TOWARDS A SCREENING MODEL FOR BREAST CANCER

by

Chrisitine S. Grover

Thesis submitted for the degree of Master of Philosophy of the University of Southampton

-1986-

CONTENTS

	page no.
ABSTRACT	iv
ACKNOWLEDGEMENTS	v
PREFACE	1
CHAPTER 1:BREAST CANCER	
1.1. Introduction	6
1.2. Incidence	7
1.3. Risk factors	13
1.4. Growth of breast cancer	22
1.5. Spread of breast cancer	33
1.6. Detection of the disease	39
CHAPTER 2: SCREENING	
2.1. General principles	44
2.2. Screening for breast cancer	50
2.3. Using risk factors for selecting women for screening	55
2.4. Screening tests for breast cancer	59
2.5. The role of BSE	61
2.6. Errors in screening tests	70
2.7. Factors affecting the frequency of screening	82

CHAPTER 3: THE AGE OF ONSET DISTRIBUTION

3.1. Introduction 87

3.2. Age distribution of patients at diagnosis of the disease 87

3.3. Tumour size at discovery 92

3.4. Tumour doubling time 94

3.5. Duration in state S2 97

3.6. The age at menopause 101

3.7. The age at onset of breast cancer 102

3.8. The false negative rate of a screening test 104

CHAPTER 4: A SCREENING MODEL

4.1. Introduction 121

4.2. The screening population 122

4.3. The two-state model 129

CHAPTER 5: SCREENING, DELAY AND SURVIVAL

5.1. Introduction 137

5.2. Duration of different stages of breast cancer 138

5.3. False negative rate according to size of tumour 139

5.4. Delay in detection 147

5.5. Survival according to tumour characterisitics 151

CHAPTER 6: TOWARDS A SCREENING MODEL FOR BREAST CANCER

6.1. Introduction 159

6.2. Screening delay:an alternative model 160

6.3. Metastases and delay 163

6.4. Screening delay and metastases 164

6.5. Benefits and costs 168

6.6. Conclusions 172

REFERENCES 175

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MATHEMATICAL STUDIES

OPERATIONAL RESEARCH

Master of Philosophy

TOWARDS A SCREENING MODEL FOR BREAST CANCER

by Christine S. Grover

A screening model is a cheap and flexible tool to facilitate the decisions involved in setting up a screening programme for the detection of breast cancer.

Before a screening model is constructed the natural history of the disease and the theory of screening must be considered.

Some of the facets of a screening model are the age to getting cancer, the growth and spread of the tumour and the sensitivity of the particular test used.

Two measures of the success of a screening programme are the quickness to discover the presence of disease and finding the disease before spread to distant sites has occurred. A screening model can indicate the delay in finding the disease and the probability of metastases when different screening tests are used at varying intervals.

ACKNOWLEDGEMENTS

I am indebted to:

Dr A. K. Shahani for his supervision during my research

Mrs S Meacock for her help and suggestions.

My husband for his encouragement and patience.

PREFACE

i. Clinical trials for the detection of breast cancer

In 1968 the Health Insurance Plan(HIP) in New York began a mass screening programme to evaluate the effects of screening on mortality from breast cancer.(Strax,1981) This consisted of two groups screened annually over 4 years. Two measures were used to evaluate the effect of the screening programme: first, the actual number of deaths in the study group versus the control group; and second the difference in case fatality rate(CFR) between the two groups. After 7 years of follow-up there were 108 deaths in the controls compared with 70 in the study group. After two more years the figures were 128 and 91 respectively. The study group results included those who were invited but refused to participate in screening. Table 1 shows the cumulative CFR for the study and control groups up to 11 years after

Table 1. Cumulative case fatality rates(per 100)

(Source:Strax,1981)

years after diagnosis					
detected	no	5	7	9	11
on screening	132	12.9	21.2	28.8	38.1
in interval	93	37.6	43.0	51.6	54.1
refusers	78	34.6	41.0	47.5	56.5
Total study	303	28.7	35.6	43.6	48.9
Controls	293	40.6	46.4	52.6	54.8

(lead time of 1 year for cases detected on screening)

diagnosis.

In Table 2 this information is presented according to age at diagnosis. There was little difference in the CFR in the under 50s but the greatest difference was displayed

Table 2. Cumulative fatality rate(per 100) by age at diagnosis

(Source:Strax,1981)

	Year 8		Year 10		Year 11	
Age	Study	Control	Study	Control	Study	Control
40-49	42.1	43.4	44.4	47.4	44.4	48.9
50-59	35.2	53.7	44.5	57.4	48.3	58.5
60+	41.4	48.1	52.7	54.3	55.4	54.3

in the 50-59 age groups.

In the 1970s, 280,000 women were involved in 27 breast cancer detection demostration projects(BCDDP). There was no control group and there was a certain amount of self-selection.(Shapiro,1978) Encouraged by the results from these two projects, screening programmes were initiated in several European countries.

In Holland, the DOM project involved four successive screens of 23,000 women aged 35 to 65 in Nijmegen (Verbeek,1984). Arnhem which had the same age-adjusted breast cancer incidence and mortality was used as a control. The odds ratio estimate of screened versus unscreened women who died to those who did not was 0.45(95%

CI =.23,1.00). The odds ratio indicates that breast cancer mortality can be reduced by nearly 50% by screening. But the upper confidence limit of 1 indicates more data is needed to ensure a difference from unity and the follow-up is too short to draw conclusions.

