

15. Jones-Lee MW (ed.) *The Value of Life and Safety*. Amsterdam, North-Holland Publishing Company, 1982.
16. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986, 5, 1–30.
17. Mehrez A, Gafni A. Quality adjusted life years, utility theory and healthy-years equivalents. *Med Decis Making* 1989, 9, 142–149.
18. Weinstein MC, Fineberg HV. *Clinical Decision Analysis*. Philadelphia, W.B. Saunders, 1980.
19. Sugden R, Williams A. *The Principles of Practical Cost-benefit Analysis*. Oxford, Oxford University Press, 1978.
20. Llevellyn-Thomas H, Sutherland JH, Tibshirani R, Ciarni A, Till JE, Boyd NF. Describing health states. Methodological issues in obtaining values for health states. *Med Care* 1984, 22, 543–552.
21. Read JL, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes. Comparison of assessment methods. *Med Decis Making* 1984, 4, 315–329.
22. Nord E. The significance of contextual factors in valuing health states. *Health Policy* 1989, 13, 189–198.
23. Loomes G, McKenzie L. The use of QALYs in health care decision making. *Soc Sci Med* 1989, 28, 299–308.

0959-8049(94)E0138-T

# To Screen or Not to Screen. How do we Decide on Which Cancer Screening Activities to Embark Upon?

J.D.F. Habbema and G.J. van Oortmarsen

## THE PUBLIC HEALTH PERSPECTIVE

SCREENING, i.e. early detection before a clinical diagnosis is made on the basis of symptoms, has been proposed for cancers at many sites (see [1] for an overview). The goal of population screening is to reduce the burden of the cancer in terms of morbidity and mortality. In this respect, it is just one of several possible interventions and should be compared with the other alternatives: primary prevention, early case finding, treatment, rehabilitation and terminal care. In judging screening, however, adverse effects should be considered especially carefully, because inevitably a number of apparently healthy persons will be subjected to medical tests, and possibly to treatment, with no real long-term improvement in health or longevity.

The comparison should be based primarily on the favourable and adverse effects and risks of the interventions. In making choices about (re)allocation of funds between types of intervention, their costs should be weighted against their health effects. Cost-effectiveness analysis, cost-utility analysis and cost-benefit analysis are the methodologies which may be applied in investigating the decision problems [2–4]. A consideration that is often overlooked is that of (re)distribution of health over the population. For example, cervical cancer is typically a disease of women in lower socioeconomic strata, while breast cancer is one of the few examples of a disease that occurs more frequently in people in higher socioeconomic strata. In view of the present inequality in the distribution of health demonstrated in many countries, cervical cancer screening should deserve some extra attention, but also taking into account that it has a less favourable cost-effectiveness ratio than breast cancer screening.

## AN ILLUSTRATION OF BREAST CANCER SCREENING

Table 1 shows the predicted favourable and adverse effects of two-yearly breast cancer screening of women aged 50–70 years in the Netherlands [5]. Nationwide screening between 1990 and 2017 is expected to result in a reduction of 6000 breast cancer deaths and a gain of 16 500 years of life per million screening examinations.

Adjustment of the years of life because of changes in quality of life brought about by other effects of screening will have only a small impact: 16 000 quality-adjusted years of life would be gained.

Health effects should be considered in relation to the resources required for establishing those effects. For example, the total cost of 15.8 million screenings for breast cancer in the 27-year period will be 489 million ECU (1 ECU = 2.30 Dfl), i.e. 31 ECU per examination. These costs are partly compensated, since screening will result in a decrease in the costs of management of breast cancer (–249 million ECU), mainly because of the

Table 1. Favourable and adverse effects of breast cancer screening in the Netherlands: implications for differences in quality-adjusted years of life per million screening examinations (no discounting)

	Number	Quality-adjusted years of life
Screening examinations	1 000 000	– 115
Years of life gained	16 500	+ 15 700
Advanced and terminal cancer	1 080	+ 730
Lead time (earlier diagnosis)	17 500	– 500
Breast conserving therapy instead of mastectomy	750	+ 40
Others		+ 145
Total		16 000

Correspondence to J.D.F. Habbema.

