18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

A microsimulation model of the BreastScreen Australia program

Nickson, CA,[1] R. Watson[2], A.M. Kavanagh[1]

[1]Key Centre for Women's Health in Society, University of Melbourne, Victoria.
[2]Department of Mathematics and Statistics, Faculty of Science, University of Melbourne, Victoria
Email: cnickson@unimelb.edu.au

Abstract: The BreastScreen Australia breast cancer screening program uses biennial mammography (breast x-rays) to screen for breast cancer in the non-symptomatic population. Breast density is a measure of how much of the mammogram is occupied by usually normal breast tissue that is distinctly radiodense. The performance of the screening program is currently inferior for women with high breast density: they are found to have larger tumours at detection and higher rates of cancers detected between scheduled screens. More frequent screening for women with higher breast density might bring some benefit.

Breast density varies substantially in the population and within women over time, and high breast density is associated with breast cancer risk. Given the complex associations between breast density, cancer risk, and individual-level factors such as age, menopause and hormone therapy use, microsimulation modelling is a useful tool to examine the potential benefit of alternative screening strategies for women with high breast density.

We specify a continuous-time, stochastic, single-cohort microsimulation model of the BreastScreen Australia program, which we name the Australian Breast Cancer Screening Simulation (ABCSS). ABCSS is driven by observed data from over 10,000 BreastScreen Australia participants. ABCSS is specified so that it predicts detection rates and tumour size at detection, in women who develop breast cancer and participate in screening. It includes an individual-level model for breast density according to age, menopause and hormone therapy use, and a sub-model for the probability of cancer detection as a function of tumour size and breast density.

We use ABCSS to simulate trials of screening strategies, where more frequent screening is offered to women with breast density above a specified threshold, and this is offset by less frequent screening for women with breast density below a specified threshold such that the number of screening services required does not change. Screening strategies are allocated according to breast density at program entry. We predict changes in program sensitivity and tumour size at detection after six years of screening, for a range of targetting strategies.

We estimate that under the optimal strategy tested, program sensitivity could improve from 68% to 75% for women with high breast density, with a concomitant reduction from 82% to 77% for women with low breast density, and a program-wide improvement from 74% to 76%. Under that strategy, the proportion of cancers that are small (<15mm) would increase from 64% to 67% for women with high breast density; reduce from 73% to 69% for women with low breast density; and the program-wide figure would not change (68%).

We conclude that BreastScreen Australia may be distributed more equitably if women with high breast density are screened more often and women with low breast density are screened less often. We are currently extending ABCSS to model the entire screened population and thereby estimate changes to false-positive rates under alternate screening strategies, which will provide a more complete picture of the potential harms of more intensive screening. ABCSS is also being extended to model alternative screening modalities such as ultrasound and magnetic resonance imaging (MRI), by modifying the sub-model for the probability of detection.

Keywords: *Breast cancer screening, mammography, health systems*

1. INTRODUCTION

BreastScreen Australia is a national program which uses biennial mammography (breast x-rays) to screen for breast cancer in the population. Screening is available free of charge to non-symptomatic women from the age of 40 years and is targeted by age to women aged 50-69 years through direct invitation and re-invitation and through media promotion. Approximately 60% of eligible women aged 50-69 years participate in the program. Mammographic breast density is a measure of how much of a mammogram is occupied by normal breast tissue that is distinctly radiodense (and hence whiter on an x-ray). Both in Australia and abroad, the performance of breast cancer screening is inferior for women with high breast density. As shown by Kerlikowske et al. (1996), van Gils et al. (1998) and Carney et al. (2003), women with high breast density tend to have larger screen-detected cancers and larger cancers detected between scheduled screens (interval cancers), and lower mammographic screening program sensitivity (the number of screen-detected cancers divided by the sum of screen-detected and interval cancers in the screening program).

As suggested by Sala et al. (1998), Yaffe (2004) and Stines et al. (2005), to improve the performance of breast cancer screening programs it may be appropriate to offer more frequent screening or an additional screening modality such as ultrasound or magnetic resonance imaging for women with very dense breasts. Identifying the likely effects of alternate screening strategies is not straightforward, because the reduced screening program performance for women with high breast density is likely to be due to the association between high breast density and a combination of the following three factors:

(i) Increased risk of breast cancer: High breast density is associated with a three to five fold increased risk of breast cancer (as demonstrated by Boyd et al. (1982), Brisson et al. (1989), Lam et al. (2000) and Maskarinec et al. (2005)).