In Utrect, 14796 women aged 50-64 were screened four times with repeated screenings at 12,18 and 24 months. Controls were matched for age(Collette,1984). The relative risk of dying from breast cancer among the screened to never screened was .30(χ^2=7.64, p=.003. 95%CI = .13,.70)

In Sweden in 1977 screening commenced involving 134,867 women aged 40-74 (Tabar,1985). The under 50s were offered screening every 2 years and the over 50s every 33 months. The average length of follow up was 6 years. There was a significant reduction(25%) in the absolute rate of Stage II and above cancers (p <.001). Overall there was a 31% reduction in mortality in the study group (p=.013). In the 50-74 age group there was a 40% reduction in mortality (RR=.6, 95%CI = .44,.84. p=.003). No reduction was observed in the 40-49 year olds and the confidence levels were wide.(RR=1.26, 95%CI =.56,2.84).

In England the DHHS initiated the Early Detection of Breast Cancer Project designed to compare detection by mammogram biannually with detection by breast self-examination(BSE). Mammography is being carried out on 45-64 year olds in Edinburgh and Guildford, while BSE education is being given in Huddersfield and Nottingham.

Four other health districts are being used as controls. Since the project has only been operated for a few years no conclusive results are available.

ii.Modelling screening for breast cancer.

The evidence for supporting screening comes from randomised controlled trials (RCTs) of a specific breast screening programme. Data is only available on the effectiveness of a specific programme using specific tests performed at a specific frequency on a specific population. Clinical trials are limited over a short period, expensive to run and limited to certain ages. in addition the rescreening intervals are arbitarily chosen and attenders are self-selecting eg they may be at greater risk.

The design of a screening programme raises many questions that can not be answered from clinical studies. It is not possible to perform RCTs to evaluate these options since it would involve 10,000s of patients and need a long follow up (say 15 to 20 years). One trial which indicates a reduction in mortality from breast cancer may not be the most effective and efficient option.

Experimental and statistical information on assessing the alternative screening strategies is largely lacking. Decisions on the proper design of screening programmes must be made despite the lack of knowledge. Therefore an attempt must be made to estimate the value of alternative screening strategies using mathematical models. All important factors

of the problem should be identified and their relationships described and manipulated.

Modelling is less costly than clinical trials, variables can be altered to change screening times and population characteristics such as age and risk. The results of such modelling is a process which can be used to explore the meaning and implications of existing research results. It enables manipulation of data and provides an estimate of expected outcomes of various screening programmes and hence enables the design of research.

The role of the model is to provide insights into questions that can not be answered directly by clinical observations. It does not generate answers and it can not replace clinical trials or substitute for clinical judgement.

This thesis attempts to consolidate the large amount of published information on breast cancer and investigates the background to the concepts related to screening and disease. Relationships are then determined that connect the important inputs of a screening model. Finally the results of a particular aspect of a screening model for breast cancer -delay in detection- are used to provide an insight into the choice of parameters of a screening programme and the expected effect on mortality.

CHAPTER 1

BREAST CANCER

1.1. Introduction

In England and Wales 12,000 women die of breast cancer each year, and there are some 21,000 new cases registered(OPCS,1983)

The death rate from all causes has been declining substantially since the beginning of the century and is most marked in the under 45s but a steady decline is shown in all other ages. Opposed to this there is a general upward trend of breast cancer deaths in the under 70s. Circulatory diseases account for 52% of all deaths, with cancer accounting for an additional 20%. One in twenty-five women die of breast cancer.

If death from circulatory disease is reduced by preventative measures, eg changes in habits like smoking, changes in diet and more exercise, then cancer, especially breast cancer, will become a more important cause of early death.

Breast cancer accounts for 17% of all deaths in the 35-54 age group and 10% in the 55-64 age group. Thus it is a major cause of early death in this country. It accounts for 22.6% of all cancer registrations in the 35-39 age groups and about 35% in the 40-49 age group.

1.2. Incidence

Breast cancer incidence rates exhibit considerable global variation with high rates in Western and Industrialised nations.(USA, Canada and Western Europe), intermediate rates in East and South Europe, and low rates in Asia, Latin America and Africa. Such difference could reflect genectic factors as well as social and dietary variations. It has also been noted that breast cancer rates are higher in urban than in rural areas.

Breast cancer rarely occurs before the age of menarche, that is the onset of puberty, but after the age of 30 the incidence of the disease increases progressively with age. Many authors have noticed a plateau, sometimes called Clemenson's hook, in the incidence around the age of 45-55 and have subsequently divided breast cancer into pre- and post-menopausal phases or disease. (DeWaard,1964. Stavraky,1974) There could be two distinct diseases with different risk indicators. In Japan and other Asian countries incidence reaches a peak around the age of the menopause and then decreases.

Several studies have shown a positive correlation between socio-economic status and incidence and mortality with higher rates in the upper classes. Migrant studies have displayed an increase in breast cancer in women originating from low-rate cultures but settling in countries with higher rates.(Petrakis, 1982)

Table 1.1 shows the registration rates per 100,000 for breast cancer in England and Wales in 1978. The third column gives the proportional increase between adjacent groups.

Incidence increases after menarche but increases more slowly after 50. After menopause, the rate of change increases until about 70 and then levels off.

Table 1.1. Registration rates for breast cancer, 1978

(Source:OPCS, 1983)

Age	Rate	Increase
15-19	0.2	-
20-24	1.8	9
25-29	6.4	3.55
30-34	20.3	3.17
35-39	52.7	2.60
40-44	94.9	1.80
45-49	146.0	1.54
50-54	149.2	1.02
55-59	160.8	1.08
60-64	176.6	1.10
65-69	201.3	1.14
70-74	212.0	1.05
75-79	225.3	1.06
80-84	237.9	1.06
85+	306.6	-

By using a 6 point Lagrangian interpolation formula (Elandt-Johnson, 1980) we can estimate the incidence rate for each age group from the age groupings given in Table 1.1. Table 1.2 shows the registration rates for each age using this method.

