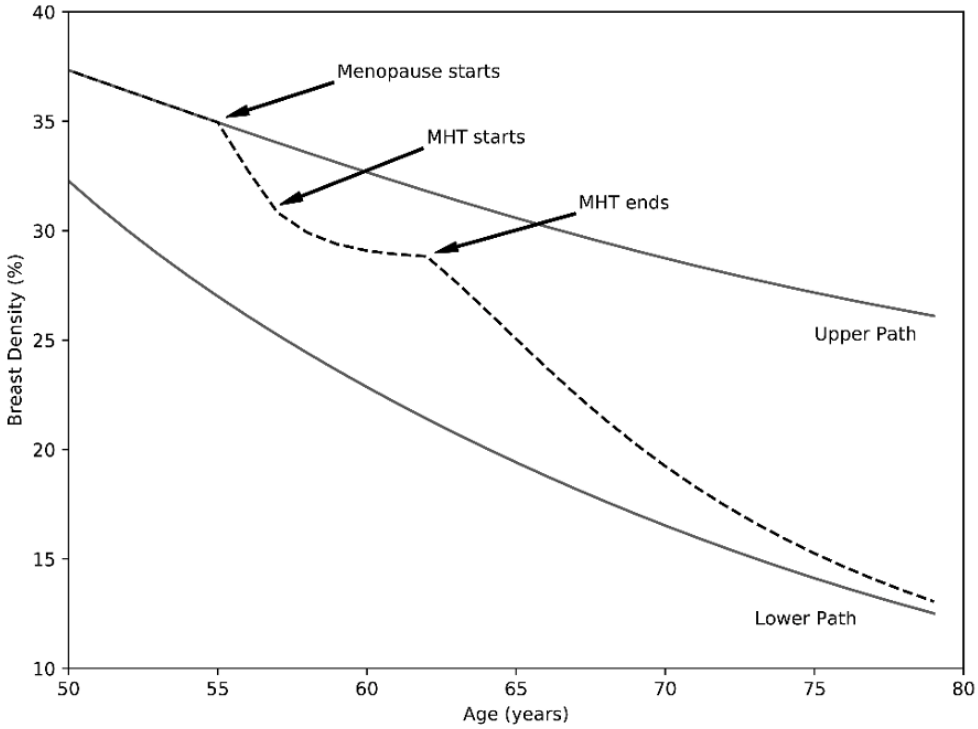

Figure 4: Example of breast density assignment to a woman in the Policy1-Breast model. Lifetime breast density rank is assigned at initialisation. The shape of the upper and lower breast density paths follows observed mean associations between breast density and age, and the gap between the paths represents the potential range of breast density values for any age

according to menopausal status and menopausal hormone therapy (MHT) use. Women are assumed to drop from their upper path after surgical or natural menopause, and then shift back towards their upper path during menopausal hormone therapy (MHT) use. This effectively simulates life-course breast density as having a strong genetic component, with a bounded degree of short-term variation due to endogenous or exogenous hormonal exposures.

The shape of the density paths is intended to capture observed mean associations between breast density and age. The particular value assigned from that range is specified by the density path locator, which is a function of age, menopause and hormone therapy use. The density path locator takes the value of 1 to specify the upper density path, and 0 to specify the lower density path. For a density path locator of 0.5 at age 55 for example, percent density is halfway between the upper and lower density paths at that age.

A woman's menopausal status and MHT use determine which breast density values she adopts from the range of values determined by her assigned upper and lower density paths for a given age.

Menopause is defined as either natural, i.e. occurring naturally or as surgical menopause, as a consequence of hysterectomy, with the latter occurring first. At initialisation each woman is assigned to undergo either natural or surgical menopause, and then she is assigned an age of onset for that menopause based on (a) the observed proportion of women who have surgical menopause, versus natural menopause and (b) the age at menopause, according to menopause type (see Table 5).

Life-course MHT use is assigned as a function of menopause type and age of menopause onset, through three parameters: (i) the proportion of women that ever use MHT, according to menopause type; (ii) the age at which MHT use begins and (iii) the duration of MHT use, according to menopause type (see Table 5). As simplifying assumptions, women use MHT for one continuous period; the formulation and dose is not modelled; the age at starting MHT and duration of MHT use are independent and any MHT use commences at the onset of the menopause (either natural and surgical, whichever comes first). No calendar effects in MHT use are modelled.

The model for breast density was calibrated using observed data for women aged 50-69 years, using screening data from the film mammographic screening epoch, with breast density measurements performed using Cumulus (Table 5). First, the breast density distribution of the cohort data was reproduced. Breast density changes over time were then estimated using a dataset of sequential screening mammograms, accounting for changes in breast density according to MHT uptake, continuation and cessation between screens. The functions were then calibrated to observed cumulative distribution functions of observed and simulated breast density according for MHT use and screening round (first/subsequent). The resulting model shows good accordance with observed associations between breast density, age and MHT.