(ii) Masking of tumours by the dense tissue, as confirmed by radiological reviews such as that by 2006 Chiarelli et al. (2006)

(iii) Faster tumour growth: For example, as shown by Sala et al. (2000), compared to women with low density, women with high breast density were more likely to have a Grade III tumour (OR=3.9, 95% CI: 1.4-10.2) and women in the second-highest density category (representing 56% of the sample) were also significantly more likely to have a Grade III tumour (OR=2.8, 95% CI: 1.2-6.2).

It is difficult to separately identify the contribution of each of these effects to the performance of screening programs on early detection of breast cancer. Furthermore, breast density varies substantially in the population and within women over time. On average, breast density decreases with age and the onset of menopause, but less so with the use of hormone therapies.

Given the complex associations between breast density, breast cancer risk, masking, tumour growth rates and individual-level factors such as age, menopause and hormone therapy use, microsimulation modelling is a useful tool to examine the potential benefit of alternative screening strategies targeted according to breast density. We use microsimulation modelling to examine whether the performance of the Australian breast cancer screening program can be improved for women with higher breast density by targetting the length of screening intervals according to breast density.

2. METHODS

We specify a continuous-time, stochastic, single-cohort microsimulation model of the BreastScreen Australia program, which we name the Australian Breast Cancer Screening Simulation (ABCSS). ABCSS is driven by observed data from over 10,000 BreastScreen Australia participants. ABCSS is specified so that it predicts detection rates and tumour size at detection in women who develop breast cancer and participate in screening. ABCSS models life histories for individual women, assigning each of the factors shown in Figure 1 as attributes, some of which can vary over time (e.g. age, hormone therapy use and mammographic density). Where possible, these factors were sampled from observed data. For example, breast density is not routinely measured in the Australian program so we utilise an existing research database of breast density measurements from over 10,000 women who participated in the BreastScreen Victoria screening program, where breast cancer cases were over-sampled to include nearly 3,000 women had a breast cancer detected either at screening or between scheduled screens. Age and hormone replacement therapy use was routinely recorded for the entire sample and for a subset of nearly four hundred women breast density was measured over two consecutive screens.

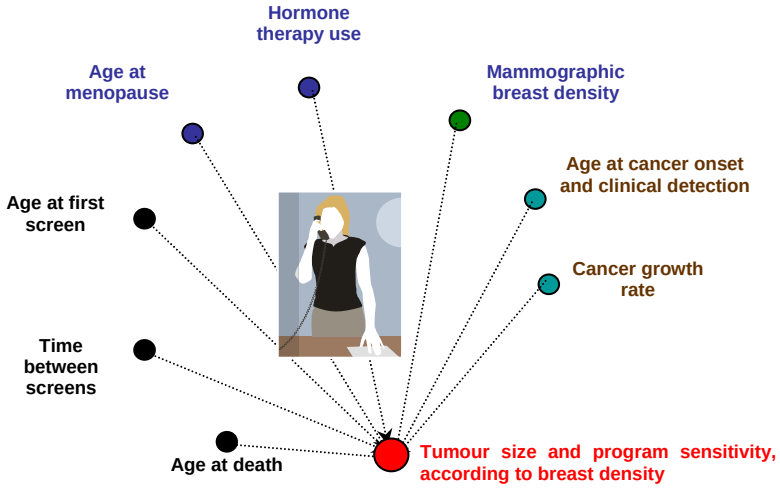

Figure 1. *A conceptual diagram of the factors included in the simulation model ABCSS.*

Breast density was measured for cranio-caudal views of both the ipsilateral cancer-affected breast and the contralateral breast for each study subject, using a semi-automated image-processing technique developed in the early 1990s called the *threshold technique*, which requires a reader to manually set a greyscale threshold that dichotomises the pixels in the digitised mammogram image as either 'dense' or 'non-dense'. This process generates continuous measures of the area (in pixels) of the *breast area*, the *dense area* and the *non-dense area*. *Percent density* is then calculated by dividing *dense area* by *breast area*. The sample and the reading method is described in detail by Kavanagh et al. (2008). Breast density was fitted using the measure *percent density*.

