

THE DOM PROJECT FOR THE EARLY DETECTION OF BREAST CANCER, UTRECHT, THE NETHERLANDS

F. DE WAARD, H. J. A. COLLETTE, J. J. ROMBACH,
E. A. BAANDERS-VAN HALEWIJN and C. HONING

PREVENTICON, Radboudkwartier 261, 3511 CK Utrecht, The Netherlands

I. DESIGN AND METHODOLOGY

1. Introduction

The Department of Epidemiology at the Institute of Social Medicine,* University of Utrecht has had a special interest in breast cancer for a number of years. In a cohort study in postmenopausal women, which involved 50 general practitioners, including H. J. A. Collette, F. de Waard and E. A. Baanders-van Halewijn [1] found evidence of an effect of nutritional status, as reflected by body height and weight, on the risk of developing breast cancer.

During a WHO fellowship in 1969 at the Johns Hopkins School of Hygiene and Public Health one of us (de Waard) became acquainted with population screening for breast cancer by means of physical examination, mammography and thermography. At the 10th International Cancer Congress in Houston (1970) de Waard saw the first examples of xeroradiograms of the breast. In 1971 Shapiro *et al.* [2] published their promising papers on the H.I.P. project in New York, which showed the benefit of screening in postmenopausal women.

We felt that further studies employing periodic screening in a population of postmenopausal women were warranted since they could lead to a better insight into the natural history of breast cancer, and in particular to a greater understanding of the nutritional determinants of growth rates of preclinical cancer in postmenopausal subjects. We had been involved in endocrinological studies at our university which had emphasized the role of extra-ovarian oestrogens in the aetiology of endometrial and breast cancer [3, 4].

Moreover, our cohort study had demonstrated a relationship between overweight and extra-ovarian oestrogen production [5].

We therefore suggested a population-based study of periodic screening in women aged 50-64 years, coupled with nutritional intervention aimed at reducing weight. The rationale was that oestrogen-dependent preclinical breast cancer would develop more slowly during and after weight reduction, thus giving support to a system of periodic screening. We hoped that Dutch public health research funds would enable us to carry out the screening and that financial support would be forthcoming from the U.S. National Cancer Institute (NCI) for a randomized trial of nutritional intervention and some endocrinological research associated with it.

*Now Institute of Public Health and Epidemiology. The study was funded by the Ministry of Health and the "Praeventiefonds".

In the event we succeeded in obtaining funds for the screening, but failed to get support from the NCI mainly because of scepticism regarding the feasibility of weight reduction on a mass scale. (A feasibility study was begun 2 years later with funds from other sources and the results will be published shortly by Baanders-van Halewijn *et al.* in *Int. J. Obesity*.)

2. Staff

Initially the scientific staff of the DOM (*Diagnostisch Onderzoek Mammacarcinoom* or diagnostic investigation of mammary cancer) project consisted of the physicians already mentioned. Later their ranks were strengthened by J. J. Rombach, Head of Diagnosis, and a team of paramedical assistants (coordinated by Mrs C. Honing) who not only took and developed the xeromammograms but also learned to read them. In addition, they performed the clinical examination (inspection, palpation) after checking data supplied by the women on personal history, risk factors and so on. Final responsibility for decisions taken at screening rested with the Head of Diagnosis, who has reported on the procedure in greater detail elsewhere [6]. Screening began in December 1974, and after a month or two of refining our techniques, we were able to screen 50–55 women a day at our facility in the “Preventicon”, which had the advantage of being near both the railway and bus stations.

3. Population

On the basis of the results in the H.I.P. study it was decided to limit screening to women aged 50–64 at intake, i.e. those born between 1911 and 1925. The catchment area consisted of the city of Utrecht, some of its suburbs and a number of nearby towns and villages.*

In the Netherlands each local government area has a population register (either on computer tape or cards) which makes it possible for official bodies to invite all people of, say, a given age, sex or residential area to participate in a specified programme, e.g. mass screening. In general cooperation with the local authorities was good to excellent, although two districts, not mentioned in our report, showed no interest in the project.

After extensive publicity in 1974 women were invited, on a street-by-street basis, to visit the Preventicon at a given time. If they did not attend the invitation was renewed. It was assumed that failure to respond after two invitations was indicative of a lack of interest in being screened for breast cancer. Each woman was asked to contribute D.fl. 12.50 (approx US \$5) to the programme but provision was made to reduce this fee.

Women who responded in the city of Utrecht were offered a total of four screenings, with an interval of 12 months between the first and second examination, of 18 months between the second and third, and of 24 months between the third and fourth. Women in the other areas were offered two screening examinations with an interval of 18 months.

4. Examination

Each woman was welcomed by a receptionist of the appropriate age (over 40) who, after dealing with a number of administrative matters, invited the patient to an audiovisual series which served to inform her about the screening procedure and reduce anxiety. Each woman was asked to fill out a questionnaire before coming to the screening centre, and this was checked by the paramedical assistant before she performed the clinical and (xero)mammographic examination. At the first screening both medio-lateral and cranio-caudal pictures were obtained; subsequent radiological examinations included only a medio-lateral picture unless there were special reasons for taking a cranio-caudal one as well, e.g. suspicious features on the first projection. Such a departure from the normal routine was necessary in less than 5% of examinations.

As stated earlier, the paramedical assistants not only made the xeroradiograms but also learnt to read them under the supervision of the Head of Diagnosis. In this regard some of these workers acquired a skill equal to that of an experienced radiologist.

*Maarssen, Maartensdijk, De Bilt, Zeist, Houten, IJsselstein, Montfoort, Kamerik, Vleuten-De Meern, Harmelen and Kockengen.

Mainly for scientific reasons measurements were made in all cases of weight, height, skinfold thickness and blood pressure. Urine was tested at the first two examinations for protein and glucose; about 200 ml was frozen and stored for future scientific work.

5. *Ionising radiation*

Particular attention was paid to radiation dosage in xeromammography. When the project was started at the end of 1974 the dose of radiation absorbed in the breast amounted to 1 rad. In mid-1977 technical modifications were implemented which reduced the amount of integrated absorbed radiation to 0.4 rad (measurements carried out by the radiation control division of TNO, Arnhem).

To minimise the radiation dose, the number of xeromammograms at the second, third and fourth screening examinations was normally limited to one per breast, i.e. in the medio-lateral projection. This seemed to be warranted in view of the Swedish experience [7] with the oblique view.

We calculated that a woman undergoing four screening examinations received an average total dose of 4.7 rad in a period of 5 years.

6. *Data collecting and processing*

All data obtained from the questionnaire and information gathered at the clinical and mammographic examination were coded by the staff, transferred to punch cards and fed into the Cyber computer at the University Computer Department.

As a service, a print-out was sent periodically to local general practitioners (GPs), which provided them with a list of their patients being screened as well as any relevant additional clinical data (e.g. use of drugs, blood pressure, weight and height).

Any changes in the population at risk due to death or migration were added to the data base. Causes of death were documented with the cooperation of the GPs.

7. *Referral system*

In building up a communication network our principal aim was to make use of the existing structure of primary health care and specialist services. Any abnormality found at screening was reported to the woman's GP, who then decided, in consultation with his patient, which surgeon should be consulted for further investigation, namely a biopsy.

From the outset we were determined to keep the number of recommendations for a biopsy at a reasonably low level. Unlike the clinical situation in which the patients seeks help for a particular symptom (e.g. a lump in the breast), one has to weigh the required degree of sensitivity against a sufficient degree of specificity. In practical terms this means that a recommendation for biopsy has a high predictive value.

In an attempt to steer a well-balanced course between these requirements the following guidelines were established:

(a) When clinical and mammographic data indicated a suspicious lesion, the patient was referred to the GP with a recommendation that biopsy be performed.

This advice was almost always followed.

(b) If clinical examination revealed some abnormality which on (xero) mammography appeared benign, the GP received a brief note, leaving him with complete freedom to decide whether to refer the patient to a surgeon or not. In practice, only a minority of such patients were biopsied.

(c) An abnormality judged by the Head of Diagnosis to be non-malignant, but which demanded follow up within the normal screening interval resulted in a message to the GP to arrange a check up after 6 months. In the beginning we left it to the GP to take the initiative, but later found it politic to assume a more active role in securing the extra check.

If none of the above guidelines was applicable (which was the case in about 95% of examinations), we sent no immediate message to the GP, but kept him posted every 3 months with a list of his screened patients. All the women undergoing screening were told

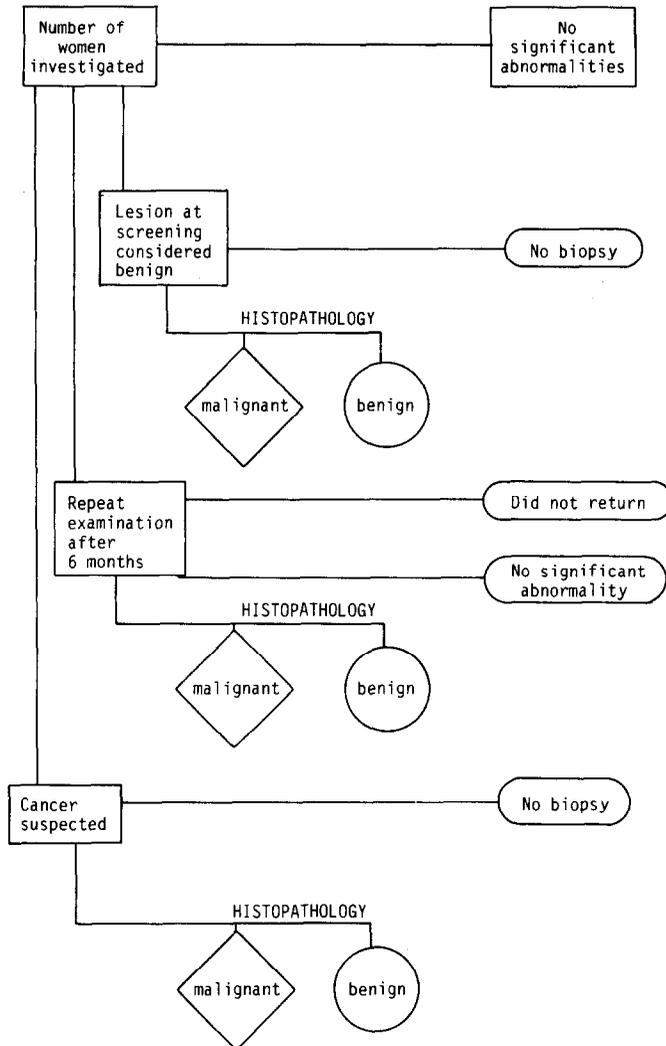


FIG. 1. "Decision tree" of the DOM project. Three types of "signals" are shown at the left side of the diagram. Each signal may lead to a biopsy, resulting in the diagnosis of a malignant or a benign lesion.

that if they were not approached by their GP within a week, then they had no cause for concern.

The course of screening events using the above-mentioned guidelines was documented in the form of "decision trees", which are presented in the next chapter. An "empty tree" is shown above (Fig. 1).

8. Definitive diagnosis

The definitive diagnosis was made by clinical specialists—surgeons, radiologists and pathologists—in various hospitals. The project triggered increased cooperation between these disciplines in the effort to diagnose small, non-palpable carcinomas.

From a histopathological point of view a distinction has been made in the present report between invasive cancer, intraductal cancer and lobular carcinoma *in situ*.

Three groups of pathologists serve Utrecht, its suburbs and surrounding areas, and their splendid cooperative efforts were invaluable when evaluating the screening project.

9. Therapy

When the project began most surgeons in the area preferred the Halsted type of radical

mastectomy. Gradually less radical operations, e.g. Patey's modification in which the pectoral muscles are left intact, became more popular. However, in the period covered by the present report (1975–1980) modern trends towards conserving the breast had yet to establish themselves in Utrecht and neighbouring hospitals.

Radiotherapy was concentrated in one hospital (the University Department of Radiotherapy) which was equipped with modern facilities, including supervoltage treatment.

10. *Psychosocial aspects*

Experience gained at screening demonstrated the importance of psychological aspects in the sudden suspicion of breast cancer. Any sign of extra attention with regard to clinical or radiological examinations produces anxiety in the patient being screened. Screening differs from the normal clinical situation inasmuch as, in the latter case, further examination tends to decrease uncertainty, following a period of disquiet.

At the start of the DOM project communicating any suspicion and providing guidance during the period of further diagnostic investigation were thought to be the exclusive responsibility of the GP. However, during the course of the project it transpired that this general rule did not always work satisfactorily. After consultation with the GPs one of us (C.H.) assumed a degree of responsibility for communication and guidance until such time as a firm diagnosis had been established.

Any woman in whom breast cancer had been diagnosed either at the previous screening or in the interval between screenings was seen at the next visit to the centre either by the Head of Diagnosis or by C.H.

11. *Scientific research*

The DOM project was designed not only to gain experience in population screening for the early detection of breast cancer, but also to try to gain further insight into the natural history of such cancer after the age of 50. The investigators were and are particularly interested in the role of extra-ovarian oestrogens in stimulating growth rates of oestrogen-dependent breast cancer. Since adipose tissue has been shown to be a source of oestrogen, any relationships between overweight and mammary cancer can be explained biologically.

Other hormonal aspects studied during the course of the project included the urinary excretion of androgen metabolites following Bulbrook's plea [8] for an investigation of their potential in defining a high-risk group. The results in this respect were rather disappointing [9]. We also studied plasma prolactin levels in women with a family history of mammary cancer and in those on rauwolfia treatment for hypertension [10].

II. POPULATION RESPONSE TO SCREENING INVITATION

1. *General remarks*

As the DOM project is a population-based study, all women who were born between 1911 and 1925 and were resident in the city of Utrecht, its suburbs or adjacent areas were invited to participate through close cooperation with the local authorities.

As mentioned earlier, those living in Utrecht itself were invited to attend four successive screenings at intervals of 12, 18 and 24 months respectively whereas women in the other areas were invited to come only twice, with an interval of 18 months between visits.

Failure to respond (after two letters) was interpreted as evidence of a lack of interest, and these women were not invited to the following screening cycle. Unlike some other screening projects, invitations to a forthcoming examination were sent only to those who had participated in a previous one. Successive screening cycles were thus performed in a population progressively selected in terms of compliance. The response rate for the initial screening examination was about 72%, while the rates for the following examinations were over 80% (see also Table 3).

TABLE 1. RESPONSE PERCENTAGES BY AGE, INITIAL SCREENING, CITY OF UTRECHT

Year of birth	Invited	Response	
		Number	Percentage
1921-1925	7392	5631	76
1916-1920	6645	4826	73
1911-1915	6518	4240	65
Total	20,555	14,697	72

2. City of Utrecht, first screening

Invitations were sent to 20,555 women. The response after one letter was 66%, but a second letter brought a further 6%, giving a total of 72% (14,697 women).

Classification according to age at first screening is presented in Table 1, which shows a somewhat lower response rate in the older women.

Classification according to residential area shows a variation in response between 57 and 81%; this cannot be explained by differences in age distribution. The municipal department of statistics provided us with an occupational classification by residential area, but no rank correlation was found between occupational grading of areas and response rates.

There was a significant trend towards decreasing response rates with time (invitations were sent out area by area). While this could be attributable in part to a return to "normal" after the wave of publicity in 1974 at the start of the project, the trend continued in other areas after 1977. A more tenable explanation may be that mammography established itself during this time, and that, with increasing use of this technique in hospitals, women came under "surveillance" by another health care system, albeit one not primarily concerned with preventive medicine.

Unmarried women had a slightly higher response rate than those who had been or were married (75 vs 71%). Those who did not come to the centre were divided into four groups on the basis of their response (or lack of it) to the invitation (Table 2). It can be seen that single women more often provided a reason for not coming (often related to their job) and were also more likely to give a straight refusal for unknown reasons.

Special attention was given to foreign nationals resident in the city (22 countries); an interpreter was available at the centre for Spanish, Italian, Greek, Turkish and Moroccan women. Of the 199 foreigners invited, 36 (30%) responded.

3. Areas outside Utrecht

The response rates after two letters show little variation. The trend towards a modest but significant decrease in these rates with time (Kendall's rank correlation, $p < 0.02$) may in part be explained by the fact that those invited became progressively older since we stuck to the original birth dates between 1911 and 1925.

