Volume 11 • Number 2 • 2008
VALUE IN HEALTH

Probabilistic Cost-Effectiveness Modeling of Different Breast Cancer Screening Policies in Slovenia

Klemen Rojnik, MPharm,[1] Klemen Naveršnik, MPharm,[2] Tatjana Mateović-Rojnik, MPharm, PhD,[3] Maja PrimicŽakelj, MD, PhD[4]

[1]Roche d.o.o. farmacevtska družba, Vodovodna, Ljubljana, Slovenia; [2]Lek Pharmaceuticals d.d., Verovskova, Ljubljana, Slovenia; [3]Faculty of Pharmacy, University of Ljubljana, Aškerčeva, Ljubljana, Slovenia; [4]Institute of Oncology Ljubljana, Zaloska, Ljubljana, Slovenia

ABSTRACT

Objectives: To determine the most cost-effective screening policy for population-based mammography breast cancer screening in Slovenia using probabilistic sensitivity analysis.

Methods: A time-dependent Markov model for breast cancer was constructed. General principles of cost-effectiveness analysis with multiple strategies were used to compare the costs and effects of 36 different screening policies. Using probability distributions for model parameters, the true effect of uncertainty across model input parameters on expected costs and effects was explored. The results from probabilistic simulation analysis are presented in a form of cost-effectiveness acceptability curves with cost-effectiveness acceptability frontier.

Results: With the presented analysis, it was shown that a 1-year screening interval in population breast cancer screening would produce less benefits at higher costs than less intensive screening and that a 2-year interval would be cost-effective only at high values of society's willingness to pay per quality-adjusted life-year (QALY). Therefore, the optimal screening policy should be chosen among 3-year-interval policies.

Conclusions: Based on commonly quoted thresholds of society's willingness to pay per QALY of $50,000, the optimal approach in the Slovenian population would be screening women aged from 40 to 80 years every 3 years.

Keywords: breast cancer screening, cost-effectiveness acceptability frontier, multiple CEAC, probabilistic cost-effectiveness analysis.

Introduction

Breast cancer is the most common cancer in the female population in Slovenia with around 100 newly diagnosed cases per 100,000 women in a year and it will afflict 1 in 15 Slovenian women by the age of 75 years [1]. The costs of this disease are high, in terms of both decreased health-related quality of life and health-care consumption. Currently, the most effective method for preventing premature mortality and morbidity due to breast cancer is through the increased use of screening programs, which enable identifying and treating breast cancer at earlier stages. Benefits arise from the more favorable prognosis associated with early-stage cancers and may prevent treatment for late-stage breast cancer with possible reductions in treatment cost. In the 1980s, the demonstration of the efficacy and effectiveness of mammography with or without clinical breast examinations in reducing mortality from breast cancer by 25% to 30% led to the adoption of guidelines in a number of countries to introduce routine screening on a population basis [2]. Commonly used guidelines recommend the initiation of mammography screening at the age of 40 to 50 years, whereas the upper limits are set at 60 to 74 years or they are left undefined; the suggested screening intervals are 1, 2, or 3 years [2].

The objective of this study was to explore the cost-effectiveness of population-based mammography breast cancer screening in Slovenia relative to no screening, and to choose the most favorable screening policy with regards to the age eligibility criteria and screening interval. A probabilistic sensitivity analysis of the model was undertaken to represent the uncertainty of the model's parameters.

Address correspondence to: Klemen Rojnik, Roche d.o.o. farmacevtska družba, Vodovodna 109, SI-1000 Ljubljana, Slovenia. E-mail: klemen.rojnik@roche.com

10.1111/j.1524-4733.2007.00223.x

Methods

Model

A time-dependent Markov model was used to compare hypothetic populations of women, one followed clinically without screening and the others underwent different screening mammography policies. The structure of the model, which is presented in Figure 1, is similar to other models already used in the evaluation of breast cancer screening programs, such as MISCAN [3,4] and MICROLIFE [5]. The difference between our

 1098-3015/08/139 139–148

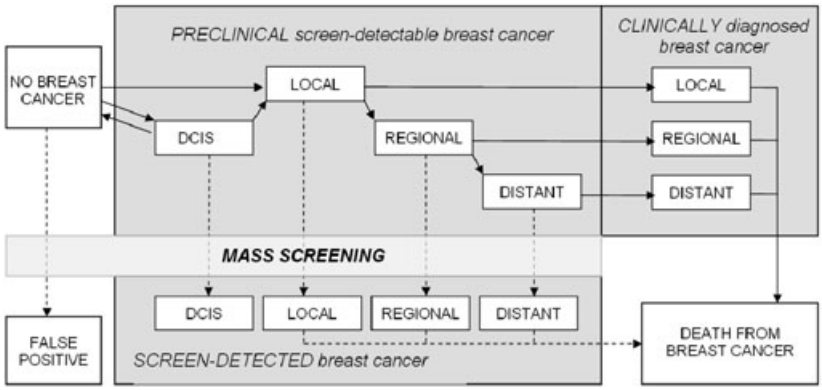

Figure 1 Structure of the model for breast cancer screening with the possible courses of the disease. The dashed lines correspond to transitions possible only by screening policies. The state "death from other causes" which can be attained from all other states is not shown. DCIS, ductal carcinoma in situ.

