DEVELOPING SIMULATION DYNAMIC MODELS OF BREAST CANCER SCREENING

Michael J Fett
Medical Advisor
Department of Health and Aged Care
PO Box 9849
CANBERRA ACT 2601
AUSTRALIA

michael.fett@health.gov.au

Within the health field, there appear to be few system dynamic models available for examining major public policy options. In the field of breast cancer screening, the results of several simulation models have been published[1,2,3,4,5,6], but none of the models uses system dynamic modelling. This paper describes the development of two system dynamic models of mammographic (X-ray) screening for breast cancer, and illustrates their use.

The purposes of system dynamic models can be grouped under four headings:
1. Elucidating systems, their components and their inter-relationships
2. Developing cognitive and group processes (eg systems thinking, shared understanding, team building)
3. Identifying information and knowledge gaps
4. Predicting and planning the future

For the first two purposes, models need not be based on data, although data-based models can be used for these purposes. For the second two purposes, the quantity and quality of data used in the model are crucial. Models that are heavily based on empirical data (i.e. quantitative information) may be used for all four purposes, while models that have a limited underpinning in data can only be used, reliably, for the first two.

Here we present attempts to build models for mammography screening that rely heavily on published data. If models can be constructed using published input data and the results of these models can replicate published output data, then a case can be made for using these models for health policy analysis. By 'input data', we mean data that inform the workings of the model, and by 'output data', we mean the results that the model calculates from the input data.

To examine the potential usefulness of system dynamic modelling for public health policy analysis, the author attempted to build two data-based models of breast cancer screening: The landmark trial of mammography known as the Swedish Two County trial, which was conducted from the late 1970's to late 1980's,[7] and the Australian BreastScreen Program[8], which is a national program offering mammographic screening to all women aged 50-69 years. The models were build in Powersim Constructor version 2.51 (4008) (Powersim AS, Isdalsto,1998).

Breast cancer screening was selected as the topic for model building for three reasons: the significance of breast cancer as a population health problem, the availability of

published research studies and extensive relevant Australian population data on breast cancer, and the presence of a large national screening program in Australia which might benefit from any policy analysis that could be undertaken with the models.

Breast cancer screening is well suited to system dynamic modelling because of the presence of many time-related phenomena. These include: trends in breast cancer incidence and mortality, trends in number of women screened, changing population numbers, sojourn time of pre-clinical breast cancer (the time between a cancer being detectable by screening and presenting clinically) and survival time following the development of cancer.

Tables 1 and 2 shows the data stocks and flows, and their methods of calculation, for both mammography models. Table 1 runs from the initial population to the development of breast cancer. Table 2 runs from the stock of breast cancers to death from breast cancer, the calculation of performance indicators and using the models for policy analysis.

Stock or *flow*	**Operation of Swedish trial model**	**Operation of Australian screening program model where it differs from Swedish trial model**
Population prior to development of breast cancer	Initial population by (Study and Control) group and age[9]. Preclinical cases are subtracted from the initial population as they occur. The population ages each year.	A dynamic model of the Australian female population by age from 1966 to 2015, from Census and projections[10,11,12,13].
Developing preclinical cancer	Breast cancer-free population multiplied by age-specific breast cancer incidence rates for Sweden 1972-75[14]. Incidence rate can be calibrated.	Australian population multiplied by age-specific breast cancer incidence rates for Australia for 1983[15].
Cases of preclinical cancer	Population of preclinical cases ages each year. Records number of years since preclinical cancer developed (up to maximum: the sojourn period[16]).	
Detecting cancer by screening	Number of preclinical cases multiplied by screening policies (age range and frequency), participation by women[16], and screening sensitivity and specificity[9]. Overdiagnosis by screening can be calibrated.	Number of preclinical cases multiplied by past (historical) screening rates[17] and by possible future screening policies (age range and screening interval) using past or target[18] participation rates.
Detecting cancer clinically after sojourn time	Preclinical cancers that reach the end of their age-specific sojourn time and have not been screen detected, are detected clinically.	

