Breast Cancer Res Treat (2007) 102:211–218
DOI 10.1007/s10549-006-9319-4

EPIDEMIOLOGY

Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands

J. H. Groenewoud · J. D. M. Otten · J. Fracheboud · G. Draisma · B. M. van Ineveld · R. Holland · A. L. M. Verbeek · H. J. de Koning · for the National Evaluation Team for Breast Cancer Screening (NETB)

Received: 30 January 2006 / Accepted: 22 June 2006 / Published online: 27 September 2006
© Springer Science+Business Media B.V. 2006

Abstract In mammography screening with double reading, different strategies can be used when the readers give discordant recommendations for referral. We investigated whether the results of the Dutch breast cancer screening programme can be optimised by replacing the standard referral strategy by consensus. Twenty-six screening radiologists independently and blinded to outcome read a test set consisting of previous screening mammograms of 250 cases (screen-detected and interval cancers) and 250 controls. Their referral recommendations were paired and, in case of discrepancy, re-read according to three referral strategies: (1) decision by one of the readers; (2) arbitration by a third reader; (3) referral if both readers agree (consensus). Data allowed studying other referral strategies, including referral if any reader suggests, as well. Double reading with referral if any reader suggests resulted in a 1.03 times higher sensitivity (76.6%) and a 1.31 times higher referral rate (1.26%) than double reading with consensus. To estimate the cost-effectiveness, the outcomes were used in a microsimulation model. Even if double reading with referral if any reader suggests results in four times as high referral rates and an accompanying increase of biopsies or other invasive procedures, the cost-effectiveness of €4,190 per life-year gained may well be in the range of acceptable cost-effectiveness for Dutch health care programmes.

Keywords Breast cancer · Cost-effectiveness · Double reading · Referral strategy · Screening

The NETB consists of H.J. de Koning (chair), J. Fracheboud, S.J. Otto, J.H. Groenewoud, G. Draisma, A.E. de Bruijn, Department of Public Health, Erasmus MC, University Medical Center Rotterdam; B.M. van Ineveld, Institute for Medical Technology Assessment, Erasmus MC, University Medical Center Rotterdam; F. van der Horst, National Expert and Training Centre for Breast Cancer Screening, Nijmegen; M.J.M. Broeders, A.L.M. Verbeek, Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Centre, Nijmegen.

J. H. Groenewoud · J. Fracheboud · G. Draisma
H. J. de Koning (✉)
Department of Public Health, Erasmus MC, University Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands
e-mail: h.dekoning@erasmusmc.nl

J. D. M. Otten · A. L. M. Verbeek
Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

B. M. van Ineveld
Institute for Medical Technology Assessment, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

R. Holland · A. L. M. Verbeek
National Expert and Training Centre for Breast Cancer Screening, Nijmegen, The Netherlands

Introduction

Mammography is widely used for breast cancer screening in postmenopausal women. The Dutch nation-wide breast cancer screening programme, targeted at women aged 50–74, provides biennial screening mammography. More than 4.5 million screens have been evaluated since the introduction of the programme in 1990 [1]. Mammographic findings leading to referral of the women were identified in 9.8 per 1,000

women screened; breast cancer was confirmed in 4.7 of every 1,000 women screened.

In the Netherlands, all screening mammograms are double read. If the two radiologists disagree about referral, the standard is that radiologists discuss the case together to reach consensus about referral. In some of the Dutch screening regions, such cases are presented to a third radiologist or a panel of radiologists to reach a decision about referral [2]. The double reading protocol demands independent, blinded reading, that is, the second radiologist should read the films without knowledge of the first reader's opinion. Two Dutch court cases in 1997 involving interval cancers addressed this reading protocol indirectly [3, 4].

Independent, blinded reading and options for referral were the subject of our study, which was occasioned by the less favourable stage distribution of screen-detected cancers in subsequent screening rounds than was expected [5]. This raised the question whether the performance of the Dutch screening programme could be optimised. In this paper, we will study the different referral strategies that can be applied in case of blinded double reading, including consensus and non-consensus double reading. In an experimental setting, 26 screening radiologists read a test set of screening mammograms, consisting of randomly distributed normal and prediagnostic mammograms [6]. We will compare the sensitivity and specificity of the different reading strategies and, by means of a microsimulation model [7], investigate which strategy would be most cost-effective if applied nation-wide.

