3-4-1992

Cost-Utility Analysis of Alternative Strategies in Screening for Breast Cancer

Mark S. Kamlet
Carnegie Mellon University, KAMLET@CMU.EDU

Giovanni Parmigiani
Duke University

A COST-UTILITY ANALYSIS

OF ALTERNATIVE STRATEGIES

IN SCREENING FOR BREAST CANCER

Giovanni Parmigiani

Institute of Statistics and Decision Sciences, Duke University

Mark S. Kamlet

Department of Social and Decision Sciences, Carnegie Mellon University

March 4, 1992

Abstract

Screening for early detection is a primary way to control breast cancer. The choice of effective screening policies involves substantial uncertainty and difficult tradeoffs among medical costs and the duration and quality of life. In this paper, we study the choice of the age at which screening should begin and the frequency of screening tests. These have been issues of important debate in the health policy community. We address these questions using the framework of cost-utility analysis, as we consider it important to evaluate the outcomes of preventive care in terms of morbidity and quality of life, in addition to survival. After a brief review of background information on breast cancer screening, we introduce cost-utility analysis and its relation with Bayesian decision-making. We then discuss modelling and prior specifications, and carry out a cost-utility evaluation of the currently recommended policy. Finally we compare those to the the Bayes-optimal decision for various values of the exchange rate between dollars and quality of life. We discuss the implications of the results for actual policy decisions.

1 Introduction

In the United States, it is projected that one in ten women will contract breast cancer in their lives. This represents approximately 150,000 new cases every year. At present, 35% of all cases of breast cancer are fatal, which makes breast cancer second only to lung cancer as the leading cause of death from cancer in women (Silverberg, Boring and Squire, 1990). Even though significant progress has been made in recent years, increased understanding of the etiology has not proven sufficient to result in a substantial reduction in incidence. Early detection through mammographic screening, however, can improve prognosis substantially, both in terms of mortality and morbidity. Therefore, mass screening for early detection is currently, and will be in the foreseeable future, the most effective way to control breast cancer.

Whereas most of the scientific, professional and governmental organizations recommend screening examinations, recommendations differ with regard to the suggested timing. For example, the American Cancer Society and the National Cancer Institute recommend a baseline test in women aged 35 to 40 years, annual or biannual tests from 40 to 49 years and annual tests in older women. On the other hand, the U.S. Preventive Services Task Force and the Council of Scientific Affairs of the American Medical Association are more cautious in recommending screening in women aged 40 to 49, withholding the judgment until more evidence about screening women under 50 years of age becomes available. The frequency of exams has also been the subject of controversy and in countries like Sweden and the Netherlands the recommended frequency is two years, while in the U.S. the frequency is annual. In this paper we analyze the cost-effectiveness of the proposed strategies and compare it to that of the Bayes-optimal solution in various different scenarios.

Screening is typically analysed in terms of reduction in mortality (Eddy et. al., 1988). We carry out an evaluation that accounts, in addition to mortality, for morbidity and other aspects of quality of life. To this purpose, the health benefits that result from a given screening schedule must be converted into measures that are meaningful in terms of the quality and length of life for the individuals receiving maintenance treatment. Additionally, the costs associated with

screening must be identified and measured. Finally, a comparison must be made between the costs and the net health impacts that result from them. To do this, we adopt a cost-utility analysis (CUA) approach. CUA evaluates a health intervention by comparing the incremental societal costs of a health intervention and the incremental health benefits that result from it. Typically, the outcome of a CUA is expressed as a ratio, with the units being dollars per quality adjusted life years ($/QALY's). CUA represents a natural methodology for applied decision problems in preventive medicine, drug design and assessment, reliability engineering, regulatory decisions and so on, where abundance of prior expertise mandates a Bayesian treatment, but where standard expected utility and cost-benefit analysis are unlikely to provide a completely satisfactory answer. In this sense, CUA can be regarded as a very important extension of the set of tools available to Bayesian analysts.

An important advantage of Bayesian analyses in the screening problem derives from the need for policy indications before the effects of the various policies can be studied empirically. In breast cancer screening, the outcome of large scale clinical trials regarding screening of woman aged 40 to 49 has been considered by some as a necessary condition to make any policy recommendation. From the point of view of Bayesian analysis, such availability is obviously desirable, but by no means necessary, to reach a decision. The same applies to the use of mammographic techniques of unprecedented sensitivity.

The discussion proceeds as follows. In Section 2 we provide a brief description of CUA and outline its relation to subjective expected utility theory. Then, in Section 3, we discuss the modelling of the natural history of breast cancer. We also derive analytic expression for costs and QALY as a function of the possible decisions. In Section 4 we discuss the elicitation of probabilities and utilities and finally, in Section 5 we present the results of the analysis.