4. Compliance

As mentioned earlier, invitations to any screening cycle were sent only to those who had participated in the previous one. Thus compliance tended to increase on the basis of this selection (Table 3).

TABLE 2. NON-RESPONDERS, INITIAL SCREENING, CITY OF UTRECHT ($n = 5858$). PERCENTAGE DISTRIBUTION ACCORDING TO CIVIL STATUS AND TYPE OF NON-RESPONSE

	Single	Married	Widowed	Divorced
No response	74	87	89	88
Not able to come, reason given	10	7	6	6
Not able to come, no reason given	4	2	2	3
Straight refusal	12	4	3	3
Total percentage	100	100	100	100
number	727	3877	941	313

TABLE 3. PERCENTAGE DISTRIBUTION OF RESPONSE AND NON-RESPONSE BY SCREENING CYCLE (ADJUSTED FOR MIGRATION AND DEATHS)

Screening cycle (Utrecht)	Response		Percentage non-response	
	Number	Percentage	Replied	No reaction
1	14,697	72	4	24
2	11,867	81	2	17
3	10,208	87	2	11
4	8617	85	3	12
Screening cycle (other areas)				
1	8814	68	5	27
2	7074	81	3	16

Age appeared to be the most important determinant of compliance in every screening cycle. Figure 2 shows graphically the fraction of those attending the fourth screening cycle divided by those attending the initial examination, according to age; the decrease is almost linear.

Part of this age factor is related to ill-health. Table 4 documents that the DOM cohort was no exception to the rule that age determines mortality. Moreover, compliance was reduced in those who were on drug therapy for heart disease or diabetes at the time of their first examination.

Relative body weight, as expressed by Quetelet's index (weight/height²), was shown to have an interesting relationship with compliance. The highest response came from the middle weights while compliance was lower in those who were severely under- or

TABLE 4. NUMBER OF DEATHS, BY AGE, UP TO THE END OF 1980 AFTER INTAKE INTO THE DOM PROJECT, CITY OF UTRECHT

Number of screening examination attended	Age at intake								Total
	50-54		55-59		60-64		> 64		
	R*	NR**	R	NR	R	NR	R	NR	
1	11	16	7	29	15	46	2	4	130
2	8	10	12	8	24	24	8	4	98
3	4	3	15	6	11	4	12	3	58
4	—	NA***	5	NA	4	NA	7	NA	16
Total	52		82		128		40		302
% of total number of women at intake	1.0		1.7		3.1		11.9		2.1

*R = responders; women deceased in interval between screenings. **NR = non-responders, women deceased after official interval time had elapsed. ***NA = not applicable.

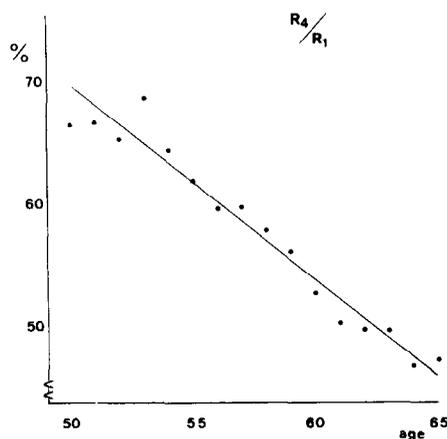


FIG. 2. Age-specific compliance expressed as the ratio of those attending the 4th screening divided by the number examined at initial screening.

overweight. Besides medical factors, obesity is more prevalent in socio-economic groups with poorer education [11] and education is thought to affect compliance.

Women with mothers who had had breast cancer had higher response rates in each successive cycle. Strangely enough, such motivation was not found when there was a positive family history in siblings.

Social factors: in the Netherlands about 70% of the population with an income below a fixed level is insured through the nationally-coordinated Sick Fund; the remainder are insured privately. Compliance was somewhat lower among women in the Sick Fund. Single women had a slightly higher compliance than those who were or had been married. Women with very large families (more than 6 children) had a lower compliance, but it seems unlikely, in view of their age, that this was due to caring for small children.

5. Repeat invitation

If a woman did not respond to the first invitation to be screened, she received a second letter about 2 months later. This improved the response rates by between 6 and 9%.

An investigation of those who responded positively to the second invitation showed that these women shared a number of similarities with the non-responders, e.g. concerning type of sickness insurance, family size and relative body weight. Those responding to the second invitation had a three-fold probability of non-response at the second screening 12 or 18 months after the initial examination.

6. Late regrets

Although women who failed to return after the first examination were automatically excluded from subsequent cycles, we guessed that about 20% of them would later come to regret their decision and be interested in further screening should the opportunity be offered. In 1980 we therefore invited all those who had failed to turn up at the second screening to come to the centre.

Of a total of 2512 invited, 372 (15%) appeared, 152 (6%) declined to come (some gave reasons, some did not), and 1988 (79%) did not react at all. It was established that all those invited were still alive and resident in the area. The 273 women who turned up were asked to fill out a questionnaire. No fewer than 72% said that their previous lack of response was attributable to their being on holiday. (We wondered if this were true in all cases since the second letter of invitation for each screening examination was always sent two months after the first.) Fourteen percent said that they were under continuous medical care for their breasts and 8% said they had not responded because a previous mammographic examination had been painful. Risk factor analysis (see Section V) revealed nothing special about this group. In none of these women was breast cancer found.

7. Financial contribution by those screened

Although it was decided to levy a fee of D.fl. 12.50 (approx. US \$5) for each examination, this sum could be waived or reduced in case of need. In the event no more than 3% of the responders applied for such help.

The percentage of women screened at a reduced fee was dependent to some extent on age (higher in older women), was higher in Utrecht than in the other areas, and, not surprisingly, was higher in the Sick Fund group than in those with private insurance. Initial screening at a reduced fee was associated with a somewhat lower response for the second screening than in the series as a whole (74 vs 81%).

8. Number of investigations

Between 11 December 1974 and 31 December 1980 a total of 61,649 breast examinations were performed. A breakdown according to screening cycle and year is presented in Table

TABLE 5. NUMBER OF SCREENING EXAMINATIONS IN THE DOM PROJECT

Year	Screening cycle	City of Utrecht				Other areas		Total
		1	2	3	4	1	2	
1975*		5327						5327
1976		6940	4248					11,188
1977		2430	5542	1432		2942		12,346
1978			2077	4620		4102	444	11,243
1979				4022	1168	1770	3846	10,806
1980				134	7449		2784	10,367**
Sub-total		14,697	11,867	10,208	8617	8814	7074	61,277
"Late regrets"								372***
Total								61,649

*Includes a few examinations in 1974; **372 "late regrets" excluded; ***Examined in 1980.

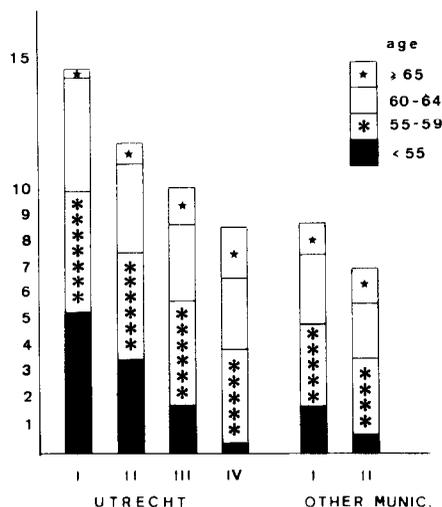


FIG. 3. Age distribution (in thousands) of women examined in 4 screening cycles (Utrecht) and in 2 screening cycles (other areas).

5. Figure 3 depicts separately the absolute number of women examined (in thousands), by age and cycle, in Utrecht and the other areas.

III. THE PROCESS OF SCREENING AND ITS IMMEDIATE RESULTS

1. Decision theory in practice

A screening project should pick out individuals with a considerably increased risk of having a disease. While breast cancer screening should detect suspicious cases, it is clearly important that one should not arouse undue anxiety in the community as a whole. The basic problem lies in making a distinction between a malignant and a benign lesion: the doctor cannot and should not burden healthy subjects with his doubts by suggesting a series of regular repeat examinations. Thus, in screening, one should decide straight away whether malignancy is suspected or not. This contrasts with the clinical situation in which it is the patient who seeks help and in which both doctor and patient may decide to postpone a decision, pending further investigation of the breast at some future date. Moreover, in a hospital setting additional investigations may be made which are no part

of a screening programme. In this way specificity and the predictive value of a recommendation for biopsy can be increased.*

When we began our project there were no clear-cut criteria for population screening for breast cancer, and we had to rely heavily in the beginning on experience in hospital practice. We were, however, careful to establish our own criteria, based on clinical examination plus xeromammography, for the delineation between suspicious and non-suspicious lesions.

We tackled the "grey area" between suspicion and no suspicion by introducing a special repeat check after 6 months [12].

We considered that women who fell into this category probably had benign lesions, but we nevertheless felt it politic to follow any possible development by instituting one additional examination during the interval between standard screening cycles.

2. Three types of signal (Utrecht, first screening)

The actual screening process can be well illustrated by using figures from the initial screening examination in the city of Utrecht.

As mentioned earlier, we distinguished between three types of "signal" and the corresponding messages which were sent to the GP of the woman in question.

- (A) An abnormality which, on the basis of combined clinical and xeromammographic investigation, aroused definite suspicion of breast cancer resulted in a *recommendation for biopsy*. 263 (1.8%)
- (B) An abnormality thought to be non malignant but which appeared to demand an additional *check after 6 months*. 201 (1.4%)
- (C) A lump in the breast which did not appear malignant on mammography resulted in a *brief note* to the GP. 252 (1.7%)
- (D) If none of the above categories was applicable, the woman was declared free of abnormalities. While no immediate message was sent to the GP, he did receive a list of his screened patients every 3 months. $\frac{13,981}{14,697}$ (95.1%) (100%)

The absolute and relative frequencies of these four diagnostic categories in the initial investigation of women from the city of Utrecht are depicted in the form of a "decision tree" (Fig. 4).

(A) *Recommendation for biopsy*. Not all recommendations were implemented: one or two women deliberately chose not to follow our advice, a few were "kept under surveillance" by the GP, and the rest were followed by the surgeon consulted. Only 10 (3.8%) of the 263 patients in this category were not biopsied.

Excluding five cases of lobular carcinoma *in situ*, 101 breast cancers were diagnosed in the 253 biopsied cases, i.e. in 39.9% of these women. This percentage—the predictive value of a positive test—constitutes an important parameter of the impact of a screening

*It should be made clear that sensitivity cannot be increased by means of additional investigations, since this characteristic of the screening test is entirely dependent on the maze of the first screen. For those not familiar with the definitions of these terms the following may be of help:

Screening test	Does the disease exist in reality?		Total
	Yes	No	
Positive	a	b	a + b
Negative	c	d	c + d
Sensitivity = $\frac{a}{a + c}$	a + c	b + d	a + b + c + d
Specificity = $\frac{d}{b + d}$	Predictive value of a positive test = $\frac{a}{a + b}$		

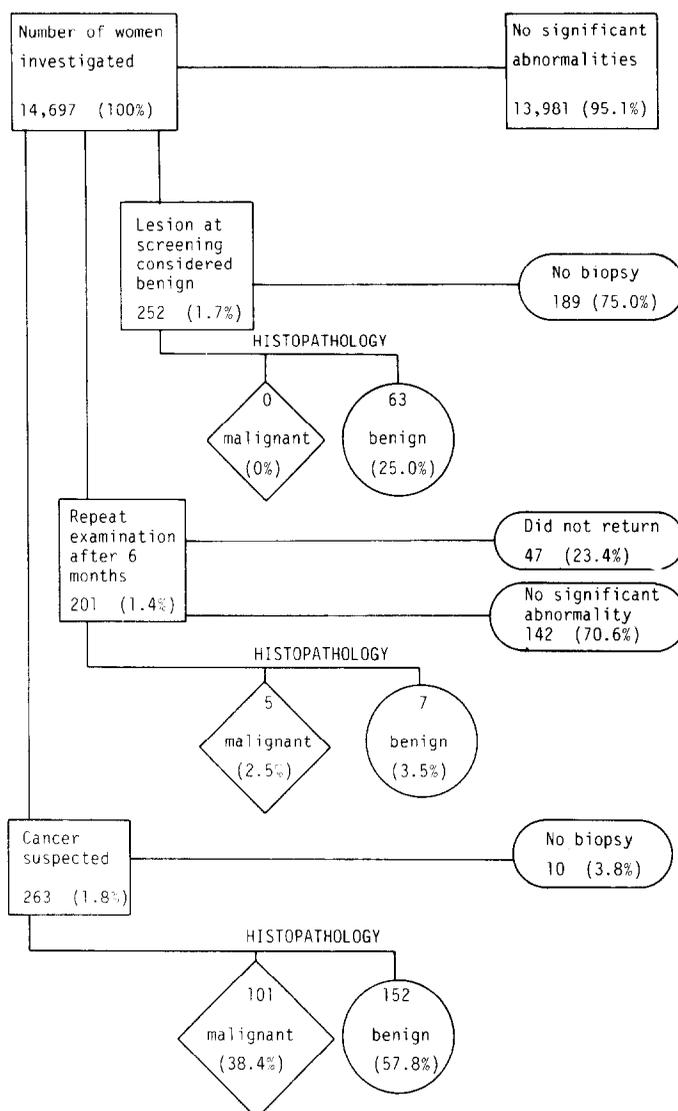


FIG. 4. "Decision tree", initial screening, city of Utrecht.

programme on the population. Our figure is fairly high on international comparison, for example, the results of the Breast Cancer Detection Demonstration Project (BCDDP) in the United States show a percentage of about 15 [13].

(B) *Check after 6 months.* In the initial investigation in the city of Utrecht, GPs were advised on 201 occasions to send their patients to the Preventicon for an additional check after 6 months. In 154 cases (76.6%) this advice was followed.

In an attempt to achieve a higher response, the staff at the centre were later encouraged to communicate this advice to the patient as well, and this resulted in greater compliance during succeeding cycles.

In the 154 women seen after 6 months, biopsy was recommended in 12, and this resulted in the detection of a further five breast cancers, giving a predictive value of 41.7%, almost identical to that in category A.

(C) *Brief note.* Following the initial examination of women from the city of Utrecht 252 brief notes were sent to GPs concerning abnormalities which were considered benign on the basis of mammography. In 63 (25.0%) of these cases, the GPs exercised their option to send the patients for biopsy, but in no case was cancer diagnosed.

TABLE 6. "PICK-UP" RATES BY AGE. INITIAL SCREENING. CITY OF UTRECHT

Age	Number of women examined	Number of breast cancer cases detected	Pick-up rate (‰)
50-54	5285	35	6.6
55-59	4871	28	5.7
> 59*	4541	43	9.5
Total	14,697	106	7.2

*Slightly more than 400 women were just over 64 years of age at the time of screening.

(D) We discuss in Section IV whether any cases of cancer were found in the interval between the initial examination and the second screening.

Thus, of the 106 cases of invasive and intraductal cancer found at the first screening in Utrecht in 14,697 women, 101 (95.3%) were detected as a result of our immediate recommendations for a biopsy and five (4.7%) as a result of the repeat check at 6 months. No case of cancer was detected when biopsy was performed after the issue of a brief note.

The yield of mammary cancer through screening can be expressed as the "pick-up rate". The numerator may signify either the number of women with cancer or the number of tumours (three women in the first Utrecht screening had bilateral cancer). The denominator represents the number of women screened. The pick-up rate in Utrecht was 7.2‰ or 7.4‰, respectively.

Table 6 shows that the rate is age-dependent, and that it dipped in the 55 to 59-yr-old group.

3. Other areas, first screening

From mid-1977 onwards the other areas mentioned in Section I were approached successively for participation in the DOM project.

The results of the first screening in these areas are depicted, in the form of a decision tree (Fig. 5), in a similar way as was done for the city of Utrecht. Table 7 shows that experience gained in the first 30 months of the project resulted in some firming up of the classification in categories A-D (Section III.2).