2.3 Screening participation

Australian women are invited to attend BreastScreen on turning 50, with most women directly invited via their records held on the Electoral Roll. This means participation rates are highest in women aged 50-54 in the established program (e.g. 64.5% in 2014). Women aged 40-49 are also eligible to attend but have lower and less regular participation than women in the target age range (28.6% participation in 2015-16). All screening episodes are modelled

from age 50, commencing with subsequent round screening for participants who first attended screening in their 40s.

The participation rate of women enrolled between 40 and 49 is used as a first order calibration tool: given observed first round participation rates and rescreening rates for women between 50 and 54 years old, we add extra participants (enrolled before they turn 50) to match the observed all-round participation rates of women between 50 and 54 years old. The number of screening rounds elapsed by the time they turn 50 is extracted from a distribution based on the observed trend followed by women of that age interval in the lifepool cohort.

Women are specified as eligible to enter the simulated BreastScreen program between the ages of 40 to 69 years provided they do not have a clinically detectable cancer at the time of their first screen; this reflects BreastScreen policy, although in practice a small proportion (around 3%) of screening participants report a breast lump or nipple discharge at their first screen which can be indicative of cancer. In line with national policy individuals diagnosed with screen-detected or interval cancers can re-join the screening program 5 years after diagnosis.

For women aged 50-69, age at first screen is assigned directly from observed national first-round participation rates by calendar year and 5-year age group (interpolated to 1-year age group). The probability of rescreening is initially assigned using observed rescreening rates in a BreastScreen Victoria research dataset (Table 5) by 5-year age group and round (according to first, second and third-and-subsequent rounds). Screening interval is assigned according to screening round (after first, second, and third-and-subsequent rounds) by drawing from observed distributions in the lifepool cohort, which has collected longitudinal screening participation from BreastScreen Victoria for over 40,000 women. Modelled individuals can cease BreastScreen participation, however, they can re-join the program later in their life mirroring 'lapsed attenders'. This simulated screening participation is then calibrated to fit to observed participation rates by age group and year, accounting for population heterogeneity in screening behaviour.

For modelling year to come, we assumed a constant participation rate until year 2030. This rate is specific for each 5-year age-group for women between 50 and 69 years and it was based on the average value of the observed participation rates over the last five years of available data.

2.4 Invasive breast cancer

To assign lifetime risk to individuals in the simulation, each woman's risk is computed correlating the lifetime risk for her age cohort (for the year in which she is aged 40) with her assigned breast density rank, so that women with higher breast density have a higher lifetime risk of breast cancer. The degree of correlation is specified using observations from the lifepool cohort (Table 2). Modelled life-course breast density increases temporarily with menopausal hormone therapy (MHT) use, so this indirectly generates an association between menopausal hormone therapy (MHT) and breast cancer risk.
The resulting lifetime risk P(BC|BD) is used to determine whether an individual will develop invasive cancer by assigning a random number $X \sim U(0,1)$ and if $X \leq P(BC)$ that individual is assigned to develop breast cancer. For women assigned to develop breast cancer, age at

onset (tumour initiation) is drawn from a negatively skewed distribution. The model does not account for cancers in women under 40 years of age, which were almost 5% of all diagnoses in Australia in 2015.

Natural histories of cancers that started developing prior to the first calendar year simulated (i.e. 1988) are modelled using a dedicated module ('Cancer History Generator' (CHG)). The CHG simulates, for each individual alive in 1988, an independent extra time loop starting in the calendar year that their breast cancer was initiated and ending at the start of main time loop in 1988, so that some women have tumours under progression at the start of the full simulation, 1 Jan 1988. The subsequent progression of those cancers is then computed as for other tumours in the model.

Tumour onset is determined by age at tumour initialisation and growth rate. Parameters for these variables were calibrated simultaneously to replicate national breast cancer diagnoses by age from the pre-screening epoch (1982-1990), with the sub-model for screening effectively turned off. Once age at inception is assigned to a woman, that is used to estimate the size at which her cancer would become symptomatic (before modelling any overlay of screening). For this purpose, we use a joint distribution of age and size data at clinical detection obtained from de-identified records of breast cancers diagnosed in 1989 in the state of Victoria, just prior to the introduction of the screening program. As currently modelled, women assigned a tumour inception age can die of other causes before reaching their inception age. Although we verified this is a rare event in this model, this could lead to a slight overestimation of the lifetime risk used to assign tumours.

To model invasive breast cancer tumour growth, the growth function coefficient is assumed to be a constant, sampled from a log-normal distribution (Table 5). The growth function is specified as either linear or parabolic, with one fifth of cancers assigned to parabolic growth to characterise substantially more aggressive tumours. Tumour size is defined by the diameter at the widest point of the dominant (largest) lesion for all cancers (as per cancer registry data). Only the dominant tumour is modelled and, as a simplifying assumption, multi-focal tumours are excluded.