Table 1.2. Registration rates per 100,000 for breast cancer, 1978, for ages 45 to 60

Age	Rate
45	116.2
46	132.7
47	141.7
48	148.4
49	152.6
50	154.5
51	148.7
52	148.5
53	148.6
54	149.3
55	150.9
56	155.5
57	158.0
58	160.7
59	163.5
60	166.4

Although this method only gives an estimation of the incidence for each age, the plateau is clearly seen from the age 51 and incidence rates do not increases to pre-50s levels until ages 55-56.

The age standardised incidence of breast cancer has increased in almost all countries. Some countries show increases in cohorts born after the turn of the century. (OPCS, 1978)

Armstrong(1976) found increases in incidence and mortality between 1950-1973 to be partly cohort specific for those born around 1899 and partly cross-sectional begining in the mid-1960s.

Registration rates prior to 1960 are scarce so we must look at death rates for some idea of incidence changes. Table 1.3 shows the increase in breast cancer deaths taking 1911-15 as the base year. The figures must be considered bearing in mind the following points:

i. Death rates may not reflect incidence rates.

ii Since 1974 there has been automatic registration of cancer cases from death certificates. Hence prior to this increases in deaths from breast cancer may not represent a true incidence but just improvements in registration.

iii. Cohort as well as cross-sectional influences must be considered. For example, environmental influences may affect certain cohorts, or affect all ages.

Table 1.3. Increases in Incidence of breast cancer deaths over the century.

1911-1915 = 100 (adapted from OPCS, 1975)

Year	Ages 30-	35-	40-	45-	50-	55-	60-	65-	70-	75-	80-
1916-20	95	102	99	106	106	105	99	103	94	108	98
1921-25	103	98	106	107	108	112	109	112	103	117	118
1926-30	115	105	108	106	117	114	119	121	103	115	118
1931-35	120	107	107	111	117	120	124	122	108	123	126
1936-40	113	102	104	106	114	118	121	126	109	121	121
1941-45	120	111	106	105	108	109	112	120	96	101	104
1946-50	135	115	111	105	105	106	109	114	100	110	108
1951-55	135	111	113	107	103	104	105	114	97	103	105
1956-60	138	110	107	112	110	102	104	110	97	101	101
1961-65	143	111	108	119	122	115	109	111	95	97	102
1966-70	150	115	119	116	124	126	118	117	95	98	94

Armstrong(1976) attempts to relate changes in incidence and mortality to changes in possible risk factors. He notes that changes coincide with falling fertility and older age at first pregnancy.

Early menarche and late menopause are known to increase the risk of breast cancer. (see 1.3). Armstrong reports that the age of menopause has remained constant this century but there has been an increase in artificial menopause (that is by surgical intervention) which may have a protective value.

The age of menarche has decreased this century as diet has improved. During the last 100 years the mean age in Europe and N. America has become progressively earlier at a rate of 3 to 4 months each decade, but there are signs of levelling off. At present the average age is 13 years (standard deviation of 1 year). (Gold & Josimouich,1980)

Since 1962 registration rates have been published. Incidence has increased but OPCS(1978) report little change in the age specific rates for 1973 to 1978. An increase in incidence must be considered in the light of three points:

i. Improvements in registration procedures would be reflected in an 'increase' in incidence.

ii. Women are more aware of breast cancer and therefore may discover it at an earlier age, or not delay in seeking medical advice. Detection twelve months prior to previously would substantially affect registration rates per age group.

iii Earlier discovery before death from other causes will also reflect in an increased incidence.

1.3.Risk factors

Many studies have tried to discover certain characteristics which indicate a high risk of developing breast cancer.

In the Netherlands, DeWaard(1964) studied 14,697 women and provides the risk ratio by comparing the women with cancers to those without in the screened population.

Stavraky(1974) compared 95 pre- and 278 post-menopausal breast cancer cases with 106 pre- and 480 post-menopausal controls who had benign breast disease or other malignant disease. These results must be considered in the light of the following problems:

i. the controls may not be representative of the population.

ii. some risk factors may be common to both groups.

iii. there is a small sample size of single and low parity women.

Brinton(1979) used the information gathered from the first screen on the BCDDP involving 405 breast cancer cases and 1156 controls. Association was measured by relative risk approximated by the relative odds. A relative risk of 1 is no different in risk. When the 95%CI did not include 1 the relative risk was significantly different at $p<.05$ level. This study consisted of self-selecting women, the participants were well educated, 1/4 reported family

history of breast cancer and 1/3 had their first birth over 27.

Tulinius(1978) in Iceland produced a retrospective study of 34,525 women born before 1945. He calculated the relative risk of various factors. The χ^2 statistic was correlated to the log likelihood function and the confidence intervals were based on a large sample. The CI are not given since they were wide as each strata contained few individuals but the significance can be judged by the χ^2 statistic.

Coombs(1979) did a prospective study of 747 women aged 15-69 with benign breast disease detected between 1957 and 1965 and followed for 12 to 20 years. The 747 controls were hospital referals who were discharged with non-malignant conditions. The advantage of this was that they were cleared of any breast disease.

Risk factors can be divided into two groups:those due to genectic susceptability and those due to reproductive factors.

1.3.1.Genectic factors

There are three main genectic factors affecting breast cancer risk:- a history of breast cancer in the family, certain benign diseases and endocrine factors.