From this Table it can be seen that fewer women were recommended for biopsy and far fewer were thought to warrant a brief note. The number of checks after 6 months showed little change. Tests of significance were not applied since they mean relatively little in such a large population. Of more significance is their change over time, which is thought to depend more on continuous adjustments in diagnostic behaviour, reflecting growing experience, than on groups screened. A graphic presentation of the three types of signal over time is given in Fig. 6. The trend towards fewer recommendations for biopsy resulted in an increased predictive value of such recommendations, as can be seen from Table 8 which compares the periods 1975 to mid-1977 and mid-1977 to 1979, when the initial examinations were carried out (see also Fig. 3). The increase in the predictive value of a recommendation for biopsy—from 40 to 56%—is significant ($p < 0.02$).

This kind of modification of decision threshold, however, has an effect on the sensitivity of the screening process. In this connection the slightly lower yield for breast cancer in the suburbs and surrounding areas—although not significant from a purely statistical point of view—may yet have import.

Utrecht: 106 cases of cancer in 14,697 women = 7.2‰
Other areas: 57 cases of cancer in 8814 women = 6.5‰.

The small difference in age between the women from Utrecht and those from the other areas cannot explain the difference in yield; on the contrary, as the women from outside Utrecht were somewhat older, one would expect that if there were a difference it would be in the opposite direction.

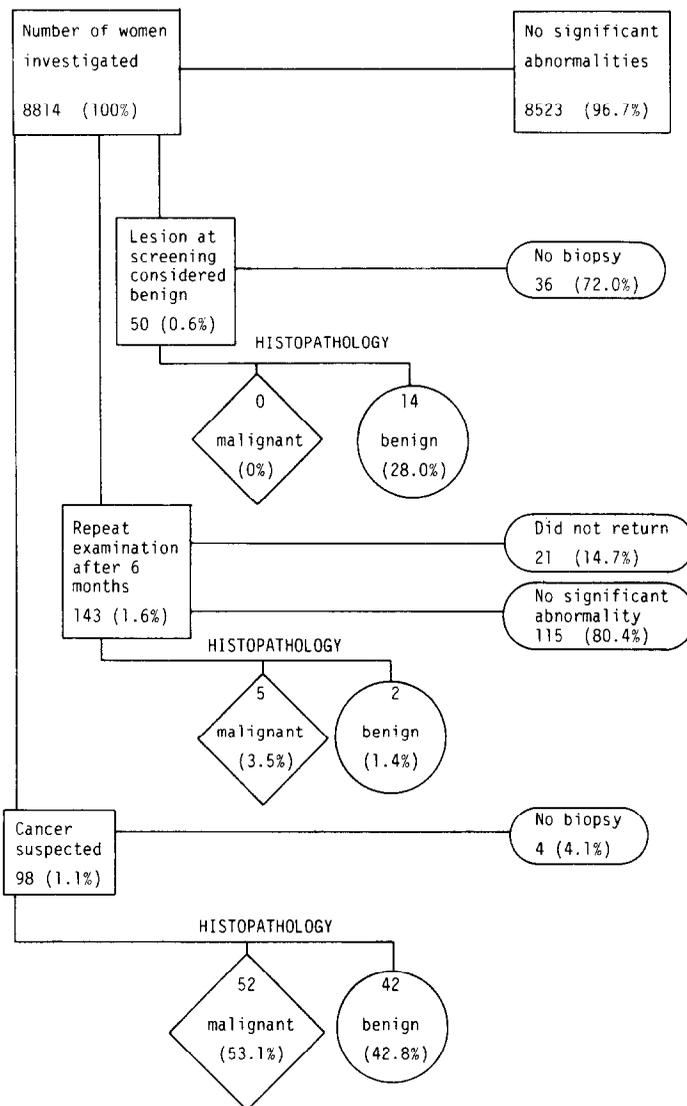


FIG. 5. "Decision tree", initial screening, other areas.

TABLE 7. TWO TIME PERIODS COMPARED

Classification at screening	1st cycle Utrecht (1975-1977)	1st cycle other areas (1977-1979)
(A) Recommendation for biopsy	263 (1.8%)	98 (1.1%)
(B) Repeat examination after 6 months	201 (1.4%)	143 (1.6%)
(C) Considered benign	252 (1.7%)	50 (0.6%)
(D) No abnormalities	13,981 (95.1%)	8523 (96.7%)
Total	14,697	8814

A further argument for assuming a slightly lower sensitivity in the period after mid-1977 is provided by the differences between Utrecht and the other areas in the yield from the second screening cycle (Table 9) and from a study of interval cancers* (Section IV).

*Breast cancers diagnosed within the designated intervals between screening examinations but not as a result of a recommendation for biopsy or a check at 6 months.

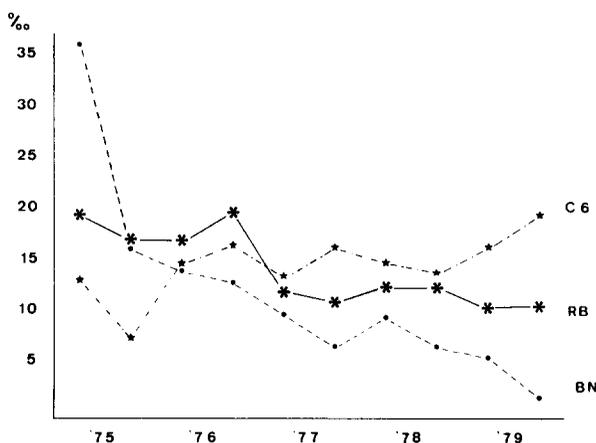


FIG. 6. Time trends for three types of signal (initial screening cycle) per 1000 examinations. RB = recommendation for biopsy; C6 = check after 6 months; BN = brief note.

TABLE 8. PREDICTIVE VALUE OF RECOMMENDATION FOR BIOPSY, 1975 TO MID-1977 AND MID-1977 TO 1979 RESPECTIVELY (1ST SCREENING)

	Utrecht (1975-1977)		Other areas (1977-1979)	
	Number of biopsies	Number of cancers	Number of biopsies	Number cancers
1st screening	253	101	94	52
At check after 6 months	12	5	7	5
Total	265	106	101	57
Predictive value	106/265 = 40.0%		57/101 = 56.4%	

In his doctoral thesis, Rombach [6] makes it clear that with a given screening test (or combination of tests) and given skills in applying such tests, an increase in specificity inevitably means a decrease in sensitivity and vice versa. Optimisation of screening necessarily means a compromise between the two.

4. Screening cycles after the initial examination

In discussing the results, we have taken the second, third and fourth examinations for Utrecht and the second examination for the other areas together ("decision tree", Fig. 7).

In evaluating the results of the later screening cycles, it should be made clear that the screening process was not identical to that in the first examination. In the later screening xeromammography was confined to the medio-lateral projection unless an abnormality was revealed. Secondly, the xeromammograms were read by comparing them with the pictures taken previously. It hardly needs to be said that the behaviour of an abnormality over time constitutes an important aspect of clinical judgement.

The pick-up rate in the initial screening was a measure of the prevalence of preclinical cancer (at varying developmental stages). By contrast, the rate in later cycles reflected the

TABLE 9. PICK-UP RATES AT 2ND-4TH SCREENING CYCLES

	Number of women examined	Number of breast cancers	Pick-up rate (%)
Utrecht, 2nd cycle	11,867	15	1.3
Utrecht, 3rd cycle	10,208	15	1.5
Utrecht, 4th cycle	8617	17	2.0
Other areas, 2nd cycle	7074	24	3.4

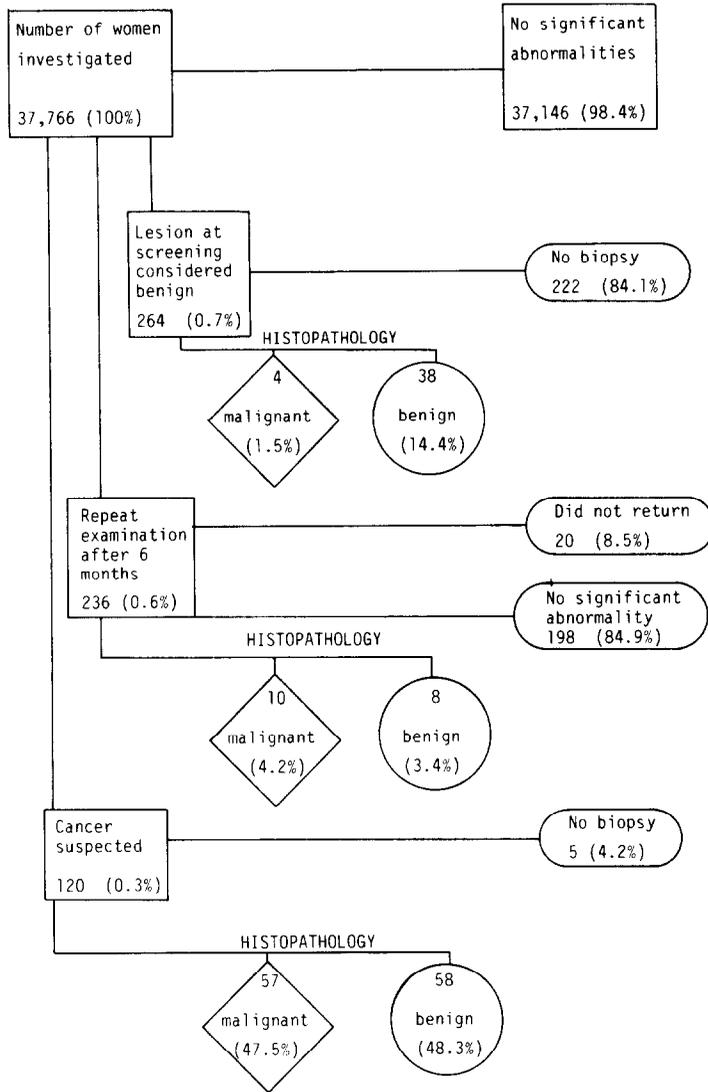


FIG. 7. "Decision tree", 2nd-4th screening cycles.

onset of new tumours (assuming a high degree of sensitivity in the first screening) and was dependent on their growth rate in relation to the interval between screening cycles. As explained in Section I, the interval was not the same between any set of two successive examination:

- | | | |
|-----------------|---|-----------------------------------|
| City of Utrecht | { | 12 months between screen 1 and 2 |
| | | 18 months between screen 2 and 3 |
| | | 24 months between screen 3 and 4* |
| Other areas | } | 18 months between screen 1 and 2* |

The pick-up rates for the second to fourth screening cycles are presented in Table 9. It is at once clear that these rates are considerably lower than those at the initial examination, and it would therefore appear that a prerequisite for successful screening is that prevalence should be much greater than incidence. This, in turn, indicates that the

*Practical necessity forced us to make the interval shorter in some sub-cohorts in Utrecht and Zeist (see Tables 10 and 11).

TABLE 10. PICK-UP RATES AFTER VARYING TIME INTERVALS AT 2ND SCREENING, OTHER AREAS

	Number of women examined	Number of breast cancers	Pick-up rate (%)
13-15 months	2428	7	2.9
16-19 months	4646	17	3.7

TABLE 11. PICK-UP RATES AFTER VARYING TIME INTERVALS AT 4TH SCREENING, CITY OF UTRECHT

	Number of women examined	Number of breast cancers	Pick-up rate (%)
13-15 months	3555	3	0.8
24-26 months	5062	14	2.8

growth rate of breast cancer is not particularly rapid and that earlier detection may result in improved prognosis.

The pick-up rates at the second, third and fourth screening cycles for the citizens of Utrecht, viz 1.3‰, 1.5‰ and 2.0‰, suggest that there is a relation between the duration of the interval between screens and the yield in terms of cancers found. Such a relationship would be far from illogical. However, this trend could also be explained by the gradual change in the decision threshold (see Section III.3). This is supported by the relatively high pick-up rate (3.4‰) in the second screening cycle for the areas outside Utrecht after an interval of no more than 18 months. This could be attributable to a slightly lower sensitivity in the initial screening compared with that in the first examination in the city of Utrecht.

We have mentioned already some minor deviations from the study protocol with regard to the duration of the intervals between screenings. Table 10 shows the results for the second screening in the areas outside Utrecht whereas Table 11 gives the results after the fourth examination in Utrecht itself. It is quite natural that pick-up rates are higher after a longer interval. Eddy [14] has shown on theoretical grounds that this relationship is not linear.

It is remarkable that in the second to fourth screening cycles it was possible to maintain high predictive values of a positive test result (expressed as the percentage of recommended biopsies which yielded evidence of cancer). From Fig. 7 it can be seen that this percentage reached 50 (67/133), a figure in between the 40 and 56% found at initial screening. On theoretical grounds (Bayes' theorem) one might have expected a lower predictive value since the number of preclinical cancers had been reduced by the initial screening. It would seem that the opportunity, in the later screening cycles, to compare xeromammograms with earlier pictures from the same woman offset the theoretical expectation of a decrease in predictive value.

A study of the decision tree (Fig. 7) of the pooled experience of all screening cycles except the first one, in all areas, reveals that four cases of breast cancer were brought to light after the GP had suggested a biopsy despite a weak signal from us in the shape of a brief note (suggesting a benign lesion). In retrospect, three of the four cases were shown to be breast cancers smaller than 1 cm in diameter with negative axillary nodes, and no mammographic signs of malignancy. In the fourth case we had found a palpable lymph node in the axilla but no further sign suggestive of cancer (Rombach [6] has reported that isolated palpable nodes have no predictive value in screening). This node proved to be cancerous but provided no lead as to the primary tumour. As a preliminary step a diagnosis of probable breast cancer was made.

A review of the total screening process during the years 1974-1980 shows that:

(a) Biopsy was performed on our recommendation in 462 cases, resulting in a diagnosis of breast cancer in 210, a predictive value of 45.5%.

(b) In 493 women who followed our advice for a check after 6 months, a further 20 cases of mammary cancer were diagnosed, yielding a predictive value of 4.1%.

TABLE 12. SCREENING MODALITY AND AXILLARY NODE STATUS (NO DATA ON AXILLA IN 14 CASES)

	Axillary nodes		Total
	Negative	Positive	
Clinical signs with or without mammographic evidence	49 (56.3%)	38 (43.7%)	87
Mammographic evidence, no clinical signs	89 (81.7%)	20 (18.3%)	109
Total	138	58	196

(c) In 566 cases a brief note was sent to the GP, but there was no suggestion that biopsy should be performed. This procedure was, however, carried out in 119 cases and four cases of breast cancer were found, a predictive value of 0.7% (4/566).

Cancers not detected by the DOM project will be discussed in Section IV.3(d) under interval cancers.

5. Stage at which breast cancer was detected

The histopathology of the axillary nodes was used to classify the cancers found at screening in stages. Axillary metastases should be considered a sign of metastasis in general and not one limited to these nodes [15].

Our analysis was confined to the 210 women in whom breast cancer was detected straight after screening. In comparing clinical examination and (xero)mammography important differences came to light. Those cases detected by mammography alone had a significantly ($p < 0.01$) lower percentage of cancerous axillary nodes than those in which there were clinical signs—most often a lump in the breast (Table 12). The figure of 82% of cancers with negative nodes may even be a low estimate since those women in whom the axilla was not surgically explored reportedly had small cancers; other cases involved intraductal cancer in which the surgeon considered axillary dissection an unnecessary procedure. From the point of view of the number of cancers detected and the stage of the cancer mammography is definitely superior to clinical (physical) examination. Nineteen cases of cancer were detected by physical examination alone: in the majority a lump was felt at palpation; the remainder were detected through skin dimpling or nipple abnormality.

6. Steps taken as a result of the DOM project

(a) *Ratio of malignant to benign abnormalities.* From time to time biopsies were performed in clinical practice in the intervals between screening cycles and after a lack of response in the second to fourth cycles. Table 13 shows the results of an investigation to find out whether there were relatively more benign lesions detected during screening.

In contrast to Table 8, the present table includes cases diagnosed after a brief note. It will be seen that the relative frequency of negative biopsies at screening (sometimes called “unnecessary biopsies”, but this is a retrospective judgement) did not differ appreciably from that in the intervals. Elimination of lobular carcinoma *in situ* (or even classification of them in the non-cancerous group) does not change the picture very much. Thus screening did not lead to an unduly large number of biopsies.