All tumours are assumed to progress and become clinically detectable (unless a woman dies of another cause prior to detection). However, assigned growth rates include some very low values, essentially replicating tumours in stasis.

Tumour growth is assigned assuming a correlation between growth rates, inception age, and clinically detectable size, so that larger clinically detected tumours and younger inception ages have higher growth rates on average.

2.5 DCIS

In the real world, some proportion of DCIS in the population will progress to invasive breast cancer if left untreated. All detected DCIS in Australia are treated, so modelling the natural history of DCIS progression to invasive breast cancer is complex. As a simplifying assumption for the current model, DCIS is assigned only to women who do not develop breast cancer, and transition from DCIS to invasive breast cancer is not yet modelled. Risk of DCIS is assigned using the same method as for invasive breast cancers, with a lower average lifetime risk. DCIS tumour growth, where tumour size is defined by the lesion length at the longest point of the dominant lesion (as per cancer registry data), is assumed to be linear, using a different growth rate distribution characterised by lower rates and not dependent on any factors (seeTable 5). As for invasive breast cancer, stasis is not directly modelled however some DCIS are assigned

to grow very slowly. The parameters of the growth rate distribution and average lifetime risk were jointly calibrated to replicate the observed incidence rates of screen-detected DCIS. Incorporating DCIS progression to invasive breast cancer will be the subject of future model development.

2.6 Risk stratification

- Policy1-Breast includes two modules related to breast cancer risk assessment and stratification: *risk_assess*, and *risk_factor*. Risk stratification allows allocation of breast screening intervals (1, 2, or 3 yearly) and screening technology according to subject risk categories. (e.g. low-risk, population-risk, moderate-risk, and high-risk). Women are assigned an underlying lifetime risk (see Invasive breast cancer) and risk category, which remains constant. Clinical risk estimates and risk categories are based on the underlying lifetime risk, but with sampled random noise (error) introduced to reflect the imperfect nature of risk assessment in practice. Clinical risk estimates may change over a woman's lifetime.
- Risk assessments are only performed during BreastScreen appointments. When risk stratification is active, the first risk assessment is scheduled for the same date as the first appointment through BreastScreen during screening history initialisation. Assignment of risk groups/strata depends on the number of strata specified within the configuration file (range of 2-4 strata), and specified values for risk group cut-offs (currently set within code, could be additional input file). An alternative method using percentiles is included, where women are instead allocated to risk groups proportionally according to where they fall within the overall distribution of breast cancer risk in the population.

The frequency of risk assessments is set according to the specified time interval in the input configuration file. Risk assessments are scheduled for the next BreastScreen appointment that is within 1 calendar year of the time interval. If the risk assessment would fall on a year that is between screens, then the model chooses the conservative option of performing the assessment 1 year earlier. fIn practice, women may present to BreastScreen and volunteer new information that modifies their clinical risk estimate (such as a new breast cancer diagnosed in a family member). This event is simulated by allowing 2% of women to have a triggered risk assessment at any BreastScreen appointment that occurs prior to their next scheduled risk assessment.

Table 1. Risk group definitions.

2 risk strata	**Risk cut-offs** Lifetime	5-year	10-year
1. Population-risk 2. High-risk	<30% ≥30%	<1.67% ≥1.67%	<9% ≥9%
3 risk strata			
1. Population-risk 2. Moderate-risk 3. High-risk	<18% 18-29% ≥30%	<1% 1-1.67% ≥1.67%	<3% 3-8% ≥9%

4 risk strata			
1. Low-risk	<7%	<0.5%	<1%
2. Population-risk	8-17%	<1%	1-3%
3. Moderate-risk	18-29%	1-1.67%	3-8%
4. High-risk	≥30%	≥1.67%	≥9%
Cut-offs in red need to be discussed (placeholder values). May need to change with age at assessment.			

Assumptions

1. Risk assessment is only performed by BreastScreen (that is, no stratification will occur for women who have opportunistic screening)
2. Risk assessment will only occur alongside BreastScreen appointment
3. Individual error for clinical risk estimates is constant for all risk assessments

2.7 Cost-effectiveness

The Policy1-Breast model health economic component combines clinical outcomes with cost estimates and health state utility scores. This enables cost and cost-effectiveness analyses of the current or alternative breast cancer screening programs, while also generating estimated clinical outcomes in detail.

Health economics outcomes will be analysed by evaluating risk-based screening options according to their potential health benefits (effectiveness) and value (cost), reporting population-level outcomes rather than potential individual benefits.

We will compare differences in costs and differences in effectiveness for each modelled scenario, using incremental cost-effectiveness ratios (ICERs), which are a common measure used in health economics for comparing interventions.

