International Journal of Technology Assessment in Health Care, **16:3** (2000), 811–821.

COST-BENEFIT ANALYSIS OF MAMMOGRAPHY SCREENING IN DENMARK BASED ON DISCRETE RANKING DATA

Dorte Gyrd-Hansen
University of Southern Denmark–Odense University

Abstract

Objective: Economic evaluations such as cost-effectiveness and cost-utility analyses generally fail to incorporate elements of intangible costs and benefits, such as anxiety and discomfort associated with the screening test and diagnostic test, as well as the magnitude of utility associated with a reduction in the risk of dying from cancer. This paper seeks to include all costs and effects incurred by introducing mammography screening through the application of discrete ranking modeling.

Methods: In the present analysis, 207 women were interviewed and asked to rank, according to priority, a number of alternative breast cancer screening setups. The alternative programs varied with respect to number of tests performed, risk reduction obtained, probability of a false-positive outcome, and extent of copayment. Using discrete ranking modeling, the stated preferences were analyzed and the relative weighting of the program attributes identified. For a range of hypothetical breast cancer programs, relative utilities and corresponding willingness-to-pay estimates were derived.

Results: A comparison of cost and willingness to pay for each of the programs suggested that net benefits are maximized when screening person aged 50–74 years biennially. More intensive screening produces lower or similar levels of utility at a higher cost.

Conclusion: Discrete ranking modeling can aid decision making by identifying inferior healthcare programs, i.e., programs that are more costly but less beneficial.

Keywords: Conjoint analysis, Breast cancer screening, Cost-benefit analysis

Various methods of screening for cancer diseases have been introduced in recent decades. The costliness of these programs varies depending on cost of screening test, cost of diagnostic test, and sensitivity and specificity of the test as well as the improvement in diagnosis by early detection. Existing and potential screening programs such as screening for cervical cancer, breast cancer, and colorectal cancer have been the focus of economic evaluations in order to determine whether the benefits justify the costs involved (6;7;10;13;15;16;17;25;26). The economic evaluations have, however, been limited to cost-effectiveness or cost-utility analyses, and have generally failed to incorporate elements of intangible costs and benefits. Moreover, the analyses have been limited by the inability of cost-effectiveness analysis to suggest the optimal program setup (screening interval and target group) among efficient programs. More knowledge of the nature of public preferences for screening programs and their attributes may contribute to the evaluation of such healthcare programs.

This paper seeks to establish a representative utility function for breast cancer screening programs that incorporates the utility and disutility associated with intangible effects. The method used is discrete ranking modeling. By establishing a utility function in which individual attributes are weighted according to their importance, relative utilities can be derived for hypothetical breast cancer screening programs. A comparison of relative utilities and program-specific costs makes it possible to identify dominated programs. Among efficient programs, relative cost-effectiveness can be calculated through a comparison of program-specific costs and willingness-to-pay (WTP) estimates.

METHODS

The Theory

In the context of healthcare services, it is of great significance that the relative importance of outcome and process attributes are identified such that the optimal healthcare service may be delivered. More specifically, in the context of cancer screening programs, there exists a vast number of program setups, the basic trade-off being frequency of screening versus screening effectiveness. In this context it is important to identify the tangible and intangible costs and benefits associated with increased frequency. As frequency of screening programs increases, more lives are saved. However, apart from being costly in monetary terms, frequent screening entails that healthy women undergo a larger number of tests over their lifetime, which may be associated with discomfort and anxiety, as well as a higher risk of receiving a false-positive diagnosis. If the increments in disutilities associated with the process of being screened overshadow the value of the extra lives being saved, frequent screening may not be worthwhile.

This paper seeks to establish the relative importance of process and outcome values through the use of conjoint analysis, in this paper termed discrete ranking modeling. The evaluation technique was originally developed by mathematical psychologists, but also has its origin in market research, where it has been employed by companies to establish what factors influence the demand for different commodities, thereby identifying the relative importance of the attributes associated with the product. Due to an increased focus on consumer sovereignty within health care, this technique has recently been applied to the field of health economics (3;18;19;21).