TABLE 13. COMPARISON OF MALIGNANT/BENIGN (MAL/BEN) RATIO BETWEEN ABNORMALITIES FOUND AT SCREENING AND DURING OR AFTER TIME INTERVALS RESP.

Diagnosis of cancer in:	Utrecht		Other areas		Total	
	Mal	Ben	Mal	Ben	Mal	Ben
1st cycle	116 (29%)	281 (71%)	58 (45%)	72 (55%)	174 (33%)	353 (67%)
2nd–4th cycles	47 (32%)	99 (68%)	24 (56%)	19 (44%)	71 (38%)	118 (62%)
Intervals and non-compliance group	59 (34%)	116 (66%)	17 (55%)	14 (45%)	76 (37%)	130 (63%)

TABLE 14. SURGICAL TREATMENT OF BREAST CANCER IN WOMEN FROM CITY OF UTRECHT (LOBULAR CANCER *IN SITU* INCLUDED)

Diagnosis of cancer	Radical mastectomy		Simple mastectomy		Minor surgery		Total
1st screening	89	126 (78%)	19	25 (15%)	8	12 (7%)	116
2nd-4th screening	37		6		4		47
Intervals		22 (65%)		5 (15%)		7 (20%)	34
"Historic control group"	41		17		5		63
Detected before invitation 1st screening	40	81 (66%)	13	30 (25%)	6	11 (9%)	59
Non-responders	31		14		13		58
Non-compliance after 1st screening	18	49 (59%)	2	16 (19%)	5	18 (22%)	25
Total		278 (69%)		76 (19%)		48 (12%)	402

(b) *Surgical treatment for cancer.* As stated in Section I, the DOM project was carried out just before so-called breast-saving treatment was introduced in the Netherlands. From a scientific point of view this may be regarded as a bonus, since the effect of changes in diagnosis could not be confounded by simultaneous changes in therapy. The only evident trend was the fairly rapid introduction of Patey's modification of mastectomy to replace Halsted's operation.

Analysis of the data is limited to the city of Utrecht, and Table 14 shows a comparison of four groups:

- cancer detected through screening;
- interval cancers;
- historic control group [cancers detected before screening was introduced, see Section IV. 3(a)];
- non-responders.

Significant differences are noted ($p < 0.005$) which can undoubtedly be explained by the different stage distribution of the cancers. In 48 women the operation involved less than breast removal, a useful practice in either very small cancers or in those at an advanced stage of development. Late stage cancer was found more often in cases not detected at screening.

(A breakdown of stages on the basis of tumour diameter and axillary node status can be found in Tables 22–25 and Figs 9 and 10 in Section IV.)

Six cases of lobular carcinoma *in situ* were detected. A radical mastectomy was performed in two cases, a simple mastectomy in one and a smaller operation in three.

(c) *Surgical procedures in non-malignant disease.* Screening produced a few therapeutic dilemmas concerning lesions which led to the surgical removal of one or both breasts despite being deemed non-malignant by the pathologist. There were four cases of bilateral disease, three of proliferative lesions and one case of recurrent palpable tumours. In six cases of unilateral mastectomy the other breast had already been removed because of malignancy. These operations were carried out after extensive consultation with the patients concerned.

(d) *Interval between screening advice and operation.* For women in whom breast cancer is suspected following screening it is of great importance that the time between screening and surgical exploration be kept short and that care be well organised. A quantitative measure of this was the interval between the centre's message to the GP (on the day of screening) and the day of biopsy and/or operation.

Figure 8 depicts a cumulative percentage distribution curve of this interval for 348 patients biopsied or operated upon in the hospitals of Utrecht. It will be observed that 68% of surgical interventions took place within 3 weeks and 95% within 6 weeks. The GPs almost invariably contacted the surgeon within one week.

(e) In the previous sections we have given the number of recommendations or procedures per screening cycle or per interval, irrespective of whether these figures refer to different women or not. The following data regarding women resident in Utrecht have been linked to individuals.

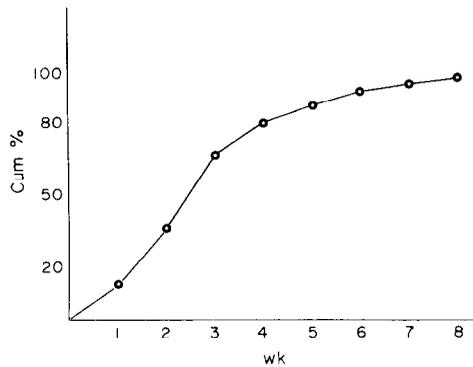


FIG. 8. Cumulative percentage distribution of time interval between screening and surgical treatment.

A total of 1198 signals were sent to GPs:

- 354 recommendations for biopsy;
- 374 recommendations for a check after 6 months;
- 470 brief notes.

While it might therefore appear that messages were sent in regard to just over 8% of the total population of 14,697 women, the actual figure was only 7%. One in eight or 12.5% of the above signals was a second message about the same woman which was sent at another screening cycle. (This is an oversimplification since not every woman responded to all screening invitations.) In the 12.5% of women for whom two or more signals were issued, breast cancer was not detected more often than in those for whom one signal was sent.

The predictive value of a recommendation for biopsy did not rise if there had been another signal in a previous screening cycle. An important finding in this respect with regard to interval cancers will be dealt with in Section IV.3(d).

Although no signal was ever sent for 93% of the women, some of these underwent a biopsy during a screening interval or after failure to respond to one of our invitations (see Table 13, bottom line). Just as some women (12.5%) rated more than one message, so others (20%) underwent more than one operation, having been submitted to surgery in a previous screening cycle.

IV. THE DOM PROJECT IN AN EPIDEMIOLOGICAL PERSPECTIVE

1. *The necessity for a cancer registry and for mortality statistics*

The DOM project is an epidemiological study in the sense that it describes a specific disease risk in a population. It is also an intervention study, but one which lacks randomisation to an intervention and a control group. In its design we acknowledged that the H.I.P. study in New York, which was randomised, had already shown a favourable effect of population screening in women aged 50–64 years.

From an epidemiological viewpoint a non-randomised design obliges the investigators to provide evidence of benefit by indirect means.

For descriptive purposes it is important that a total picture is obtained by registering not only the cases detected by screening but also cases diagnosed in the intervals between screening cycles and cases in women who did not respond to an invitation for screening.

Scientifically speaking, these various groups might have undergone some form of selection bias, and, by ignoring this, incorrect conclusions could be drawn from the data.

The reader will at once realise that failure to respond to an invitation for screening might be associated with illness or negative health-care behaviour; these factors might themselves influence the prognosis of breast cancer. Thus, it seemed appropriate to follow up all women from the cohort born between 1911 and 1925, irrespective of their degree of participation in the project.

Interval cancers might also form a selective group. On theoretical grounds it is possible to show that slow-growing tumours have a greater probability of being detected by periodic screening than fast-growing ones (the so-called length bias). Some sceptics maintain that screening merely picks up the less malignant ones, which means that prognosis would be determined long beforehand.

To be able to answer this kind of criticism, it is necessary to document all cases of breast cancer in the study cohort. Moreover, it would be as well to have a record of all cases of mammary cancer in the years before the project started.

A further problem in evaluating a non-randomised study is that of lead time, viz. the advancement of diagnosis by screening on the time axis. Since its magnitude (and its distribution) is unknown it is virtually impossible to adjust survival curves of screened populations in comparison with unscreened populations for it. Thus mortality rates rather than case-fatality rates should be the ultimate means of evaluating whether screening does any good. If it does, comparisons between breast cancer mortality in the cities of Utrecht and Nijmegen (where screening is being conducted) vs other areas of the Netherlands will in the not too distant future have to show separate tracks.

2. Cooperation with medical practitioners

When the project began there was no cancer registry in the Utrecht region since the system of registration financed by the Queen Wilhelmina Fund (National Cancer Fund) had ended in 1968. The staff of the DOM project therefore suggested that a new registry be set up, and for practical reasons this was limited to the city of Utrecht.

For this system to work it was essential to enlist the cooperation of the surgeons, the radiotherapists and the three groups of pathologists serving the area, cooperation which was swiftly forthcoming. While data collection was begun on special forms, it proved to be easier in practice for the various specialists to send copies of their routine medical reports to the medical staff at the Preventicon. Steps were taken to ensure that all data were handled and stored with due regard to privacy.

All deaths in the cohort of women born between 1911 and 1925 were notified systematically by the municipal authorities to the staff of the DOM project. On request, GPs provided data on causes of death in the study population.

3. Cancer registry data

A synopsis of all breast cancers registered in the population of Utrecht during the years 1974–1980 is given in Table 15. For our evaluation the cohort born between 1911 and 1925 is of special importance. The Table shows a fairly constant level of new breast cancers in those born before 1911 or after 1925. However, in the cohort born between these years there was a peak during 1976, followed by a marked decline after mid-1977, when the initial screening cycle had been completed.

TABLE 15. BREAST CANCER REGISTRY, CITY OF UTRECHT (1974–1980)

	Cohort born 1911–1925	Other years of birth	Total
1974	56	No registration	(56)
1975	66	83	149
1976	94	79	173
1977	58	87	145
1978	24	84	108
1979	34	85	119
1980	43	75	118

TABLE 16. CLASSIFICATION OF BREAST CANCERS DIAGNOSED IN THE COHORT BORN 1911-1925

	Utrecht	Other areas
Detected through screening	155*	82
Not diagnosed through screening		
(a) Historic control group (1974)	56	Not included in cancer registry
(b) Diagnosed before screening invitation	55	
(c) Cancer after non-response	81	
(d) Interval cancer	30	

*2 patients operated on in 1981.

In the first part of this section we mentioned various categories of registered patients whose breast cancer was not diagnosed as a result of screening. Four groups could be distinguished and these are listed in Table 16.

Group (a): Historic control group. Women born between 1909* and 1925 in whom breast cancer was diagnosed before the start of the DOM project in 1974, were classed as a "historic control group" and showed many similarities with group b (mentioned below). The incidence in 1974 was 2.5 per 1000 per yr, compared with a figure of about 1.8 per 1000 per yr in most of the rest of Western Europe. The high incidence in Utrecht reflects, *inter alia*, the thoroughness with which cancer was registered.

Group (b): Cases diagnosed after the start of the DOM project but before invitation for initial screening. For the purposes of evaluation it would have been ideal if all women could have received their initial examination within a very short time. However, this was not to be. The first screening cycle for the 14,697 women from the city of Utrecht took $2\frac{1}{2}$ years to complete, because we could only handle a total of about 10,000 examinations a year (in 1975 we managed only 5300) and we had to adhere as strictly as possible to the fixed intervals between cycles (one year between the first and second screening). Thus, some women had their second examination before others had their first.

As far as we can see group b is an unselected group of breast cancer patients and therefore as useful for purposes of comparison as group a.

Group (c): Mammary carcinoma diagnosed after failure to respond to screening invitation. As explained in Sections I and II a woman who showed no interest in the first screening cycle was not invited to subsequent cycles. Moreover, there were women who failed to comply in the second or third cycle. Women in whom breast cancer was diagnosed after failure to respond to a screening invitation were classified as a special group to be distinguished from those in group d.

Group (d): Interval cancers. These are defined as breast cancers diagnosed within the designed intervals between screening examinations but not as a result of a recommendation for biopsy or a check at 6 months. Also included in this group are cancers diagnosed at least 6 months after the issue of a brief note (as defined in Section III).

The probability of having an interval cancer depends, *inter alia*, on the length of the interval between screening examinations. In Table 17 a distinction is made between interval cancers diagnosed within 12 months after a screening and those diagnosed thereafter.

TABLE 17. INTERVAL CANCERS ACCORDING TO DURATION OF INTERVAL

	Time since last screening	
	≤ 1 yr	> 1 yr
City of Utrecht		
After 1st cycle	3	*
After 2nd cycle	7	2
After 3rd cycle	5	12
After 4th cycle	1	0**
Other areas		
After 1st cycle	6	3
After 2nd cycle	2	4**

*Impossible by design. **Period of observation limited to end of 1980.

*The years 1909 and 1910 are included to have comparable groups according to age.

TABLE 18. RELATION BETWEEN PRIOR BRIEF NOTE FROM PREVENTICON TO GP, AND REASON FOR BIOPSY LEADING TO DIAGNOSIS OF INTERVAL CANCER, CITY OF UTRECHT

	Lump felt	Nipple abnormality	Surveillance by GP	Other	Total
No brief note	10	5	—	2**	17
Prior brief note	6	2	2	—	10
			1*		1*
Recommendation for biopsy	—	1*	1*	—	2*
Total	16	8	4	2	30

*Brief note mentioned abnormality in breast contralateral to the one in which an interval cancer emerged. **One woman consulted her doctor about an axillary swelling, the other consulted a hospital service elsewhere with unspecified complaints.

Interval cancers may be seen as a kind of yardstick for determining the diagnostic boundaries of screening. In the first place the Head of Diagnosis must consider whether he “missed” the case in the previous examination. While man remains fallible, the system is not without its limitations as some cancers, even those with a diameter or more than 1 cm, can be invisible on the mammogram.

In his doctoral thesis Rombach [6] reported how he compared his diagnostic skills with those of a number of internationally renowned mammographers. He concluded that differences between members of this review panel in “blindly” reading his interval cases mainly reflected differences in diagnostic threshold levels, and that his own threshold was not very different from that of most of the panelists, as judged on the basis of the predictive value of their recommendations for a biopsy.

From a statistical point of view the number of interval cancers is of value in computing the sensitivity of the screening procedure. Many authors arbitrarily consider an interval cancer diagnosed within 12 months after a negative screen as “missed”. Based on this criterion the sensitivity of screening in the DOM project was 94%. The few cancers diagnosed after action by the GP following a brief note from our staff have already been mentioned in paragraph 4 of Section III.

Interval cancers are detected in most cases by the woman herself. Of the 29 women with interval cancer in the city of Utrecht (30 cancers), 16 had felt a lump in the breast and eight had noticed some nipple abnormality. An investigation into the screening history of these women showed that 10 of the 30 cancers occurred in a breast about which a brief note had ever been sent (Table 18). Table 18 was simplified (and “purified”, see below*) to provide Table 19, which documents the rate of interval cancers per 1000 women. It can be seen that the sending of a brief note after one or more previous examinations was associated with a more than 20-fold increased risk of interval cancer ($25.9:1.2 = 21.6$). These brief notes mentioned in all but one case abnormalities in the same breast where the cancer was found later.

This important finding drove us to consider whether we had been aggressive enough in recommending biopsy. What would have been the consequences of recommending a biopsy instead of sending a brief note?

Up to and including the third screening cycle we sent out a total of 434 brief notes. If all these had recommended biopsy, and this procedure had been performed, the situation

TABLE 19. RELATION BETWEEN RISK OF INTERVAL CANCER AND PREVIOUS BRIEF NOTE FROM PREVENTICON, CITY OF UTRECHT (3RD CYCLE INCLUDED)

	Number of women	Interval cancer	Rate of occurrence (‰)
No brief note	13,630	17	1.2
Prior brief note	347	9	25.9

*Since the interval after the fourth screening had not been completed at the time of the present report, the numbers in Table 19 are somewhat restricted compared with those of Table 18.

TABLE 20. ASSUMED CHANGE IN PREDICTIVE VALUE OF RECOMMENDATION FOR BIOPSY IF ALL BRIEF NOTES HAD RECOMMENDED THIS PROCEDURE

	Recommendations	Breast cancers	Predictive value
Recommendations for biopsy (carried out)	329	134	40.7%
Assumed extra recommendations for biopsy	434	11*	2.5%
	763	145	19.0%

*Nine true interval cancers plus two cancers found at biopsy after a brief note.

would have been as shown in Table 20. A study of the Table shows that 11 extra cancers could have been detected at the expense of a large number (434) of biopsies, and this would have meant a decrease in the predictive value of a recommendation for biopsy from 40.7 to 19.0%.

The decision to limit radiological screening after the first cycle to the medio-lateral projection might cause one to consider whether an interval cancer occurring after the second examination could have been missed because of the lack of a cranio-caudal picture. While nothing constructive can be said about this on an individual basis, pictures in the cranio-caudal projection were taken in 10% of cases (five of 50 breasts; 25 women) presenting with interval cancer after the second or third screening. Since, overall, cranio-caudal pictures were taken in only 3% of the women during the second and third cycles, one cannot conclude that the policy of reducing the number of such pictures was linked with the late detection of these cancers. On the contrary, it seems as though women who went on to develop interval cancers got extra attention. It is probable that there are risk factors for interval cancer which are associated with the decision to make pictures in a second projection. (For a further study of risk factors see Section V.)