2 Cost-Utility Analysis and Bayesian Decisions

CUA has roots both in economic analysis and decision analysis. Its earliest development grew out of cost-benefit analysis in economics, under the rubric of "cost-effectiveness analysis" (CEA).

The concept of cost-effectiveness has taken on a multiple of meanings over time, however. These range from "saving money while not impairing health," to "the lowest cost way of achieving a given set of health outcomes" (Gramlich, 1981; Office of Technology Assessment, 1980; Rapopart, Robertson, and Stuart, 1982), to "the best way to achieve whatever objectives a decisionmaker is pursuing" (Warner, 1983; Hatziandreu et al., 1989).

A more formal decision-theoretic formulation of cost-effectiveness, using the $/QALY framework is proposed by Pliskin, Shepard and Weinstein (1980). See also Weinstein and Fineberg (1980). In this formulation, quality of life is formally equivalent to (von Neumann - Morgenstern) utility. A utility value of 1 is assigned to "full health" and a utility value of 0 to "death." Intermediate values are interpreted accordingly. An individual is assigned a utility value in this fashion as a function of his or her health state at each point in time. The quality-adjusted life year measure for a health intervention is then determined by integrating this utility-weighted utility over an individual's life subsequent to the health intervention.

Weinstein et al. (1982) provide the formal conditions under which such a $/QALY outcome measure is consistent with subjective expected utility theory. In order for CUA and its outcome measure, $/QALY, to be consistent with a utility-based decision analytic framework, several properties must be fulfilled concerning utility functions over health states. First, utility over health states must display utility independence between length of life and quality of life. This means that tradeoffs between length of life are not affected by the quality of life experienced by the individual. Similarly, tradeoffs among levels of morbidity are not affected by the length of life that the individual will live. Second, utility over health states must display proportional trade-offs. This means the following. Suppose a person is indifferent between X years of life in health state A and Y years of life in health state B, where X is less than Y and A is a better health state than B. Then he or she must also be indifferent between spending aX years of life in health state A and aY years of life in health state B. Finally, the individual must display risk neutrality with respect to years of life.

Following the lead of Anderson et al. (1985), Torrance (1986), and Drummond et al. (1987), we use the term "cost-utility analysis" to distinguish that approach from the other meanings of

CEA. Nonetheless, as shorthand we will sometimes use the term "cost-effective" to mean that a given screening schedule is desirable from a CUA perspective. Similarly, we will sometimes use the term "cost-effectiveness ratio" to refer to the ratio of costs to health impacts.

While CUA has evolved over time towards a well-defined set of methodological procedures for analyzing the efficacy of health interventions from an economic perspective, several important controversies remain in the literature. One concerns the use of discount rates. The issue is not only what discount rate to use but also whether health outcomes and costs should be discounted or only the latter. We choose several discount rates, 0%, 3%, and 5%, and examine the results obtained when discounting both costs and health outcomes and when discounting only costs.

Another controversy in the CUA literature occurs when, as is the case here, the impact of a health intervention involves more than direct medical costs and more than health narrowly conceived. Here, there are impacts of breast cancer on the social functioning of the individual, her work, her family life, and so forth. As discussed in Kamlet (1991), these impacts can in principle either be costed out and included in the costs of the health intervention (the numerator of the CUA cost-effectiveness ratio), or measured in terms of their quality of life impact and included in the quality of life measure (the denominator of the cost-effectiveness ratio). We include direct medical costs and the money equivalent of leisure devoted to treatment (measured in terms of the willingness to pay of the individual for the time involved) as costs in the numerator of the cost-effectiveness ratio. We therefore consider the remaining impacts, including the so-called indirect costs of the illness on ability to work and productivity at work, in terms of quality of life in the denominator of the cost effectiveness ratio.