model and the reported models is in the breast cancer classification approach. While the majority of other models use tumor size as the classification criterion, a simplified TNM (Tumor Node Metastasis Classification) cancer stage [6] of breast cancers was used in our model. The model characterizes the natural history of the disease as having four preclinical stages when breast cancer can be detected by screening but shows no clinical symptoms. Approximately 60% of the invasive breast cancers are assumed not to be preceded by ductal carcinoma in situ (DCIS), which is screen-detectable and from which a 65% progression to invasive cancer is assumed [7]. The invasive stages are defined as follows. In the localized stage of breast cancer, neither regional node involvement nor distant metastases are found (T1-3 and N0, M0 after TNM Classification). The regional stage includes tumors classified as T3 and T4 and/or regional node metastases (N1), if no metastases in distant lymph nodes or organs (M0) are found. The disease with metastases in distant nodes or organs is classified as distant stage (M1) [1].

The transitions to clinically diagnosed local, regional, and distant states are governed by the rate of the incidence, clinical-stage distribution data, and sojourn time. In the case of early detection by screening, the women enter the corresponding screen-detected DCIS, local, regional, or distant states. The state "false positives" refers to women with positive screening examination in whom no breast cancer is found at further invasive assessment. The two absorbing end-states of the model are death from breast cancer and death from other causes.

The cohort simulation approach with a cycle length of 1 week was used for running the Markov model [8,9], which was developed with the SAS System for Windows Release 8.02 (SAS Institute Inc., Cary, NC, USA) [10]. The cycle length of 1 week was chosen to enable the modeling of the sojourn times and treatment durations with sufficient precision. Breast cancer incidence, mammography sensitivity, mortality, and breast cancer relative survival were modeled as time-dependent transition probabilities.

The perspective for the evaluation was that of health-care sector and the time horizon covered the full lifetime of the patients from age 40 years onward. A discount rate of 3% for costs and effects was applied to the analyses, in accordance with recommendations from the Panel on Cost Effectiveness in Health and Medicine [11]. All costs are expressed in 2004 European euros. Throughout the modeling process, the principles of good practice for decision analytic modeling in health-care evaluation were followed, as proposed by the ISPOR Task Force on Good Research Practices-Modeling Studies [12].

Data

All data regarding age-dependent cancer incidence, clinical-stage distribution, treatments, and survival were obtained from the Cancer Registry of Slovenia [1]. Because of the introduction of preventive mammography examinations (though unorganized) in the early 1990s and consequently the slight shift of clinical-stage distribution toward the earlier stages of breast cancer, the database for the time period of 1980–1990 was used for estimating a clinical-stage distribution of breast cancer. In that period, 41% of breast cancers were detected in the local stage, 47% in the regional stage, and the remaining 12% in the distant stage. The database for the period 1999–2001 was used for the estimation of age-dependent incidence. The incidence in this period is slightly higher than it was in the period 1980–1990; here the assumption was made that the effect of unorganized screening on higher incidence is negligible in respect to the effect of higher risk factors in recent years, which are consequences of less healthy lifestyle, lower fertility rates, lower average number of children, higher age at first birth, and higher awareness of the disease [13].

The survival of women with preclinical stages of breast cancer, of women with screen-detected DCIS, and of women with false positive results was assumed to be equal to the survival of women with no breast cancer. Although this assumption is obvious for women with false positive results, the choice for the same assumption in the other two women groups

needs further comment. It was reported that operational removal of DCIS cures at least 98% of lesions [14] and that the 5-year mortality because of DCIS is 1% [6], which justifies this assumption also for screen-detected DCIS. The determination of survival of women with preclinical stages of breast cancer is quite problematical if not impossible. One possible way to justify the assumption is by comparing the number of women that died because of the breast cancer and the number of women in which the breast cancer was registered only on the death certificate. The number of women, in which breast cancer was registered only on a death certificate represents only approximately 1% to 2% of incidence of breast cancer and only 2% to 4% of deaths because of breast cancer [1]. The assumption is further justified by the fact that the durations of the preclinical stages are quite short when compared to the durations in the clinical stages.

Stage-specific relative mortality from breast cancer (i.e., the proportion of women that die because of breast cancer after specific period) was obtained in the following manner. A relative 1–7 years survival of breast cancer patients (i.e., the proportion of women with breast cancer alive after specific period) diagnosed in the period 1991–1995 was subtracted from the relative survival of the general population (i.e., the proportion of women alive after specific period) in corresponding age groups. Because the obtained relative stage-specific mortality was quite similar among different ages at diagnosis, it was assumed to be independent of age at diagnosis.

Various patterns of stage-specific treatments were taken from the Cancer Registry of Slovenia for the same groups of women that were included in the relative survival calculation, to properly estimate the effect of treatment on survival. Treatments consisted of four basic interventions: surgery, hormonal therapy, radiotherapy, and chemotherapy. For screen-detected DCIS, surgery with no further treatment was presumed [14].