Three cases occurring between cycles were classified as cancers detected at screening rather than interval cancers: they were recognized by the staff of the project during examination, but the woman or her doctor deliberately took no action to pursue the suggestion that a biopsy be performed.

Table 21 gives an overview of the prevalence and incidence of breast cancer in the city of Utrecht and in the other areas. In the calculation of the "pick-up" rates the population

TABLE 21. BREAST CANCER PREVALENCE AND INCIDENCE IN THE DOM PROJECT (1974-1980)

Diagnosis of cancer	Women or women-years*	Number of breast cancers	Pick-up rate or annual incidence per 1000
Utrecht			
1974, historic control group	21,700**	55**	2.5
Before invitation			
1st screening	26,200	55	2.1
1st screening cycle	14,697	108	7.3
2nd screening cycle	11,867	15	1.3
3rd screening cycle	10,208	15	1.5
4th screening cycle	8617	17	2.0
Interval between 1-2	15,200	3	0.2
Interval between 2-3	18,000	9	0.5
Interval between 3-4	16,900	17	1.0
Interval between 4-5	4400	1	0.2
Non-response 1st screening	30,000	56	1.9
Non-compliance			
at 2nd cycle	9600	20	2.1
at 3rd cycle	3100	2	0.7
at 4th cycle	850	3***	3.5
Other areas****			
1st cycle	8814	58	6.6
2nd cycle	7074	24	3.4
Interval between 1-2	12,400	9	0.7
Interval between 2-3	?	6	?

*Rounded off. **Women 50-64 years of age. ***2 women with 3 breast cancers (1 bilateral). ****Preliminary data.

TABLE 22. TUMOUR DIAMETER ACCORDING TO MANNER OR TIME OF DIAGNOSIS

	Microscopically small	< 1.0 cm	1.0-2.9 cm	≥ 3.0 cm	Not measured (inoperable)	Unknown	Total*
1974	5	4	27	19	2	2	59
After 1974, but before 1st invitation	—	4	32	15	2	2	55
1st cycle	10	32	60	3	—	—	105**
2nd cycle	1	6	7	1	—	—	15
3rd cycle	1	5	7	2	—	—	15
4th cycle	5	3	8	1	—	—	17
Non-response							
1st screening	1	3	33	11	6	2	56
Non-compliance							
2nd, 3rd or 4th cycle	—	2	14	7	2	—	25
Interval	3	6	18	3	—	—	30
Total	26	65	206	62	12	6	377

*Number of cancers, not women. **Three patients excluded since operation was deliberately postponed for a long time.

at risk posed no problem. Deaths and migration to places outside the Utrecht area necessitated calculations on the basis of a changing population at risk, and this laborious work was accomplished with the help of the staff of the local authority population registers.

Interpretation of the figures in Table 21 is only possible in a wider context (see paragraph 6 on this section).

4. Prognostic parameters. Tumour size

Tumour size and the occurrence of metastases in the axillary lymph nodes have a significant bearing on the prognosis of breast cancer.

Size, which is usually expressed as the largest diameter of the tumour, may be estimated by (a) palpation; (b) measuring tumour "shadows" on the mammograms; (c) measuring the tumour after its surgical removal. The last two methods yield highly comparable results [6], whereas palpation is associated with exaggerated estimates.

Data supplied by the pathologists (method c) were used for our evaluation of tumour size and are presented in Tables 22 and 23. Table 22 gives a detailed account while Table 23 is a somewhat more condensed version which can be interpreted more easily (we have removed six cancers of unknown diameter).

From these Tables it can be seen that the distribution of tumour size in non-responders was very similar to that in the historic control group. Cancers found through screening were significantly smaller.

Interestingly, the distribution of tumour diameters in patients with interval cancer was also favourable when compared with that in the historic controls. We have already made mention of the suggestion that screening selects tumours with relatively low growth rates and leaves the more malignant ones untouched; but if this were true the sum of cancers detected through screening plus interval cancers would not be very different with regard to size from those diagnosed before screening was introduced. Our findings show a significant difference ($p < 0.0001$) and lead us to the conclusion that this supposition does not hold. On average the interval cancers were smaller than those of the historic control

TABLE 23. TUMOUR DIAMETER. CITY OF UTRECHT

	< 1.0 cm	1.0-2.9 cm	≥ 3.0 cm*	Total
Cancer before screening	13 (11.8%)	59 (53.6%)	38 (34.6%)	110 (100%)
Non-response to invitation for screening; cancer found afterwards	6 (7.6%)	47 (59.5%)	26 (32.9%)	79 (100%)
Cancer detected through screening	63 (41.5%)	82 (53.9%)	7 (4.6%)	152 (100%)
Interval cancer	9 (30.0%)	18 (60.0%)	3 (10.0%)	30 (100%)

*Inoperable cancers included.

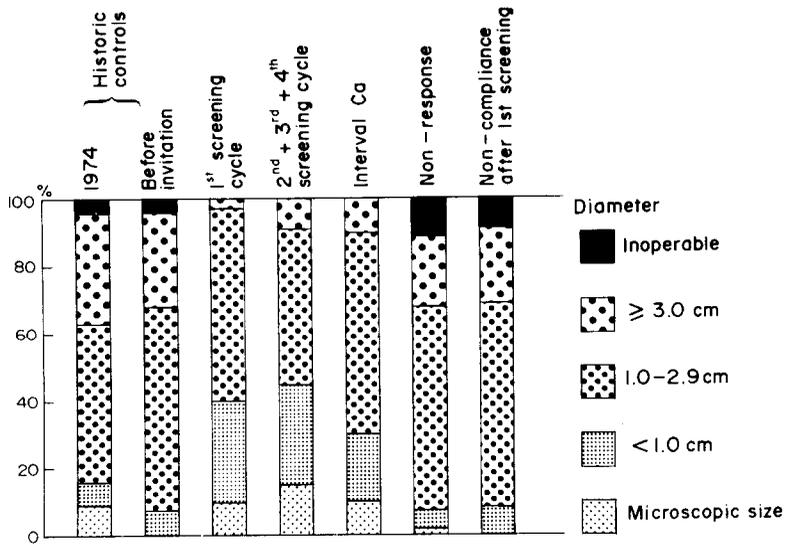


FIG. 9. Percentage distribution of breast cancer diameter in various classes according to manner or time of diagnosis.

TABLE 24. TUMOUR DIAMETER, OTHER AREAS

	< 1.0 cm	1.0- 2.9 cm	≥ 3.0 cm	Total
Cancer found through screening				
1st cycle	13 (23.3%)	38 (67.9%)	5 (8.9%)	56 (100%)
Same, 2nd cycle	7 (30.4%)	12 (52.2%)	4 (17.4%)	23 (100%)
Interval cancers	1 (7.7%)	10 (76.9%)	2 (15.4%)	13 (100%)

group, a finding which might be explained on the grounds that participation in the DOM project made these women more alert to the possibility of detecting breast cancer early. In Fig. 9 the differences in size are displayed graphically.

Table 24 provides similar information for the other areas. Five cancers of unknown diameter could not be included. Once again the distribution of tumour diameters in patients with interval cancer does not appear to be less favourable when compared with that in women with cancer detected through screening.

TABLE 25. AXILLARY NODE STATUS ACCORDING TO MANNER OR TIME OF DIAGNOSIS, DOM PROJECT

	Axillary nodes negative	Axillary nodes positive	Distant metastases* (incl. inoperable cases)	Axillary nodes not investigated	Total
1974	24	26	3	6	59
After 1974, but before 1st invitation	26	25	3	1	55
1st cycle	69	30	—	6	105**
2nd cycle	9	5	—	1	15
3rd cycle	10	3	—	2	15
4th cycle	13	2	1	1	17
Non-response 1st screening	22	22	8	4	56
Non-compliance 2nd, 3rd or 4th cycle	11	10	3	1	25
Interval	10	16	—	4	30
Total	194	139	18	26	377

*Irrespective of axillary node status. **3 patients excluded since operation was deliberately postponed for a long time.

TABLE 26. AXILLARY NODE STATUS ("POSITIVE" = CANCER PRESENT). CITY OF UTRECHT

	Axilla negative	Axilla positive	Total
Cancer before screening	50 (46.7%)	57 (53.3%)	107 (100%)
Cancer after non-response or non-compliance	33 (43.4%)	43 (56.6%)	76 (100%)
Cancer detected through screening	101 (71.1%)	41 (28.9%)	142 (100%)
Interval cancer	10 (38.5%)	16 (61.5%)	26 (100%)

5. Axillary node status

The same categories of breast cancer mentioned in the previous section were compared with regard to the presence or absence of metastatic deposits in the homolateral axillary lymph nodes (Table 25).

If distant metastases were present they got preference over axillary node status in the tabulation. In 26 of 377 cases (6.9%) no axillary dissection or biopsy was performed, so that no histopathological information was available. Since reasons for not exploring the axilla can be associated with very different stages of the disease, this does not necessarily produce a selective bias in either a favourable or an unfavourable direction.

In Table 36 the information has been simplified so as to enable interpretation. Four categories of cancer patient are compared. Cases with distant metastases are included with the positives; cases in which the status was unknown are excluded.

Patients with interval cancer had a slightly poorer axillary node status than the controls. However, if interval cancers are added to those detected through screening their joint distribution with regard to axillary node status is significantly more favourable than that of the control group ($p < 0.01$; see also Fig. 10).

Table 27 provides information on axillary node status in the other areas. Eight cases

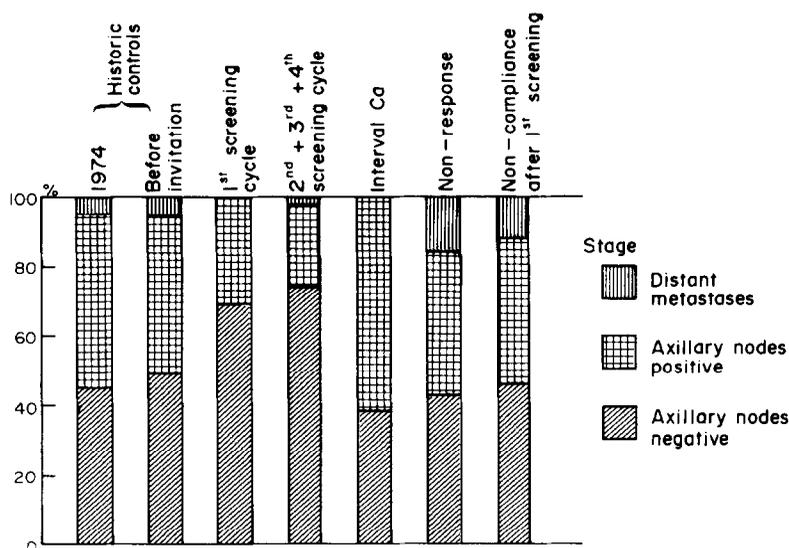


FIG. 10. Percentage distribution of metastatic breast cancer, DOM project.

TABLE 27. AXILLARY NODE STATUS, OTHER AREAS

	Axilla negative	Axilla positive	Total
Cancer detected through screening	36 (69.2%)	16 (30.8%)	52 (100%)
Same, 2nd cycle	19 (79.2%)	5 (20.8%)	24 (100%)
Interval cancer	8 (61.5%)	5 (38.5%)	13 (100%)

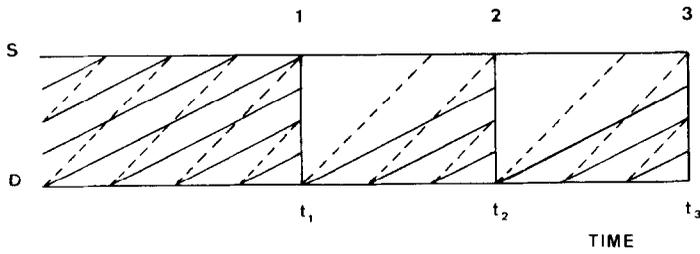


FIG. 11. Model of screening. Vertical axis represents growth stage, horizontal axis time. Preclinical detectable phase is between levels D and S. Cancers enter this phase by crossing level D, and leave by crossing S (= symptoms appearing). Solid diagonal line is slow-growing type of cancer, broken line is faster growing type. Screening cycles take place at t_1 , t_2 and t_3 . Incidence at level D is equal to incidence at level S before intervention.

of unknown status have been excluded. It will be seen that the proportion of negative node cases among the interval cancers is more favourable than in the city of Utrecht.

6. *Natural history of breast cancer*

In this section we present models (Figs 11 and 12) with the purpose of discussing some aspects of the natural history of breast cancer in women aged 50 years and over. The figures presented in Table 21 on prevalence and incidence provide the factual basis for this discussion.

Figure 11 shows two horizontal lines depicting a *level D* in the natural history of breast cancer at which detection by special means (e.g. mammography) is possible, and a *level S* at which symptoms cause the patient to seek medical help.

The time axis runs from left to right. In the vertical axis one should envisage a logarithmic scale with a continuum of development stages of breast cancer, beginning from a base lower than level D. If one considers tumour growth as an exponential process (1 cell–2 cells–4–8–16 etc.), growth occurs along straight lines. Slow-growing tumours need a considerable time to go from D to S whereas fast-growing ones will do this in a relatively short period. In the literature on screening theory this is known as the sojourn time in the preclinical detectable phase.

In Fig. 11 a population is screened at three points in time: t_1 , t_2 and t_3 , the intervals between these points being t_2-t_1 and t_3-t_2 respectively. Inspection of Fig. 11 makes it clear that slow-growing tumours have a greater probability of being picked up at screening than faster growing ones (inversely proportional to the slope of the growth line).

After the first screening (and assuming a sensitivity of 100%), the number of interval cancers will depend on the incidence of “new” cancers crossing level D first (and thereafter level S) and the ratio between the rate of tumour growth and the length of the time interval.

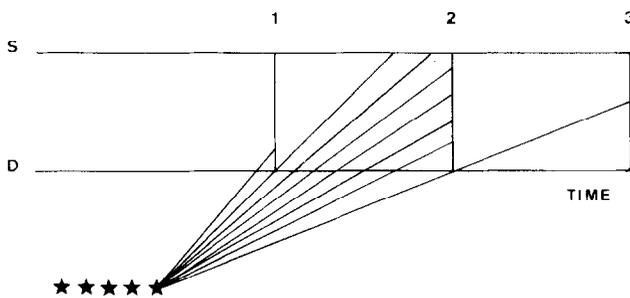


FIG. 12. Alternative model of screening. Stars represent beginning of cancerous process. It is assumed that menopause occurs just to the right of right star, and that new cancers are not formed after menopause. Notice spectrum of growth rates. At older ages cancer tends to be of slower growth type.

From this model it follows that the average sojourn time of a tumour in the preclinical detectable phase (i.e. between levels D and S) equals P/I where P = the prevalence rate of preclinical tumours—estimated as the cancer pick-up rate in the initial screening; and I = the incidence rate—estimated from cancer registry data. It should be borne in mind that this incidence estimate refers to level S. In our figure the estimate at level S equals that at level D, an essential feature of the “stable disease model” of Zelen and Feinleib [16].

If all tumours grew with equal speed a “full recovery” of the original situation would be attained after a period equal to P/I years; in other words, the space between levels D and S would again be “saturated” with breast cancer.

In the DOM project the prevalence rate of preclinical breast cancer is about 7.3 per 1000* women (assuming a 100% sensitivity at the first screening). For incidence rates we have two estimates: 2.1 and 2.5 per 1000 women-years respectively. Thus the average sojourn time of breast cancer in the preclinical detectable phase in our population is 7.3/2.3, or 3 years and 2 months. However, in real life all cancers do not grow with equal speed, and we have only a rough estimate of the frequency distribution of their growth rates. On the basis of mammographic studies Kusama *et al.* [17] suggested that this distribution is lognormal, which means that sojourn times cannot be distributed symmetrically around the mean of 3 years 2 months.