The strength of discrete ranking modeling is that the focus is on the utility associated with the attributes that constitute a commodity, rather than merely the utility of the commodity as a whole. This is a major strength when a healthcare service such as a cancer screening program can be offered in a large variety of setups with differing attribute values. The random utility model describes person i's utility from choice j out of J choices as:

$$U_{ij} = V(s_i, x_{ij}) + \epsilon_{ij} = V_{ij} + \epsilon_{ij}$$

where V_{ij} is the deterministic and observable component of utility associated with participation in a screening program described by x_{ij}, and ϵ_{ij} is the stochastic and unobservable component to the analyst.

In the usual discrete choice case, we want to determine the probability that U_{ij} exceeds U_{ik} $(k \neq j)$. With ordinal ranking information, we want to determine the probability of a particular rank order, such as $\Pr[U_{i1} > U_{i2} > \ldots > U_{iJ}]$, to exploit the additional information in rank, rather than only choice data, and hence obtain precise estimates.

Pragmatic considerations limit the number of probability distribution functions we can choose from for the stochastic component, and we have little choice but to assume ϵ_{ij} to be independently and identically distributed, each with an extreme value distribution. This

https://doi.org/10.1017/S0266462300102089 Published online by Cambridge University Press

leads to the ordered logit model. It has the same limitations, known as the independence of irrelevant alternatives, in the regular logit models but allows for a relatively simple and computationally robust likelihood function for estimation purposes, as derived by Beggs et al. (2, equation 12). We use the Newton-Raphson algorithm in LIMDEP 7.0 to estimate the parameters of the ordered logit model. The model allows for different J_i (number of choices) in the sample.

The deterministic part of utility V_{ij} is modeled as a function of characteristics of the decision maker (s_i) and of a vector of attributes (x_{ij}) for each choice. Socioeconomic characteristics (s_i) must be modeled as interaction terms with choice attributes (x_{ij}), including choice-specific constants.

In this analysis the reference option is the no-screening scenario, which means that per definition the utility associated with this alternative is zero. The utility associated with participating in a screening program is a function of the risk reduction obtained by the program plus a series of other attributes affecting the utility derived from participating. An additive utility function can be presented in the following manner:

$$dU = \frac{\partial U}{\partial x_1} * dx_1 + \frac{\partial U}{\partial x_2} * dx_2 \cdots + \frac{\partial U}{\partial x_n} * dx_n$$

where dU denotes the change in utility relative to the null option of no screening, $dx_1, dx_2 \ldots dx_n$ denotes the changes in attributes x_1 to x_n relative to no screening and $\partial U/\partial x_1$ to $\partial U/\partial x_n$ denotes the estimated coefficients for variables x_1 to x_n. If one chooses to include a cost variable in the utility function, one has the option of deriving WTP estimates by dividing the estimated dU by the cost coefficient $\delta U/\delta COST$.

The Survey

A random sample of the Danish female population (all 50 years of age) was drawn from the national registry, and subsequently invited by letter and contacted by phone. For those women who agreed to participate, a meeting was arranged for a personal interview with a professional interviewer. A total of 255 women were invited to an interview; of these 207 were successfully interviewed, 23 did not wish to participate, 13 individuals were incapable of responding, and 12 could not be contacted. A participation rate of 81.2% was achieved.

The objective of the survey was to identify the respondent's preferences for screening program attributes. Each respondent was given a description of three alternative screening setups and the consequences of not entering a screening program. Prior to presentation of the card, respondents were given descriptions of the screening test (a visit to the mammography unit) and the diagnostic test in the form of a clinical mammography.

Table 1 illustrates one of a large series of cards, of which each respondent was only introduced to one. For the breast cancer screening interview, 24 different cards were designed. The initial risk level in a no-screening scenario was held constant at 340 of 10,000, equivalent to the overall risk of dying from breast cancer between the ages of 50 to 80 years. The individual cards differed only on variable values. Each card was constructed in a similar manner on the following principles. The respondent is presented with four alternatives, of which the first option is no screening. The second option (alternative A) is the least intensive but also the least effective program. The third option (alternative B) is the program that uses improved technology; hence, effectiveness increases without increasing the number of tests or increasing the risk of a false-positive diagnosis. Instead, the participant is required to pay out of pocket. Finally, the fourth option (alternative C) can attain a similar effectiveness as alternative B but without an out-of-pocket expense. Here the "cost" is an increase in the number of screening tests or an increase in the risk of a false-positive

https://doi.org/10.1017/S0266462300102089 Published online by Cambridge University Press