One of the key assumptions of the screening, and hence of this model, is that mammography makes it possible to capture a sizable amount of cancer cases that would otherwise have gone undetected until the appearance of clinical symptoms. The model thereby needs to incorporate an estimate of sojourn time, that is, the period when the cancer is screen-detectable but shows no clinical symptoms. We assumed a value of 3 to 7 years with a mean of 5 years [8,15] for DCIS and a value of 2 to 3 years with a mean of 2.5 years [16–18] for invasive carcinoma in preclinical local stage. The sojourn time in preclinical regional and in preclinical distant states was estimated from the approximate growth rates of tumors. Approximately one half of the breast cancers in the Cancer Registry of Slovenia have a defined TNM stage from which approximate tumor sizes were calculated for local, regional, and distant carcinomas. Then the cancer cells' doubling times from 60 to 180 days were presumed to obtain the ranges for a sojourn time of 0.36 to 1.08 years and 0.35 to 1.04 years for regional and distant stages, respectively; this method was adopted from Michaelson et al. [19] and Kopans et al. [20].

The assumed sensitivity of the mammography is 86.7% for ages between 40 and 49 years, 93.6% for ages between 50 and 59 years, 94.1% for ages between 60 and 69 years, and 91.2% for ages 70 years and older [21]. Based on the results from mass screening in other countries, a reasonable estimate of the attendance is 75% [22] and the estimate of the recall rate is 7% [22,23]. About 20% of the women who are recalled for additional diagnostic procedures undergo invasive diagnostic procedures such as fine-needle aspiration and surgical biopsy. The remaining 80% undergo noninvasive imaging (further mammography, ultrasound) and clinical examination [22,23]. Diagnostic procedures for clinically detected breast cancers include noninvasive (mammography, ultrasound, clinical examination) and invasive techniques (fine-needle aspiration, surgical biopsy).

The costs for mammography examination, the costs for diagnostic interventions for clinically detected breast cancer, the costs for invasive and noninvasive diagnostics at recall, and the costs for treatment interventions were obtained from the Institute of Oncology Ljubljana [1].

In order to capture the difference in mortality and morbidity due to screening, health improvement was measured in quality-adjusted life-years (QALYs). QALYs for treatment and the corresponding durations of treatments were obtained from the literature [24]. The quality of life for DCIS, local and regional breast cancers after treatment was weighted according to the treatment interventions. The quality of life for distant cancer was weighted with 0.515 [25].

The quality of life for women with false positive result was also reduced according to the diagnostic duration and QALY weight [24]. In the case of death from breast cancer, a terminal illness lasting 1 month with QALY weight of 0.288 was taken into account [24].

Screening Policies

Different screening policies were considered with respect to the following eligibility criteria: age at the beginning of the screening, age at the end of the screening, and the interval between two screenings. All possible combinations of starting ages 40, 45 and 55 years, ending ages 65, 70, 75 and 80 years, and screening intervals of 1, 2 and 3 years were considered, thus giving 36 different screening policies, listed in Table 1.

Currently, only opportunistic screening activity takes place in Slovenia. Because there is no register of the women that underwent such screening, the extent

Table 1 List of screening policies taken into consideration and their labeling in the article

	Screening period											
Interval	40–65	40–70	40–75	40–80	45–65	45–70	45–75	45–80	50–65	50–70	50–75	50–80
1	1	2	3	4	5	6	7	8	9	10	11	12
2	13	14	15	16	17	18	19	20	21	22	23	24
3	25	26	27	28	29	30	31	32	33	34	35	36

of screening would be difficult to estimate suitably. Therefore, the present situation was omitted from the analysis and the "null option" chosen for screening policies was the no-screen option.

Deterministic Cost-Effectiveness Analysis

The cost-effectiveness model for the evaluation of breast cancer screening was built in several phases. The model was carefully tested in each phase to ensure that the mathematical calculations were accurate and consistent with the specification of the model (internal validity of the model). First, a deterministic model for a cohort without screening was built and calibrated with the breast cancer incidence data and general survival. Then, the model was upgraded to include the screening policies and finally, the costs and effects in terms of QALY were incorporated in the model. The cross validation of the model was carried out with the published results from screening programs in other countries. A deterministic evaluation of the model with the input parameters at their baseline values was performed and incremental cost-effectiveness ratios (ICERs) were calculated among nondominated alternatives. Additionally, simple one-way sensitivity analysis was performed on all model parameters to test the robustness of the results [26]. For each of the 82 univariate sensitivity analysis calculations (41 with parameters on their upper limit and 41 with parameters on their lower limit), an efficiency frontier with corresponding ICER was determined thus representing the impact of parameter value change on the selection of the optimal policy.

Probabilistic Sensitivity Analysis

Probability distributions were defined for all the model parameters except for breast cancer incidence and the cost of mammography examination, as the incidence of breast cancer has been quite constant in recent years and the cost of mammography examination is fixed. Beta distributions were fitted using the method of moments [27] with mean and standard error drawn from the literature to represent the uncertainties surrounding probabilities (attendance, recall rate, proportion of invasive diagnostics, sensitivities, clinical-stage distribution, and progression from DCIS) and QALYs. A negative exponential function was fitted to stage-specific breast cancer relative mortality, and fitted parameter's distributions were used to describe the distributions of relative mortality at specific times after the diagnosis. In order to gain distributions of costs for chemotherapy, hormonal therapy, radiotherapy, and surgery, an approximate number of patients for various subtypes of each of the four treatment interventions (i.e., various types of chemotherapy, hormonal therapy, radiotherapy, and surgery) along with treatment costs were obtained from the Institute of Oncology Ljubljana. Log-normal distributions were fitted to different treatment costs and, in a similar manner, log-normal distributions were fitted to costs for invasive and noninvasive diagnostic examinations and for examinations after clinically detected breast cancer. Finally, a log-normal distribution was assumed for sojourn times to assure that generated times were all positive.