Materials and methods

Reading and referral strategies

Blinded double reading can lead to the following outcomes: concordant *negative* referral recommendation (both radiologists recommend that the woman should not be referred); concordant *positive* referral recommendation (both radiologists recommend that the woman should be referred); discrepant reading (initial disagreement about referral). We studied three double reading referral strategies in case of discrepant reading:

1. decision by one of the readers (in the knowledge that the other reader gave a different referral advice);
2. decision by consensus. The radiologists discuss the case to reach consensus about referral;
3. decision by arbitration. A third radiologist decides whether the woman should be referred.

Without any further experiments, we were able to study the following additional reading strategies:

– single reading;
– double reading with referral of the woman if any reader suggests (i.e. referral in all instances of concordant *positive* referral recommendation or discrepant reading);
– double reading with referral only if both radiologists directly agreed about referral of the woman (i.e. referral in all instances of concordant *positive* referral recommendation).

Test set

We used a test set consisting of 500 series of screening mammograms [6]. The mammograms were obtained from five of nine Dutch screening regions across the country. The regional administrations were asked to provide 60 recent, non-selected consecutive series of mammograms of controls, 30 of women with screen-detected cancer and 30 of women with interval cancer. *Controls* were defined as women in whom no breast cancer was diagnosed in none of three most recent screening rounds. Women with screen-detected cancers—cancers diagnosed as the result of screening—must have had at least two negative screen examinations before the cancer was detected. Women with interval cancers—breast cancers diagnosed in the interval since the last screen examination—must have had at least two negative screen examinations prior to the diagnosis of breast cancer. Of all 619 series that were received, 500 of good technical quality were selected and divided randomly into 10 subsets of 50 mammograms. The final test set consisted of:

– Two hundred and fifty series mammograms of controls. The series included the mammograms that were made in the last but one screening round (defined as the *index mammograms*) and the mammograms made in the second last but one screening round (defined as the *prior mammograms*).
– Two hundred and fifty series mammograms of cases, including 125 women with screen-detected breast cancers and 125 women with interval cancers. The series of women with screen-detected cancer involved the last two screening rounds prior to the round in which breast cancer was detected. The series of women with interval cancers involved the last two screening rounds before the interval

cancer was diagnosed. Here, the *index mammograms* were the mammograms of the screen examination preceding the examination that led to the diagnosis of breast cancer (screen-detected cancer) or preceding the diagnosis of interval cancer, respectively. The *prior mammograms* were the mammograms from the screening round that preceded the index round.

Diagnostic mammograms and histo-pathological reports of the cases were obtained as well. For screen-detected cancers, the diagnostic mammograms were the screening mammograms that resulted in detection of breast cancer; for interval cancers these were the clinical mammograms.

Four screen-detected cancers and one interval cancer were excluded from the analysis because the diagnostic mammograms (four cases) or the histo-pathological report (one case) were missing.

Setting

In total, 26 Dutch radiologists who were specialised in screening mammography, participated in our study. They were selected after having volunteered to participate, taking into account their representation of all screening regions. The radiologists were—in shifts—invited for a two-stage experiment.

First stage

Independent reading of the index mammograms. Similar to daily screening practice, radiologists were offered all available films of the index and prior screening round. They did not know the number of cases in which cancer was subsequently diagnosed, nor were they given any clinical information. Ten radiologists read all subsets of mammograms in a 2-day session. They were asked to fill in a form for each series of mammograms and to denote whether any mammographic finding was visible, its location, its characteristics (mass, calcifications, architectural distortion, asymmetry or other). They were also asked to assess the probability of malignancy on a 13-point scale from < 0.5 to > 95%, and whether they would refer the woman for that particular finding. Sixteen other radiologists each read the mammograms of five of ten subsets (a total of 250 mammograms, divided into two groups). For each index mammogram, the radiologists had to fill in a form. They were asked to classify both breasts according to the BI-RADS classification system (scale 1–5, with 1 'normal', 2 'benign', 3 'probably benign', 4 'probably malignant', 5 'definitely malignant'); record their recommendation for referral for diagnostic evaluation; the localisation of any abnormalities, that is, suspected lesions; whether the abnormality had been visible on the prior mammogram; and whether they thought a cranio-caudal view should be available, and if so, for what reason. Thus, each series of mammograms in the test set was ultimately read by 18 radiologists.