3 Model

Consider a patient, with either no cancer or undetectable cancer, facing the choice of what screening recommendations to follow. A convenient and adequately general way to model the natural history of chronic diseases for the purpose of screening is discussed in detail in Parmigiani (1990). In brief, we consider a stochastic process with four states: one with no cancer, or

undetectable cancer (termed, for brevity, pre-detectable); one with detectable asymptomatic cancer (called pre-clinical); one with symptomatic cancer (called clinical), and one representing death. Transitions can occur from pre-detectable to pre-clinical, from pre-clinical to clinical and from any state to death. The time spent in the various states is random. Let Y be the the sojourn time in the pre-detectable state, beginning at birth, and U the sojourn time in the pre-clinical state. If no screening takes place, $Y + U$ is the age of the patient at the time of the surfacing of symptoms, and consequent treatment. Also, let $f(y)$ represent the density of transitions from pre-detectable to pre-clinical, $h(y)$ the density of transitions from pre-detectable to death and $g(u|y)$ and the conditional transition density from pre-clinical to clinical, given an arrival in the pre-clinical state at time y. It is important to allow for a dependence between Y and U as younger women tend to contract faster growing tumors. Transitions from pre-clinical to death are not considered in this model. The probability of dying from a cause other than breast cancer while in the pre-clinical state is between .01 and .02 depending on the age of the patient, (Parmigiani, 1991). Therefore, such omission should have negligible consequences on the conclusions.

The sensitivity of mammographies will be denoted by β. The probabilities of a false negative result in successive mammographies on the same patient are, somewhat restrictively, assumed to be independent. Specificity is easier to handle, as it can be factored in as part of the cost of examination. Usually, positive mammograms are followed by a highly specific biopsy. Therefore, screening does not terminate unless the illness is actually present. False positives of the initial mammogram that lead to biopsies will, however, represent an additional cost.

Screening examinations are scheduled for asymptomatic patients, and terminate as soon as the clinical stage is reached (interval detection), the disease is detected by screening in the pre-clinical stage (screen detection), or the patient dies. An examination strategy (or schedule) will consist of an age α at which examinations begin, and of an interval δ at which rate examinations continue. So the i-th screening examination takes place at age $\alpha + (i-1)\delta$. We confine attention to periodic schedules (i.e. constant δ). Policies with age-dependent δ have been studied by Kirch and Klein (1974) and Parmigiani (1991), who showed that the additional benefits of

age-dependence probably do not outweigh, at least in the case of breast cancer, the additional operational difficulties.

The main advantage associated with screening is the ability to detect the disease at an early stage. This has been documented to entail longer life expectancy (see Habbema et al., 1986). In addition there are gains in quality of life in case of early detection. Here we assume that patients are treated according to the guidelines developed by the NIH Consensus Conference on Early Breast Cancer (1991). The main difference in treatment resulting from an early detection stems from the fact that screen detected cases present a much lower percentage of cases with positive axillary node involvement. While adjuvant chemoterapy is recommended in case of positive node involvment, there is still not enough evidence to make a clear recommendation in case of negative node involvment and the choice regarding chemoterapy is left to the patient. Consequently, screen detected cases will receive chemoteratpy less often, with a resulting improvement in quality of life. Most other treatment decisions will not depend decisevely on screen detection, and therefore will not be considered.

We assume that each examination has a fixed cost C_m. This amount includes the direct medical costs of mammography as well as the opportunity cost value of the patient's time. There is a small loss in quality of life due to false positive examinations which can be neglected. Also, Let C_e (for early) and C_l (for late) be the total costs of treatment in case of screen detection and interval detection respectively.

The quality of life enjoyed by the patient while in the pre-detectable state is not affected by the decisions about screening. Therefore, for the purpose of comparing strategies, we only need to consider quality of life from Y on. Let $Q_e(y,u)$ and $Q_l(y,u)$ be the expected QALY given a transition to the pre-clinical state at age y, and a sojourn time in the pre-clinical stage of u, in case of screen detection and interval detection respectively. Finally, let r_1 and r_2 be the discount rates for monetary payments and health outcome respectively.

In the remainder of the section we give analytic expression for three critical quantities in the analysis: the expected number of examinations, $\mathcal{I}_{\alpha,\delta}$, the expected treatment cost, $\mathcal{C}_{\alpha,\delta}$ and the expected QALY, $\mathcal{Q}_{\alpha,\delta}$, for a fixed choice of α and δ. If $\mathcal{C}_0$ and $\mathcal{Q}_0$ are the cost and QALY

associated with no screening, the cost-utility ratio is the given by $(C_m\mathcal{I}_{\alpha,\delta}+\mathcal{C}_{\alpha,\delta}-\mathcal{C}_0)/(\mathcal{Q}_{\alpha,\delta}-\mathcal{Q}_0)$. A standard expected utility analysis can be carried out based on a linear combination of the three quantities. Let $p_{i0}(y) = G(\alpha+i\delta-y|y)$, $p_{ij}(y) = G(\alpha+(i+j)\delta-y|y) - G(\alpha+(i+j-1)\delta-y|y)$ and $q_{ij}(y) = 1-G(\alpha+(i+j)\delta-y|y)$, where G is the c.d.f. of g. Then:

$$
\begin{aligned}
\mathcal{I}_{\alpha,\delta} &= \int_0^\alpha \sum_{j=0}^{\infty}(j+1)(1-\beta)^j[p_{0j}(y)+\beta q_{0j}(y)]f(y)e^{-r_1 y}dy \\
&\quad + \sum_{i=1}^{\infty}\int_{\alpha+(i-1)\delta}^{\alpha+i\delta}\left[ih(y)+\sum_{j=0}^{\infty}(i+j+1)(1-\beta)^j[p_{ij}(y)+\beta q_{ij}(y)]f(y)\right]e^{-r_1 y}dy \\
\mathcal{C}_{\alpha,\delta} &= \int_0^\alpha \sum_{j=0}^{\infty}(1-\beta)^j[C_l p_{0j}(y)+C_e\beta q_{0j}(y)]f(y)e^{-r_1 y}dy \\
&\quad + \sum_{i=1}^{\infty}\int_{\alpha+(i-1)\delta}^{\alpha+i\delta}\sum_{j=0}^{\infty}(1-\beta)^j[C_l p_{ij}(y)+C_e\beta q_{ij}(y)]f(y)e^{-r_1 y}dy \\
\mathcal{Q}_{\alpha,\delta} &= \int_0^\alpha\left[\beta\int_{\alpha-y}^{\infty}Q_e(y,u)g(u|y)du+\int_0^{\alpha-y}Q_l(y,u)g(u|y)du\right]f(y)e^{-r_2 y}dy \\
&\quad \int_0^\alpha\sum_{j=1}^{\infty}(1-\beta)^j\left[\beta\int_{\alpha+j\delta-y}^{\infty}Q_e(y,u)g(u|y)du+\int_{\alpha+(j-1)\delta-y}^{\alpha+j\delta-y}Q_l(y,u)g(u|y)du\right]f(y)e^{-r_2 y}dy \\
&\quad + \sum_{i=1}^{\infty}\int_{\alpha+(i-1)\delta}^{\alpha+i\delta}\sum_{j=0}^{\infty}(1-\beta)^j\left[\beta\int_{\alpha+(i+j)\delta-y}^{\infty}Q_e(y,u)g(u|y)du\right. \\
&\qquad\qquad \left.+\int_{\alpha+(i+j-1)\delta-y}^{\alpha+(i+j)\delta-y}Q_l(y,u)g(u|y)du\right]f(y)e^{-r_2 y}dy.
\end{aligned}
$$

The derivation of the above expectations follows Parmigiani (1992).

4 Evaluation of Probabilities, Utilities and Costs

The densities described in this section represent our marginal prior densities on the sojourn times for a generic patient with no information available about risk factors. Such information can, however, be incorporated by changing densities when risk factors, such as history of breast cancer in the family, become known.

The evaluation of the transition densities is based on several existing studies. Moolgavkar, Stevens and Lee (1979) developed estimates of the the incidence of breast cancer, accounting for both age and cohort effect. Based on these one can estimate the density of $Y + U$. Deaths from other causes can be derived from life tables and used to evaluate the density h. We assumed h to be a Weibull density, that is: $h(y) = \frac{a}{b}\left(\frac{y}{b}\right)^{a-1} \exp\{-\left(\frac{y}{b}\right)^{a}\}$ Least squares yield parameter estimates of $a = 7.233$ and $b = 82.651$.

Spratt, Greenberg and Heuser (1986) obtained estimates of the sojourn time in the pre-clinical stage that depend on the age of the patient, of great importance in this problem. At each fixed age, the authors postulated a lognormal distribution, given by: $p(u) = \frac{1}{\sqrt{2\pi}su} \exp\{-\frac{1}{2s^2}(\log u - m)^2\}$ The choice of the lognormal is motivated by knowledge regarding growth patterns of breast tumours. The authors obtained evaluations at 8 different ages. It is convenient to specify a sojourn time distribution indexed by an arbitrary age. To this end, we modelled the location parameter m of the lognormal as a logistic function of the age y, that is: $\log(m_0 - m(y)) = m_1 + m_2 y$. We specified m_0 to approach linearity and constant variance in the regression; a convenient choice was $m_0 = 1.4$. This yields least squares estimates of $m_1 = 1.6$ and $m_2 = -0.038$. Due to the highly noisy information, a second stage prior distribution was assigned to the parameter S. We chose a Inverse Gamma density with parameters $a = 6.33$ and $b = 3.36$, determined based on evidence from Spratt, Greenberg and Heuser (1986). After marginalization over S, the distribution of U given $Y = y$ is a scale mixture of lognormals, given by:

$$g(u|y) = \frac{B(a, .5)}{u}\sqrt{\frac{b}{2}}\left[1 + \frac{b}{2}\left(\log u - m_0 + e^{m_1 + m_2 y}\right)\right]^{-\frac{2a+1}{2}}.$$

Further, this density can be used for estimating the number of individuals moving from the pre-detectable to the pre-clinical stage at any given age. First, the number of cases was estimated using a cubic spline curve. Then a Weibull curve was fitted to the derived table. The resulting values of the parameters are $a = 4.48$ and $b = 65.52$. Further details on the procedure followed in fitting can be found in Parmigiani (1991).