The net benefit framework was used for analyzing the results from probabilistic sensitivity analysis to derive cost-effectiveness acceptability curves (CEACs) for each screening policy at different values of willingness to pay per QALY (λ) [28,29]. The policies with the highest expected net benefits (i.e., the policies of choice) at different λ's were plotted on CEACs in the form of cost-effectiveness acceptability frontier [30].

Results and Discussion

Deterministic Analysis

During the building phase of the model, the calibration was performed with the observed breast cancer incidence and general population survival of women. There was a reasonably good fit of the model's incidence to the actual data. A good fit was obtained for the complete life table of Slovenian women as well (data not shown). Because there is currently no population screening program in Slovenia, the results of the screening part of the model were cross-validated with the results from pilot and population-based screening programs in other countries. Mortality reduction from screening programs is a suitable parameter for comparison, even though incidence, treatment, life tables, and other factors may differ from the modeled Slovenian population. In the health insurance plan breast cancer screening trial, where women aged from 40 to 64 years had been screened, a mortality reduction of 24% has been reported. Our model, assuming the same screening policy but based on the Slovenian population, estimates a mortality reduction of 27%.

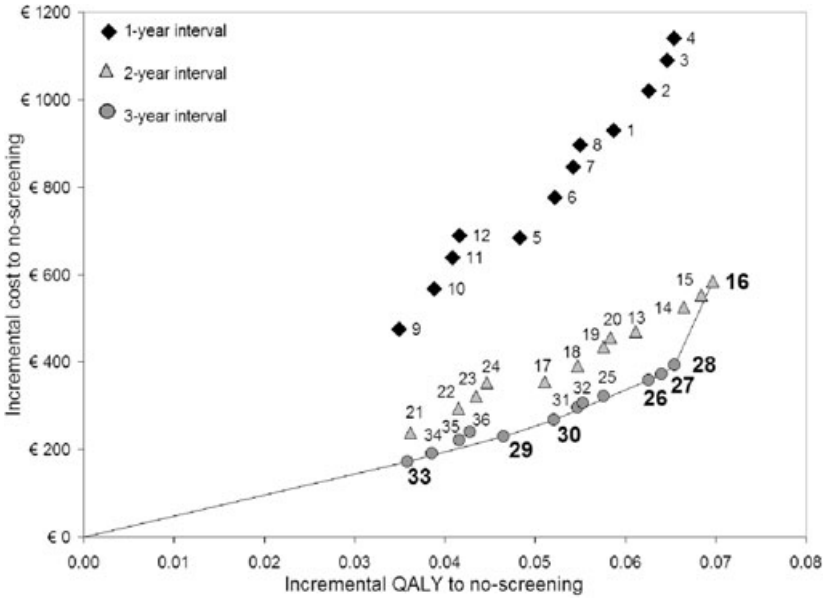

Figure 2 Cost-effectiveness plane. Baseline incremental costs and effects of screening policies relative to a no-screening on the cost-effectiveness plane. The efficiency frontier consisting of policies 33, 29, 30, 26, 27, 28, and 16 is outlined. The "no-screening" policy has a cost of €231 and effect of 23.0 quality-adjusted life-years (QALYs).

Analogically, reported mortality reductions in other studies were compared to the results obtained by our model: Sweden Two Country Trial—reported 32% and modeled 34%; Stockholm reported 26% and modeled 25%; Malmö reported 19% and modeled 29%; Edinburgh reported 21% and modeled 17% [31]. We found the results agreeable enough to apprise the model as satisfactory. In addition, a modeled detection of DCIS was compared to the observed portion of DCIS in all detected carcinomas to verify the appropriateness of the modeled progression to invasive cancer through DCIS. In the UK screening program, approximately 18% of all carcinomas were DCIS; our model, regarding the same screening policy, gives an estimate of 15%. In The Netherlands the proportion of DCIS is 13%, while our model gives an estimate of 18% [7].

In order to implement the principles of ICERs and dominance in economic evaluation of multiple interventions, mutual exclusivity of those interventions must be ensured [32,33]. As the screening has an important effect on mortality, the time horizon for comparison of costs and effects covered the full lifetime from age 40 years onward, to quantify the differential impact on life expectancy [34]. The screening policies noted as, for example, 50 to 75 years with 2-year interval mean that women are followed clinically without screening from ages 40 to 50 years, are undergoing screening every 2 years from 50 to 75 years, and are again followed clinically until their death. Hence, the screening policies are mutually exclusive, as a woman can enter only one screening program.

Estimates of the differences in costs and differences in QALYs between the population screened with a specific screening policy and the nonscreened population are presented in Table 4a in the supplementary material. The base-case values of the costs and the QALYs are expressed per one 40-year-old woman. The cost for a woman from the nonscreened population is €231 and the effect is 23.0 QALYs. Screening policy 33 (screening from ages 50 to 65 years every 3 years) has the lowest cost per QALY (€173 for approximately 2 quality-adjusted life-weeks) incrementally to no screening. The cost-effectiveness results are displayed graphically in Figure 2. The efficiency frontier, represented by policies 33, 29, 30, 26, 27, 28, and 16, is outlined as well.