After these sessions, referral recommendations were compared for all potential pairs of radiologists. For each radiologist a selection of discrepant readings (readings where the radiologists did not agree about referral) was made for the second stage of our study.

Second stage

The discrepant readings were re-assessed according to three referral strategies: (1) decision by one of the readers: the radiologists re-assessed 631 series of mammograms (451 of cases and 180 of controls) in the knowledge that at least one other reader gave a different referral recommendation; (2) decision by consensus: in pairs, the radiologists re-assessed 339 series of mammograms (244 of cases and 95 of controls) where they had disagreed about referral; (3) decision by arbitration: radiologists assessed 681 discrepant readings of mammograms (488 of cases and 193 of controls) that had not been read by the radiologist in the first stage.

Extrapolation to the nation-wide screening programme

The distribution of concordant and discrepant readings obtained in our experimental setting will not reflect daily screening practice. Because no actual or official national figures are available, we used published regional data to estimate the distribution of concordant and discrepant readings instead (Table 1) [2].

Table 1 Assumed distribution of concordant and discrepant readings in the nation-wide screening programme, per 10,000 screen examinations

Double reading	Cases	Controls	Total
Concordant *positive*	45	31	76
Discrepant	5	46	50
Concordant *negative*	15	9,859	9,874
Total	64	9,936	10,000

Based on data from the southern breast cancer screening region of the Netherlands (Stichting Bevolkingsonderzoek Borstkanker Zuid) [2]

Table 2 Referral recommendations (%) of discrepant readings (experimental setting)

Referral strategy	Controls[a]	Cases			All readings
		All cases	Screen-detected cancers[b]	Interval cancers[c]	
	Referral rate (%)				
Decision by consensus	37.9	73.8	69.7	77.9	63.7
Decision by one of the readers	35.6	57.4	53.0	61.5	51.2
Decision by arbitration	25.9	52.7	49.6	55.7	45.1
P-value (chi-square)[d]	0.054	<0.001	0.001	<0.001	<0.001

[a] The percentages concern 180 decisions by one of the readers, 95 decisions by consensus and 193 decisions by arbitration, respectively

[b] The percentages concern 217 decisions by one of the readers, 122 decisions by consensus and 242 decisions by arbitration, respectively

[c] The percentages concern 234 decisions by one of the readers, 122 decisions by consensus and 246 decisions by arbitration, respectively

[d] *P*-values concern the comparison of referral rates for the three referral strategies, per subgroup

We then applied the referral rates for discrepant readings that were found in our experimental setting (Table 2) to the estimated proportions of discrepant readings in daily screening practice. Referral rates were 100% for concordant positive readings and 0% for concordant negative readings. For each of the various referral strategies, we estimated the sensitivity if applied in the nation-wide breast cancer screening programme, assuming that each referral of a case would lead to the diagnosis of cancer.