As of errors in mammography-based diagnosis, the probability of a false positive is about $.03 - .05$ (Eddy et. al. 1988). Estimates of the sensitivity based on technology of the late 1970's give an expected value around .8 (Brookmeyer, Day and Moss, 1986). New and much more sensitive technologies have been developed since then, and we find it of interest to consider the case of sensitivity equal to 1. Using data from Shapiro et. al. (1988), the probabilities of axillary node involvement are estimated to be .227 in case of screen detection versus .424 in case of interval detection. The improvement in life expectancy due to early detection is taken to be 22%, based on data from the HIP (Health Insurance Plan of Greater New York) study; (see Habbema et al., 1986). These estimates are based on the therapy choices prevalent when the HIP study was carried out, and may need to be revised when evidence regarding screening under the new NIH recommendations becomes available.

Moving to financial charges, a mammography can cost from \$50 to \$200, and the biopsy that follows it in case of positive outcome costs around \$1000. Therefore, a reasonable choice for C_m is \$200. Following Hillner and Smith (1991), the cost of adjuvant chemoterapy is evaluated to be \$13.000, inclusive of treatment cost as well as costs related to complications. We assumed that 80% of node negative women will recieve adjuvant chemoterapy, but the final results are not sensitive to the choice of this number in a range from 40% to 95%. Patients saved from cancer by early detection may incur further medical expenses later on in their lives. This additional cost implied by screening is not factored into the present analysis.

Despite the important quality of life/morbidity effects associated with breast cancer, there has been relatively little attention devoted in the literature to quality-of-life measures for breast cancer outcomes, at least with regard to measures that can be interpreted as utility in a decision-theoretic sense. Eddy et. al. (1988), for instance, focuses only on the mortality consequences of screening in examining the value of mammography screening in women under age 50. Similarly, Miller (1991), in a recent review of the literature on mammography screening with an express intent of determining whether such screening is appropriate for women in their forties, restricted attention purely to mortality reduction. The loss of quality of life in health states following the detection of breast cancer has been recently discussed in deHaes et al. (1991) (see also Sackett

and Torrance, 1978). We have used estimates adapted from deHaes et al. Patients will be assumed to experience some ongoing reduction in quality of life: quality of life is .2 during two months of perioperative period, .5 during the remaining of the the first year if chemoterapy is administered, and .9 for the remainder of disease free survival.

5 Results

Consider a person who is known to be in the pre-detectable state at age 35, with no information available on risk factors. Suppose that mammographies have sensitivity equal to unity —we will move to the case of smaller sensitivity later.

We begin with the evaluation of the cost per QALY of the four screening policies obtained by setting α equal to 40 and 50, and δ equal to 1 and 2. Table 1 summarizes the results when both r_1 and r_2 are zero. Costs are expressed in thousands of 1989 dollars. The results show that the difference in \$/QALY between annual and biannual policies in substantial. However, all strategies have a \$/QALY well below the typical cutoff for a cost-utility value of \$ 50.000 (see for example Frybeck and Thornbury, 1991) typically used in the literature to judge wether a given health intervention represents a reasonable use of resources.

	$\alpha = 40, \delta = 1$	$\alpha = 40, \delta = 2$	$\alpha = 50, \delta = 1$	$\alpha = 50, \delta = 2$
Additional Cost	6.976	3.508	5.073	2.558
Gain in QALY	0.219	0.191	0.170	0.154
Ratio	31.772	18.302	29.702	16.576

Table 1: \$/QALY evaluation for four basic strategies. $r_1 = r_2 = 0$.

Table 2 shows how the result change depending on the assumptions on the interest rates. Discounting both dollars payments and health outcomes increases the \$/QALY ratio, as health outcomes are expected to occur at a later time.