The ICER is calculated by dividing the difference in costs by the difference in effectiveness $\left(\frac{C_{intervention}-C_{comparator}}{E_{intervention}-E_{comparator}}\right)$.

The ICER can be presented visually on the cost-effectiveness plane (Figure 5).

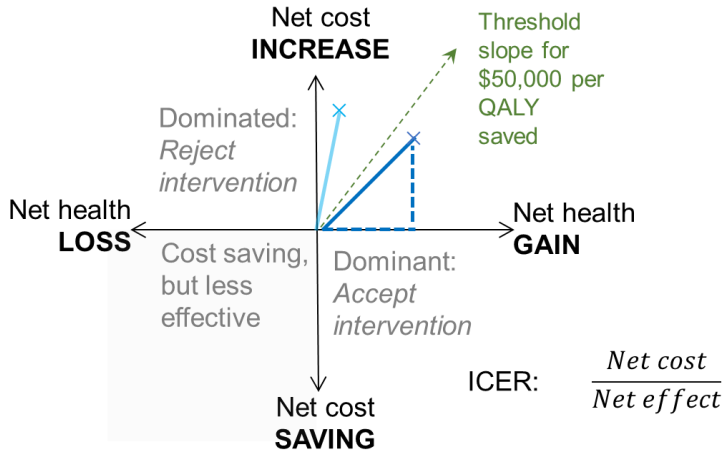

Figure 5. Incremental cost-effectiveness plane. If the ICER falls below the slope for the willingness-to-pay threshold, then the intervention is likely to be considered cost-effective. There is no fixed threshold for funding of health interventions in Australia,

however there is thought to be an implied willingness-to-pay threshold of $50,000 per quality-adjusted life year saved (Wang et al. 2018).[3]

Effectiveness will be reported either as life years saved, or as quality-adjusted life years saved (QALYs). QALYs are the number of life years saved adjusted for any reduction in quality of life, such as a temporary decrease after receiving a false positive screening result, or a prolonged decrease due to a breast cancer diagnosis.

Costs for economic evaluations can include:

- healthcare costs related to direct medical expenses (e.g. medication, hospital admissions)
- indirect medical expenses (e.g. overhead building costs, administration)
- productivity losses/gains as a result of changes in morbidity and mortality
- patient out-of-pocket costs and time, and other related costs.

Inclusion of a specific costs depends on the perspective taken, such as a healthcare payer perspective (medical costs only), or a societal perspective (medical costs, patient costs, productivity losses), as well as data availability.

To help determine the approach to be used in ROSA modelling, we note the range of methods described in a recent systematic review of cost-effectiveness studies for risk-based breast screening by Khan et al, 2021. There were 10 studies included, which used a variety of risk stratification methods and risk group classifications (Figure 6). All studies used simulation models for their evaluations.

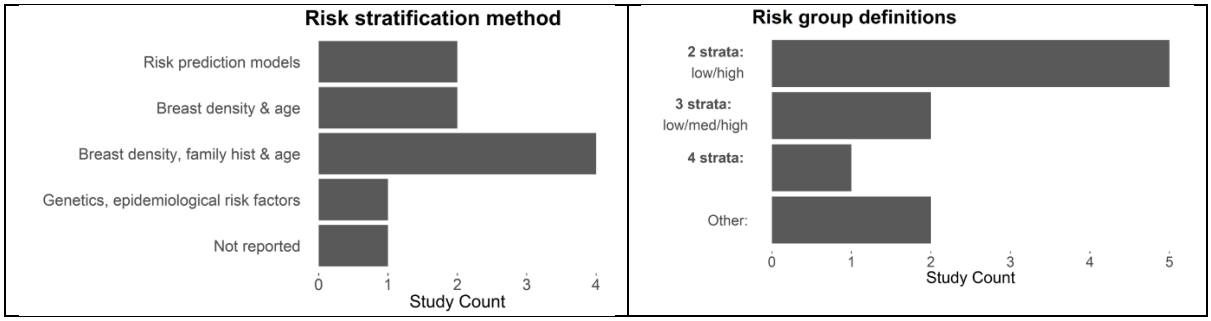

Figure 6. Methods used for risk stratification, and risk group strata, as used in cost-effectiveness studies reported in Khan et al.

The ten studies evaluated a wide range of screening protocols, population characteristics, and screening outcomes. Digital mammography was the most common screening technology evaluated, either alone or in combination with film mammography, ultrasound and/or breast MRI (Figure 7). Most studies (7/10) included interventions related to screening start/stop age, and/or differing screen intervals. The remaining studies evaluated using a different screening technology (1/10), or a different technology combined with screening intervals (2/10).