We used the MIcrosimulation SCreening ANalysis (MISCAN) model to estimate the cost-effectiveness of screening with a potentially more effective double reading strategy replacing double reading with consensus. The MISCAN model was developed to estimate the costs and effects of the introduction of a nation-wide screening programme in the Netherlands and has been described in detail previously [8–10]. In the model, individual life histories are generated as a Markov process of stages and transitions. The natural history of breast cancer is modelled as progression through four invasive, screen-detectable, preclinical stages or transformation into a clinical stage. It is assumed that some of the invasive cancers are preceded by DCIS. Without screening, a preclinical cancer may either become clinically diagnosed or progress to the next preclinical stage. Screening will result in earlier detection, treatment and improved prognosis for some of the women. Screening characteristics have been validated using the results of the Dutch nation-wide breast cancer screening programme. The improvement in prognosis after screen detection has been validated by the results of the five Swedish randomised trials [11]. The output of the model yields a number of estimates of the effect of screening and subsequent outcome. The numbers of women invited and screened are assessed, as well as the resulting number of referred women and of screen-detected cancers, as well as the number of cancers diagnosed outside the screening programme. The output also includes the stage distributions of all breast cancers, and the number of breast cancer treatments. Moreover, the model calculates the number of breast cancer deaths and life-years lost due to cancer. Cost-effectiveness was calculated by adding a profile of diagnostic and treatment costs over time to each disease state.

To estimate the cost-effectiveness of decision by consensus (the current referral strategy) and of any referral strategy with a potentially higher detection rate, we had to change the sensitivity of the 'screening test' (which is related to the detection rate) and the positive predictive value of referral (which is related to the referral rate) in the model.

Results

Referral rates and discrepant readings

For all readings of the test set, the referral rate was 16.5% (individual radiologists' range 4.5–31.7). The mean referral rate was 27.8% for the cases (range 8.7–45.7) and 5.4% for the controls (range 0.0–18.1). The proportion of concordant referral recommendations by all 153 radiologist pairs that can be formed with 18 radiologists was 90.2% for controls: in 89.4% the radiologists both recommended that the woman not be referred, in 0.8% they both recommended referral. For the cases, the proportion of concordant readings was 75.2%: 59.3% were concordant negative referral recommendations, 15.9% were concordant positive recommendations. Of all readings by the 153 radiologist pairs, 17.7% were discrepant.

Re-assessment of discrepant readings

Referral rates after re-assessment of discrepant readings are presented in Table 2. For all three referral

strategies, re-assessment of discrepant readings in controls resulted in lower referral percentages than in cases. Referral rates were highest with decision-making by consensus; for cases this referral rate was 73.8% compared to 57.4% if the decision was made by one of the readers, and 52.7% in case of decision-making by arbitration.

Referral rate and sensitivity nation-wide; extrapolation

From the nation-wide extrapolation, it was estimated that single reading resulted in a lower sensitivity (73%) than double reading in general, except for double reading with referral only if recommended by both radiologists (sensitivity 69%) (Table 3). Of double reading strategies, only double reading with referral if any reader suggests resulted in a higher sensitivity than double reading with consensus (current situation), but this was accompanied by a 1.31 times higher referral rate.

Cost-effectiveness

Table 4 shows the effects and costs of screening if double reading with consensus would be replaced by double reading with referral if any reader suggests, assuming a relative increase of the detection rate by 2% (that is, from 4.42 to 4.51 per 1,000 women screened), and a relative increase of the referral rate by 30% (that is, from 0.93 to 1.21%). In the new situation, 700 women less will die from breast cancer, resulting in a gain of 12,000 life-years. In this scenario, double reading with referral if any reader suggests is comparably cost-effective to double reading with consensus (€2,168 and €2,207 per life-year gained, respectively). The extra costs related to the extra referrals offset the costs of palliative treatment that will be saved.

Model calculations show that any relative increase of the detection rate by up to 5% (that is, from 4.42 to 4.64 per 1,000 women screened) at the cost of a referral rate up to double the current referral rate is equally or even more cost-effective as long as the additional referrals only result in an increase of diagnostic mammograms, without an increase of biopsies or other invasive diagnostic procedures (Table 5). If the additional referrals result in extra invasive diagnostic procedures (the unfavourable scenario), the costs per life-year gained may double to €4,190.

Discussion

Reading strategies in the context of breast cancer screening programmes have been studied before. Mostly, double reading and single reading were compared. Most studies showed that the sensitivity of mammography screening increases with both non-consensus [12–16] and consensus [16–19] double reading compared to single reading. The effect of double reading on the specificity is not as unambiguous [18], but it tends to a decrease [13, 14, 19, 20].