Table 1 Data stocks, flows and calculations to simulate the development of breast cancer

Stock or *flow*	**Operation of Swedish trial model**	**Operation of Australian screening program model where it differs from Swedish trial model**
Cases of screen detected and clinically detected breast cancer	Population of clinical cases ages each year. Records number of years since cancer was detected.	Breast cancer cases developing or detected each year are summed into 5 y age groups and divided by population to calculate 5 y age group incidence rates.
Applying mortality curves for screen detected and clinically detected cancer	Use proportion dying each year from breast cancer to calculate number of breast cancer deaths each year. Different mortality curves are used for screen and clinically detected cancer[19]. Curves can be calibrated.	
Breast cancer deaths	Breast cancer deaths each year summed to give cumulative breast cancer deaths in Study and Control groups.	Breast cancer deaths each year are summed within 5 y age groups and divided by population to calculate 5 y age group death rates.
Performance indicators are calculated	Cumulative number of deaths from breast cancer in Study and Control groups are divided by respective initial populations to calculate simulated cumulative death rates over 15 y study period. Cumulative death rate of Study group are divided by cumulative death rate of Control group to calculate simulated mortality rate ratios. These simulated mortality data are compared with observed mortality data[9].	Death rates and incidence rates are calculated for each 5 y age group. These are combined into age standardised death rates and incidence rates using a standard population. These simulated data are compared with observed age standardised breast cancer death rates and incidence rates[20].
Model used for policy analysis	Confirmed quantitative approach to building system dynamic models of breast cancer screening.	Vary age range and screening interval to assess impact of different screening policies on age standardised breast cancer death rates.

Table 2 Data stocks, flows and calculations to simulate transition from cases of breast cancer to death from breast cancer, calculation of performance indicators and policy analysis

The listings of stocks and flows in Tables 1 and 2 show that the mammography models are linear.

In the Swedish trial, communities were randomly assigned to either the Study or Control group. At the start of the trial (the time of randomisation), women in the Study group were offered breast cancer screening and those in the Control group were not. After eight years, the trial ceased and women in the Control group were then offered screening[16]. Mortality in the two groups was followed for 14 years from the time of randomisation. The Swedish trial simulation presented here is confined to women aged 50 to 64 years at entry because of the availability of cumulative mortality data for women in this age range [21]. In the 50-64 y age group, there were approximately 35,300 women in the Study group and 25,000 in the Control group.

Without calibration, there is quantitatively good agreement between the observed and simulated cumulative breast cancer death rates (Figure 1). That is, the observed and simulated rates are well within an order of magnitude and the Study/ Control death rate ratios are similar for the observed and simulated data.

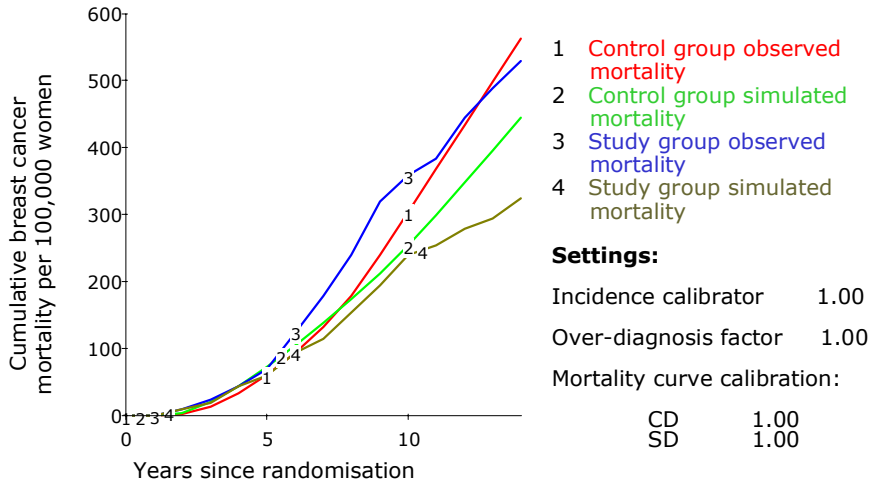

Figure 1 Observed and simulated cumulative breast cancer death rate Swedish trial model - uncalibrated

Four parameters can be calibrated: the breast cancer incidence rate, the degree of over-diagnosis of breast cancer by screening (whereby screening detects 'cancers' that would not otherwise have become clinically apparent) and the mortality (survival) curves for clinically detected and screen detected cancers. It should be noted that in both the Study and Control groups, both clinically detected and screen detected cancers occur, but in different proportions and at different times.

The survival curves that were subject to calibration come from an analysis of the survival of women with breast cancer in South Australia[19]. The screen detected mortality curve was derived from information on breast cancer tumour characteristics

of women with cancers that were diagnosed by the South Australian Breast X-ray Service combined with information on survival by tumour characteristics from the Swedish trial[22,23].