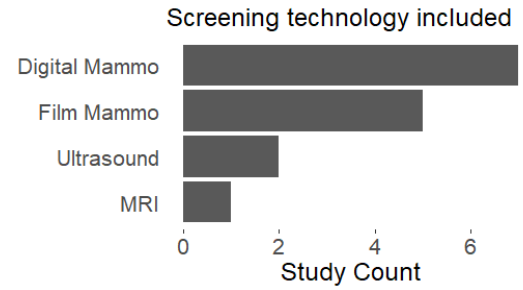

Figure 7. Screening technologies included in cost-effectiveness studies reported in Khan et al.

Generally, tailoring screening frequencies to different risk strata was considered cost-effective compared to both no screening or age-based screening (Table 2). Supplemental ultrasound for high-risk women improved health outcomes, and MRI for women with high breast density was considered unlikely to be cost-effective. Direct comparisons between studies is problematic due to the variation in screening protocols and risk group definitions used.

Table 2. Cost-effectiveness outcomes for risk-based screening studies.

Risk-stratification method	**Comparators**	**Outcome summary**
Mammographic density and age **3 studies**	No screening / Age-based screening	Low-frequency screening (3 yearly) in lower MD with higher-frequency screening (1 or 2 yearly) of higher MD women was cost-effective.
Multiple risk factors **7 studies**		Low-frequency (3-4 yearly) for low-risk women, with 2-yearly for medium-risk, annual for high-risk was cost-effective. Addition of ultrasound to DM yielded higher health benefits compared to DM alone in high-risk women. Addition of MRI for high MD women increases cost considerably, unlikely to be cost-effective.

Cost-effectiveness will differ between health systems, populations, and established cost thresholds for each QALY saved.

2.8 Modelled clinical and health economic outputs

Outputs that can be reported include:

- Tumour stage at diagnosis (size, grade, nodal involvement)
- Mode of detection (screen-detected, interval, other)
- Recalls to assessment, false positive recall rates, program sensitivity
- Mortality
- Treatment patterns
- Overdiagnosis
- Costs and cost-effectiveness (including quality adjusted life years)

[can be described in terms of age groups, time periods, birth cohorts etc]

3 Model validation

Policy1-Breast has been validated against observed data for various outcomes by age group and calendar time, including screening participation, population cancer diagnosis rates, and cancer diagnoses by mode of detection (screen-detected, interval or other). For example, modelled (simulated) BreastScreen participation trends accord well with observed data by screening round and by age group (Figure 8).

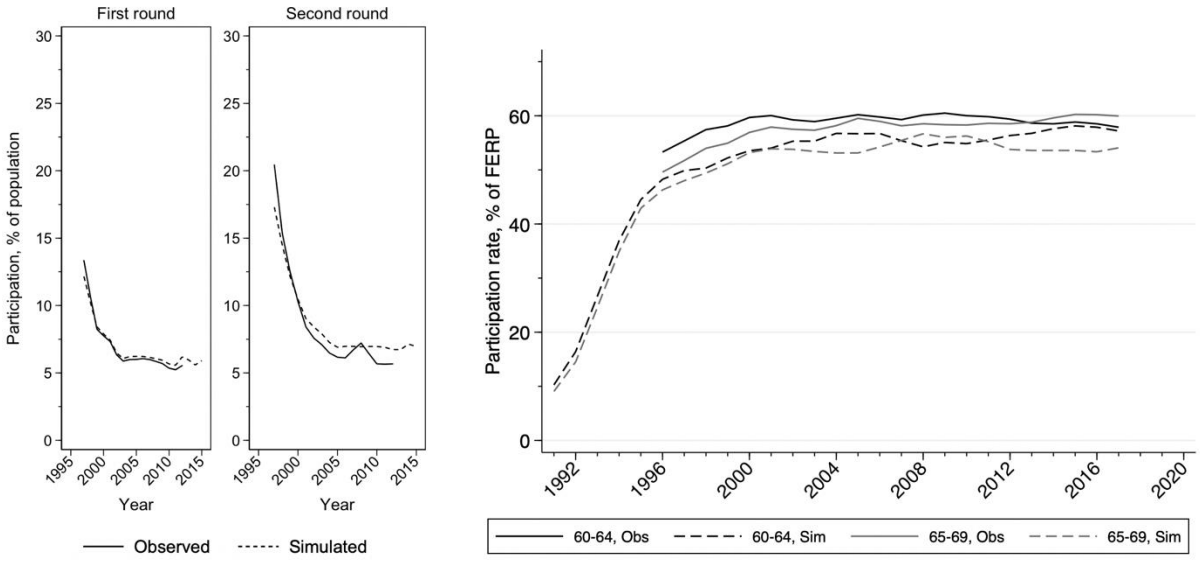

Figure 8. Comparison between observed (solid lines) and simulated (dashed lines) participation rates. The left two plots show the comparison of observed versus simulated participation rates by round (all ages). The right plot shows observed and simulated participation (all rounds) by 5-year age group (60-64 years and 65-69 years as examples).