We move now to the consideration of the case in which the sensitivity of mammographies is .8. Table 3 summarizes the results. In general, sensitivity of mammographies has an impact on

	$\alpha = 40, \delta = 1$	$\alpha = 40, \delta = 2$	$\alpha = 50, \delta = 1$	$\alpha = 50, \delta = 2$
$r_1 = .03$, $r_2 = 0$	8.586	4.932	7.599	4.229
$r_1 = .05$, $r_2 = 0$	3.821	2.188	3.217	1.785
$r_1 = .03$, $r_2 = .03$	34.073	19.774	30.756	17.264

Table 2: \$/QALY ratios for four basic strategies at varying discount rates.

cost-effectiveness, mostly through a lower gain in QALY. Also, when the sensitivity is lower, the cost-effectiveness gap between strategies with $\alpha = 50$ and $\alpha = 40$ is substantially reduced, as frequent examination contribute to the detection of cases missed by previous examinations.

	$\alpha = 40, \delta = 1$	$\alpha = 40, \delta = 2$	$\alpha = 50, \delta = 1$	$\alpha = 50, \delta = 2$
Additional Cost	6.983	3.521	5.079	2.568
Gain in QALY	0.203	0.159	0.159	0.129
Ratio	34.370	22.056	31.829	19.834

Table 3: \$/QALY evaluation for four basic strategies.$r_1 = r_2 = 0$ $r_1 = r_2 = 0$ $\beta = .8$.

Finally, we consider the choice of α and δ, with $\beta = 1$. Combining costs and QALY into a single objective function requires specification of an exchange rate λ between dollars and QALY units. Figure 1 displays the trade-off curves for varying λ.

One way to interpret the curves is the following. For a fixed value of the expected cost of examination and treatment on the vertical axis, the curve gives the highest QALY that can be obtained within the class of policies considered. Likewise, for a fixed value of QALY on the horizontal axis, the curve gives the lowest cost at which that can be reached. For example, the optimal way to spend 3.65 thousand additional dollars is to start screening at age $\alpha = 38.7$ and continue every $\delta = 1.98$ years.

The advantage of optimal schedules over some of the current recommendations may be significant. In particular, it appears that biannual policies approach optimality at both $\alpha = 40$ and $\alpha = 50$ for appropriate values of λ. On the other hand, annual policies show a significantly higher cost not rewarded by an adequate increase in QALY. For example if $\alpha = 40$ and exams

are annual, the marginal cost —that is the cost of a very small additional increment in QALY— is over \$300.000/QALY. Also, moving from $\alpha = 40$ and biannual exams to $\alpha = 40$ and annual exams implies a spending of \$3,470 for an increase of .028 QALY, or well over \$100,000/QALY. Finally, early screening seems to be worthwile, as the incremental cost of moving from $\alpha = 50$ to $\alpha = 40$ with biannual exams is around \$25/QALY.

6 Conclusions

Screening for breast cancer is effective in improving quality of life as well as in reducing mortality. Skepticism towards mammographic screening of women aged 40 to 49 has been supported by studies where the endpoint of the analysis is mortality. When evaluated in terms of QALY, screening of women aged 40 to 49 appears to be cost-effective. Early gains in quality of life, stretching over a long period of time, account for this result.

A further controversial issue is the frequency of examinations. According to our analysis, if mammographic technology with high sensitivity is available, annual examinations may be unnecessary (reflecting in the high marginal cost). Thus a higher sensitivity makes effective screening more affordable and easier to implement on a large scale than it has been in the past.

Finally, sensitivity analysis is necessary to corroborate the validity of the conclusions of this study. The number of parameters and assumptions involved makes it prohibitive to attempt a systematic account of our findings. In summary, some of the numerical evaluations can be sensitive to important parameters —like the QALY measurement, and assumptions —like the form of the sojourn time distribution. Assumptions on treatment are also crucial. In a scenario in which early detection implies milder surgery and and persistent gain in quality of life, the cost-effectiveness is greater. However, we found that the relative ranking of the strategies, the form of the trade-off curves and therefore the main qualitative conclusions of this study, appear to be robust.

Acknowledgement: We thank Karen Bash and the reviewers for very helpful comments.

References

[1] A. M. A.; Council on Scientific Affairs: (1989) "Mammographic Screening in Asymptomatic Women Aged 40 Years and Older," *J.A.M.A.* **261** 2535–2542.

[2] Brookmeyer R., S. D. Day and S. Moss: (1986) "Case Control Studies for the Estimation of the Natural History of Preclinical Disease from Screening Data," *Statistics in Medicine* **5** 127–138.

[3] de Haes, J. C. J. M. et al. (1991) "The Impact of Breast Cancer Screening Programme on Quality-Adjusted Life-Years" *Int. J. Cancer* **49** 538–544.