Statisticians have attempted to gain insight into the distribution of sojourn times empirically, and until recently the H.I.P. study in New York, which was conducted in the 1960s, was their only source of information. Using mathematical models, Zelen and Feinleib [16] and Walter and Day [18] found the best approximation to be a negative exponential distribution. The relevance of this to the DOM project population can best be illustrated by preliminary computations carried out by Day (personal communication) with the Walter–Day model. It transpires that had we wished to postpone a fifth screening cycle in the Utrecht women until the original situation had been re-established (i.e. the pick-up rate was equal to that of the first screening), we should have to have waited until 1990†.

It will now be clear that in trying to determine an optimum interval between screening cycles, it is of considerable importance to obtain more information about the incidence rate of breast cancer after a lengthy interval and about the prevalence rate at the end of such an interval. We shall be trying to predict the “pick-up rate” in a fifth cycle (mid-1983 to the end of 1984) to be carried out exactly 4 years after the fourth screening of the Utrecht women born between 1911 and 1925. This will be done in close cooperation with Dr N. E. Day of the International Agency for Research on Cancer, Lyon.

In Section I we mentioned that the H.I.P. project in New York had demonstrated that screening was associated with a favourable effect in women aged 50 and over. There appeared to be no such benefit in younger women. Most commentators have explained this finding by suggesting that mammograms of the premenopausal breast are more difficult to read than those taken in women of postmenopausal age. While this might seem a reasonable explanation—at least in respect of the situation in the 1960s—other possibilities should also be considered. It could be that there are differences between the pre- and postmenopausal state with regard to the cancer becoming clinically manifest.

A variety of epidemiological considerations outside the scope of the DOM project could be pieced together with the findings on screening discussed above to form a model of the natural history of breast cancer (Fig. 12).

A condition of the “stable disease model” of Zelen and Feinleib [16] (Fig. 11) is the continuing risk of new cancers being formed. One could, however, also envisage a model in which the genesis of new carcinomas in the breast becomes rare after the age of 50. This

*A slight deviation from the rate given in Section III is due to the exclusion of a case of recurrent cancer detected at screening.

†Aging of the population has been deliberately neglected in these preliminary computations. This factor might further decelerate growth rates.

would imply that all cancers diagnosed (with or without screening) postmenopausally are a consequence of processes initiated long before. In such a model (Fig. 12) the cancers successively crossing level D show a trend towards progressively slower growth rates. If this were true, it could justify the progressive spacing out of further screening cycles, that is, increasing the interval between screenings as the rank order of these became higher and higher. This would be a radical departure from the present advice to screen at fixed intervals of 1 or 2 years. It would have important consequences in terms of such issues as cost and radiation exposure.

V. RISK FACTORS FOR BREAST CANCER

1. Introduction

The DOM project enabled us to estimate the effect of a number of risk factors* mentioned in the literature. Data on these factors were collected at the initial screening examination by means of a questionnaire or by direct measurement. Although risk factor status was updated at each screening cycle, the present estimates were based on data from the first examination. Information on the city of Utrecht and the other areas was pooled.

Although the study can be considered a cohort study in which each individual experiences risk for a variable period of time, from the day of entry (at the first screening examination) until the end of 1980, the calculations are not based on the concept of person-years at risk, since we cannot apply incidence rates to the number of person-years. The reason for this is that the intervention of screening has an effect on the natural history of breast cancer, the "unnatural" peak of occurrence at the first screening being the result.

We have therefore limited the computations to a direct comparison between cases and non-cases, as if we had analysed a case-control study.

The reference group for a dichotomous risk factor was always the group in the study population without the risk factor under investigation. If a risk factor was polytomous (more than two classes being distinguished) we usually, but not invariably, chose the class with the lowest risk as the reference. Relative risk was determined as the odds ratio ab/dc in the following scheme:

		Breast cancer	
		present	absent
Risk factor	present	a	b
	absent	c	d

In determining the significance of any association with dichotomous factors 90% confidence intervals were calculated. With polytomous factors a test on linear trends was performed (according to Rothman and Boice [19]).

The total number of breast cancer cases included in the analysis was 296, viz. 198 from the city of Utrecht and 98 from the other areas. In this context cases mean patients with breast cancer and not breast cancers themselves. We excluded lobular carcinoma *in situ* since this was a more or less accidental finding. Intraductal carcinoma is included with the invasive cancers.

The 296 cancer cases could be divided into four groups on the basis of the way in which they were detected:

- 163 cases detected at the initial screening;
- 69 cases† detected at one of the following screening cycles (2nd, 3rd or 4th in Utrecht; 2nd in other areas;
- 42 cases diagnosed in the intervals between cycles and not on account of screening;

*The concept of risk factors is merely a statistical expression of variables associated with increased risk. Causality is difficult to prove. With a single exception we shall not discuss any evidence regarding biological relationships.

†Excluding two cases diagnosed in 1981, 4th screening, city of Utrecht.

TABLE 28. MARITAL STATUS AND NUMBER OF CHILDREN AS RISK FACTOR

	Mca*	Pop min Mca*	RR	(= relative risk or risk ratio)
No children, unmarried	28	2133	1.00	(reference group)
No children, married**	30	2225	1.03	
1-2 children	97	7594	0.97	
3-4 children	97	7035	1.05	
≥ 5 children	33	4228	0.79	
	296	23,215***		

*Mca = mammary cancer patients; Pop min Mca = population minus mammary cancer patients. **Includes women who were or had been married. ***Total numbers may differ slightly since on a few occasions data were considered unreliable.

22 cases diagnosed after failing to respond to an invitation for screening after the initial examination.

In a few instances where marked differences in risk between these four subgroups were observed, pertinent details have been included in the tables.

The following risk factors were studied:

- marital status, number of children, age at birth of first child;
- personal medical history regarding the breasts: benign and malignant lesions;
- family history of breast cancer in mother and/or sister(s);
- age at menopause and way in which menopause was experienced (natural or artificial);
- body weight and height, Quetelet's index of overweight and body surface area;
- use of drugs, in particular oestrogens, antihypertensive medication, tranquillizers, thyroid preparations;
- hormone profile;
- radiologically measured risk factors (parenchymal patterns).

The analytic approach was to initially study factor after factor in an univariate manner, viz. without control for any confounding variables, and to subsequently perform multivariate analysis of many factors at the same time.

2. Results. Marital status, number of children, age at birth of first child

Marital status and number of children. The distribution is given in Table 28, which divides the nulliparas into those who were or had been married and those who had never married; these subgroups seem to have equal risk. Women with five or more children have a somewhat lower risk, but no linear trend towards decreasing risk with an increasing number of children can be discerned.

Age at birth of first child. The data are presented in Table 29. Taking the nulliparous women as the reference group one can discern a gradual trend in which these women have a risk intermediate between that of those who had their first baby when young, and those who gave birth for the first time after entering the fourth decade. This risk pattern has been described by MacMahon *et al.* [20]. The trend found in Utrecht is just short of formal significance ($p = 0.07$).

TABLE 29. AGE AT FIRST BIRTH AS RISK FACTOR

	Mca	Pop min Mca	RR
No children	58	4362	1.00 (ref.)
≤ 23 years	45	4225	0.80
24-27 years	79	6999	0.85
28-31 years	67	4874	1.03
≥ 32 years	47	2755	1.28
	296	23,215	

TABLE 30. PERSONAL HISTORY OF BREAST ABNORMALITIES LEADING TO BIOPSY (BUT NOT CANCER) AS RISK FACTOR

	Mca	Pop min	Mca	RR*
No mammary disease	274	22,072		1.00
Benign breast disease	22	1143		1.55
	296	23,215		(1.07-2.26)

*90% confidence limits in parentheses.

TABLE 31. PERSONAL HISTORY OF PREVIOUS MASTECTOMY (LIKELY TO BE CANCER) AS RISK FACTOR

	Mca	Pop min	Mca	RR*
No malignant mammary tumour	293	23,067		1.00
Malignant mammary tumour	3	148		1.60
	296	23,215		(0.45-5.36)

*90% confidence limits in parentheses.

3. Personal history regarding the breasts

In the questionnaire which was filled out at each screening examination, a distinction was made between breast abnormalities leading to biopsy alone (with an apparently benign outcome) and lesions leading to radical or less than radical mastectomy; the latter were considered malignant.

The relative risks associated with these abnormalities are shown in Tables 30 and 31 respectively. It will be observed that both types of lesion, i.e. benign and malignant, are associated with a somewhat increased risk of breast cancer. The finding in the latter case fails to attain formal significance in view of the small number of patients. We did not attempt to obtain detailed histopathological information on any breast lesions in the personal history since this would have entailed intensive searching. Studies specifically designed to examine this topic can be found in the literature [21].

With regard to probable breast malignancies in the personal history the calculations were based on the number of persons at risk. If we had made them on the basis of the number of "breasts at risk", we should have to have taken into account the reduction of breast tissue following the previous removal of one breast.

Although the number of women who had undergone mastectomy was rather small, this was quite predictable since it is common for such women to remain under medical supervision for a number of years and therefore to feel no necessity to have another examination by others.

4. Family history of breast cancer

The family history was limited to breast pathology (known or suspected to be cancer) in a woman's mother and sister(s). Our previous experience had shown that this kind of information was generally reliable [22].

A total of 23,511 women were invited to participate in the DOM project. At the time of the first screening 78% had already lost their mother and only 10% of the mothers who remained alive were under 75 years of age. It is therefore clear that the maternal population had almost experienced its full breast cancer risk. In this population of mothers breast cancer was reported 938 times, or in about 4%, which is somewhat lower than the figure of 7% usually quoted for the life-time of breast cancer. It should, however, be borne in mind that these mothers had lived the greater part of their lives in the first half of this century, when the incidence of breast cancer was lower than it is at present (if we base our estimate on mortality figures).

A similar kind of reasoning can be followed with regard to the population of 49,691 sisters, all of whom were at least 30 years old. On average each woman screened (proposita) had 2.1 sisters, although 20% had none. The age distribution of the sisters could be equated

TABLE 32. FAMILY HISTORY OF BREAST CANCER IN MOTHER AS A RISK FACTOR

	Mca in proposita	Pop min Mca	RR*
No such history	272	22,301	1.00
Mother with mammary cancer	24 296	914 23,215	2.16 (1.50-3.09)

*90% confidence limits in parentheses.

TABLE 33. FAMILY HISTORY OF BREAST CANCER IN SISTER(S) AS A RISK FACTOR

	Mca in proposita	Pop min Mca	RR*
No such history**	210	17,517	1.00
Sister(s) with mammary cancer	21 231	1075 18,592	1.63 (1.09-2.43)

*90% confidence limits in parentheses. **Negative family history in those with one or more sisters.

to that of the propositae. In the population of sisters a total of 1182 cases of breast cancer were reported by the propositae*.

The relative risk of breast cancer associated with a positive family history is shown in Tables 32 and 33. These Tables confirm the findings of others (e.g. [23]) of a degree of familial risk, but the statistics do not allow one to ascertain whether the underlying mechanisms are genetic or are determined by a shared environment.

5. Menopause

Two aspects were studied: the age at menopause and the way in which the postmenopausal state was attained—naturally or artificially (iatrogenic).

Age at menopause. The investigation of the age at menopause is beset with some methodological problems. When determined retrospectively, one is dependent on the human memory. Unlike menarche, which is a well-remembered event in the life of a woman, the occurrence of the last menstrual period can only be recognized after an interval of, say, one year. Thus it is engraved less sharply in the mind than the establishment of the menses.

From a statistical point of view, a retrospective inquiry demands that all women should have attained the postmenopausal state. If investigators ignore this requirement, they will encounter women who are still menstruating, and, in discarding these from their statistics, will introduce a bias towards a low estimate of the age at menopause.

A better solution is to employ the "status quo" method which requires an answer as to the present state of the woman, i.e. pre- or postmenopausal. In collecting answers from large populations according to age (with one-year classification), one obtains percentage curves of menstruant women by age which appear similar, but are not equal to cumulative percentage distributions constructed from retrospective studies. The only disadvantage of the status quo method is that the study population should be fairly large if the statistics are not to suffer from random errors in the age-stratified estimates.

The results of using the retrospective approach are depicted in Fig. 13. Two cumulative percentage distributions are presented, one for the total population ($n = 8841$) aged 60 or over at the time of inquiry (initial screening), the other for the women who went on to develop breast cancer ($n = 126$). The postmenopausal state was defined as a 12-month menstruation-free period.

*Among these cases there were 78 of two or more sisters with breast cancer (excluding the proposita) and 64 cases of a mother and (one or more) sisters with breast cancer.

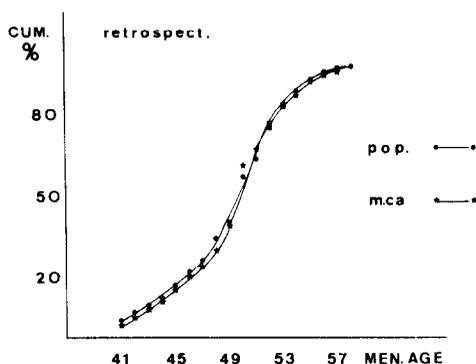


FIG. 13. Cumulative percentage distribution of postmenopausal women according to age in retrospective inquiry among women aged 60 and over, compared with 126 breast cancer patients of similar age.

The course of both curves is almost identical. The estimates of women experiencing menopause at the age of 50 are too high (both for the cases and the total population), a well-known phenomenon due to an inaccurate memory.

The results of the other approach, viz. the status quo method, are different. Figure 14 presents the status quo curve of 14,275 women aged 50–61 years at the time of inquiry (initial screening), who did not develop breast cancer during the study period. Also plotted in the same graph are the percentages of postmenopausal women by age in the 203 breast cancer patients found in the population aged 50–61 years at intake.

Dividing these 203 cases by age (12 one-year classes) naturally creates the problem of random fluctuation of estimates. Nevertheless, 10 or 12 points (stars) are under the curve of the study population and only two are above it. It should be remembered that these 12 points are independent estimates, so that the result of a one-tailed test suggests a significantly later age at menopause in breast cancer patients than in non-patients, an observation in keeping with the literature [24]. This finding should be considered in conjunction with our results on the effect of body weight on breast cancer risk (see paragraph 6 of this section) as obese women have a later menopause than lean women [25].

Way in which menopause was attained. Systematically the women undergoing screening were asked whether the menopause was a natural or an artificial one induced by iatrogenic means. A distinction was made between hysterectomy alone, hysterectomy plus oophorectomy, oophorectomy or X-ray castration alone. A very small proportion of women did not know what had happened.

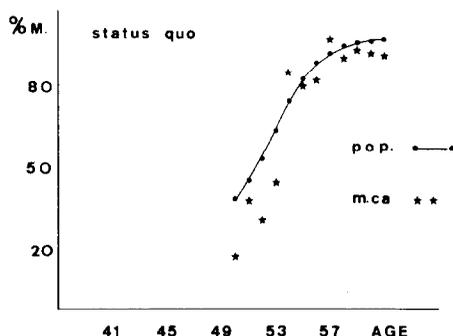


FIG. 14. Percentage of postmenopausal women by age (1-year classes) in population aged 50–61 years. Same procedure with 203 breast cancer patients (plotted as stars).

TABLE 34. NATURE OF MENOPAUSE IN POSTMENOPAUSAL WOMEN AS RISK FACTOR

	Mammary cancer			Total	Pop min	Mca	RR*
	Detected at screening	Interval cancer	Non-compliance				
Natural	165	24	17	206	16,195		1.00
Ovaries removed or irradiated	10	5	3	18	1833		0.77
Hysterectomy (without ovariectomy)	15	3	1	19	1520		(0.52-1.16) 0.98
Unknown	6	3	2	11	756		1.14
	196	35	23	254	20,304		(0.69-1.89)

*90% confidence limits in parentheses.

In Table 34 the results of this inquiry are presented. Among those who were already postmenopausal at intake in the study (initial screening) minor differences are observed between those who were to develop breast cancer in the study period ($n = 254$) and the remainder of the population. Surgical or radiological interference with ovarian function appeared to be less frequent in the breast cancer patients which confirms the protective role of ovarian ablation reported in the literature [26]. Five of 18 such cancers were found in the interval between screening cycles; this is in keeping with the fact that interval cancers infrequently contain oestrogen receptors, a feature typical of the castrated status [27].