Figure 9 shows good concordance between simulated and observed cancer detection rates (invasive breast cancers) by mode of detection and year up to 2015.

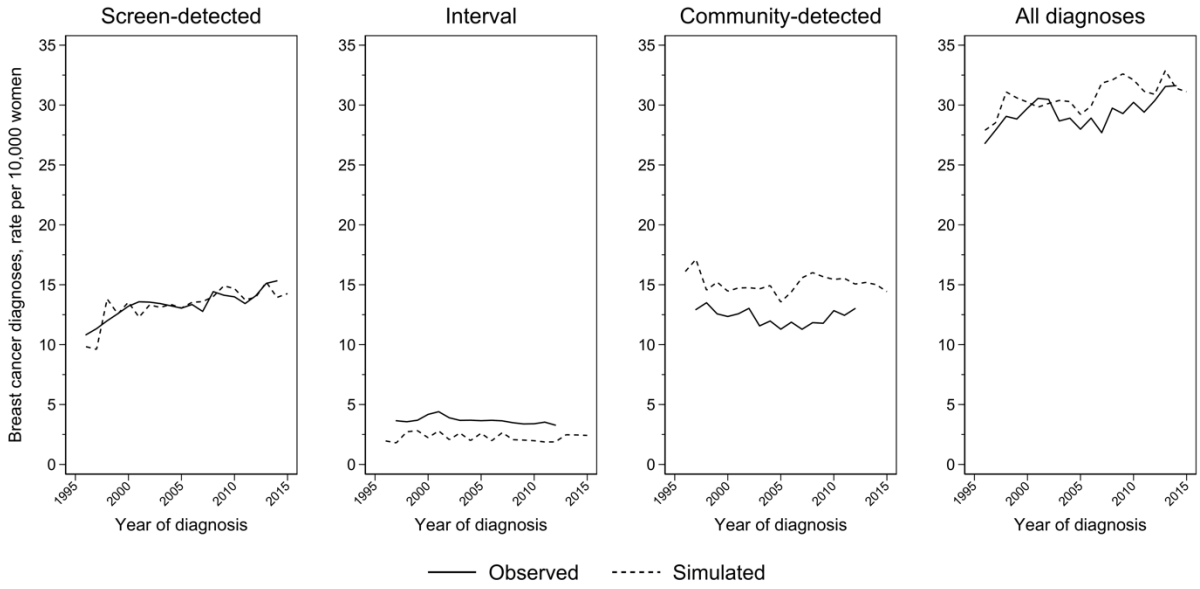

Figure 9. *Comparison between observed (solid lines) and simulated (dashed lines) breast cancer detection rates, by modality of detection (from left to right: screen-detected, interval and community-detected cancers) and for all detection modes (right plot).*

The model also produces a good fit between simulated and observed false positive screens by round (Figure 10).

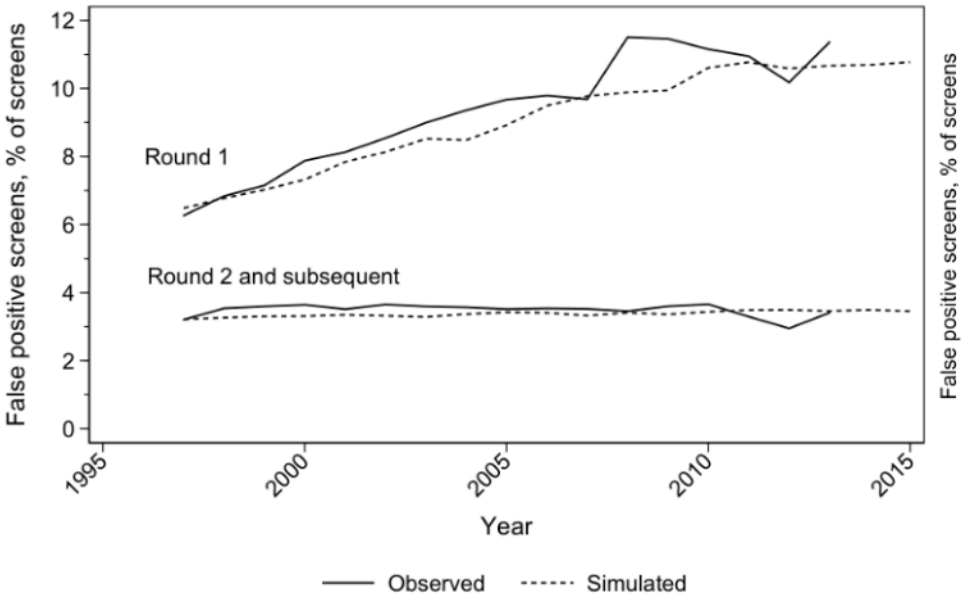

Figure 10. Observed (solid line) and simulated (dashed line) false positive screens by round (first and subsequent) as a percentage of the total number of screens performed in any calendar year.