A history of breast cancer in the family deems a woman

at greater risk. Table 1.4. gives the relative risk according to the relative involved. This risk was also examined according to age(less than 50, 50 to 59 and

Table 1.4. Relative risk of breast cancer according to family history

(Source:Brinton,1979)

relative with breast cancer	RR	95%CI	no of patients	no of controls
none	1	-	314	930
mother	3.88	2.2- 6.8(i)	33	26
grandmother	4.82	2.1-11.1(ii)	18	11
both	4.87	1 -26.1(ii)	5	3

i p <.01 ii p <.05

over 60). A history of breast cancer in a mother or grandmother was associated with a six-fold increases for patients under 50, as compared to a 3 or 4 fold relative risk for the over 50s.

Certain benign diseases are precursors of breast cancer although the actual risk attached to different benign diseases vary between studies.

Table 1.5 shows the incidence rates and the relative risk of women with benign breast diseases(BBD) compared with women diagnosed as having non-malignant conditions(NMC) and used as controls.

Table 1.5. Age adjusted incidence rates by cause per 1000 person years of follow-up and the RR for women with BBD compared with women with NMC

(Source:Coombs,1979)

Site	BBD		NMC		RR
	No	Rate	No	Rate	
All	59	6.27	42	4.16	1.5
B.C.	23	2.43	9	0.81	3.0*
Not B.C	36	3.84	33	3.26	1.2

* p <.05

The relative rate of 3.0 for breast cancer among women who had benign disease was significant. Of the 32 cases of breast cancer, 12 were less than 50, including 2 in the control group. Of the other 20 cases in the over 50s, 7 were in the control group. No further analysis was performed due to small numbers.

Table 1.6. shows the relationship between BBD and the occurrance of breast cancer, and Table 1.7 shows them in relation to age at menopause.

The relationship between benign breast disease and breast cancer is not clear. The prevalance of benign disease decreases after the age of 50. One explaination of association between the two is that BBD is a precancerous lesion or represents an intermediate stage between

Table 1.6. Observed and expected numbers of incident breast cancers cases among BBD women.

(Source:Coombs,1979)

Interval between BBD and BC	No followed	incidence obs.	incidence exp*
0-4	646	5	1.0
5-9	629	12	3.1
10-14	605	5	5.1
15+	416	1	0

*based on distribution of breast cancer in controls.

Table 1.7. RR of breast cancer incidence on follow-up according to BBD and age at menopause

(Source:Coombs,1979)

	Menopause	
	less than 50	over 50
No BBD	1.00(i)	2.45
BBD	3.73	9.55*

* p< .05 (i) one person placed here for calculation

causative factors and breast cancer. If this association is direct then breast cancer would be expected to occur in the same breast as the disease but Haagensen & Donnely(Coombs,1979) reported that 50% occur in the opposite breast. Alternatively BBD and breast cancer is the result of abnormal hormonal status and BBD is an earlier manifestation of this abnormal state.

Certain endocrine factors are associated with enhanced risk of breast cancer. These are defective production of progesterone in part of the menstrual cycle, raised level of plasma prolactin in post-menopausal women, a subnormal secretion of androgen metabolites, and a degree of hypothyroidism in menopausal women who also have a family history of breast cancer. it has also been discovered that breast cancer patients have twice as much available oestradiol than controls.(Imperial Cancer Research, 1982)

1.3.2.Reproductive factors.

There are three main reproductive factors that affect risk- age at menarche, age at first pregnancy and age at menopause.

The relationship between age at menarche and breast cancer risk was discussed by Brinton(1979) and Tulinius(1978). The results are shown in Table 1.8. Breast cancer risk appears to significantly decrease with later age at start of menarche.

1.8. Breast cancer risk and age at menarche

a. (Source:Brinton,1979)

Age	RR	95%CI	no patients	no controls
under 12	1	-	64	165
12	0.97	0.7-1.4	110	291
13	0.85	0.6-1.2	108	335
14+	0.82	0.6-1.2	113	345

linear trend p=.20

b. (Source:Tulinius,1978)

under 12	1.86
12-13	1.39
14-15	1.08
16+	1.00

χ^2 =4.7,p=.19

Stavraky(1974) divided the cases up into pre- and post-menopausal cancers. She found a significant decreases in the risk with late menarche (over 13) in the AMC cases. See Table 1.9.

Table 1.9. RR of breast cancer for women according to age and menopausal status

(Source:Stavraky,1974)

Age	pre-menopause	post-menopause
under 12	1.00	1.00
12-13	0.91	0.90
13+	0.45	1.28

χ^2 = 4.28,df=1,p <.05 χ^2 =3.60,df=1,p <.1

Table 1.10 shows the increased risk of breast cancer with delay in first birth.

Table 1.10 RR of breast cancer and age at first birth.

a. (Source:Brinton,1979)

Age	RR	95%CI	no patients	no controls
under 20	1	-	26	116
20-24	1.37	0.8-2.3	133	423
25-29	1.53	0.9-2.6	98	288
30+	2.15*	1.2-4.0	76	134

linear trend p=.001 *p <.05

b. (Source:Tulinius,1978)

under 20	1
20-24	1.63
25-29	2.61
30-34	2.53
35+	4.12
nill	3.76

χ^2 =29.2, p <.001

Stavraky divided patients into AMC and PMC. Increased risk of breast cancer for older age at first pregnancy appeared in both groups, but the trend was only significant in PMC. When cervical cancer patients were excluded from controls the AMC trend disappears.

Table 1.11. RR of breast cancer and age at first pregnancy.

(Source:Stavraky,1974)

Age	Pre-menopausal	Post-menopausal
under 20	1.00	1.00
20-24	2.36	1.10
25-29	2.23	1.41
30+	2.29	3.00

Linear trend =2.44,df=1, $p<.2$ χ^2 =17.9,df=1,$p<.005$

Brinton(1974) found a significant trend toward an increase of breast cancer risk with older women at natural menopause. Table 1.12. shows that the risk for women having natural menopause after 55 was 2.5 times greater than for those less than 40 at menopause.