The authors are at the Department of Public Health, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

prevention of advanced disease. Women who benefit from screening will, however, have other medical costs during their additional years of life. An important complication is that, in general, costs and adverse effects occur long before favourable effects and savings are seen. Time preference is taken into account by using discount rates, calculating present values for effects, costs and the cost-effectiveness ratio (Table 2).

The time lag between costs and effects may differ considerably between screening programmes, e.g. those for cervical cancer, which has a very long preclinical stage, and those for breast cancer, which has a much shorter preclinical stage. Thus, discounting will influence the cost-effectiveness of cervical cancer screening in a more unfavourable way than for breast cancer screening.

The outcomes of a cost-effectiveness study are highly dependent on the specific assumptions on the various aspects of the problem. Hence, results of a sensitivity analysis, in which assumptions about all aspects are varied in order to investigate their impact on the outcomes, are more informative than just the most plausible outcomes. The results of the sensitivity analysis will indicate how robust the conclusions about the screening programme are.

### REALISTIC SCENARIOS

In making predictions about the potential effects of a screening test, a number of complicating factors should not be overlooked.

First, the coverage of the population has been found to vary considerably between countries, screening tests and risk groups (e.g. age). Coverage also strongly depends on the way in which screening is organised: highest coverage is achieved with centralised programmes involving an invitation scheme based on population registries (Nordic countries, the Netherlands). In the Netherlands, a considerable drop in attendance occurred when cervical cancer screening ceased to be organised centrally.

Secondly, thorough quality control of all stages of the screening process, i.e. of the test procedure (mammograms, Pap smear), of the interpretation of the results, and of further assessment of women with positive results at screening, is mandatory in order to prevent high proportions of false-positive and false-negative results and unnecessary, invasive diagnostic tests. Remarkable differences in referral rates after a positive mammography are found between the U.S.A. and the U.K. on the one hand, and Sweden, Finland and the Netherlands on the other. Similar differences might occur within one country in the absence of good quality control.

Table 2. Costs and effects of breast cancer screening, the Netherlands (see Table 1): comparison of outcomes with and without discounting

	Discount rate	
	0%	5%
Years of life gained (crude)	260 000	61 000
Quality-adjusted years of life gained	252 000	57 500
I. Screening costs*	489	260
II. Costs of breast cancer management*	- 249	- 58
III. Other medical costs*	+ 920	+ 181
Costs (I + II) per year of life <sup>†</sup>	920	3300
Costs (I + II) per quality-adjusted year of life gained <sup>†</sup>	960	3500
Costs (I + II + III) per year of life <sup>†</sup>	4 440	6 300

\*Million ECU. <sup>†</sup>ECU.

Thirdly, recommending a screening test for specific age groups (or other risk categories) will always raise the demand from other groups. Experience with cervical cancer screening shows that centrally organised screening with a test that does not require specialised equipment results in large-scale opportunistic screening outside the programme. In the case of cervical cancer, such opportunistic screening is inefficient, since it involves predominantly women who are too young and intervals between tests that are too short. Coverage, quality control and opportunistic screening deserve special attention when results and conclusions from an experimental scientific project are transferred to a routine population screening service.

### VERIFYING THE EFFECTIVENESS OF SCREENING

A screening test must undergo rigorous testing and evaluation before it can be judged to be eligible for mass implementation. Basic criteria for judging a screening test were specified by Wilson and Jungner in 1966 [6]. Only a small number of tests fulfil the minimal requirements for sensitivity and specificity, and elicit general agreement on the assessment of abnormalities and on the type of treatment to be used for diagnosed (pre-)invasive cancer. These requirements may be investigated on the basis of small-scale (clinical) experiments.