The resulting simulation model ABCSS includes several sub-models which were each specified based on literature reviews and analysis of unit data where available, and then calibrated and validated against observed data. A full description of this process is beyond the scope of this paper. The primary sub-models are:

- The *Breast Density sub-model*, which specifies individual life-course breast density as a continuous function of age, menopause and hormone therapy use. Each woman in the model is assigned upper and lower bounds of her life course breast density, and the location of her breast density between these paths at any particular age is a function of her menopausal status and hormone therapy use.
- The *Breast Cancer natural history sub-model*, which models tumour growth backwards from an assigned clinically detectable tumour which is specified according to tumour size and woman age at detection and sampled directly from observed data from Victoria, Australia prior to the introduction of the screening program. Like Tan et al (2006), we assume that the rate of tumour growth is positively correlated with the assigned tumour size at clinical detection.
- The *Screening Participation sub-model*, which models screening participation using the observed variation in screening intervals. As a simplifying assumption, this model does not incorporate the small amount of targeted annual screening currently in place, and it assumes that all women participate in all scheduled screens.
- The *Detection sub-model*, which specifies the probability of detection by screening mammography as a function of breast density and tumour size.

Key calibration targets in the model were program sensitivity according to breast density and screening round, and tumour size according to breast density and screening round. After achieving a suitably well-fitting model we used ABCSS to simulate trials of screening strategies, where more frequent screening is offered to women with breast density above a specified threshold and this is offset by less frequent screening for women with breast density below a specified threshold such that the number of screening services required does not change. Screening strategies are allocated according to breast density at program entry. We predicted changes in program sensitivity and tumour size at detection after six years of screening, for a range of targetting strategies.

For example, Figure 2 shows an example of targeted screening intervals simulated in our model. Here, numbers in boxes show screening rounds (i.e. first, second, third round screening). The first row shows the

biennial screening intervals of the current Australian program, the second shows how we modelled more frequent screening for women with high breast density, and the third row shows how we modelled less frequent screening for an equal number of women with low breast density, effectively offsetting the additional screens required for women with higher breast density. The additional screen for the high breast density group is allocated soon after the first round of screening in response the particularly high rates of interval cancers detected after first-round screening in that group, theoretically due to a greater problem with masking by breast density in the absence of a comparison mammogram.

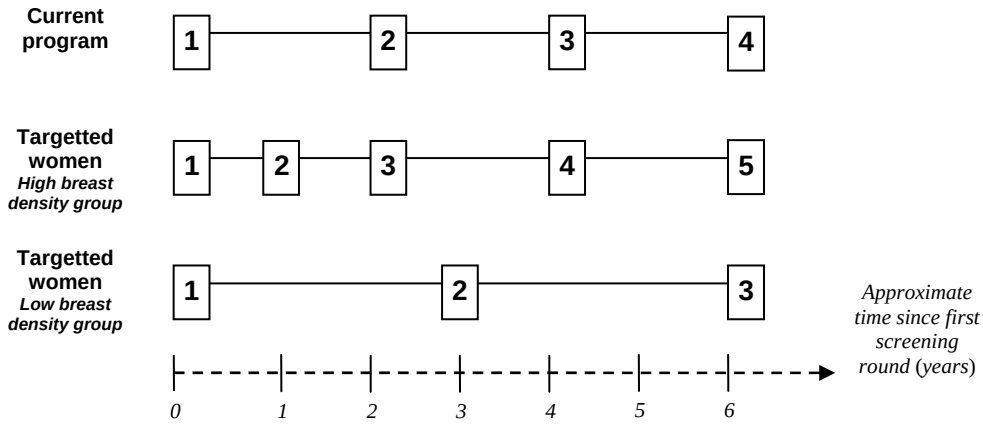

Figure 2: An example of targeted screening intervals simulated in our model, for the first six years of screening participation. Numbers in boxes show screening rounds (i.e. first, second, third round screening).

We assessed strategies where the minimum breast density value for entry into the high breast density group (denoted T_{HIGH}) was the 50th, 60th, 70th, 80th and 90th percentiles in the screened population. Then, for example, where T_{HIGH} = the 90th percentile, the maximum breast density value for entry into the low breast density group was the 10th percentile, and the remainder of the population screened (women with breast density between the 10th and 90th percentiles) were allocated the current screening strategy.