Table 1. An Example of a Card Presented to the Interviewee

	No participation	Program A	Program B	Program C
Number of mammographies performed over the next 25 years	0	17	17	25
Your risk of dying of breast cancer over the next 30 years	340 of 10,000 (3.40%)	220 of 10,000 (2.20%)	210 of 10,000 (2.10%)	210 of 10,000 (2.10%)
Your risk of being called in for an unnecessary clinical mammography	0	3,500 of 10,000 (35%)	3,500 of 10,000 (35%)	3,500 of 10,000 (35%)
Your out-of-pocket expense	—	DKK 0	DKK 2,000 per test In total: DKK 34,000 over 25 years	DKK 0

diagnosis. In order to present the respondents with plausible options, it was not possible to construct the cards on the basis of a full factorial or a systematic fractional factorial design.

Realistic parameter values were chosen with respect to number of screening tests performed over lifetime, risk of a false-positive diagnosis over lifetime, and risk reduction over lifetime. Realistic risk reductions were estimated using the Day and Walter model (5) and knowledge of lead time and sensitivity of the screening tests (22). Mortality reductions were based on experiences from the Swedish two-county randomized trial (12), whereas false-positive rates were based on specificity levels as observed in a Danish setting. For the cost attribute, realistic as well as extreme values were chosen such that *maximum* WTP estimates could be derived.

Individuals were asked to rank the four options presented to them, and subsequently asked to qualify their choices by checking off a list of possible motivations. Finally, respondents were asked questions regarding their income and education.

RESULTS

The Discrete Ranking Analysis

With four choice attributes as well as a number of personal characteristics, one faces a large number of parameters, including all possible interaction combinations. We therefore chose selectively to include characteristics through interaction with attributes and choice specific constants that seemed plausible or theoretically justifiable. For a more thorough discussion of included candidate variables and corresponding hypotheses, the reader is referred to a different paper by the author (11).

Backward (one-) stepwise regression was used to limit the number of parameters. Table 2 illustrates the results of the restricted model for breast cancer, including only variables with significant coefficients. The result is an explanatory model that produces a good fit to the preference data for breast cancer screening. Significant parameters in the models are COST/ln(Y), PAY, and PAY*Y, which demonstrate that an out-of-pocket expense produces two forms of disutility: one associated with the degree of out-of-pocket expense, and one associated with the fact that an out-of-pocket expense is required at all. That COST/ln(Y) is included in the restricted model signifies that income has a significant

Table 2. Results of the Discrete Ranking Model

Variable	Variable description	Coefficients, restricted model	*p* Value
RED	Risk reduction over lifetime (out of 10,000)	0.01642	.0000
REDNEDU	Risk reduction over lifetime*nedu; nedu = 1 if no professional training	0.00692	.0000
FP*FP	Risk of false-positive diagnosis over lifetime (out of 100)	−0.000297	.0162
COST/ln(Y)	Out-of-pocket expense over lifetime/ ln(monthly net of tax household income)	−0.0001989	.0000
PAY	Pay = 1 if out-of-pocket expense > 0	−1.1386	.0000
PAY*Y	Pay*monthly net of tax household income	0.000016183	.0184

impact on the disutility associated with an out-of-pocket expense. The impact of a cost element is negative, but to a lesser degree when household income net of tax is high. That PAY has a negative parameter value, while PAY*Y demonstrates a positive impact, illustrates that an out-of-pocket payment *per se* decreases utility, but to a lesser extent for higher income groups.

The parameter value of RED is positive and highly significant. The utility derived from a risk reduction is influenced by level of education. No professional training entails an increase in the utility associated with a risk reduction (RED*NEDU), implying that uneducated individuals may be more susceptible to judgment biases. Interpretation of risks are subject to various forms of biases, such as availability bias, whereby perception of probabilities are influenced by the degree of media coverage and probability assessment bias, whereby people tend to overestimate small probabilities. A significant negative coefficient for the FP*FP parameter was disclosed, which implies that risk of a false-positive diagnosis has an influence on choice of breast cancer screening program and that the effect is characterized by increasing marginal disutility as programs are intensified. Number of screening tests over lifetime had no impact on preferences, signifying that the inconvenience of frequent screening is perceived as minor.