Table 1.12. Breast cancer risk and age at menopause.

a.

(Source:Brinton,1974)

Age	RR	95%CI	no patients	no controls
under 45	1.06	0.5-2.1	17	54
45-49	1.00	-	47	159
50-54	1.22	0.8-1.9	72	199
55+	2.17*	1.1-4.2	24	37

$p<.05$ linear trend $\chi^2=4.43,df=1,\ p<.05$

1.3.3.

Other factors may affect the risk of developing breast cancer:

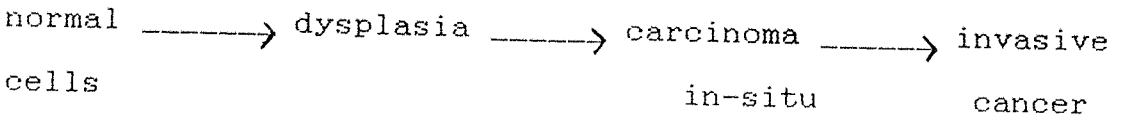
i. some studies have associated breast cancer rates with total fat and animal protein consumption and increased prevalence of obesity.(DeWaard,1964)

ii. increased risk has been noted where patients have been exposed to radiation eg fluroscopy for pulmonary tuberculosis. The latent period is about 15 years. (Petrakis,1982)

1.4. The growth of breast cancer

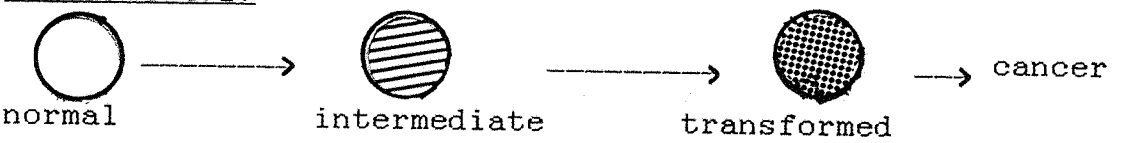
A tumour develops from one or more cells which show signs of dysplasia, that is abnormal characteristics. Some cells may develop malignant characteristics and grow in an abnormal way. The growth of these abnormal cells may be contained in-situ or may break out and infiltrate surrounding tissue. When this happens, the tumour becomes invasive. Prior to this the tumour may regress. Malignant cells enter the bloodstream and travel to other parts of the body and result in the establishment of secondary tumours. Figure 1.4.1. shows these stages.

Figure1.4.1. Stages of tumour growth



Movlgavkar(1980) postulated a two-stage model for carcinogenisis in which a normal cell undergoes two changes to reach a cancer cell. He assumed a malignant transformation of a single cell leads to the development

Figure 1.4.2.



a malignant tumour, any susceptible cell is likely to be transformed independently of another.

Epidemiologic behaviour of breast cancer follows logically from simple assumptions about hormonal changes in breast tissue at menarche and menopause. With menarche there is an increase in the number of cells and in menopause a decrease in both the susceptible and intermediate stage cells. With early menarche the increase occurs early in life and with a late menopause there is a delay in the involution of the breast and hence an increased risk in both cases.

This idea of a two stage model can also be used to explain other facets of breast cancer. Where there is a family history of breast cancer, ie an inherited gene, cells are at an intermediate stage and hence only one transformation is required which usually happens at an early age. Lobular carcinoma in-situ may represent an intermediate stage where the cells are benign for many

years. With irradiation, the effect is greater as the age between radiation and menopause increases, that is where the intermediate cells are in existence longer. High risk mammographic patterns show atypia cells as high risk benign conditions which could be associated with the intermediate stage.

Tumours have a complex structure composed of proliferating and non-proliferating fractions, the relative size of which may vary considerably. This fraction not only varies according to the type and location of tumour but also between individuals. Progression and regression may give the impression of a constant growth rate over small periods. (Smithers, 1968)

Effective growth is the net result of the cell dividing rate and growth inhibiting factors. There are three parameters which affect net growth rates:

i. the cell cycle time of proliferating cells.

ii. the proportion of proliferating cells.

iii. the extent of cell loss which may be due to exfoliation, metastases, or cell death.

If the destruction of cells occur on the tumour surface then the larger the tumour the less cells are lost in proportion to the volume. Steel(1967) states that cell loss may exceed 50%.

There is evidence to suggest that there may be a

difference in the proportion of proliferating cells in tumours of pre- and post-menopausal women and the proportion may also vary according to the stage of the menstrual cycle.(Mayer & Baum,1976)

Various growth equations have been postulated to explain the pattern of growth within a tumour.(Spratt & Spratt,1979)

The linear growth model postulates that linear dimensions of the tumour are increased by a specific increment each day regardless of the size of the tumour. This type of growth has been used to model rat sarcoma and some lung cancers.

The exponential growth model assumes a steady increase in a cell population by binary division with negative or steady cell loss. The tumour grows exponentially with a steady increase in volume per unit of time. The exponetial growth constant is given by:

$$b = \frac{\text{Ln } V1 - \text{Ln } V0}{t} \qquad 1.4.1$$

where V0 and V1 are the volumes at first and second observations taken at t time apart.

The actual doubling time of the tumour DT is given by:

$$DT = Ln2/b \qquad 1.4.2$$

The Shell model was first proposed by Mayneord in 1932 to take account of the deceleration of tumour growth rate with increased size. It postulates that only cells in the outer shell of the tumour divide exponentially. This change in rate can also be modellled by using a Gompertz equation.