4. Data sources

Policy1-Breast is based on detailed observed data; selected data sources are presented in Table 3. A summary of parameter dependencies and specifications for the model are shown in *Table 4* with additional parameters in Table 5 and key model calibrations and observations are presented in Table 6.

Where possible, parameter values are assigned by direct replication of observed distributions rather than modelling from summary statistics of those distributions. This is done by random sampling from observed individual or aggregate data values. These observed distributions can also be adjusted to simulate alternative scenarios while maintaining similar distribution patterns. For example, to simulate a policy of targeted annual screening, the observed distribution of time between biennial screens could be scaled by 0.5 and constrained to a minimum of 12 months, reflecting real-world scenarios of screening appointment scheduling and client adherence to recommended screening intervals.

Table 3. Data sources for key model input parameters.

Sub-model	**Data Source**
Demographics	Female estimated resident population by 5-year age group and calendar years 1971-2018 (2)
Mortality	Australian national mortality rates by age and year (all causes mortality) (3) Breast cancer deaths in women aged 50-69, by detection status (4)
Screening	BreastScreen Australia Monitoring Reports up to 2019 version [data from 1997 to 2017] The *lifepool* cohort (5) including breast density measurements and linked BreastScreen and cancer registry outcomes. A cohort of BreastScreen Victoria clients with breast density measurements from 2003-2007 and follow-up data to 2014.

Table 4. A summary of parameter dependencies and specifications in the Policy1-Breast model.

Parameter	**Method**	**Range**
Breast Density	Random extraction from observed breast density distribution.	0-100
Lifetime risk	Random extraction of chance of being assigned BC	P(BC) = 1/8 = 0.125 (+0.0025 per year, from 2004)
Association between lifetime risk and breast density	Lifetime risk + (BD – 0.5)*(0.053/0.5)	P(BC) = 0.072-0.178
Tumour inception age	Extracted from lognormal distribution (mu=3.8,sigma=0.2)	Allowed values: 18-85
Tumour growth rate	Extracted from lognormal distribution (mu=0.95+[10/Inception age], sigma=1.)	NA
Probability of breast cancer death	Polynomial function (es: $-3.67\cdot10^{-7}x^4+2\cdot10^{-5}x^3-3.86\cdot10^{-4}x^2+0.0025x+0.0032$)	0-max, where max is < 1††
Time between screens	Random extraction from observed distributions	NA
Probability of detection	$Pr(s, d) = 1 – [\lambda d\ q_m(s) + (1- \lambda d)\ q_b(s)]$*, where $\lambda = 3$	0-1
Program adherence	Probability of rescreening by round and age group	$P(R_2\|R_1) = 0.83 – 0.94$

*d= breast density; qb(s)=baseline detection function; qm(s)=masking detection function; s=tumour size

†Please note that the programming code is protected intellectual property and not available for distribution.

†† The function shown is for screen-detected cancers and for a cancer of size 10mm at detection. The polynomial function changes in function of the cancer size at detection, hence its max value changes as well. For interval and community detected cancers similar formulas are computed. x is the number of years since cancer detection.

Table 5. Additional parameters and parameter values for the Policy-1-Breast sub-models.