Three well-known biases (lead time, length and self-selection) preclude final testing of a screening test in a non-randomised design. The theoretical problems of the case-control design have been confirmed empirically [7]. The effectiveness of a screening test in reducing mortality and morbidity should be demonstrated in one or more population-based randomised controlled trials. The absence of evidence based on such trials for the effectiveness of cervical cancer screening by Pap smears has always troubled decision makers.

Randomised controlled trials are, however, very costly and they cannot solve all questions. A single trial generally allows only direct conclusions about the effectiveness of the specific policy used and, depending on the size of the trial, about differences in effectiveness across age groups. Another inevitable problem is that conclusions can be drawn from a randomised controlled trial only after some 10 years, by which time the screening test or the treatment may have changed considerably.

### CHOOSING A SCREENING POLICY

Once effectiveness has been demonstrated, further evidence of the effectiveness in different age groups or with different screening intervals can be obtained from case-control studies. For a comprehensive evaluation of screening policies (age groups, intervals, risk groups) with respect to health effects, risks and costs, simulation models are powerful tools, as all important factors can be taken into account. The predictions shown in Tables 1 and 2 were obtained using a MISCAN model for breast cancer screening, which has been validated against data from randomised trials in Sweden and from the Dutch pilot studies [8, 9]. Another MISCAN model was used to evaluate cervical cancer screening in the Netherlands: the best cost-effectiveness ratio (just over 10 000 ECU per year of life gained, 5% discounted) was found using a programme that starts around age 30 and continues up to age 70, with an interval between successive screenings of about 6 years. Policies with much shorter intervals show a much less favourable balance between benefits, risks and costs. When screening starts at a very young age, risks and costs increase considerably with only a small improvement in beneficial effects [10, 11, 12].

The decision-orientated approach is still useful after the

screening test has been implemented in a population-based early detection programme. The results of the programme (participation, detection rate, stage distribution, interval cancers, costs) should be carefully registered and monitored and can be compared with expectations made at the outset. Major discrepancies or new developments in diagnosis and therapy can be investigated by computer simulation, which will lead to recommendations about adaptations of the screening policy, conducting new pilot studies or sometimes even randomised controlled trials.

1. Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PC (eds). *Cancer Screening*. Cambridge, Cambridge University Press, 1991.
2. Drummond MF, Stoddart GL, Torrance GW, *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, Oxford University Press, 1987.
3. Warner KE, Luce B. *Cost-benefit and Cost-effectiveness Analysis in Health Care*. Ann Arbor, Health Administration Press, 1982.
4. Habbema JDF, van Ineveld BM, De Koning HJ. A cost-effectiveness approach to breast cancer screening. In Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PG, eds. *Cancer Screening*. Cambridge, Cambridge University Press, 1991, 68-77.
5. 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, 41, 531-537.
6. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Public Health Papers no. 34. Geneva, WHO, 1968.
7. Janzon L, Andersson I. The Malmö mammographic screening trial. In Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PC, eds. *Cancer Screening*. Cambridge, Cambridge University Press, 1991, 37-44.
8. Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comp Meth Progr Biomed* 1984, 20, 79-93.
9. van Oortmarssen GJ, Habbema JDF, van der Maas PJ, et al. A model for breast cancer screening. *Cancer* 1990, 66, 1601-1612.
10. Koopmanschap MA, Lubbe JTN, van Oortmarssen GJ, van Agt HME, van Ballegooijen M, Habbema JDF. Economic aspects of cervical cancer screening. *Soc Sci Med* 1990, 30, 1081-1087.
11. van Ballegooijen M, Koopmanschap MA, van Oortmarssen GJ, Habbema JDF, Lubbe JTN, van Agt HMA. Diagnostic and treatment procedures induced by cervical cancer screening. *Eur J Cancer* 1990, 26, 941-945.
12. van Ballegooyen M, Habbema JDF, van Oortmarssen GJ, Koopmanschap MA, Lubbe JTHN, van Agt HMA. Preventive Papsmears: striking the balance between costs, risks, and benefits. *Br J Cancer* 1992, 65, 930-933.