In 1962, Mendelsohn introduced the concept of a growth fraction, postulating that there is a difference between the potential and the actual doubling time of a tumour. Mendelsohn explained this by showing that not all cells divide at any given time. He defined the growth fraction as the proportion of proliferating or growing cells to the total cell population. Studies of breast cancer in mice have given this to be 0.4 and in the cheek pouches of hamsters to be 0.36.

The pattern of growth can be given as

$$\frac{Vt}{Vo} = \frac{Exp(Ln2.g)t}{Tc} \qquad 1.4.3$$

where Vt is the volume at time t, Tc is the mitiotic time cycle and g is the fraction of cells able to produce viable daughter cells and the growth fraction is G= 2g-1.

Thus if g is smaller than 0.5 the tumour will regress, while a value of 0.5 represents equilibrium.

The tumour doubling time is given by:

$$DT = \frac{Tc.Ln\ 2}{Ln2g} \qquad 1.4.4$$

Hall & Laing(1968) proceed to develop equation 1.4.3. to take into account a changing growth fraction. The viable fraction is represented by gt at time t where:

$$Ln\ 2gt = \frac{Ln\ 2g0(1-e^{-at})}{Tc\ at} \qquad 1.4.5$$

and hence they obtain:

$$Vt/Vo = \frac{Ln\ 2go(1-e^{-at})}{Tc\ at} \qquad 1.4.6$$

Cells live for a finite time and during this time will produce 0,1 or 2 daughter cells. Depending on the fraction of cells producing this number of daughter cells, the tumour will regress, remain constant, or increase in size at varying rates. Cell loss (not through death) can vary considerably. Hence the growth fraction does not depend on the viable fraction.

Blumenson & Bross(1969) give the volume of the tumour at time t as:

$$Vt = Vo.2^{t/\theta} \qquad 1.4.7$$

where Vo is the volume of a single cell and θ is the tumour doubling time.

The diameter of a single cell is 10u or 10^{-3} cm.

Hence:

$$Vo = \frac{4\Pi R0^3}{3} = \frac{\Pi 10^{-9}}{6} \qquad 1.4.8$$

If t is the time to detection then;

$$Vt = \Pi dt^3/6$$

and substitution gives:

$$dt^3 = 10^{-9}.2^{t/\theta}$$

$$= 2^{t/\theta}\,10^{-30} \quad \text{since } 10^{-9} \sim 2^{-30} \qquad 1.4.9$$

and thus:

$$dt = 2^{t/3\theta}\,10^{-30} \qquad 1.4.10$$

Table 1.13 shows the number of doublings it takes for a tumour to reach a certain diameter. Thirty doublings have occured before a tumour has reached 1cm. in diameter. As many as 23 doublings have taken place before the tumour is detectable by some screening test.

Table 1.13. Number of doublings required for a tumour to reach a certain diameter.

diameter(cm)	no of doublings*
0.2	23
0.5	27
1	30
2	33
3	35
4	36
5	37

*to nearest whole no.

Blumenson & Bross(1969) state that the net doubling time is an average which takes into account the cell cycle time, the proportion of cells proliferating, host defences and cell deaths. Their equations can be extended to incorporate a growth factor whichreflect s the growth fraction.

We let t/θ

$$Vt = Vo.\ D \qquad 1.4.11$$

where θ is the potential doubling time in the abscence of cell loss. D is the growth factor which equals 1+d where d is the proportional increase(or decrease) after cell doublings and loss.

After substitution of Vo and cancellation we obtain:

$$\frac{9}{10}\ DT^{\frac{3}{t/\theta}} = D \qquad 1.4.12$$

and $$t = \frac{\theta(3\mathrm{Ln}(10^3 dt)}{\mathrm{Ln}\ D} \qquad 1.4.13$$

Table 1.14. shows the number of cell cycles required for a tumour to reach a certain diameter for various growth factors. A growth factor of two represents exponential growth as displayed in equation 1.4.10.

Table 1.14. Number of cell cycles required for a tumour to reach a certain diameter for various growth factors.

Diameter(cm)	Growth factor				
	2	1.8	1.6	1.4	1.2
0.2	23	27	34	47	87
0.5	27	32	40	55	102
1	30	35	44	62	114
2	33	39	49	68	125
3	35	41	51	71	132
4	36	42	53	74	136

With exponential growth it takes 33 cell cycles for a tumour to reach 2cm, but as it is more likely that the growth factor is 1.6 or 1.4 then 49 or 68 cell cycles respectively are required to reach this size.

Many studies have tried to measure cell growth by in-vito methods, serial mammograms or X-rays of secondary

tumours.

Heuser(1979) reports 64 women with 2 or more serial mammograms of which 32 tumours were detectable on the previous screen. 23 had a doubling time from 109 to 944 days and the other 9 were too slow to measure. The measurements along the major and minor axes and the volume was calculated. There was a high degree of correlation with a log-normal distribution in the major axis with mean 0.003mm/day (SD= .39, p =.6 to .7). The minor axis was not normally distributed. This is a small sample whose results do not record fast growing tum ours and thus whose distribution is skewed towards slower growing tumours.

Fournier(1980) found a range of doubling times of 44 to 1869 days in 147 women with 388 serial mammograms. The mean doubling time was 212 days(95%CI= 191,235). He also notes that observed doubling time is variable during the lifetime of a tumour and quotes one case of doubling times of 63, 384 and 174 days between mammograms.

Kusuma(1972) calcualted the doubling time of 163 primary and 36 secondary tumours. The median time was 108 days and range from 6 to 558 days. The doubling time of the primary tumours were longer than for the secondary ones. 45% of the breast tumours had a doubling time of over 124 days.

Malaise(1973) estimated the doubling time by Tritiated thymidine labelling(TTL) which measures the division of cells by considering the need of thymine in the production of DNA. This permits a measurement of the percentage of

cells in mitosis and thus gives the potential doubling time of the cell population which Malaise estimated as 23.8 days.