The utility model presented in Table 1 can also be expressed in the form of an equation:

$$\begin{aligned} dU = {} & 0.01642 * dRED + 0.00692 * d(RED * NEDU) - 0.000297 * d(FP * FP) \\ & - 0.0001989 * d(COST/\ln(Y)) - 1.1386 * dPAY + 0.000016183 * d(PAY * Y) \end{aligned} \tag{1}$$

where dRED is the risk reduction obtained by the program (out of 10,000), FP is the probability of a false-positive diagnosis over lifetime (in percentage), COST is the total out-of-pocket expense over lifetime, Y is the monthly household income net of tax, and PAY is a dummy variable that equals one if an out-of-pocket expense is required. The model is simplified if no out-of-pocket expense is prevalent:

$$dU = 0.01642 * dRED + 0.00692 * d(RED * NEDU) - 0.000297 * d(FP * FP) \tag{2}$$

In the study, individuals were asked to explain their choices by checking off a list of motivations. The main motivation for participation was, not surprisingly, to reduce the risk of dying from cancer. Other frequently observed motivations were to eliminate potential feelings of regret, gain information, and a tendency to accept what is offered or, in other words, do what is recommended.

https://doi.org/10.1017/S0266462300102089 Published online by Cambridge University Press

Table 3. Calculated Utilities Derived from Participating in a Breast Cancer Screening Program

Program age group; screening interval	Number of tests performed over lifetime	Risk of false-positive diagnosis (out of 100)	Absolute risk reduction (out of 10,000)	Estimated utility (dU) (professional training/ no training)	Estimated participation rate (professional training/ no training)
50–64;3	5	9.6	56	0.892/1.279	0.709/0.782
50–64;2	7	13.2	62	0.966/1.394 *d*[a]	0.724/0.801
50–64;$1^1/_2$	10	18.3	72	1.083/1.570 *d*[a]	0.747/0.829
50–64;1	15	26.2	79	1.093/1.638 *d*[a]	0.749/0.837
450–69;3	7	13.2	79	1.245/1.790	0.777/0.857
50–69;2	10	18.3	91	1.395/2.022 *d*[a]	0.801/0.883
50–69;$1^1/_2$	14	24.6	100	1.462/2.152 *d*[a]	0.812/0.896
50–69;1	20	33.2	104	1.380/2.098 *d*[a]	0.799/0.891
50–74;3	9	16.6	93	1.445/2.087	0.809/0.890
50–74;2	13	23.1	112	1.681/2.453	0.843/0.921
50–74;$1^1/_2$	17	29.1	117	1.670/2.477 *d*[a]	0.842/0.922
50–74;1	25	39.7	125	1.584/2.447 *d*[a]	0.830/0.920

[a] d = dominated alternative.

Calculating the Utility of Screening Programs

The parameter coefficients presented in Table 2, and hence equation 2, can be used to estimate the relative utility of alternative screening setups. Below we have estimated the utility levels associated with different screening setups for breast cancer. In this analysis false-positive rates were calculated assuming a specificity of 98% for breast cancer, a value that is in accordance with Danish experience. Risk reduction estimates were estimated by modeling. The Day and Walter model (5) was used to predict the cancer rate at each screening round for a given screening interval and age group. Input variables into the model are lead time, sensitivity, and Danish incidence data. Lead time (3.5 years) and sensitivity (92.8%) values were based on the experience of the Swedish two-county trial (22). Similarly, excess survival due to early detection of breast cancer was based on the Swedish experience (23), producing an excess survival rate of 0.24 for screen-detected cancers (8). The reader should, however, be reminded that the explicit weighting of attributes provides the opportunity for estimating relative utilities, based on alternative assumptions regarding specificity and risk reduction.

Number of tests performed over lifetime, risk of a false-positive diagnosis over lifetime, and risk reduction over lifetime were calculated and multiplied by the relevant coefficients. It was assumed that neither of the programs involved an out-of-pocket expense.

In Table 3 the utilities are calculated for a series of possible screening setups. Number of tests performed over lifetime can be used as a proxy for the costliness of the program. Screening costs increase almost proportionately to number of screening tests performed, since in the long term most costs are variable. An increase in the number of tests performed will increase costs directly, and also indirectly through an increase in risk of a false-positive diagnosis. Treatment costs are reduced significantly as a result of screening (1). These cost savings rise proportionately with risk reduction. Hence, inclusion of this derived cost effect does not influence the conclusion that dominated programs can be identified merely by focusing on utility estimates and number of screening tests over lifetime.