3. RESULTS

The predicted program sensitivity after the first six years of screening is shown in Figure 3, for each decile of the expected baseline (first-round) breast density values. Predicted program sensitivity after six years under the current program varies substantially according to breast density; this is in keeping with the observed data which is available only up to the first 2-3 years of screening. The predicted absolute improvement in program sensitivity resulting from more frequent screening is similar for all breast density groups; however, the reduction in program sensitivity with less frequent screening is less for women with low breast density.

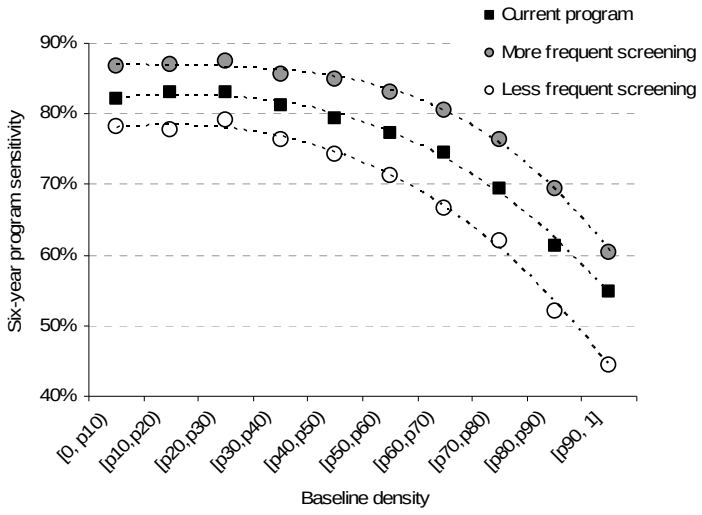

Figure 3. Program sensitivity after six years according to baseline breast density deciles and frequency of screening, fitted with non-increasing lines.

The predicted program sensitivity within targeted groups for various strategies is shown in Figure 4. For each strategy, the model indicates that more intensive screening for women in the high breast density group would improve program sensitivity within that group, and that this would be offset by a reduction in program sensitivity for women in the low breast density group (who would receive less intensive screening). The six-year program sensitivity for the whole screening program would improve slightly with the introduction of targeted screening services, while improving program sensitivity for women with high breast density. Program sensitivity for women allocated to less frequent screening would reduce but would be no lower than the current six-year program sensitivity for all women and no lower than that expected for the high breast density group under more intensive screening.

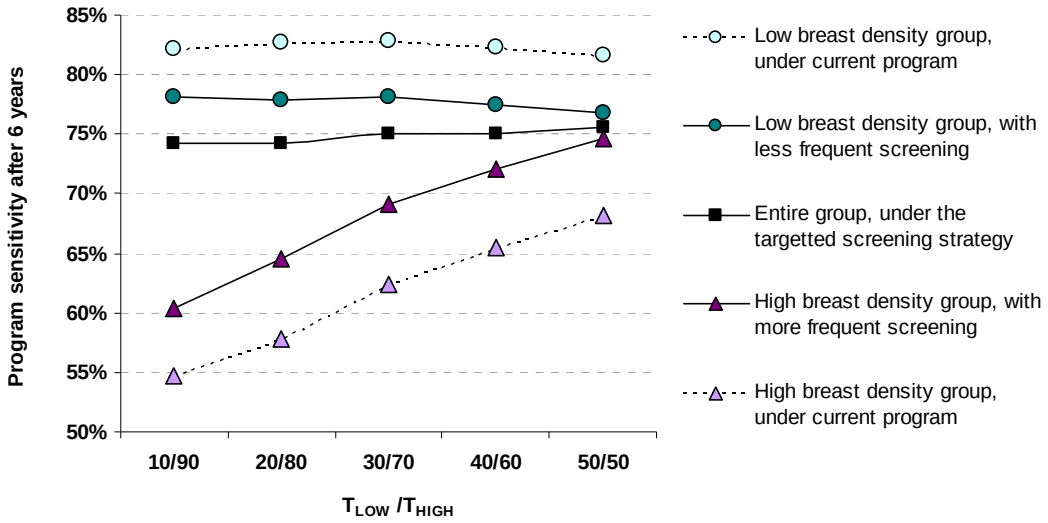

Figure 4. Predicted program sensitivity in a cohort of women screened for six years, according to allocation of more/less frequent screening based on breast density at first round screening. T_{LOW}/T_{HIGH} denotes the screening strategy; for example T_{LOW}/T_{HIGH} = 10/90 describes a strategy where women in the lowest 10% of breast density measures are allocated to less frequent screening, and women in the highest 10% of breast density measures (i.e. above the 90th percentile) are allocated to more frequent screening. Note that under strategy 10/90 the high breast density group comprises women above the 90th percentile of breast density, and under strategy 20/80 90 the high breast density group comprises women above the 80th percentile etc.