The univariate sensitivity analysis was performed with upper and lower limits of 95% credibility intervals of parameter distributions and with 1% or 5% discounting, ±10% cancer incidence, and ±20% the cost of mammography examination. The results indicated that the model is quite robust. In virtually all of the cases (altogether 82 different univariate analyses), the efficiency frontiers consisted of all seven policies stated in the base-case deterministic analysis. On the average, the greatest impacts on ICER comparing to base-case deterministic analysis were observed with discounting, percent of DCIS progression to invasive cancer, recall rate, relative mortality in regional stage, percent of invasive diagnostics, cost of mammography examination, and percent of invasive cancers preceded by DCIS.

Probabilistic Sensitivity Analysis

Probabilistic sensitivity analysis of the model was performed with 10,000 Monte Carlo simulations. The simulation of the cohort without screening and all 36 cohorts with different screening policies took approxi-

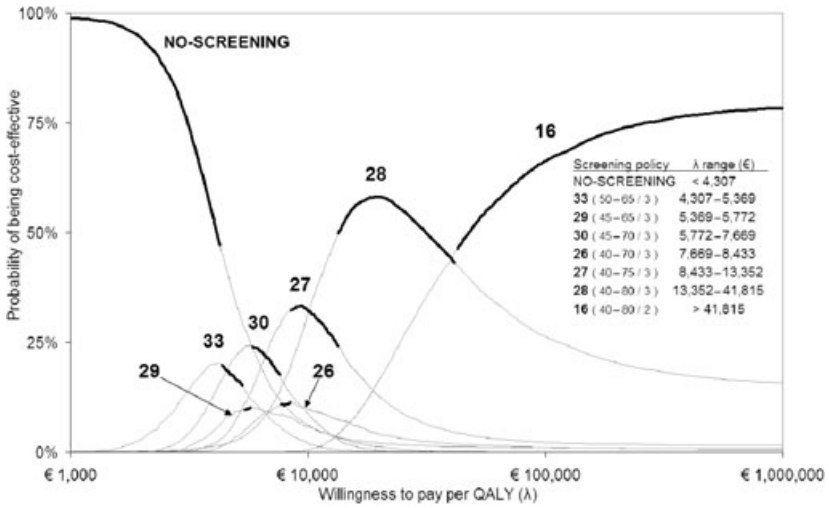

Figure 3 Cost-effectiveness acceptability curves (CEACs). CEACs (gray lines) with cost-effectiveness acceptability frontier (CEAF; bold black line). CEAF illustrates the uncertainty associated with the optimal policy over a range of λ. CEAF is also presented numerically. CEACs of other screening policies never reach the 3% probability of being the most cost-effective. QALY, quality-adjusted life-year.

mately 400 hours on a Pentium IV 3.0 GHz computer. The results of probabilistic sensitivity analysis can be presented graphically on a plane showing cost and effect pairings, but because of the large number of screening variants, a graphical presentation of all the screening variants would be very unclear. To indicate the amount of variation, three screening policies (4, 16, and 33) were compared on a cost-effectiveness plane. The simulation results revealed quite a large variation around the base-case analysis. Moreover, a very high correlation between the costs of different policies and a very high correlation between the effects of different policies was observed. This was expected because all the models are simulated simultaneously with the same parameters and the models differ only in the number of screenings and the patients' ages at screenings.

A cost-effectiveness acceptability frontier illustrating the uncertainty over a range of values of λ shows the error probability associated with the optimal strategy at each level of λ (Fig. 3). The frontier does not follow the outer limit of the family of CEAC because of skewness in the distributions of net benefits. Therefore, even though the policies 33, 29, and 26 never have the highest probability of being cost-effective, they are still the policies of choice at some values of λ, as indicated on Figure 3. Nevertheless, the λ ranges at which these policies are the policies of choice are very narrow. This is also the case for the policies 30 and 27, which are policies of choice for ceiling ratios between €5772 and €7669 per QALY, and €8433 and €13,352 per QALY, respectively. Policy 28 is the policy of choice at a much wider range of λ between €13,352 and €41,815 per QALY, while for ceiling ratios above €41,815 per QALY, policy 16 is the policy of choice. No screening is cost-effective at λ-values less than €4307. Consistent with the results of deterministic analysis, the screening policies with 1-year screening interval do not feature on CEACs, indicating that they are never contenders for cost-effectiveness. Among policies with 2-year screening interval, only the policy 16 is featured on CEAC and is the policy of choice at high vales of λ.

Using the commonly quoted threshold of $50,000 (approximately €38,500) per QALY [35], the optimal screening policy would be screening women aged from 40 to 80 years every 3 years (screening policy 28). In fact, this policy remains optimal throughout the wide range of threshold values between €13,352 and €41,815 per QALY. The threshold can also be gleaned from a retrospective analysis of previous resource allocation decisions. Towse and Pritchard analyzed the first 41 decisions made by the UK National Institute for Clinical Excellence and concluded that the thresholds for likely reimbursement and likely rejection of therapies were approximately 1.4 and 2.1 times the gross domestic product (GDP) per capita [36]. Relative to the Slovenian GDP per capita, these limits in Slovenia would be approximately €17,000 and €26,000 per QALY. Applying these thresholds also confirms the screening policy 28 as the policy of choice.