A “d” symbolizes a dominated alternative, i.e., a program setup for which another program can be found that produces more utility at a similar or lower level of cost or, alternatively, the same amount of utility at less cost. Figure 1 illustrates the results listed

https://doi.org/10.1017/S0266462300102089 Published online by Cambridge University Press

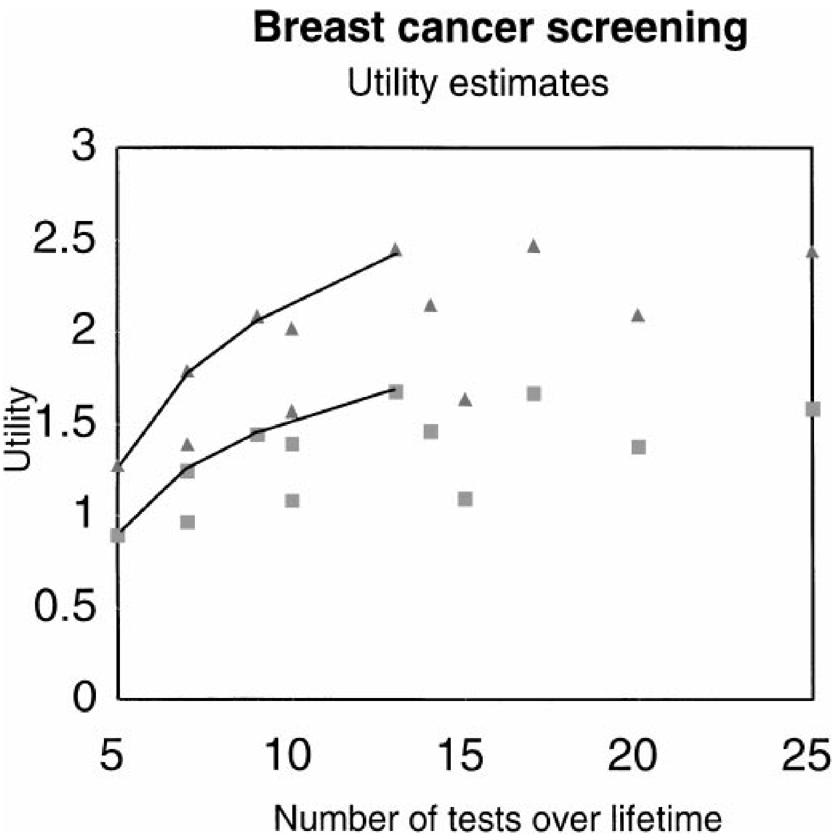

Figure 1. Hypothetical programs plotted according to estimated utility and number of tests performed over a lifetime. Lines indicate the efficiency curve on which lie the dominating programs. Squares indicate the utility levels for individuals with professional training. Triangles indicate utility levels for those without professional training.

in Table 3 in an alternative manner. Here, number of tests over lifetime is indicated on the x-axis, utility is indicated on the y-axis, and an efficiency curve is drawn by connecting dominating programs. The curve is an approximation since it is based on the estimates from 12 hypothetical programs only. In reality, an infinite number of alternative programs exist, and extending the analysis to encompass a large range of programs would produce a more precise curve. Nevertheless, Figure 1 illustrates the overall picture well: marginal utility decreases as programs are intensified. For programs that exceed 13 tests over lifetime, marginal utility is close to zero.

Participation Rates

The discrete choice model is derived under a specific assumption regarding the distribution of the unobserved portion of utility (random utility). It is assumed that the random utilities are independent and identically distributed according to a type I extreme value distribution in standard form (sometimes also called a Weibull or Gumbell distribution). On the basis of the characteristics of this distribution, a formula for the probability of participation can be calculated as follows: $p = \exp(dU)/(1 + \exp(dU))$, where dU is the utility estimated when applying equation 2 (4;24). The model implies that if utility of a program is zero, participation will be 50%; $\exp(0)/(1 + \exp(0)) = 0.5$. In the last column of Table 3, participation rates are calculated for various breast cancer screening programs, ranging from 70.9% to

84.3% among individuals with professional training. Among the uneducated, estimated participation rates are higher: 78.2%–92.2%. For a program screening the 50–69-year-olds every second year, the estimated participation rates are 80.1% and 88.3%, respectively. This program represents the program that has been initiated in the County of Funen, Denmark, where a participation rate of approximately 88% was observed. Model estimates of participation rates cannot be directly compared to actual participation rates due to information bias, but the results lie within a reasonable range of participation rates observed for mammography screening in Denmark, indicating that the preferences elicited are in line with observed behavior.