Such measurements using in-vito growth are unreliable in determining rates of growth, cell proliferation and cell loss since the body immune defence are not represented. Methods of measuring tumour nucleus shadows in past mammograms may be critisised- first, fast and slow growing tumours over the time interval were excluded, and second the measurement is subject to error and nucleus shadow can not be precisely defined.

Table 1.15. gives the number of years for a tumour to reach a certain diameter for various doubling times. For a tumour doubling time of 100 days it takes over 6 years for a tumour to grow to 0.2cm and for a doubling time of 200

Table 1.15. Number of years for a tumour to reach a certain diameter for various doubling times.

Size(cm)	doubling times				
	50	100	150	200	300
0.2	3.15	6.30	9.45	12.60	18.90
0.5	3.70	7.40	11.10	14.79	22.19
1	4.11	8.22	12.33	16.44	24.66
2	4.52	9.04	13.56	18.08	27.12
3	4.79	9.59	14.38	19.18	28.77
4	4.93	9.86	14.79	19.73	29.59
5	5.06	10.14	15.21	20.27	30.41

days this figure is over 12 years. For a doubling time of 100 days it takes 23 months for a tumour to grow from 0.2cm to 1cm and over 19 months from 1cm to 4cm. When the doubling time increases to 150 days the intervals becomes 34 1/2 and 29 1/2 months respectively.

1.5. The spread of breast cancer.

The progress of spread from the tumour to lymph nodes and more distant sites is not fully understood. Metastatic deposits in the nodes are either confined or extended into surrounding tissue. Extra-nodal extension is more likely when 4 or more nodes are affected and in infiltating ductal type cancers.(Stroll,1977)

Spread depends on the site of the primary tumour.(Baum,1981). Involvement of one or more lymph nodes can lead to involvement of more distant lymph nodes and into the base of the neck. Tumours in the inner quadrant may spread to the chest cavity. Malignant cells that enter the bloodstream may become lodged in the bones, liver, lungs or brain. Seidman(1972) found that 71% of tumours in 939 cases occured in the outer quadrants, and of these 50% had negative nodes compared with 82% of the inner quadrant's tumours.

Recent evidence (Imperial Cancer Fund,1983) questions the belief in a stepwise progression from the tumour to surrounding lymph nodes and thus to more distant sites where favourable conditions exist. One-fifth of patients

with localised disease at diagnosis later developed distant metastases, whereas one-half of patients with nodal involvement remained disease free after surgery. The author concludes that the ability to metastise is a random event. No mention is made of how long the patients who remain disease free were followed up. The estimation of 50% may be rather high when slow growing metastases are considered.

It is not known what proportion of cases metastise before symptoms appear, or indeed before the tumour becomes detectable by a screening test. Efforts to estimate the time of establishment of metastases have not been successful. Methods using tumour doubling times have indicated inititation of the secondary tumour before the patient's birth.

Alvord(Spratt,1979) working on Kusuma's data tried to estimate when carcinomas begin to metastise. He concluded that probably this does not occur before the 21st doubling and as many as half do not do so before the 33rd doubling. That is, few do so before the tumour is just under 0.2cm in diameter and half not before it is 2cm. There may be a type of breast cancer that does not metastise whatever the size. Donegan(Spratt,1979) identified some women with tumours greater than 9cm who had no lymph node metastases and whose 5 year survival was 70%.

Pickren(Spratt,1979) reviewed 200 cancers of which 19 were 1cm or smaller. Of these, 6 less than 8mm showed no metastases, whereas 30% of those measuring 8 to 10mm had

metastised.

Campbell(Spratt,1979) discovered 30% of those with Stage 1 cancers and 35% with tumours less than 2cm had positive bone scans. Heuser(1979) found that of 7 cancers less than 1cm, 3 developed distant metastases. The above evidence indicates that some tumours could metastise between 1 and 8mm. Gershon & Cohen(1963) postulated a 30% increase in the likelihood of metastases in cancers over 1.5cm in diameter.

Several studies have published details of spread to nodes according to tumour size. These are given in Tables 1.16, 1.17 and 1.18 for 1105, 89 and 81 cases respectively. In the Rombach study the mean size of negative node tumours was 1.4cm compared with 1.9cm where there was nodal involvement. This difference was significant with p=0.02.

Table 1.16. Nodal involvement according to tumour size

(Source:Seidman,1972)

Size (cm)	% involvement for each size group according to nodal status			% in each size group
	negative	1 to 3	4 or more	
0.1-1.9	60	27	13	15
2.0-2.9	59	23	18	22
3.0-3.9	46	27	29	22
4.0-4.9	50	21	29	14
5.0-5.9	33	33	33	12
6.0+	40	20	40	15
% of cases	49	25	26	100

Table 1.17. % of cancers with positive and negative nodes according to size.

(Source: Thomas, 1983)

Size(mm)	negative	positive	no
non-invasive	100	0	14
under 5	100	0	11
6-10	88	12	17
11-15	79	21	19
16-20	82	18	11
21-50	56	44	16
50+	-	-	1(neg)

Table 1.18. % of cancers with positive or negative nodes according to size.

(Source: Rombach, 1980)

Size(mm)	negative	positive	no
under 5	100	0	7
5-9	71	29	14
10-19	62	38	42
20-29	40	60	15
30-39	-	-	0
40-49	-	-	2(pos)
50+	-	-	1(neg)

Bartosynski(1982) questions the traditional view that the primary tumour disseminates potential metastases at a rate proportional to size. He suggests a model in which a systemic mechanism, which gives a small constant increase to the probability of metastases, is added to the traditional metastatic mechanism. That is the probability of metastases forming consists of a random element and an element related to tumour size.