Symbol	**Parameter**	**Parameter Values**
A	Individual's path parameter Assigned as a function of first screen density $d_w^{(1)}$ and first screen age $a_w^{(k)}$ via $b_w^{(k)}$	$A = d_w^{(1)} / (v + (1\text{-}v)\, b_w^{(k)})$, where if $L^{(1)} = 1$, v = 0.5 and if $L^{(1)} = 0$, v = 0.2
$b_w^{(k)}$	Age-dependent parameter defining the gap between upper and lower density paths (modifiable parameter):	$b_w^{(k)} = \text{expit}(z)$ where $z = 3(50 - \alpha_w^{(k)})/50$ and expit(z) = exp(z)/(1+exp(z))
$L^{(1)}$	Path locator at first screen. $L^{(1)}$ is in the set {0,1}, where 0 represents the lower density path and 1 represents the upper density path.	If $a_w^{(k)} < \text{Mw}$, then $L^{(1)} = 1$ (pre-menopausal) If $a_w^{(k)} > \text{Mw}$ and $\varphi w = 0$, then $L^{(1)} = 0$ (post-menopausal, not using MHT) If $a_w^{(k)} < \text{Mw}$ and $\varphi w = 1$, then $L^{(1)} = 1$ (post-menopausal, using MHT)
D_u	Upper density path at screen k:	$Du = A(u + (1 - u)\, b_w^{(k)})$ Where u is the reduction rate of the upper path (u = 0.4) And $_l$ is the reduction rate of the lower path ($_l$= 0.1)
D_l	Lower density path at screen k:	$D_l = A(l + (1 - l)\, b_w^{(k)})$
$L^{(k)}$	Path locator at screen k, defined as some number in the range [0,1], where 0 represents the lower density path, 1 represents the upper density path and if $0 < L^{(k)} < 1$, $d_w^{(k)}$ is between D_L and D_U.	If $a_w^{(k)} < M_w$, then $L^{(k)} = 1$ If $a_w^{(k)} \geq M_w$ and $\theta_w = 0$ and $\varphi w = 0$, then $L^{(k)} = L_P = \exp(-0.2(a_w^{(k)} - M_w)$ If $a_w^{(k)} \geq M_w$ and $\theta_w = 1$ and $\varphi w = 1$, then $L^{(k)} = L_U = 1 - (1 - L_P) \exp(-0.2(a_w^{(k)} - H_U)$ If $a_w^{(k)} \geq M_w$ and $\theta_w = 1$ and $\varphi w = 0$, then $L^{(k)} = L_U \exp(-0.2(a_w^{(k)} - H_c)) + L_P(1 - \exp(-0.2(a_w^{(k)} - H_c)$
$d_w^{(k)}$	Breast density at screen k.	$d_w^{(k)} = L^{(k)} D_u + (1 - L^{(k)}) D_l$
ps	The proportion of women that have surgical menopause, versus natural menopause	Natural menopause Surgical Menopause ps = 20%
an	Age at menopause	$an \overset{d}{\sim} N(51,5)$
hp	The proportion of women that ever use MHT. Randomly assigned from the estimated probability of ever using hormone therapy, according to menopause type.	Probability of MHT use following natural menopause: hp = 0.50
ha	Age at starting MHT use Assigned as a function of menopause age.	MHT use starts with menopause ma: $ha = ma$
hd	Duration of MHT use Sampled from the observed distribution, as a function of menopause type.	For CDF of observed duration of hormone therapy use (years) N(8.0,4.6),and for upper censoring adjustment c = 3, and $p \overset{d}{\sim} U(0,1)$: $hd = F^{-1}(p) + c$ $hd = \text{F-1 (p)} + c + \max(50\text{-aw}^{(k)}, 0)$
$q_b(s)$	Baseline detection function	$q_b(s) = 1 - F(7.5, 4.5)$
$q_m(s)$	Masking detection function	$q_m(s) = \min(1, e^{-0.0002 s^3})$
$Pr(s,d)$	Probability of detection	$Pr(s, d) = 1 - [\lambda d\, q_m(s) + (1 - \lambda d)\, q_b(s)]$, where $\lambda = 3$

$a_w^{(k)}$: age at screen k; $d_w^{(k)}$: percent density at screen k; M^w: age at menopause (surgical or natural); ϑw: MHT ever use indicator (Never uses = 0, Ever uses = 1); H_u: MHT uptake age; H_c : MHT cessation age; φw: MHT current use indicator (Not using = 0, Current use = 1)

Table 6. Overview of key model calibrations and validations.

Modelling stage	**Relevant inputs**	**Parameter tuned**	**Comparison between simulated and observed data**
Calibration	Lifetime risk	Lifetime risk	Cancer incidence by year (pre-screening era)
Calibration	Lifetime risk correlated with breast density, size at clinical detection	Tumour growth rate Inception age	Cancer incidence by year and 5yr age group (pre-screening era)
Calibration	First round participation rate, rescreening rates by round	Rescreening rates by round	Screening participation rate by year and 5-year age group
Calibration	Observed survival according to mode of detection and tumour size	BC survival according to mode of detection and tumour size at detection	Hazard ratios of mortality by mode of detection
Calibration	Observed breast density by MHT use, age at menopause	Upper and lower density path Path locator	CDFs of breast density by MHT use and screening round
Validation	All of the above	Not applicable	Cancer detection rates by year, 5yr age group and modality of detection
Validation	HRs of BC mortality by mode of detection	Not applicable	Breast cancer mortality by year and 10yr age group

CDF = Cumulative Distribution Function.

References

1. Pathways to a cancer free future. https://www.cancercouncil.com.au/research/research-we-conduct/about-pathways/
2. Australian Bureau of Statistics (ABS) 2019. Australian Demographic Statistics 'Table 59. Estimated Resident Population by Single Year of Age, Australia', time series spreadsheet, cat. no. 3101.0, viewed 28 January 2020, https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3101.0Sep%202019?OpenDocument
3. Australian Institute of Health and Welfare (AIHW) 2019a. GRIM (General Record of Incidence of Mortality) books 2018: All causes combined. Canberra: AIHW.
4. Australian Institute of Health and Welfare (AIHW) 2019b. GRIM (General Record of Incidence of Mortality) books 2018: Breast cancer. Canberra: AIHW.
5. Lifepool Study. Available from: http://www.lifepool.org/