DISCUSSION

Table 3 and Figure 1 demonstrate the importance of including intangibles in economic evaluations. Although the two latter programs in Table 3 (screening persons aged 50–74 years or every year to $1^1/_2$ years) are more effective, as indicated by the risk reduction, the overall utility of these programs is similar to the utility levels derived from participation in less intensive programs (e.g., screening 50–74-year-olds every second year). This result demonstrates the importance of analyses such as the one presented in this paper, if inoptimal resource allocations are to be avoided. Discrete ranking modeling can be used as a tool in identifying inferior program options. Based on a specificity of 98% and risk reductions similar to those observed in the Swedish two-county trial, the model results infer that only 4 of 12 hypothetical programs are viable choices. These are the four programs on the efficiency curve, i.e., the programs found not to be dominated by other options in Table 3. Which of these four programs, if any, should be implemented may be determined through cost-benefit analyses. WTP per program can be determined by exploring the information that lies in the estimated utility function.

Based on the parameter values presented in Table 2, WTP per statistical life saved can be estimated by focusing on the marginal rate of substitution between RED and COST/ln(Y). Assuming a median net of tax household income of 25,000 Danish kroners (DKK), the WTP for an increase in risk reduction of 1/10,000 can be calculated by the ratio of the respective parameter values: 0.01642/(0.0001989*(1/ln(25,000))) = DKK 836. With linear extrapolation (assuming constant marginal rate of substitution), this amounts to an implied WTP of DKK 8.4 million for a statistical life. Among those with no professional training, the WTP for a statistical life amounts to DKK 11.9 million. In the literature, estimates of WTP per statistical life typically lie in the region of DKK 7–30 million (14), implying that the estimates derived from the model presented in this analysis correspond well with estimates presented by other researchers.

Recently, a detailed cost study on breast cancer screening of 50–69-year-olds biennially was performed based on Danish data (1). The study included direct and indirect costs of screening, hence including the costs of setting up and running a mammography unit as well as travel and time costs incurred by the women participating. Detailed estimations on the implications of screening activity relative to no screening on cost of diagnosing and treating breast cancer were incorporated in the analysis. The results was a net cost of DKK 208 per screening test performed (including time and travel costs) and a cost of DKK 128 per test if women's time and travel costs were excluded. The latter estimate is relevant in the context of this cost-benefit analysis, since time and travel costs are already included in the utility function. The cost information provides an opportunity for looking at the marginal cost per statistical life saved as one moves from one efficient program to another. Assuming that the net cost per screening test remains approximately constant for alternative program setups, the cost per statistical life saved can be derived by: (number of test*cost per test)/risk reduction. For the efficient programs, the direct cost per statistical

Table 4. Willingness to Pay per Program

Program	Number of tests over lifetime	Estimated utility, dU (professional training/no professional training)	WTP (DKK) (professional training/no professional training)	Marginal WTP per extra test performed (DKK) (professional training/ no professional training)
50–64;3	5	0.892/1.279	45,417/65,121	
50–69;3	7	1.245/1.790	63,390/91,139	8,987/13,009
50–74;3	9	1.445/2.087	73,573/106,261	5,091/7,561
50–74;2	13	1.681/2.477	85,589/126,118	3,004/4,964

life saved lies in the range of DKK 200,000 and DKK 480,000. The results clearly indicate that WTP per statistical life exceeds the direct cost per statistical life by a significant amount.

Indirect costs such as unrelated healthcare costs and costs of general consumption generated by prolonging of life were not included in the analysis. Production gains were likewise ignored. Under the conservative assumption that production gains are likely to be negligible for the age group in question, the net effect of including indirect effects would be a net present value in the range of DKK 80,000 per life-year gained (9). Assuming that 20 to 25 life-years are gained for every life saved, indirect costs would amount to a maximum of DKK 2 million per statistical life saved. Clearly the inclusion of indirect costs would not alter the conclusion that WTP exceeds costs.

That WTP per statistical life exceeds cost per statistical life does not conclusively infer that the programs in question are net beneficial, since only a subset of intangibles are valued. In order to estimate WTP for the programs as a whole, the program-specific utility (dU) is divided by $\delta U/\delta(COST/\ln(Y))$, assuming constant marginal utility of income.