Negative Aspects of Breast Cancer Screening

The opinion that the screening is an effective way of reducing the breast cancer mortality is widely accepted, although not entirely unified. The opposite viewpoints are based on the belief that screening causes more harm than good because of the unnecessary workup in the population [7]. At 7% recall rate and 20% invasive diagnostics at recall, approximately 1.4% of the screened women will unnecessarily go through the invasive diagnostic procedures, whereas (assuming 0.2% yearly incidence and 3-year screening interval) approximately 0.6% of the women population will be correctly diagnosed with the breast cancer. The

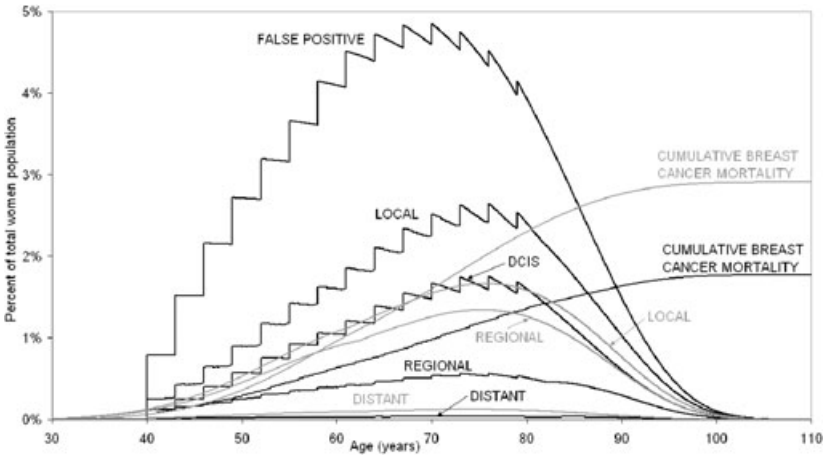

Figure 4 Effects of breast cancer screening. Graphical comparison of the breast cancer stage distribution, breast cancer cumulative mortality and portion of women who were diagnosed as false positive during their life-course in population that is screened from 40 to 80 years of age by a 3-year interval (black lines) and the population that is not screened (gray lines). In the screened population, the percent of woman in local, regional, and distant stages is the sum of screen-detected and clinically detected cancers in corresponding stages. The percentage of women refers to the total number of women alive at the age of 40 years. DCIS, ductal carcinoma in situ.

graphical comparison of women in the false positive state and different cancer states between screened and nonscreened population is presented in Figure 4. Although the magnitude of false positives is quite obvious, there is also clear conformation of the screening benefits, that is, the decreased breast cancer mortality and the shift of cancer's clinical stage distribution toward earlier stages.

Another issue that has been raised since the initiation of the mammography screening programs is the possibility of overdiagnosis of the breast cancer, in particular of the DCIS [7]. Various viewpoints regarding this problem have appeared in the literature ranging from the opinion that this mainly represents overdiagnosis and is likely to cause more harm than benefit, to the position that detection of DCIS is the ideal target of early detection and that a high rate of DCIS represents a large number of invasive cancers avoided [7]. Based on literature estimates of progression from DCIS to invasive cancers and estimates of nonprogressive DCIS [7], both aspects of this dilemma were included in the model.

The downsides of mass screening—overdiagnosis of the DCIS and the false positives, become more expressed in screening policies with a shorter interval between two screenings. As could be expected, more frequent screening would produce higher benefits, at of course higher costs. But, one has to consider also the added negative effects that reduce the incremental benefits. The model incorporates the negative aspects of screening by applying the disutility to false positive state because of invasive diagnostic procedures and by applying the disutility to "overdiagnosed" part of screen-detected DCIS state because of unnecessary treatment. The added negative effect of the screening because of shortening the screening interval from 3 to 2 years is exceeded by the positive effect in the form of increased survival. Therefore, the benefits of screening policies with a 2-year interval are greater than the benefits of screening policies with a 3-year interval at the same starting and ending ages. But when the intensity of screening is increased to yearly screenings, the incremental negative effects prevail over the incremental benefits from increased survival. This results in the inability of the yearly screening programs to reach the effectiveness of equivalent screening policies with a 2-year interval. The comparison of the portions of women which were diagnosed as false positive during their life-course for screening policies with 1-, 2-, and 3-year intervals from 40 to 80 years of age was investigated graphically. Also, the comparison of cumulative breast cancer mortality was compared in relationship to cumulative breast cancer mortality in population that is not screened. The cumulative mortality of the screening policy with 1-year interval is lower than the cumulative mortality for 2-year screening policy, but the reduction of mortality is much smaller than that between a 3-year and a 2-year screening policy. Hence, the positive effect is outweighed by a much larger increase of false positive women.