The clinically detected mortality curve was derived from information on breast cancer tumour characteristics of women with breast cancer that was diagnosed outside the South Australian Breast X-ray Service. An unknown proportion of these may have been detected by mammography outside the screening program. Thus, this curve is likely to be intermediate between the survival curves of cancers that are solely screen detected and cancers that are solely clinically detected.

With calibration, the correspondence between the observed and simulated cumulative breast cancer death rates could be improved substantially by adjusting the survival curves for clinically detected and screen detected breast cancers (Figure 2).

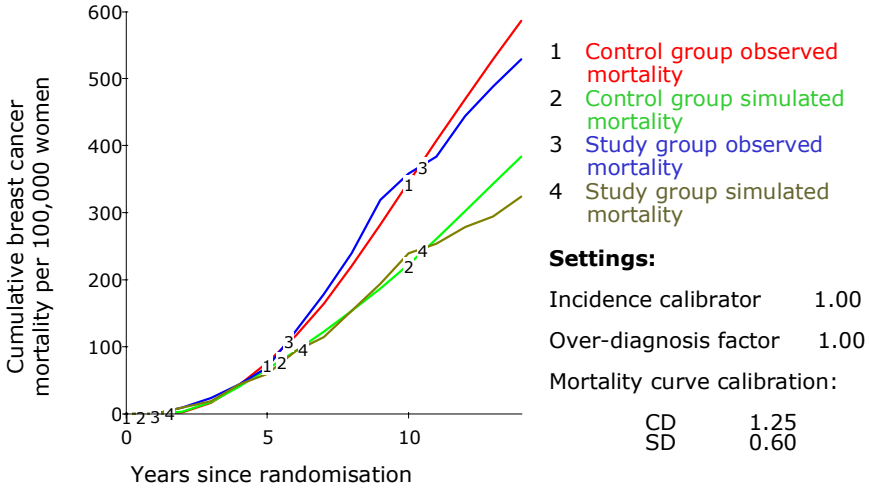

Figure 2 Observed and simulated cumulative breast cancer death rate Swedish trial model - calibrated

This calibration was performed by adjusting the values of the calibration factors and observing by eye the degree of correspondence between the observed and simulated rates. No calibration of incidence rate, nor calibration for over-diagnosis, was required. The calibration factor of 1.25 for clinically detected cancers is consistent with the possibility of the uncalibrated mortality curve for clinically detected cancers being 'contaminated' with screen detected cancers and thus underestimating the death rate for clinically detected cancers. The calibration factor of 0.6 for screen detected cancers may be attributable to different survival expectations for screen detected cancers in different settings. Alternatively, the calibration factors may be necessary to compensate for the effects of the assumptions that have been made in constructing the model or to adjust for integration errors in the model's algorithms.

This model demonstrates that is it possible to use system dynamic modelling to quantitatively replicate the results of clinical trials using information about characteristics of the trial (such as number of subjects, their ages and screening participation rates) and 'intermediate' statistical results (such as screening test characteristics and mortality curves). Furthermore, models of trials can be used to calibrate parameters, such as the intermediate statistical results, to improve the fit of simulations with observed results, and thereby provide new information on the values of the intermediate statistics. In effect, the model is suggesting what the actual mortality curves may have been, in the absence of published data.

The capacity of system dynamic modelling to replicate a trial of breast cancer screening suggested that replication of the Austrlian national breast cancer screening program BreastScreen Australia may be practicable. There were doubts about its feasibility, given the large amount of population data (e.g. numbers of women in each age group each year over a 50 y period) and computation required. The first stage of the BreastScreen model building process focussed on simulating breast cancer incidence (i.e. the onset of breast cancer) for comparison with published breast cancer incidence data. Figure 3 presents uncalibrated simulated incidence compared with observed incidence. Two calibration factors were incorporated: baseline breast cancer incidence calibrator and over-diagnosis calibrator. With these calibration factors set at 1 (i.e. no calibration) there is a substantial difference between observed and simulated incidence.