6. Body weight and height, Quetelet's index and body surface area

As a large body weight and a tall stature have been described as risk factors for breast cancer, especially in women over 50 years of age (see [28]), the effect of these factors was studied in the DOM project. From weight and height two variables were computed: Quetelet's index (weight in kg divided by the square of height in metres) and body surface area according to Gehan and George's [29] formula in which both weight and height work in the same direction.

The results with regard to body weight are given in Table 35. The effect is minor and without a clear "dose-response" relationship. Further scrutiny reveals two opposite trends: with increasing weight the risk of cancer at screening increases whereas that of interval cancer decreases (women in the non-response group hold an intermediate position).

With regard to height the reverse applies: women at increased risk of developing breast cancer in the interval tend to be tall (Table 36).

TABLE 35. BODY WEIGHT AND BREAST CANCER RISK

	Interval cancer		Other Mca*		Total Mca		Pop min
	Number	RR	Number	RR	Number	RR	Mca Number
< 60 kg	8	1.00	37	1.00	45	1.00	4287
60-69 kg	18	1.10	92	1.21	110	1.19	8799
70-79 kg	12	0.99	82	1.46	94	1.37	6513
≥ 80 kg	4	0.59	43	1.38	47	1.24	3607
	42		254		296		23,206
Significance	NS		$p = 0.03$		$p = 0.07$		

*Non-compliance group plus those detected through screening.

TABLE 36. BODY HEIGHT AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min
	Number	RR	Number	RR	Number	RR	Mca Number
< 155 cm	4	1.00	30	1.00	34	1.00	2437
155-159 cm	7	0.80	57	0.87	64	0.86	5325
160-164 cm	9	0.74	84	0.92	93	0.90	7433
165-169 cm	12	1.35	56	0.84	68	0.90	5403
≥ 170 cm	10	2.34	27	0.84	37	1.02	2605
	42		254		296		23,203
Significance	$p = 0.01$		NS		NS		

TABLE 37. QUETELET'S INDEX AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca
	Number	RR	Number	RR	Number	RR	Number
<23	13	1.00	33	1.00	46	1.00	4480
23-26	20	0.63	112	1.49	132	1.26	10,194
27-30	8	0.46	70	1.60	78	1.28	5945
≥31	1	0.13	39	2.05	40	1.51	2581
	42		254		296		23,200
Significance	$p < 0.01$		$p < 0.01$		$p < 0.01$		

TABLE 38. RELATION BETWEEN QUETELET'S INDEX (AS AN OBESITY INDEX) AND OESTROGEN RECEPTOR STATUS (ER) OF BREAST CANCER. WEIGHT AND HEIGHT AS DETERMINED AT INTAKE IN THE DOM PROJECT

Quetelet's index	Population in which no breast cancer was detected from 1975 to the end of 1980	Breast cancer patients*					
		ER +		ER -		Total	
		Abs	RR	Abs	RR	Abs	RR
<23	4482	18	1.00	10	1.00	28	1.00
23-26	10,191	50	1.22	33	1.45	83	1.30
27-30	5942	35	1.47	11	0.83	46	1.24
≥31	2581	28	2.70	1	0.17	29	1.80
	23,196	131		55		186	

*21 patients with dubious ER status were excluded (decision by Department of Endocrinology independent of our study). In remaining 89 patients ER not determined; they were unselected with regard to Quetelet's index.

It is clear that these opposite trends will produce an interesting contrast when exploring the effect of Quetelet's index. Indeed, Table 37 shows that the obese tend to have a higher breast cancer risk in general, with the exception of interval cancers which show a preference for the lean individual. These trends are statistically significant.

Interval cancers contain relatively few oestrogen receptors, and in our material information about oestrogen receptor (ER) status was available in 207 of 296 cases. Table 38 contains data on the relationship between Quetelet's index and ER status; the latter was divided into ER-positive and ER-negative on the basis of Scatchard plots [30]. Two facts emerged: ER-positive cancers were associated with fatness and ER-negative cancers with leanness. In fact, women with ER-negative breast cancer were on average leaner than the population at large.

How can these statistical relationship be explained? It has now been established beyond doubt that a number of obese women produce oestrone in adipose tissue [31], which is the main source of oestrogens in postmenopausal women. Therefore ER-positive cancer cells thrive and multiply in fat women.

On the other hand, in lean postmenopausal women ER-positive breast cancer cells find no stimulus for multiplication. Any clones of cells not containing oestrogen receptors will have a relative growth advantage, so that, by a process of clonal selection, these cells will be in the majority after a number of doublings.

Table 39 provides data on the effect of body surface, as determined by a modern version of the old Du Bois formula. This kind of amalgamation of individual weight and height

TABLE 39. BODY SURFACE AREA AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca
	Number	RR	Number	RR	Number	RR	Number
< 1.64 m ²	6	1.00	42	1.00	48	1.00	4168
1.64-1.75	16	1.64	65	0.95	81	1.04	6809
1.76-1.87	13	1.36	82	1.23	95	1.25	6644
1.88-1.99	4	0.77	37	1.01	41	0.98	3641
≥ 2.00 m ²	3	1.08	28	1.44	31	1.40	1935
	42		254		296		23,197
Significance	NS		NS		NS		

data apparently does not produce trends comparable with those of an obesity index like Quetelet's.

7. Use of drugs

In our evaluation of the effect of drugs on breast cancer risk we studied the following groups:

(a) *Oestrogens*: as medication in the treatment of climacteric and postmenopausal problems. In the Netherlands these substances have not (yet) found general acceptance as preventive therapy of postmenopausal disturbances. Oestrogens are also, of course, a major constituent of oral contraceptive agents, but in our middle-aged population the use of the "pill" could have been stopped some years before the project began.

(b) *Antihypertensives*: in particular the effect of rauwolfia alkaloids (reserpine) which received much publicity after three papers had appeared in the *Lancet* [32]. Later publications tended to be less disquieting.

(c) *Drugs for alleviating anxiety states*: Rumours circulated in the lay press that studies by Horrobin *et al.* [33] had shown that diazepam was associated with some risk in experimental animals. These observations could not be confirmed [34, 35].

(d) *Thyroid preparations*: a paper by Kapdi and Wolfe [36] aroused some concern (see [37]). In the older literature unexplained relationships between breast cancer and thyroid disease (hypothyroidism) have been reported. A relationship between cancer and a drug could, however, be attributable to a relationship between cancer and the disease for which the drug was given.

The questions put to all women during screening were of a qualitative nature; we had no data on dose and only limited data on the duration of drug use (from successive individually-linked screening cycles). Each question was concerned with drug use during the 12 months preceding the screening visit of the Preventicon.

Tables 40–43 provide data on the risk of mammary cancer associated with drug use as disclosed at the initial screening examination. These Tables also throw light on the widespread use of a number of drugs.

With regard to oestrogens no increase of risk was encountered in the city of Utrecht and its neighbouring municipalities. This is reassuring since some American reports [38] have indicated that there is a small risk.

Our overall findings notwithstanding, oestrogens seemed to confer some degree of risk in respect of interval cancers. This is puzzling since most interval cancers contain no oestrogen receptors. We can offer no explanation for this result.

TABLE 40. OESTROGEN USE AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca Number
	Number	RR*	Number	RR*	Number	RR*	
No oestrogens	35	1.00	243	1.00	278	1.00	21,514
Oestrogens	7	2.53 (1.31–4.88)	11	0.57 (0.35–0.94)	18	0.82 (0.54–1.24)	1700
	42		254		296		23,214

*90% confidence limits in parentheses.

TABLE 41. ORAL CONTRACEPTIVES AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca Number
	Number	RR*	Number	RR*	Number	RR*	
Never on "pill"	38	1.00	234	1.00	272	1.00	21,303
Ever on "pill"	4	1.17 (0.50–2.74)	20	0.95 (0.65–1.42)	24	0.98 (0.69–1.39)	1912
	42		254		296		23,215

*90% confidence limits in parentheses.

TABLE 42. USE OF ANTIHYPERTENSIVE DRUGS AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca Number
	Number	RR*	Number	RR*	Number	RR*	
No such drugs, syst. blood pressure ≥ 160 mm	9	1.00	121	1.00	130	1.00	9005
No such drugs, syst. blood pressure < 160 mm	18	1.87 (0.79-4.44)	87	0.67	105	0.76 (0.61-0.94)	9604
Reserpine	2	4.76 (1.49-15.26)	4	0.71	6	0.99 (0.50-1.96)	420
β-Blockers	2	1.27	20	0.95	22	0.97 (0.67-1.42)	1567
Diuretics	3	2.00	11	0.55	14	0.65 (0.40-1.04)	1497
Other	0	0.00	2	0.54	2	0.50 (0.14-1.74)	277
	34		245		279		22,370**

*90% confidence limits in parentheses. **During first few months of project blood pressure was not measured.

TABLE 43. USE OF TRANQUILLIZERS AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min Mca Number
	Number	RR	Number	RR	Number	RR*	
No such drugs	40	1.00	220	1.00	260	1.00	19,838
Tranquillizers	1	0.23	20	0.85	21	0.75 (0.52-1.09)	2124
Other drugs for relief of anxiety	1	0.40	14	1.01	15	0.91 (0.58-1.43)	1253
	42		254		296		23,215

*90% confidence limits in parentheses.

As far as oral contraceptives were concerned, we were unable to make a detailed analysis and had to content ourselves with a comparison of those who had never used them and those who had ever used them (Table 41). In accordance with more specific studies (e.g. [39]) no risk was revealed.

In analysing the data on the effect of antihypertensive drugs we chose as a reference group (relative risk = 1.00) those women who had a systolic blood pressure of 160 mmHg or more but who were not receiving specific drug treatment for this. While there seems to be no reason for great concern with regard to the rauwolfia alkaloids (Table 42), an increased risk for interval cancers was found which proved significant despite small numbers. We are inclined to relate this to our finding [10] that women on reserpine tend to have increased prolactin levels in their blood; this could form a stimulus for breast cancer growth.

The somewhat low risk of breast cancer associated with normal blood pressure (compared to untreated hypertension) should very probably be explained by a confounding effect of overweight: in our material (as in so many other studies) overweight was strongly correlated with blood pressure. Since interval cancers show a preference for lean women, they are preferentially found in the normotensive group. The effect of tranquillizers (mainly benzodiazepines, including diazepam) is shown in Table 43: there is no indication of any increased risk of breast cancer. Similarly, we found no risk associated with the use of thyroid medication (Table 44).

8. Endocrinological studies

Extensive studies were undertaken with regard to the urinary excretion of 11-desoxy-17-oxosteroids (11-DOOS). Those women who were soon detected as having (asymptomatic) breast cancer did not differ in their excretion of these androgen metabolites from large groups of population controls. Neither did we find any statistical relationships

TABLE 44. THYROID MEDICATION AND BREAST CANCER RISK

	Interval cancer		Other Mca		Total Mca		Pop min
	Number	RR	Number	RR	Number	RR*	Mca Number
No thyroid medication	41	1.00	251	1.00	292	1.00	22,841
Thyroid medication	1	1.50	3	0.74	4	0.84 (0.37–1.92)	371
	42		254		296		23,212

*90% confidence limits in parentheses.

between 11-DOOS/creatinine ratios on the one hand and a number of known risk factors on the other. Thus we could find no support for the suggestion made by Bulbrook that measurement of these hormone metabolites could help in defining a high-risk group [9].

This investigation led to the establishment of a large "bank" of urinary samples (12-hr overnight specimens), collected at the first and second screening of women from the city of Utrecht.

These specimens are being kept at -20°C for research purposes.

9. Radiological risk factors (breast parenchymal patterns)

Thanks mainly to the work of Wolfe [40] attention has been focused on possible relationships between breast parenchymal patterns, as seen on (xero)mammograms, and the risk of breast cancer. Wolfe distinguishes several classes on the basis of two criteria, viz.:

- dysplasia (not to be confused with the histopathological abnormality bearing the same name);
- prominent duct pattern (PDP).

We classified all our xeromammograms to these characteristics; for criteria see Rombach's thesis [6].*

The analysis of our data—like that of the other risk factors—was aimed at determining the risk of mammary cancer during the study period (1975–1980) associated with the parenchymal pattern observed at the initial examination. Risk was determined not only in terms of the individual but also in terms of cancer in the homolateral breast (patterns were not always identical in both breasts).

In Table 45 the effect of dysplasia and PDP is presented on the basis of laterality (left or right breast), expressed as the cancer risk in an individual (irrespective of laterality). The reference group—not mentioned in the Table—is in each case comprised of those without the parenchymal characteristic mentioned.

With regard to the effect of PDP there are no differences of any significance in the risk estimate between cancers detected through screening, interval cancers or those in women failing to comply with screening after the initial examination.

However, there was a difference between these groups with regard to dysplasia, the risk of interval cancer being increased in the dysplastic breast (Table 46). The larger number

TABLE 45. EFFECT OF DYSPLASIA AND PROMINENT DUCT PATTERN (PDP) ON BREAST CANCER RISK IN INDIVIDUALS. NOTE THAT RR IS GIVEN WITHOUT SPECIFICALLY MENTIONING A REFERENCE GROUP WITH RR = 1.00

Effect of	RR
Dysplasia in right breast	1.09
Dysplasia in left breast	1.10
PDP in right breast	1.38
PDP in left breast	1.11

*Copies in English available on request.

TABLE 46. EFFECT OF RADIOLOGICAL PATTERN OF DYSPLASIA ON RISK OF BREAST CANCER IN SAME BREAST

	Number of breasts at risk minus breasts with cancer	Interval cancer		Other Mca		Total Mca	
		Number	RR*	Number	RR	Number	RR*
Dysplasia absent	30,269	21	1.00	171	1.00	192	1.00
Dysplasia present	16,270	23	2.04	92	1.00	115	1.11
			(1.26-3.32)				(0.92-1.35)

*90% confidence limits in parentheses.

of breast cancers in this table is due to cases of bilateral cancer being counted as double cancers (instead of merely logging the cancer which appeared first, an event which changes a woman at risk into a cancer case).

The association between roentgenologic dysplasia and interval cancers might be explained on the ground that X-ray images showing dysplasia impede recognition of cancer. However, if this were the whole answer one would expect dysplasia to be underrepresented among the cancers found at screening, and this was not the case. If recognition bias occurs, it may perhaps be neutralised in the non-interval cases by a weak, but real risk of biological phenomena underlying the dysplasia. Such a real risk would, of course, add to the apparent risk due to recognition bias in the interval cancers.

The weak effects of these radiological risk factors seem to be at variance with Wolfe's findings and resemble Swedish experience in population screening [41, 42]. In this context one should remember that our study is largely confined to postmenopausal women in whom parenchymal patterns tend to be less pronounced than in premenopausal women.

On the basis of the evidence in this chapter, we can confirm the existence of a number of risk factors for breast cancer reported in the literature. Some features with regard to interval cancer clearly need further attention. A search for other possible risk factors among routinely collected data did not reveal anything worth further exploration.

10. Multivariate analysis

In the univariate analysis the study of confounding and interaction effects by means of standardization techniques becomes problematic if several factors are involved. Multivariate analysis is a powerful tool in assessing both the interplay and the joint effect of a number of factors simultaneously.

We have employed stepwise logistic regression using the BMDPLR program of the University of California, Los Angeles*. Because of restrictions in memory capacity, and in view of the large number of possible combinations between variables and the limited number of breast cancer cases ($n = 296$ including 42 interval cancers) the study population apart from these cases was randomly halved for computation purposes.

Since univariate analysis had shown that the group of interval cancer cases differed in its risk profile from the other cases of breast cancer it was decided to analyse two sets of data:

- (a) all 296 cases in the study population;
- (b) 254 cases (excluding interval cancers) in the study population.

In the first part of the analysis the following risk factors were studied simultaneously: family history (dichotomous), number of children and age at first birth (polytomous factors). If a woman had no children her age at first birth was artificially set at 30 years for calculation purposes; this seemed reasonable in view of MacMahon's findings [20].

The analysis showed in the first place that the strongest independent risk factor was a family history of breast cancer (in mother and/or sister); furthermore, that age at first birth got preference over the number of children as an explanatory variable for breast cancer; its contribution to the logistic function was significant in both data sets ($p < 0.05$). After allowing for the effect of the former factor no independent effect of the latter remained.