Nevertheless, by representing the effectiveness results of the model in the form of (quality-unadjusted) life-years saved (LYS) and thereby "ignoring" also the effect of false positives and DCIS overdiagnosis, the screening policies with a 1-year interval produce higher effects than the screening policies with a 2-year interval. The ICER of screening women from 40 to 80 years of age by a 1-year interval (policy 4) over the same policy with a 2-year interval (policy 16) is €395,384 per LYS. The ICERs for the policies 28 and 16 are €11,604 and €31,486 per LYS, respectively, and are lower than the ICERs for cost per QALY. This is expected, as the negative effects of screening are ignored in the cost per LYS estimates, which leads to underestimation of screening's cost-effectiveness. Therefore, the inclusion of the QALYs in the model proves crucial, as the model is then able to capture the negative and positive effects of screening more accurately.

Comparison with Other Economic Evaluations

As already mentioned, numerous studies on the economic evaluation of breast cancer screening were performed. Although the majority of economic evaluations are cost-effectiveness analyses, most of them do not adhere to basic economic evaluation principles, which make them difficult to compare. Additionally, because of the lack of standardization among the studies, a wide range of results can be found across breast cancer screening studies [37]. Published estimates of the screening cost per life-year gained vary from $2844 to $114,300 [38]. Therefore, a suitable comparison can only be made with similar models, such as MISCAN [3,4]. De Koning et al. estimated the cost gained for various screening policies in The Netherlands from €2620 to €4482 per QALY [39]. For the same policies, our model gives estimates of the cost gained from €4560 to €7368 per QALY, which is reasonably comparable.

Model Limitations

As with all modeling studies, a number of limitations of the present study are noticeable. First, it is difficult to model a very heterogenic disease, such as breast cancer. Therefore, the choice of appropriate health states must be considered with great care to conveniently capture the progression of breast cancer. We chose a simplified TNM stage of breast cancer as the classification criterion of health stages and sacrificed comparability with other studies to model the Slovenian population of women as closely as possible. Moreover, because of the disease characteristics, where the effects of treatment are not observable in a short period, the effect of current treatment cannot be estimated precisely. In the model, treatment options for women diagnosed with breast cancer from 1991 to 1995 were used to capture the 7-year mortality from breast cancer. It is anticipated that because of improvements in breast cancer treatment (new drugs, better surgery techniques, guidelines, etc.), better survival is expected from present treatment options. Nevertheless, we believe that observed stage-specific mortality for "old" treatments captures the mortality reduction of breast cancer screening more adequately than just the assumed value (i.e., of 25–30%) that most of the other studies use.

Questions also arise about applicability of the model inputs based on data obtained from foreign countries. As some factors like attendance, recall, and portion of invasive diagnostics on recall can be country-specific [22,23], great care was taken in estimation of those model inputs. Also, as the prediction of those parameters is quite uncertain, distributions with relatively large ranges were applied to those input parameters (i.e., 95% confidence intervals for attendance were 60–87%, for recall 1–17%, and for invasive diagnostics 5–43%) to implement the lack of knowledge about those parameters. Furthermore, since this is a country-specific economic evaluation, the transferability of results and conclusions to other countries is questionable. This problem, or more appropriately this limitation, is solved by program characteristics that enable easy incorporation of country-specific breast cancer incidence, life tables, costs, survival, and other inputs. This feature also allows us to adopt the model to new evidence as it becomes available thus reducing the degree of uncertainty in model inputs and consequently in model results.

Further limitation of the model is also the choice of no-screening as the null option. Only opportunistic screening activity goes on in Slovenia in accordance with the instructions for preventive health care, which state that women aged 50 years and older are liable for monographic examination every 2 years. Since the extent of the opportunistic screening in Slovenia could not be presently estimated, the ICER for present situation to no-screening cannot be calculated. But as we have shown that, for example, the screening from ages 50 to 70 years every 2 years is cost-ineffective, the screening policies of choice in our analysis can only be more cost-effective when compared to the present situation. Therefore, the results of our analysis are most probably overestimates of the current situation in Slovenia. The extent of this overestimation is unfortunately not known.

Conclusions

In the present cost-effectiveness analysis, the question regarding the breast cancer screening in Slovenia was addressed. It was shown that a 1-year screening interval in population breast cancer screening would produce fewer benefits at higher costs than less intensive screening and that screening policies with a 2-year interval would be cost-effective only at high values of society's willingness to pay per QALY. This can be attributed to the increased costs and decreased benefits because of false positives and DCIS overdiagnosis. But these two negative aspects of screening are prevailed in a 3-year interval screening policies by the reduced breast cancer mortality and by the shift of the breast cancer stages toward earlier stages. Therefore, based on commonly quoted thresholds of society's willingness to pay per QALY, the policy of choice for breast cancer screening in the Slovenian population would be screening women aged from 40 to 80 years every 3 years. With this policy, the optimal ratio between the positive and the negative effects of breast cancer screening per invested costs is obtained.

To access supplementary material for this article, please go to: http://www.ispor.org/publications/value/ViHsupplementary.asp

The authors would like to thank Ana Žličar and Janez Žgajnar from the Institute of Oncology Ljubljana for valuable help and the four anonymous reviewers for helpful suggestions.