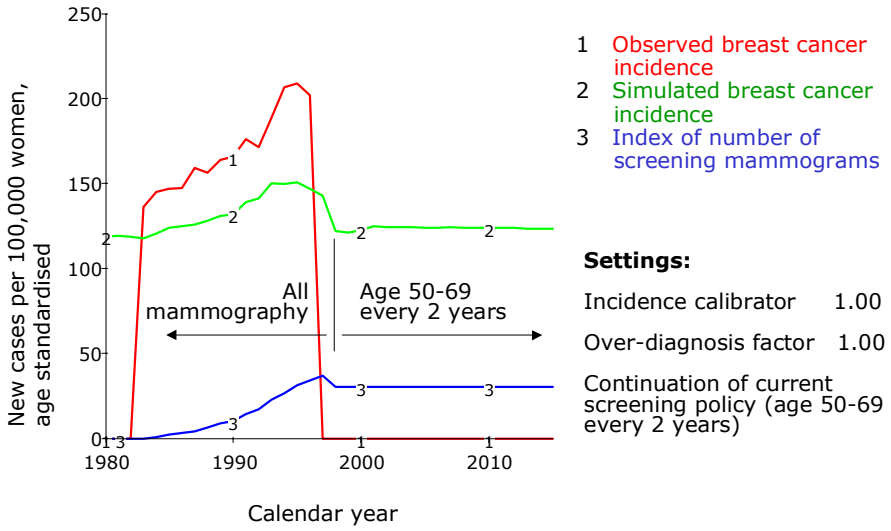

Figure 3 Observed and simulated breast cancer incidence rates in Australian breast cancer screening program - uncalibrated

In contrast to the method of optimising the Swedish trial model, the simulated incidence in the Australian program model was optimised by systematically searching for values of the incidence and over-diagnosis calibration factors that would minimise the sum of squares between the observed and simulated incidence. The simulated

incidence corresponding to this minimum is shown in Figure 4, along with the values of the calibration factors.

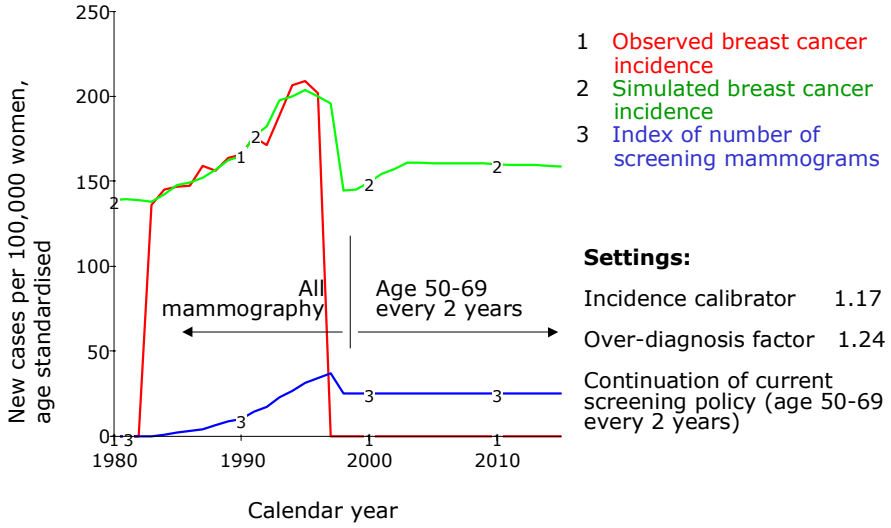

Figure 4 Observed and simulated breast cancer incidence rates in Australian breast cancer screening program - calibrated

The optimum incidence calibration factor is 1.17 (i.e. an incidence rate 17% greater than that originally chosen). The optimum over-diagnosis factor is 1.24, suggesting that in the Australian program, screening may detect 24% more cancers in screened women than would have become clinically apparent without screening. A substantial proportion of these cancers may be *in situ* cancers, most of which do not become invasive, but which, once detected, need to be treated because of their invasive potential.

Using simulated incidence optimised to correspond most closely with observed incidence (Figure 4), the second stage of simulating breast cancer mortality was undertaken. With optimised incidence simulation but without calibration of the mortality curves for clinically and screen detected breast cancer, there is again a substantial difference between the observed and simulated results (Figure 5).