*Mr B. J. Slotboom performed the computer analysis.

TABLE 47. BREAST PARENCHYMAL PATTERN AND AGE AT FIRST BIRTH

Age at first birth	Right breast			Left breast		
	Number of breasts examined	% breasts with DY	% breasts with PDP	Number of breasts examined	% breasts with DY	% breasts with PDP
<20	565	10.4	4.4	566	11.1	3.9
20-23	3695	13.1	5.0	3692	13.4	5.4
24-27	7065	16.1	6.2	7054	16.5	6.8
28-31	4927	18.1	6.7	4922	17.7	7.0
≥32	2789	21.9	7.7	2793	21.8	8.8
No children						
Single	2146	37.0	12.5	2143	36.0	13.1
Married	2244	32.8	13.8	2246	31.8	13.7

The next investigation explored the joint effect of family history, age at first pregnancy (= first birth) and Quetelet's (obesity) index. It emerged that in data set 1 (employing all 296 breast cancer cases) the obesity index did not contribute significantly to the cancer risk; however, in data set 2, viz. excluding the interval cancers, it did ($p < 0.02$). Thus, there is no ground for Bulbrook's suggestion (see discussion, [25]) that the effect of obesity might be confounded by factors like parity or age at first birth. On the basis of trends shown in the univariate analysis one can understand that the effect of obesity in the whole series is being diluted by an opposite trend in the interval cancer cases.

Further analysis was done of the simultaneous effect of Wolfe's breast parenchymal patterns and the proven risk factors of family history, age at first pregnancy and obesity index. A striking statistical relationship was discovered between both aspects of parenchymal pattern and age at first baby (Table 47) and parity (not shown). Notwithstanding these relationships, all these factors displayed significant and independent contributions to breast cancer risk.

None of the risk factors found proved to be of really practical importance in the sense that one could concentrate on women with a certain risk factor and leave those without it unscreened.

11. Practical implications of risk group analysis

Apart from the scientific concept of high risk groups for our understanding of the etiology of breast cancer there is the relevant question whether any of the risk factors would be of sufficient magnitude to identify a group with such high risk that screening should be strongly recommended. (The converse, naturally, looks for a low risk group in whom screening seems unnecessary).

Our analysis seems to show that none of the risk factors so far identified is of sufficient strength to warrant special attention for screening policy. Our estimates are of the same magnitude as those in a number of case-control studies and we do not expect entirely different estimates during further study of the breast cancer experience in the DOM-cohort.

The question could be raised whether a combination of risk factors, e.g. a risk profile on the basis of multivariate analysis would be in order. We are inclined to reject such an idea since the women would have great difficulty to understand why some of them are at "high" and others at "low" risk. With our present knowledge, any dichotomy in the Netherlands would leave a population with a substantial cumulative breast cancer risk unscreened.

SUMMARY

This report describes the short-term results of a population-based screening programme—the DOM* project—during the years 1975–1980. Screening was offered to all women in the city of Utrecht (The Netherlands) and 11 neighbouring areas who were born between 1911 and 1925.

*DOM = Diagnostisch Onderzoek Mammacarcinoom or diagnostic investigation of mammary cancer.

Further evaluation will be reported in future publications.

The programme was organised by the Department of Epidemiology, Institute of Social Medicine (now renamed The Institute of Public Health and Epidemiology), University of Utrecht, which established a screening centre—The Preventicon—in a new commercial and shopping centre with good public transport facilities.

The study population was between 50 and 65 years old at intake, a decision based on results obtained in the H.I.P. study in New York. Women resident in the city of Utrecht ($n = 14,697$ at intake) were screened four times, with intervals of 12, 18 and 24 months respectively between the first and fourth examinations. Women resident in the other areas ($n = 8814$) were screened twice with an interval of 18 months.

Screening consisted of a clinical examination (inspection, palpation) and xero-mammography. An important role was played by paramedical personnel, who performed the examination and had a degree of responsibility in reading the mammograms. Initially the radiation dose per breast was about 1 rad, but from 1977 onwards this was reduced to 0.4 rad per exposure; for women living in Utrecht (4 screenings) the total dose was less than 5 rad over the 5-year period.

Referral policy was based on a set of three types of "signal" to general practitioners (who coordinated specialist services):

1. Recommendation for biopsy;
2. Suggested repeat examination after 6 months;
3. Brief note on palpable abnormalities, considered likely to be benign.

In Section II the response of the invited population is analysed. The response rate for the first screening was 72% in the city of Utrecht and 68% in the other areas. Variation in response rates can only be explained to a minor degree by measured demographic and sociological variables, age being the most important determinant. Women were asked to contribute D.fl. 12.50 (approx. US \$5) per screening cycle to meet part of the expenses, the rest being paid by public health research funds.

A total number of 61,649 examinations was performed in 23,511 women.

In Section III the immediate results of screening are presented. An analysis is made of the cancer yield of the 3 signals mentioned above. The decision rules led to a high degree of specificity (99%), with a high predictive value of a positive test (40% in Utrecht, 57% in the other areas screened subsequently). "Pick-up" rates of breast cancer at the initial screening were 7.2 per 1000 in Utrecht and 6.5 per 1000 in the other areas. In the second, third and fourth screenings these detection rates varied between 1.3 and 3.4 per 1000. The high predictive value of a recommendation for biopsy was maintained.

Mammographic examination plays a far greater role in detecting breast cancer than clinical examination: only 6 of 210 cancers detected by screening were found solely on the basis of clinical examination. And cancers detected through mammography alone were found at a more favourable stage in terms of axillary node status. Analysis reveals that screening did not lead to an increased percentage of women biopsied for lesions that proved to be non-cancerous ("unnecessary" biopsies).

Radical surgery was performed in 69% of cases in the period 1975–1980 (major changes towards less radical procedures were initiated in 1981). Two-thirds of the patients were operated upon within 3 weeks after detection and 95% within 6 weeks.

Section IV stresses the importance of a population-based cancer registry when conducting intervention studies like the DOM project. With the help of three groups of pathologists (and others) it was possible to attain full coverage in the city of Utrecht. Cancer in the non-response group could be assessed, interval cancers could be counted, and the sensitivity of screening could be estimated (94% in the initial examination).

Some interval cancers (10 cases) occurred in women who were considered to have a benign lump in their breast at a previous screening cycle (brief note, signal 3, see Section I). Their risk of interval cancer was shown to be increased 20-fold; however, if we had suggested that all 347 women with such lumps should go for biopsy the predictive value

of a recommendation for this procedure would have dropped (Utrecht women) from 40 to 20%. This illustrates the kind of decision-rule dilemmas inherent in screening.

A comparison of the stage at diagnosis, as determined by tumour size (largest diameter), and axillary node status reveals the following:

(a) Cancers found at screening were smaller than those diagnosed before the start of the DOM project (1974) and also smaller than those of the non-response group.

(b) Interval cancers were also smaller than those under (a) not detected through screening. This seems to refute the suggestion that screening merely detects the slow-growing cancers and leaves undetected the faster growing ones with a poor prognosis.

(c) More than 70% of cancers detected through screening had no axillary metastases compared with less than 50% in the other groups mentioned under (a).

(d) The axillary node status of interval cancers in Utrecht was similar to that of the cancers not detected through screening and mentioned under (a); in the other areas it more resembled that of cancers found at screening.

While the overall picture gives reason for hope, a number of years of follow-up are still necessary to evaluate the impact of screening in the total population.

The last part of Section IV deals with some screening theory which may help programmes such as the DOM project to gain new insight into the natural history of breast cancer. The "stable disease model" is contrasted with an alternative model which assumes that the genesis of new breast cancers becomes rare after the age of 50. Further understanding will be obtained by the application of mathematical models and from a fifth screening cycle which is to be undertaken in the city of Utrecht four years after the fourth examination. If the alternative model turns out to be the more plausible one, it would not only be of biological interest but also of practical importance in determining screening strategy with regard to the length of intervals between screening cycles.

In Section V the effects of known risk factors are evaluated. The risk of the following factors is confirmed:

- relative late age at first pregnancy;
- a personal history of benign breast lesions;
- a family history of breast cancer in mother and/or sister(s);
- a late menopause;
- overweight.

As far as the last factor is concerned analysis reveals striking differences with regard to the presence or absence of oestrogen receptors in the tumour. Cancers with such receptors show a preference for obese women whereas those without such receptors are associated with the lean. An explanation of these findings is based on our knowledge of extra-ovarian oestrogen production in obese women and the concept of clonal selection in cancer cell kinetics.

Exploring the possible effects of drugs, we found on the whole that there seemed to be no increased risk of breast cancer with the use of oestrogens, oral contraceptives, reserpine, benzodiazepines and thyroid preparations. Oestrogens and reserpine had some influence on the risk of interval cancer.

Multivariate analysis was applied to answer a few questions as to the interdependence of several risk factors. Family history of breast cancer, age at first birth and obesity were found to have independent effects on breast cancer risk. Breast parenchymal pattern was found to depend on parity and age at first birth, but nevertheless had some independent effect on the risk of mammary cancer.

Determination of androgen metabolites did not prove of value in predicting risk in these (mainly postmenopausal) women. Nor was the classification of xeromammograms according to parenchymal pattern of much help.

It was not possible to define a group with a sufficiently high or low risk on which to base a selective screening policy.

REFERENCES

1. Waard F de, Baanders-v. Halewijn EA: A prospective study in general practice on breast cancer risk in postmenopausal women. *Int J Cancer* 14: 153-160, 1974
2. Shapiro S, Strax Ph, Venet L: Periodic breast cancer screening in reducing mortality from breast cancer. *JAMA* 215: 1777-1785, 1971
3. Waard F de, Thijssen JHH, Veeman W, Sander PC: Steroid hormone excretion pattern in women with endometrial carcinoma. *Cancer* 22: 988-993, 1968
4. Poortman J, Thijssen JHH, Schwarz F: Androgen production and conversion to estrogen in normal postmenopausal women and in selected breast cancer patients. *J Clin Endocr Metab* 37: 101-109, 1973
5. Waard F de, Baanders-v. Halewijn EA: Cross-sectional data on estrogenic smears in a postmenopausal population. *Acta Cytol (Philad.)* 13: 675-678, 1969
6. Rombach JJ: Breast Cancer Screening, results and implications for diagnostic decision-making. **Doctoral Dissertation** Utrecht: Stafleu, Alphen a/d Rijn, The Netherlands, 1980
7. Lundgren B, Jakobsson S: Single view mammography, a simple and efficient approach to breast cancer screening. *Cancer* 38: 1124-1129, 1976
8. Bulbrook RD: Urinary androgen excretion and the etiology of breast cancer. *J Natn Cancer Inst* 48: 1039-1042, 1972
9. Poortman J, v.d. Smissen J, Collette HJA, De Waard F: Ratio of 11-desoxy 17-oxosteroids to creatinine in a population screened for breast cancer. *Br J Cancer* 39: 688-695, 1979
10. Waard F de, Kwa HG, Poortman J: Plasma prolactin levels in women at postmenopausal age with a family history of breast cancer or a prescription of antihypertensive rauwolfia treatment. *Oncology* 37: 33-36, 1980
11. Waard F de, Baanders-v. Halewijn EA, Collette HJA, Vroegindewey-Jie D: Overgewicht in Overvecht. *Tijdschr Soc Geneesk* 53: 70-75, 1975
12. Waard F de, Rombach JJ, Collette HJA: The DOM project for the early diagnosis of breast cancer in the city of Utrecht, The Netherlands. In **Screening in Cancer**, Miller AB (Ed). UICC Technical Report Series 40: 183-200, 1978
13. Baker LH: Breast cancer detection demonstration project: five-year summary report. **Ca—a Cancer Journal for Clinicians** 32: 194-225, 1982
14. Eddy DM: **Screening for Cancer: Theory, Analysis and Design**. Englewood Cliffs, NJ: Prentice-Hall, 1980
15. Fisher B: Breast cancer management, alternatives to radical mastectomy. *N Engl J Med* 301: 325-328, 1979
16. Zelen M, Feinleib M: On the theory of screening for chronic diseases. *Biometrika* 56: 601-614, 1969
17. Kusama S, Spratt JS, Donegan WL, Watson FR, Cunningham C: The gross rates of growth of human mammary carcinoma. *Cancer* 30: 549-599, 1972
18. Walter SD, Day NE: Estimation of the duration of a preclinical disease state using screening data. *Am J Epid*. In press
19. Rothman KJ, Boice JD: Epidemiologic analysis with a programmable calculator. **NIH publ** 79-1649, US Dept of Health, Education and Welfare, 1979
20. MacMahon B, *et al*: Age at first birth and breast cancer risk. *WHO* 43: 209-221, 1970
21. Hutchinson WB, *et al*: Risk of breast cancer in women with benign breast disease. *J Natn Cancer Inst* 65: 13-20, 1980
22. Waard F de, Baanders-v. Halewijn EA, Huizinga J: The bimodal age distribution of patients with mammary carcinoma. *Cancer* 17: 141-151, 1964
23. Macklin MT: Comparison of the number of breast cancer deaths observed in relatives of breast cancer patients and the number expected on the basis of mortality rates. *J Natn Cancer Inst* 22: 927-951, 1959
24. Trichopoulos D., MacMahon B, Cole Ph: Menopause and breast cancer risk. *J Natn Cancer Inst* 48: 605-613, 1972
25. Waard F de: Body size and breast cancer risk. In **Banbury Report 8: Hormones and Breast Cancer**. Pike MC, *et al*. (Eds). New York: Cold Spring Harbor Laboratory, 1981, pp. 21-32
26. Feinleib M: Breast cancer and artificial menopause; A cohort study. *J Natn Cancer Inst* 41: 315-329, 1968
27. Sluysen M, Evers SG, De Goeij CCJ: Sex hormone receptors in mammary tumours of GR mice. *Nature* 263: 386-389, 1976
28. Waard F de: Nutritional etiology of breast cancer. *Nutr Cancer* 4: 85-89, 1982
29. Gehan EA, George SL: Estimation of human body surface area from height and weight. *Cancer Chemother Rep* 54: 225-230, 1970
30. Waard F de, Poortman J, Collette HJA: Relationship of weight to the promotion of breast cancer after menopause. *Nutr Cancer* 2: 237-240, 1981
31. MacDonald PC, Edman CD, Hemsell DL, Porter JL, Siiteri PK: Effect of obesity on conversion of plasma androstenedione to estrone in postmenopausal women with and without endometrial cancer. *Am J Obstet Gynecol* 130: 448-452, 1978
32. Heinonen OP, Shapiro S, Tuominen L, Turunen MJ: Reserpine use in relation to breast cancer. *Lancet* 2: 675-677, 1974
33. Horrobin DF, Ghatur T, Karmali RA: Letter to the editor. *Lancet* 1: 978, 1979
34. Wallace RB, Sherman BM, Bean JA: A case-control study of breast cancer and psychotropic drug use. *Oncology* 39: 279-283, 1982
35. Danielson DA, Jick H, Hunter JR, Stergachis A., Madsen S: Non-estrogenic drugs and breast cancer. *Am J Epid* 116: 329-332, 1982
36. Kapdi CC, Wolfe JN: Breast cancer: relationship to thyroid supplements for hypothyroidism. *JAMA* 236: 1124-1127, 1976
37. Gorman CA, *et al*: Breast cancer and thyroid therapy. *JAMA* 237: 1459-1460, 1977
38. Hoover R, Gray LA, Cole Ph, MacMahon B: Menopausal estrogens and breast cancer. *N Engl J Med* 295: 401-405, 1976

39. Vessey MP, Doll R, Jones K, McPherson K, Yeates D: An epidemiological study of oral contraceptives and breast cancer. **Br Med J** 1: 1757-1760, 1979
40. Wolfe JN: Risk for breast cancer determined by mammographic parenchymal pattern. **Cancer** 37: 2486-2494, 1976
41. Andersson J: Mammographic screening for breast carcinoma. **Acad Thesis University of Lund**, Malmö, 1980
42. Tabar L, Dean PB: Mammographic parenchymal patterns, risk indicators for breast cancer? **JAMA** 247: 185-189, 1982