[4] Eddy D. M.: (1983) "A Mathematical Model for Timing Repeated Medical Tests," *Medical Decision Making* **3** 34–62.

[5] Eddy D. M., V. Hasselblad, W. McGivney and W. Hendee: (1989) "The Value of Mammography Screening in Women Under Age 50 Years," *Journal of the American Medical Association* **259 n. 10** 1512–1519. Philadelphia.

[6] Eddy D. M. and M. Shwartz:(1982) "Mathematical Models in Screening," *Cancer Epidemiology and Prevention* (D. Schottenfeld and J. F. Jr Fraumeni ed.) 1075–1090. Saunders, Philadelphia.

[7] Edwards, M.J. et. al.: (1988) "Economic Impact of Reducing Hospitalization for Mastectomy Patients," *Annals of Surgery* **208 (3)** 330–336.

[8] Fryback, S. C. and J. R. Thronbury: (1991) "The Efficacy of Diagnostic Imaging," *Medical Decision Making*, **11**, 86-95.

[9] Habbema J. D. F. et al.: (1986) "Age-Specific Reduction in Breast Cancer Mortality by Screening" *JNCI* **77-2** 317–320.

[10] Hillner, B. E. and T. J. Smith: (1991) "Efficacy and Cost Effectiveness of Adjuvant Chemoterapy in Women With Node-Negative Breast Cancer," *New England J. Medicine*, **324 (3)**, 160–168.

[11] Kamlet M. S.: (1991) *A Framework for Cost-Utility Analysis of Government Health Care Programs.* U.S. Department of Health and Human Services, Washington.

[12] Kirch, R. L. A. and M. Klein: (1974) "Examination Schedules for Breast Cancer," *Cancer* **33** 1444–1450.

[13] Louis, T. A, A. Albert and S. Heghinian: (1978) "Screening for the Early Detection of Cancer, III: Estimation of Disease Natural History," *Mathematical Biosciences* **40** 111–144.

[14] Miller, A. B.: (1991) "Is Routine Mammography Screening Appropriate for Women 40–49 Years of Age?," *American Journal of Preventive Medicine* **7** 55–62.

[15] Moolgavkar S. H., R. G. Stevens and J. A. H. Lee: (1979) "Effect of Age on Incidence of Breast Cancer in Females," *Journal of the National Cancer Institute* **62** 493–501.

[16] Moskowitz M.: (1986) "Breast Cancer: Age-Specific Growth Rates and Screening Strategies," *Radiology* **161** 37–41.

[17] NIH Consensus Development Conference.: (1991) "Early Stage Breast Cancer," *JAMA* **265 (3)** 391–395.

[18] Parmigiani G.: (1990) "On Optimal Screening Ages," *Technical Report n. 497, Dept. of Statistics, CMU* Pittsburgh, PA.

[19] Parmigiani G.: (1991) "Optimal Scheduling of Medical Exams: the Case of Breast Cancer," *ISDS, Duke University, Discussion Paper* **91-A16**.

[20] Pliskin, J.S., D.S. Shepard and M.C. Weinstein (1980) "Utility Functions for Life Years and Health Status: Theory, Assessment, and Application", *Op. Res.*, **28** 206–224.

[21] Prorok, P. C., B. F. Hankey and B. N. Bundy: (1981) "Concepts and Problems in the Evaluation of Screening Progams," *Journal of Chronic Diseases* **34** 159–171.

[22] Rapopart, J. F., R. Robertson and B. Stuart: (1982) *Understanding Health Economics.* Aspen, Rockwille MD.

[23] Sackett, D. L. and G. W. Torrance: (1978) "The Utility of Different Health Staes as Perceived by the General Public," *Journal of Chronic Disease* **31** 697–704.

[24] Shapiro, S, W. Venet, P. Strax and L.Venet: (1988) *Periodic Screening for Breast Cancer.* Johns Hopkins University Press, Baltimore.

[25] Silverberg, E., C. C. Boring and T. S. Squires: (1990) "Cancer Statistics 1990," *Ca–A Cancer Journal for Clinicians* **40** 9–26.

[26] Spratt, J.S., R.A. Greenberg and L.S. Heuser: (1986) "Geometry, Growth Rates, and Duration of Cancer and Carcinoma in Situ of the Breast before Detection by Screening," *Cancer Research* **46** 970–974.

[27] Walter S. D. and N. E. Day: (1983) "Estimation of the Duration of a Pre-clinical Disease State Using Screening Data," *American Journal of Epidemiology* **118** 865–886.