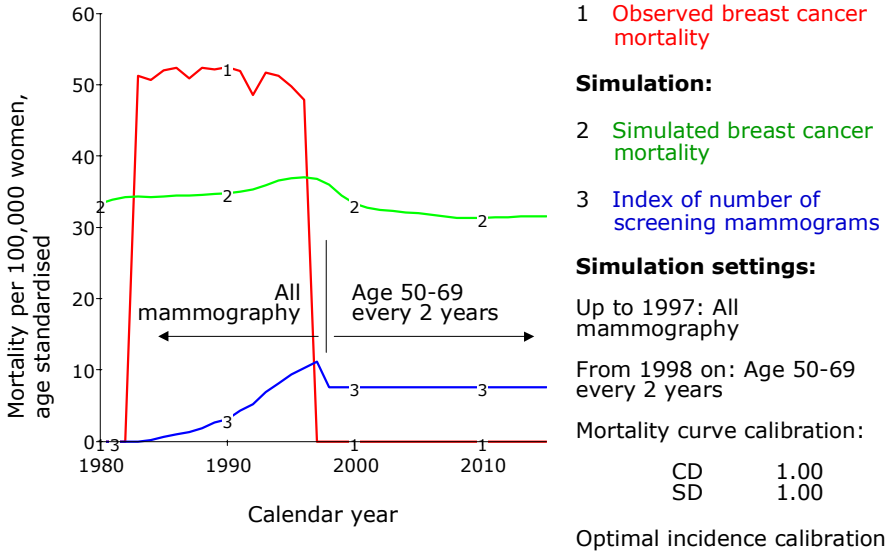

Figure 5 Observed and simulated breast cancer death rates in Australian breast cancer screening program - uncalibrated

Systematically searching for mortality calibration factors to minimise the sum of squares between observed and simulated age standardised death rates yielded values of 1.52 for the clinically detected cancer mortality calibration factor and 1.04 for the screen detected cancer mortality calibration factor. The optimised simulated age standardised mortality rate for the Australian BreastScreen program is shown in Figure 6. The value of 1.52 is consistent with the observation in the Swedish model that the clinically detected survival curve may underestimate the true clinically detected survival curve because of contamination by screen detected cancers. The small degree of calibration required of the screen detected mortality curve (1.04) suggests that the South Australian curve may be representative of the survival experienced by Australian women with screen detected cancer.

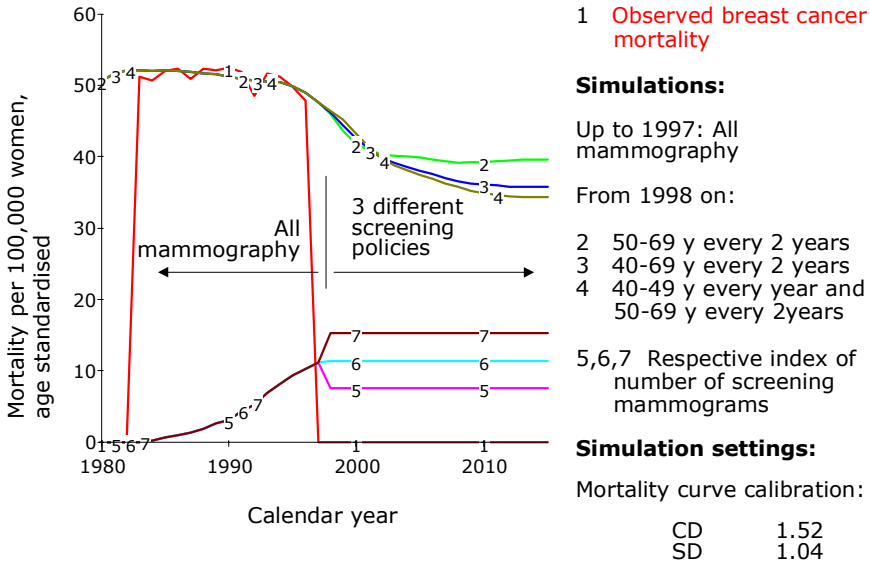

Figure 6 Observed and simulated breast cancer death rates in Australian breast cancer screening program - calibrated

Figure 6 also presents expected future age standardised mortality rates under different screening policies. These curves provide an indication of the policy uses that could be made of this model: As the quantity of screening increases (by including women in their forties in the program) the breast cancer mortality rate drops. Moving from the current policy of screening women aged 50 - 69 y every two years by adding biennial screening of women in their forties produces an increase in the amount of screening required and a reduction in the death rate from breast cancer. Shortening the screening interval for women in their forties from two to one year requires a similar increase in the quantity of screening but yields a proportionately smaller decrease in the death rate. This suggests that the marginal cost-effectiveness of adding biennial screening of women in their forties is greater than moving to annual screening of women in their forties.