In the present study, we compared different double reading strategies. We found that the sensitivity of the Dutch nation-wide breast cancer screening programme may increase if double reading with consensus is replaced by non-consensus double reading, that is, double reading with referral if any reader suggests. A higher detection rate, almost inevitably, can only be reached at the cost of an increased referral rate then. For that reason, the balance between true positive and false positive screen results becomes extremely important as well. We estimated

Table 3 Extrapolated sensitivity values and referral rates

Referral strategy in case of discrepant reading	Sensitivity[a]		Referral rate[b]	
	%	Relative increase by a factor	%	Relative increase by a factor
Decision by consensus (current strategy)	74.7	1.00	0.97	1.00
Referral if recommended by at least one of the readers	76.6	1.03	1.26	1.31
Decision by one of the readers	73.5	0.98	0.95	0.98
Decision by arbitration	73.1	0.98	0.90	0.93
Single reading[b]	72.9	0.98	1.01	1.05
Referral only if recommended by both radiologists	69.3	0.93	0.76	0.78

[a] Sensitivity is calculated as the proportion of cases referred for further assessment, assuming that breast cancer is diagnosed in all referred cases

[b] For 'single reading' we assumed that all concordant positive readings are referred, and half of discrepant readings

Table 4 Effects and costs of screening (model calculations)

	Current situation	Optimised; 2% relative increase of detection rate; 30% relative increase of referral rate
Effects[a]		
Number of screen examinations ($\times 10^6$)	26.4	26.4
Number of referred women ($\times 10^3$)	244.2	318.9
Referral rate (%)	*0.93*	*1.21*
Number of screen-detected cancers ($\times 10^3$)	116.7	119.0
Detection rate (per 1,000 screened women)	*4.42*	*4.51*
Breast cancer deaths ($\times 10^3$)	320.1	319.4
Life-years lost due to breast cancer ($\times 10^3$)	5,860.1	5,848.1
Quality-adjusted life-years lost ($\times 10^3$)	6,302.5	6,289.9
Costs[b] ($\times 10^6$ €)		
Screening	629	629
Diagnostics	1,268	1,273
Treatment	2,006	2,008
Follow-up	704	705
Palliative care	2,201	2,194
Total	6,807	6,810
Cost-effectiveness[b] (€)		
Costs per prevented death	27,602	27,125
Costs per life-year gained	2,207	2,168

In MISCAN, a 27-year breast cancer screening programme is assumed; the effects are calculated for a total period of 100 years (until all women in the cohort have died)

[a] Effects are 0% discounted

[b] To take account of time preference, the costs estimates over time were adjusted with 3% annual discount rate. To compute the cost-effectiveness, both effects and costs were 3% discounted

that double reading with referral if any reader suggests results in a 2% relative increase of the sensitivity, or an increased detection rate from 4.42 to 4.51%, whereas the referral rate increases by almost 30% compared to double reading with consensus. A computer-based simulation showed that the costs per life-year gained for double reading with referral if any reader suggests were comparable to the costs per life-year gained of consensus double reading, even if referral rates are double the current referral rate. If the additional referrals result in extra invasive diagnostic procedures (the unfavourable scenario), the costs per life-year gained will become twice as high as in the current situation, but may well be in the range

Table 5 Cost per life-year gained (sensitivity analysis)

Relative increase of the detection rate by a factor	Relative increase of the referral rate by a factor	Diagnostic follow-up[a]	
		Favourable scenario	Unfavourable scenario
1.01	1.3	€2,200	€2,390
	1.5	€2,210	€2,540
	2.0	€2,250	€2,920
1.02	1.3	€2,170	€2,350
	1.5	€2,180	€2,500
	2.0	€2,220	€2,870
	4.0	€2,380	€4,350
1.05	1.3	€2,080	€2,240
	1.5	€2,100	€2,390
	2.0	€2,140	€2,750
	4.0	€2,290	€4,190

To take account of time preference, both effects and costs were 3% discounted. Costs were rounded to the nearest €10

[a] In the favourable scenario, all women who are extra referred will have normal diagnostic mammograms (and thus will not face biopsy or other invasive diagnostic procedures). In the unfavourable scenario, the extra referrals will lead to a proportional increase of the number of biopsies or other invasive diagnostic procedures

of acceptable cost-effectiveness for Dutch health care programmes.