There is, however, a problem involved in determining this value, since it will represent the mean maximum WTP estimate across all respondents. Participation in a screening program must be seen as a private good rather than a public good. This allows individuals to choose not to participate in the program, thereby avoiding negative utility. The WTP estimate that is elicited from a utility function based on all individuals' preferences will contain these negative values and result in an underestimated WTP estimate, albeit the ranking of programs remains unaffected. If this biased estimate turns out to be higher than the costs incurred so that benefits exceed costs, it may be ample to conclude that there is a net benefit of the program. If, however, one wishes to establish the exact value of the net benefit or if the biased WTP estimate is not great enough to warrant this conclusion, an adjustment of the WTP estimate must be performed.

In Table 4, total WTP and marginal WTP estimates for each extra test performed are listed for the four efficient breast cancer programs. When comparing cost per test (DKK 128) and marginal WTP per test, the programs clearly produce marginal net benefits greater than zero. This conclusion also holds when indirect costs of approximately DKK 2 million per statistical life saved are included in the cost per test (equivalent to an extra cost of DKK 2,240, DKK 2,300, DKK 1,400, and DKK 950 per test for the four efficient programs). If the decision rule is to choose the program, among the four mutually exclusive programs, that produces maximum net benefits, the optimal program involves screening 50–74-year-olds biennially. However, this decision rule is unlikely to contribute to optimizing resource allocation if we are operating within a fixed healthcare budget in which prioritization should be made among alternative healthcare services. In this case a comparison of benefit-cost ratios across independent programs should determine which of the four programs should be implemented.

Apart from the considerations with respect to decision rule, there are other problems that should be considered before making recommendations based on the above results. Respondents' motivations for participating have shown that a major benefit of cancer screening programs is the elimination of regret plus a sense of being part of a group, i.e., doing what is recommended. However, these elements of benefit are not true benefits of the program, since the introduction of the program produces potential disutilities (e.g., feelings of regret) that are merely eliminated through participation. This causes estimated utilities, and hence WTP estimates, to be overestimated, but unfortunately it is not possible to determine the extent to which these motivations influence preferences.

It is also important to emphasize that the utilities estimated for the hypothetical programs are not all statistically different. Confidence intervals were derived for the utility estimates calculated for the 12 hypothetical programs listed in Table 2. At the 95% level the four programs in which 50–64-year-olds are invited all produce significantly less utility than those programs that invite 50–74-year-olds. Within these two groups of programs utility levels are not significantly different, nor are the utilities for the programs involving 50–69-year-olds significantly different from utility levels produced by any other of the 12 listed programs. This weakness does not, however, markedly alter the conclusions of this analysis. Assuming that those programs in which utilities are not significantly different are in fact identical, a new list of efficient programs can be identified. The new list includes the same programs as those indicated in Table 2, with the exception of the program that invites 50–74-year-olds biennially. This program is now dominated by the program that invites 50–74-year-olds every 3 years.

Finally, one should be wary in applying the results of this analysis for policy purposes, since the analysis is highly exploratory. The practical and theoretical applicability of the methods applied in this paper have not been fully investigated, as is emphasized by Ryan (20). In the present analysis, choices were framed to enforce realism, eliminating the possibility of using a full factorial design or a systematic fractional factorial design. Hence, statistical design has been compromised somewhat for realism. Whether this produces less valid results compared to strong statistical designs accompanied by lack of realism is yet unanswered. A more general problem, which at present remains unsolved, is whether the utility function extracted by conjoint analyses can be interpreted as having cardinal properties. If this is not the case, utility levels can only indicate the ranking of alternatives, and discrete choice analysis will be limited to identifying a list of efficient programs without subsequent identification of the program that is most net beneficial.

The focus of this paper was to demonstrate the potential usefulness of conjoint analysis (discrete ranking modeling). In the context of cancer screening programs for which an abundance of alternative screening setups (target group, screening interval, choice of screening test, etc.) exist, it is important to identify the relative importance of the program attributes to identify and exclude programs that are more resource-consuming and potentially provide less utility to the consumers.