This model demonstrates that system dynamic modelling can be used to quantitatively simulate significant public health programs and their impact on population health. This model has also enabled estimation of the extent of over-diagnosis of breast cancer in the Australian screening program and estimation of the expected timing and degree of mortality reduction from BreastScreen. It can also be used to investigate the impact of different screening policies on program costs and the number of lives saved. The model also leaves significant computational capacity unused, allowing for further development of the model. These developments could include modelling type of cancer and cancer stage and their respective survival curves, incorporating stochastic processes and confidence bands, automation of sensitivity testing and optimisation, and duplication of the basic model to allow head-to-head comparisons of different screening policies in terms of cost and health gain.

[1] Knox EG. Evaluation of a proposed breast cancer screening regimen. BMJ 1988; 297: 650-654.

[2] Van Oortmarssen GJ, Habbema JDF, van der Maas PJ et al. A model for breast cancer screening. Cancer 1990; 66: 1601-1612.

[3] Brown ML. Sensitivity analysis in the cost-effectiveness of breast cancer screening. Cancer supplement 1992; 69: 1963-1967.

[4] Lindfors KK, Rosenquist CJ. The cost-effectiveness of mammographic screening strategies. JAMA 1995; 274: 881-884.

[5] van Oortmarssen GJ, Boer R, Habbema JDF. Modelling issues in cancer screening. Statistical Methods in Medical Research 1995; 4: 33-54.

[6] Boer R, de Koning H, Threfall A et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening program: computer simulation study. BMJ 1998: 317: 376-379.

[7] Tabar L, Fagerberg CJG, Gad A et al. Reduction of mortality from breast cancer after mass screening with mammography. Lancet 1985; i: 829-832.

[8] BreastScreen Australia. BreastScreen Australia statistical report 1996. Canberra: BreastScreen Australia, 1996.

[9] Chen HH, Tabar L, Fagerberg G, Duffy SW. Effect of breast cancer screening after age 65. Journal of Medical Screening 1995; 2: 10-14.

[10] d'Espaignet ET, van Ommeren M, Taylor F, Briscoe N, Pentony P. Trends in Australian mortality 1921 – 1988. Australian Institute of Health (AIH) Mortality Series No 1. Canberra: Australian Government Publishing Service, 1991.

[11] Australian Bureau of Statistics. Estimated resident population by age and sex: Australian States and Territories, June 1987 to June 1992. ABS Cat No 3201.0. Canberra: ABS, 1993.

[12] Australian Bureau of Statistics. Estimated resident population by age and sex: Australian States and Territories, June 1992 to June 1997. ABS Cat No 3201.0. Canberra: ABS, 1997.

[13] Australian Bureau of Statistics. Population projections 1997 to 2051. ABS Cat No 3222.0. Canberra: ABS, 1998.

[14] International Agency for Research on Cancer. Cancer incidence in five continents. Volume IV. Lyon: IARC, 1982.

[15] Australian Institute of Health and Welfare. Breast and cervical cancer screening in Australia 1996–1997. AIHW Cat No CAN 3. Canberra: AIHW, 1998.

[16] De Koning HJ, Boer R, Warmerdam PG, Beenisterboer PMM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995; 87: 1217-1223.

[17] Personal communication: Department of Health and Aged Care.

[18] Australian Institute of Health and Welfare. Breast cancer screening in Australia: future directions. AIH Prevention Program Evaluation Series No 1. Canberra: AIH, 1990.

[19] South Australian Cancer Registry. Prognostic features of breast cancers detected in mammographically screened women in South Australia: a pilot study. Chapter 4 in: Epidemiology of cancer in South Australia. Adelaide: South Australian Health Commission, 1996.

[20] Australian Institute of Health and Welfare. Breast and cervical cancer screening in Australia 1996–1997. AIHW Cat No CAN 3. Canberra: AIHW, 1998.

[21] Chen HH, Tabar L, Fagerberg G, Duffy SW. Effect of breast cancer screening after age 65. Journal of Medical Screening 1995; 2: 10-14.

[22] Tabar L, Fagerberg G, Chen HH et al. Efficacy of breast cancer screening by age. New results from the Swedish two-county trial. Cancer 1995; 75: 2507-2517.

[23] Day NE. Surrogate measures in the design of breast screening trials. In: Miller AB, Chamerlain J, Day NE et al. (eds). Cancer screening. Cambridge: Cambridge University Press, 1991.