A comparison between UK National Health Service breast screening programmes [21] showed that referral rates were highest with double reading with arbitration (4.0%), higher than with single reading (3.6%), and higher than with non-consensus (3.4%) or consensus double reading (3.1%). Like in our study, however, higher referral rates do not necessarily imply higher detection rates. In the NHS study, detection rates were lowest for single reading and highest for double reading with arbitration [21]. These differences may be explained by the circumstances in which results were obtained. We obtained our results in an experimental setting; the participating screening radiologists fulfilled the role of arbiter in the arbitration scenario. In the UK study, the comparison of reading protocols was a comparison between programmes and regions, and the arbiters may have been greatly experienced radiologists. Another study in the NHS breast screening programme showed that referral rates were lower with consensus double reading than with non-consensus double reading (4.2 and 9.9%, respectively), but differences in the detection rate were not significant [16].

Remarkably, Ciatto et al. suggest double reading with arbitration instead of non-consensus double reading in case of high referral rates [22]. In the Florence screening programme, this would substantially reduce referral rates by 32.1%, that is, from 3.82 to 2.59%, at the cost of a 1.7% relative decrease of the detection rate, that is, from 4.58 to 4.50 per 1,000 women screened.

On the one hand, there may be some practical advantages of double reading with referral if any reader suggests, too. In the Netherlands, the majority of screening radiologists spend most of their time working in a hospital. It may not always be easy to bring both radiologists together to discuss discrepant readings. In future, however, digital mammography might facilitate quick discussion. On the other hand, the need for additional consultations and diagnostics may contribute to capacity problems.

If it is decided to implement double reading with referral if any reader suggests as the standard reading strategy into the Dutch nation-wide breast cancer screening programme, the effects should be monitored closely. For this evaluation, the intermediate outcomes of the screening programme, including referral rates, detection rates and interval cancer rates, can be used.

Acknowledgements We would like to acknowledge H. Rijken and M. Jacobs of the National Expert and Training Centre for breast cancer screening for their help with the study. Rob Boer of the Department of Public Health, is gratefully acknowledged for his contribution to the study design and the model calculations. We also would like to acknowledge the radiologists who participated in the study: Y.T. van Aardenne, D. Beijerinck, C. Boetes, J.H.B. Boomsma, W.G.J. Bors, A.C.W. Borstlap, P.A.M. Bun ev. Sevenstern, A. van Dalen, J.J.M. Deurenberg, H.A.J. Dijkstra, F.H. Jansen, B.A.E. van der Lande, T.H. Lauw, S.J. Liem, R.A. Manoliu, W.R. Obermann, E. Paalman, S.T. Parabirsing, R.M. Pijnappel, D.J. Reiding, J.L. Schreutelkamp, J. Schut, C.M. Stassen, P.T. Thung, P. Tirajoh, E.B.W. Treu. The following screening regions provided the mammograms for our test set: Preventicon, Stichting Bevolkingsonderzoek Borstkanker Noord-Nederland (BBNN), Stichting Vroege Opsporing Kanker Oost-Nederland (SVOKON), Stichting Bevolkingsonderzoek Borstkanker Zuid (BOBZ), Stichting Kankerpreventie en -screening Limburg (SKsL). This paper was completed in remembrance of Dr Jan H.C.L. Hendriks, who died 16 June 2004. This study was supported by a grant from the Health Care Insurance Board.