POLICY IMPLICATIONS

A cost-benefit analysis of mammography screening was performed, including effects such as risk reduction of dying from breast cancer, risk of a false-positive diagnosis, and number of tests performed. The analysis was performed on 12 hypothetical programs, and eight of these were identified as being dominated by the other four options. Among the four efficient options, a program that screens 50–74-year-olds every second year produces the highest level of utility. If screening frequency is increased beyond 13 tests over lifetime, marginal utility levels off as increments in disutilities associated with the process of screening overshadow the value of the extra lives being saved. This analysis suggests that discrete ranking modeling can contribute to optimal resource allocation, primarily by identifying programs that are inferior.

REFERENCES

1. Bech M, Gyrd-Hansen D. Cost implications of routine mammography screening of women 50-69 years of age in the county of Funen, Denmark. *Health Policy*. In press.
2. Beggs S, Cardell S, Hausman J. Assessing the potential demand for electric cars. *J Econ*. 1981;16:1-19.
3. Chakraborty G, Gaeth GJ, Cunningham M. Understanding consumers' preference for dental service. *J Health Care Mark.* 1993;13:48-58.
4. Cramer JS. The logit model for economists. London: Edward Arnold; 1991.
5. Day NE, Walter SD. Simplified models for screening for chronic disease: Estimation procedures for mass screening. *Biometrics*. 1984;40:1-14.
6. De Koning HJ, van Ineveld BM, van Oortmarrsen 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;49:531-537.
7. Eddy DM. Screening for cervical cancer. *Ann Intern Med.* 1990;113:373-384.
8. Gyrd-Hansen D. *Breast cancer screening: A sensitivity analysis. CHS working paper*. Denmark: Odense University; 1997;5.
9. Gyrd-Hansen D. Consequences of including unrelated future costs and earnings in economic evaluations. In: *Modelling the cost effectiveness of cancer screening in Denmark*. PhD thesis from the Faculty of Social Science, University of Southern Denmark, 1998.
10. Gyrd-Hansen D, Hølund B, Andersen P. A cost-effectiveness analysis of cervical cancer screening: Health policy implications. *Health Policy.* 1995;34:35-51.
11. Gyrd-Hansen D, Søgaard J. Analysing public preferences for cancer screening programmes. Presented at the iHEA Conference, Rotterdam, June 1999.
12. Gyrd-Hansen D, Søgaard J, Kronborg O. Analysis of screening data: Colorectal cancer. *Int J Epidemiol*. 1997;26:1172-1181.
13. Gyrd-Hansen D, Søgaard J, Kronborg O. Colorectal cancer screening: Efficiency and cost effectiveness. *Health Econ*. 1998;7:9-20.
14. Kidholm K. Estimation of willingness to pay for prevention of traffic accidents [in Danish]. PhD dissertation. Odense University, Denmark; 1995.
15. Koopmanschap MA, Lubbe KTN, Oortmarssen GJ, et al. Economic aspects of cervical cancer screening. *Soc Sci Med.* 1990;30:1081-1087.
16. Koopmanschap MA, van Oortmarssen GJ, van Aagt HMA, et al. Cervical cancer screening: Attendance and cost-effectiveness. *Int J Cancer*. 1990;45:410-415.
17. Mushlin AI, Fintor L. Is screening for breast cancer cost-effective? *Cancer Suppl*. 1992;69:1957-1962.
18. Propper C. Contingent valuation of time spent on NHS waiting lists. *Econ J.* 1991;100:193-199.
19. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: An application to in vitro fertilisation. *Soc Sci Med*. 1999;48:535-546.
20. Ryan M. A Role for conjoint analysis in technology assessment in health care? *Int J Technol Assess Health Care.* 1999;15:443-457.
21. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Econ.* 1997;6:261-274.
22. Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer by age: New results from the Swedish two-county trial. *Cancer*. 1995;75:2507-2517.
23. Tabar L, Fagerberg C, Duffy SW, et al. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am*. 1992;30:187-210.
24. Train K. *Qualitative choice analysis*. Cambridge, MA: MIT Press; 1986.
25. Van der Maas PJ, de Koning HJ, van Ineveld BM, et al. The cost-effectiveness of breast cancer screening. *Int J Cancer.* 1989;43:1055-1060.
26. Wagner JL, Tunis S, Brown M, Ching A, Almeida R. Cost-effectiveness of colorectal cancer screening in average-risk adults. In: Young GP, Rozen P, and Levin B, eds. *Prevention and early detection of colorectal cancer*. London: W.B. Saunders; 1996.