If the threshold was set at the median value (p_{50}, which is equivalent to 10% *percent density*) the resulting program sensitivity would be similar for women with breast density above and below the threshold value. Under that strategy:

- Program sensitivity could improve from 68% to 75% for women with high breast density, with a concomitant reduction from 82% to 77% for women with low breast density, and a program-wide improvement from 74% to 76%.
- The proportion of cancers that are small (<15mm) would increase from 64% to 67% for women with high breast density; reduce from 73% to 69% for women with low breast density; and the program-wide figure would not change (68%).
- Women with breast density above approximately 20% *percent density* (the 70th percentile) would require additional improvements in detection to experience the program performance currently achieved for women with lower breast density (as indicated by Figure 3).

4. DISCUSSION

Like all simulation models, ABCSS includes many simplifying assumptions. These are not limitations *per se*, as some assumptions are required in order to specify a parsimonious model of the real-world system. However some assumptions are made due to limited availability of evidence. For example, we assume that breast cancers do not regress or enter stasis. ABCSS would be further refined if stasis and regression were incorporated into the model.

ABCSS has a number of important strengths, namely:

- ABCSS incorporates a continuous individual natural history model of breast density. This enabled the prediction of outcomes for alternative screening strategies across the full range of breast density values. For example, we were able to simulate 'offset' screening strategies where resources are allocated according to percentiles in the population.
- ABCSS is driven by data from the Australian screening program and the Australian population. It captures the screening participation behaviour of Australian women, and the predictions are based on cancer detection rates from the Australian population. Breast density distributions are based on the Australian population, as are estimates of hormone therapy use and age at menopause.
- ABCSS models tumour size continuously, in interaction with the continuously defined *detection sub-model*, which is a function of tumour size and breast density.
- Several variables in ABCSS are directly sampled from observed distributions. For example, age at first screen and the distribution of time between screens are sampled from distributions observed in the existing biennial program. This method simplifies the application of ABCSS for other populations.

Limitations of ABCSS include the following:

- ABCSS does not model false positive screening outcomes. The increased number of screens proposed under the alternate screening strategies modelled will increase false-positive outcomes, but it is not possible to predict the extent of this increase with the current model.
- ABCSS does not model the natural history of *Ductal Carcinoma In Situ*, which is a non-invasive form of breast cancer which is considered to be a precursor lesion for invasive breast cancer that comprises an increasing proportion of cancers detected in the screening program (currently approximately one fifth) and its detection leads to exclusion from the program and a likely reduction in the probability of future invasive breast cancer.
- ABCSS is based on breast density readings taken mostly from first and second round screening, allowing reasonable prediction of outcomes just prior to six years after the first screen as described here. Prediction of longer-term outcomes would require further evidence about breast density and screening program performance in later-round screening. Longer-term targeting would also need to account for within-women variation in life-course breast density.
- ABCSS would benefit from being validated against further observed data than that used to date. Such data are sparse, since no other large sets of continuous breast density measurements exist in Australia except in studies with samples that are not representative all screened women.

Our findings suggest that the current biennial screening program could be delivered with more equity if it is targeted according to breast density. These improvements are expected to be offset to some extent by a reduction in program specificity. In its current form, ABCSS cannot be used to predict program specificity because it models only women who get breast cancer during the course of their participation and so the current findings need to be considered on balance with the likely reduction in program specificity.

In order to implement targeted strategies such as those modelled here, the screening program would need to measure breast density from the first screening round and then provide women with advice about future screening intervals based on that reading. At this time, breast density is measured in the Australian program only for research purposes, and this requires the retrieval of mammograms followed by *threshold technique* measurement of mammographic density by trained readers. Measuring breast density for all women attending first-round screening will require an efficient and reliable reading method that has very high inter-reader reliability, requires minimal training and generates fast results. Screening epidemiologists and program managers should collaborate with image processing specialists to develop workable, automated, affordable systems for measuring breast density.