References

1. Fracheboud J, de Koning HJ, Boer R, Groenewoud JH, Verbeek AL, Broeders MJ, van Ineveld BM, Hendriks JH, de Bruyn AE, Holland R, van der Maas PJ (2001) Nationwide breast cancer screening programme fully implemented in The Netherlands. Breast 10(1):6–11
2. Duijm LE, Groenewoud JH, Hendriks JH, de Koning HJ (2004) Independent double reading of screening mammograms in The Netherlands: effect of arbitration following reader disagreements. Radiology 231(2):564–570
3. Uitspraak Medisch Tuchtcollege 's-Gravenhage (1998) Verdict Medical Disciplinary Board 's-Gravenhage [in Dutch]. Med Contact 53:661–662
4. Uitspraak Medisch Tuchtcollege 's-Gravenhage (1998) Verdict Medical Disciplinary Board 's-Gravenhage [in Dutch]. Staatscourant 20–21
5. Boer R, de Koning H, van Oortmarssen G, Warmerdam P, van der Maas P (1999) Stage distribution at first and repeat examinations in breast cancer screening. J Med Screen 6(3):132–138
6. Otten JD, Karssemeijer N, Hendriks JH, Groenewoud JH, Fracheboud J, Verbeek AL, de Koning HJ, Holland R (2005) Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. J Natl Cancer Inst 97(10):748–754
7. van den Akker-van Marle E, de Koning H, Boer R, van der Maas P (1999) Reduction in breast cancer mortality due to the introduction of mass screening in the Netherlands: comparison with the United Kingdom. J Med Screen 6(1):30–34
8. de Koning HJ, Fracheboud J, Boer R, Verbeek AL, Collette HJ, Hendriks JH, van Ineveld BM, de Bruyn AE, van der Maas PJ (1995) Nation-wide breast cancer screening in The Netherlands: support for breast-cancer mortality reduction. National Evaluation Team for Breast Cancer Screening (NETB). Int J Cancer 60(6):777–780
9. Groenewoud JH, Pijnappel RM, van den Akker-Van Marle ME, Birnie E, Buijs-van der Woude T, Mali WP, de Koning HJ, Buskens E (2004) Cost-effectiveness of stereotactic large-core needle biopsy for nonpalpable breast lesions compared to open-breast biopsy. Br J Cancer 90(2):383–392

10. Rijnsburger AJ, Essink-Bot ML, van Dooren S, Borsboom GJ, Seynaeve C, Bartels CC, Klijn JG, Tibben A, de Koning HJ (2004) Impact of screening for breast cancer in high-risk women on health-related quality of life. Br J Cancer 91(1):69–76
11. Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L, Larsson LG (1993) Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 341(8851):973–978
12. Denton ER, Field S (1997) Just how valuable is double reporting in screening mammography? Clin Radiol 52:466–468
13. Séradour B, Wait S, Jacquemier J, Dubuc M (1996) Dual reading in a non-specialized breast cancer screening programme. Breast 5:398–403
14. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE (1994) The efficacy of double reading mammograms in breast screening. Clin Radiol 49:248–251
15. Thurfjell EL, Lernevall KA, Taube AA (1994) Benefit of independent double reading in a population-based mammography screening program. Radiology 191:241–244
16. Brown J, Bryan S, Warren R (1996) Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. BMJ 312:809–812
17. Warren RM, Duffy SW (1995) Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. Br J Radiol 68(813):958–962
18. Anttinen I, Pamilo M, Soiva M, Roiha M (1993) Double reading of mammography screening films—one radiologist or two? Clin Radiol 48:414–421
19. Ciatto S, Ambrogetti D, Bonardi R, Catarzi S, Risso G, Rosselli Del Turco M, Mantellini P (2005) Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme. J Med Screen 12(2):103–106
20. Pauli R, Hammond S, Cooke J, Ansell J (1996) Comparison of radiographer/radiologist double film reading with single reading in breast cancer screening. J Med Screen 3(1):18–22
21. Blanks RG, Wallis MG, Moss SM (1998) A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two view mammography: results from the UK National Health Service breast screening programme. J Med Screen 5(4):195–201
22. Ciatto S, Ambrogetti D, Risso G, Catarzi S, Morrone D, Mantellini P, Rosselli Del Turco M (2005) The role of arbitration of discordant reports at double reading of screening mammograms. J Med Screen 12:125–127