[28] Weinstein M. C. and H. Feinberg: (1980) *Clinical Decision Analysis* Saunders, Philadelphia.

[29] Weinstein M. C. and W. B. Stason: (1977) "Foundations of Cost-effectiveness Analysis for Health and Medical Practices," *New England Journal of Medicine* **296 (13)** 716–721.

[30] Williams, C. J. and R. B. Buchanan: (1987) *The Medical Management of Breast Cancer.* The Johns Hopkins University Press, Baltimore, MD.

[31] Zelen M.: (1992) *Optimal Scheduling of Examinations for the Early Detection of Disease. Biometrika*, forthcoming.

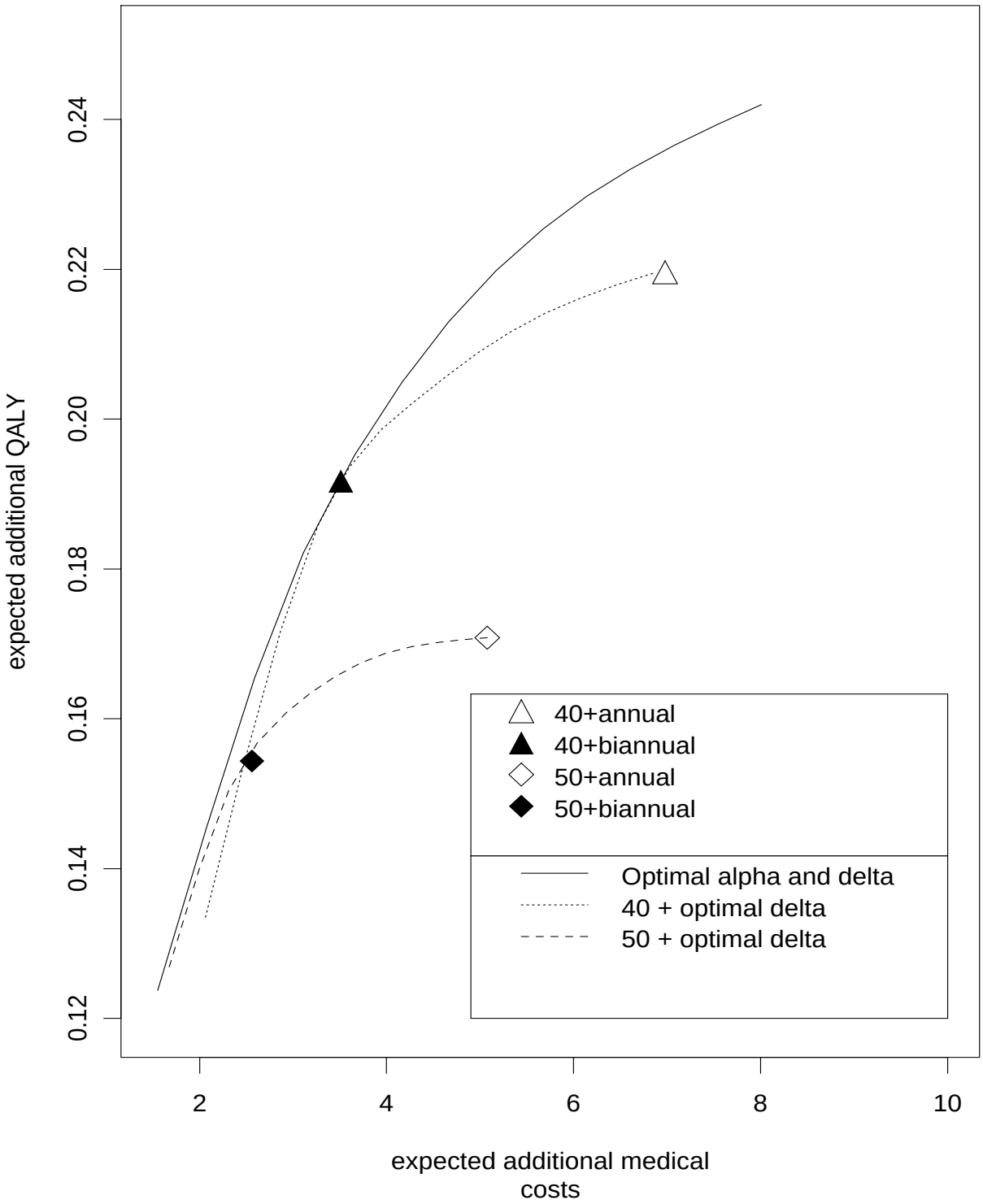

Figure 1: Trade-off analysis (1989 dollars).