ABCSS as currently specified is a starting point for further modelling of alternative designs for the Australian breast cancer screening program. For example:

- With modification of the *Detection sub-model*, ABCSS could be extended to simulate the effect of targeted screening using modalities other than mammography, such as ultrasound or MRI.
- ABCSS could be extended to simulate screening in women younger than 50 years, with some further evidence about the relative contribution of faster tumour growth and higher breast density to screening

program performance. ABCSS could also be extended to simulate screening in women older than 70 years, with the addition of a *sub-model* representing the impact of competing risks.

- Breast density is modelled in ABCSS using *percent density*; however, ABCSS may be further refined by modelling masking as a function of percent density and risk as a function of some other measure of breast density, such as *dense area*.
- ABCSS could be used to predict the impact of introducing digital mammography, through the modification of the *Detection sub-model* to reflect improvements in detection observed in studies comparing film and digital mammography.
- The current outcomes generated by ABCSS could be used to predict relative survival using published estimates for survival time according to tumour size at detection and screen-detection or interval detection. ABCSS could also be extended to cost-effectiveness estimates.
- ABCSS currently models a single cohort. It could be modified to model multiple cohorts, which would capture, for example, changes in hormone therapy use and screening participation over time.

5. CONCLUSIONS AND RECOMMENDATIONS

BreastScreen Australia may be distributed more equitably if women with high breast density are screened more often and women with low breast density are screened less often. The shorter screening intervals modelled would benefit women with high breast density, and the additional screening required can be offset by longer intervals for women with low breast density without a reduction in overall program performance.

ACKNOWLEDGEMENTS

This work was undertaken through a PhD thesis funded by a National Breast Cancer Foundation *Women in Super* Postgraduate Scholarship, with additional funding from the Victorian Health Promotion Foundation and in-kind support from BreastScreen Victoria.

REFERENCES

Boyd, N. F., H. Guo, et al. (2007). "Mammographic density and the risk and detection of breast cancer." N Engl J Med **356**(3): 227-36.

Brisson, J., R. Verreault, et al. (1989). "Diet, mammographic features of breast tissue, and breast cancer risk." Am J Epidemiol **130**(1): 14-24.

Carney, P. A., D. L. Miglioretti, et al. (2003). "Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography." Ann Intern Med **138**(3): 168-75.

Chiarelli, A. M., V. A. Kirsh, et al. (2006). "Influence of patterns of hormone replacement therapy use and mammographic density on breast cancer detection." Cancer Epidemiol Biomarkers Prev **15**(10): 1856-62.

Haybittle, J. L., R. W. Blamey, et al. (1982). "A prognostic index in primary breast cancer." Br J Cancer **45**(3): 361-6.

Kavanagh, A. M., G. B. Byrnes, et al. (2008). "Using mammographic density to improve breast cancer screening outcomes." Cancer Epidemiol Biomarkers Prev **17**(10): 2818-24.

Kerlikowske, K., D. Grady, et al. (1996). "Effect of age, breast density, and family history on the sensitivity of first screening mammography." Jama **276**(1): 33-8.

Lam, P. B., P. M. Vacek, et al. (2000). "The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont." Cancer **89**(2): 369-75.

Maskarinec, G., I. Pagano, et al. (2005). "Mammographic density and breast cancer risk: the multiethnic cohort study." Am J Epidemiol. **162**(8): 743-52. Epub 2005 Sep 8.

Sala, E., R. Warren, et al. (1998). "Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme." J Med Screen **5**(4): 207-12.

Stines, J. and H. Tristant (2005). "The normal breast and its variations in mammography." Eur J Radiol **54**(1): 26-36.

Tan, S. Y., G. J. van Oortmarssen, et al. (2006). "The MISCAN-Fadia continuous tumor growth model for breast cancer." J Natl Cancer Inst Monogr(36): 56-65.

van Gils, C. H., J. D. Otten, et al. (1998). "Mammographic breast density and risk of breast cancer: masking bias or causality?" Eur J Epidemiol **14**(4): 315-20.

Yaffe, M. J. (2004). "What should the burden of proof be for acceptance of a new breast-cancer screening technique?" Lancet **364**(9440): 1111-2.