References

1 Cancer Registry of Slovenia. Epidemiology and Cancer Registry. Slovenia: Institute of Oncology Ljubljana.
2 Shapiro S, Coleman EA, Broeders M, et al. Breast cancer screening programmes in 22 countries: current policies, administration and guidelines. International Breast Cancer Screening Network (IBSN) and the European Network of Pilot Projects for Breast Cancer Screening. Int J Epidemiol 1998;27:735–42.
3 Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed 1985;20:79–93.
4 van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al. A model for breast cancer screening. Cancer 1990;66:1601–12.
5 Szeto KL, Devlin NJ. The cost-effectiveness of mammography screening: evidence from a microsimulation model for New Zealand. Health Policy 1996;38:101–15.
6 Lippman ME. Breast cancer. In: Kasper DL, ed. Harrison's Principles of Internal Medicine (16th ed.). New York: McGraw-Hill, Medical Publishing Division, 2005.
7 Yen MF, Tabar L, Vitak B, et al. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. Eur J Cancer 2003;39:1746–54.
8 Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993;13:322–38.
9 Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 1998;13:397–409.
10 SAS OnlineDoc, version 8. SAS Institute Inc., Cary, NC, USA, 1999. Available from: http://v8doc.sas.com/sashtml/ [Accessed December 17, 2004].
11 Siegel JE, Torrance GW, Russell LB, et al. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on Cost Effectiveness in Health and Medicine. Pharmacoeconomics 1997;11:159–68.
12 Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices-Modeling Studies. Value Health 2003;6:9–17.
13 Pompe-Kirn V, Japelj B, Primic-Žakelj M, et al. Which risk factors could have affected breast cancer incidence in Slovenia in the past, and what are the predictions for this decade. Zdrav Vestn 2001;70:341–5.
14 Burstein HJ, Polyak K, Wong JS, et al. Ductal carcinoma in situ of the breast. N Engl J Med 2004;350:1430–41.
15 Cady B, Chung MA. The prevention of invasive breast carcinoma. Cancer 2004;101:2147–51.
16 Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive value in the Florence District Programme (1975–1986). Int J Epidemiol 1991;20:852–8.
17 Duffy SW, Chen HH, Tabar L, et al. Sojourn time, sensitivity and positive predictive value of mammography screening for breast cancer in women aged 40–49. Int J Epidemiol 1996;25:1139–45.
18 Brekelmans CT, Westers P, Faber JA, et al. Age specific sensitivity and sojourn time in a breast cancer screening programme (DOM) in the Netherlands: a comparison of different methods. Epidemiol Community Health 1996;50:68–71.
19 Michaelson JS, Halpern E, Kopans DB. Breast cancer: computer simulation method for estimating optimal intervals for screening. Radiology 1999;212:551–60.
20 Kopans DB, Rafferty E, Georgian-Smith D, et al. A simple model of breast carcinoma growth may provide explanations for observations of apparently complex phenomena. Cancer 2003;97:2951–9.
21 Kerlikowske K, Grady D, Barclay J, et al. Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation. JAMA 1996;276:39–43.
22 Perry N, Broeders M, deWolf C, Tornberg S. European Guidelines for Quality Assurance in Mammography Screening. Luxemburg: Office For Official Publications of the European Communities, 2001.
23 Smith-Bindman R, Chu PW, Miglioretti DL, et al. Comparison of screening mammography in the United States and the United Kingdom. JAMA 2003;290:2129–37.
24 de Haes JC, de Koning HJ, van Oortmarssen GJ, et al. The impact of a breast cancer screening programme on quality-adjusted life-years. Int J Cancer 1991;49:538–44.
25 Hutton J, Brown R, Borowitz M, et al. A new decision model for cost-utility comparisons of chemotherapy in recurrent metastatic breast cancer. Pharmacoeconomics 1996;9(Suppl. 2):S8–22.
26 Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russel LB, et al., eds. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, 1996.
27 Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000;17:479–500.
28 Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998;18(Suppl. 2):S68–80.
29 Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. Health Econ 1999;8:257–61.
30 Fenwick E, Claxton K, Schulper M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001;10:779–87.
31 Smith RA, Saslow D, Sawyer KA, et al. American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast cancer

screening: update 2003. CA Cancer J Clin 2003; 53:141–69.
32 Drummond FM. Methods for the Economic Evaluation of Health Care Programmes (2nd ed.). New York: Oxford University Press, 1997.
33 Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. Pharmacoeconomics 1996; 9:113–20.
34 Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision-making: recent developments in, and challenges for, cost-effectiveness research. In: Smith P, Sculpher M, Ginnelly L, eds. Health Policy and Economics: Opportunities and Challenges. London: McGraw-Hill Education, 2004.
35 Eichler HG, Kong SX, Gerth WC, et al. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health 2004; 7:518–28.
36 Towse A, Pritchard C. Does NICE have a threshold? An external view. In: Towse A, Pritchard C, Devlin N, eds. Cost-Effectiveness Thresholds. Economic and Ethical Issues. London: King's Fund and Office of Health Economics, 2002.
37 Brown DW, French MT, Schweitzer ME, et al. Economic evaluation of breast cancer screening: a review. Cancer Pract 1999;7:28–33.
38 Arveux P, Wait S, Schaffer P. Building a model to determine the cost-effectiveness of breast cancer screening in France. Eur J Cancer Care 2003;12:143–53.
39 de Koning HJ, van Ineveld BM, van Oortmarssen GJ, et al. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49:531–7.