Results from 116 patients in Warsaw suggest a tumour doubling time of 2.2 months and a median time from origination to detection of 59.2 months. For each patient, the tumour was localised. The primary was removed but no masectomy was performed, lymph nodes were not affected and no other treatment was given. Table 1.19 shows the probability of secondary tumours from the two mechanisms.

Table 1.19. The probability of secondary tumours.

(Source:Bartosynski,1982)

Time(years)	Metastsatic mechanism	Systemic mechanism
1	2.26×10^{-8}	0.0354
2	9.34×10^{-7}	0.0695
3	3.85×10^{-5}	0.1020
4	1.57×10^{-3}	0.1340
5	0.0423	0.1650
6	0.0688	0.1940
10	0.0688	0.3020

For screening we are only interested in tumours over 0.2. Bartosynski's data suggests that 15% of tumours have metastised to secondary sites before this and 28% do so before the tumour reaches 3cm.

Koscielny(1984) studied 2648 patients who underwent similar treatment with no chemotherapy to try to establish a relationship between the establishment of metastases and tumour size. He found that the median delay between treatment time and the appearance of metastases was shorter when tumours were larger with 50% of metastases in tumour size 1 to 2.5cm appearing within 42 months but when the diameter was over 8.5cm in only 4 months. He found a lognormal relationship between tumour volume and the probability of metastases. The median was 3.56cm(95% CI 0.14,4000ml.).The mean was 3.16cm and SD of 2.62 (χ^2 = 1.56,df=5).Table 1.20 shows the observed and fitted proportion of metastases as a function of tumour size.

Table 1.20 Observed and expected proportion of metastases

(Source:Koscielny,1984)

size(cm)	observed	lognormal	no.
1 up to 2.5	.271	.240	317
2.5 up to 3.5	.420	.450	496
3.5 up to 4.5	.567	.572	544
4.5 up to 5.5	.665	.664	422
5.5 up to 6.5	.728	.735	329
6.5 up to 7.5	.838	.789	192
7.5 up to 8.5	.813	.829	136
over 8.5	.920	.903	212

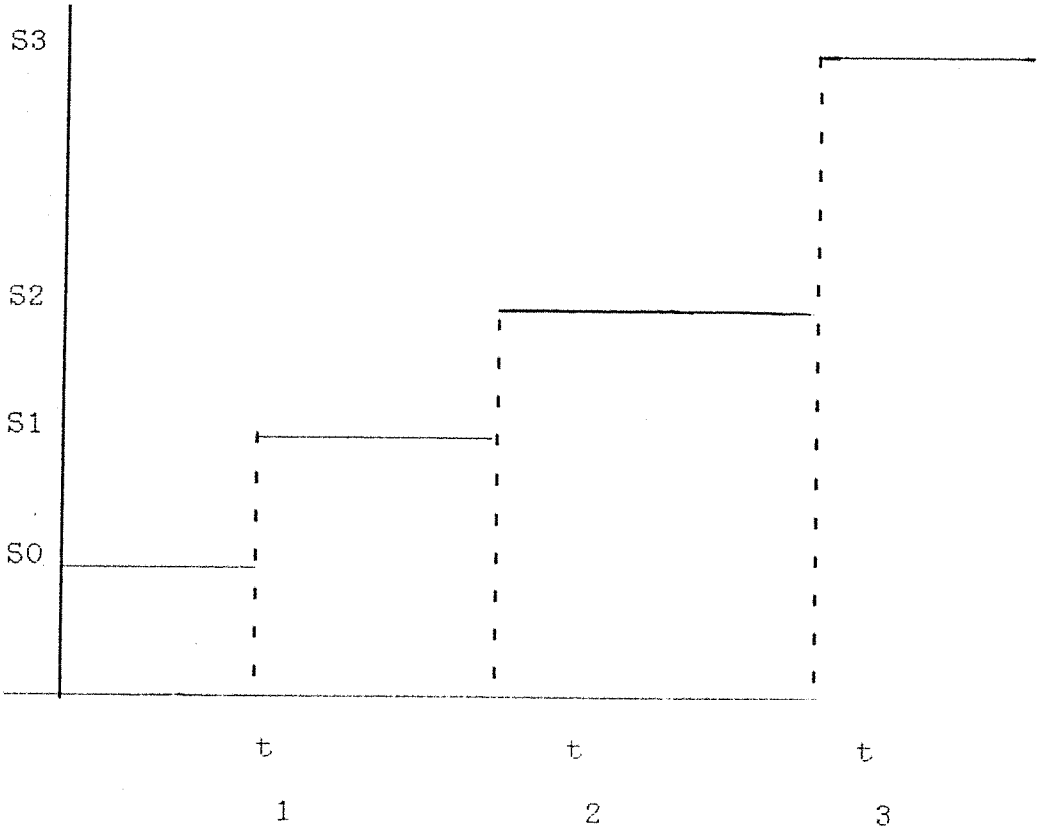
1.6. Detection of the disease

A woman deemed to have breast cancer enters various states from initiation to detection of the disease. Although this thesis deals with screening for breast cancer, the term disease will be used since the application can be extended to any disease

and screening programme.

Figure 1.6.1 shows the various states a patient with disease enters. Let us assume there are four stages through which a woman who gets breast cancer may pass in the absence of screening.

Figure 1.6.1. The stages of disease



A woman in S0 is disease free. In state S1, the disease is present but undetectable by a screening test until a woman enters state S2 where the tumour has grown to such a size or is displaying signs which makes it detectable by a test. If a screening test is not applied in S2 or if the test fails to detect the disease, then the disease is not apparent until symptoms appear- state S3 and the patient may seek advice.

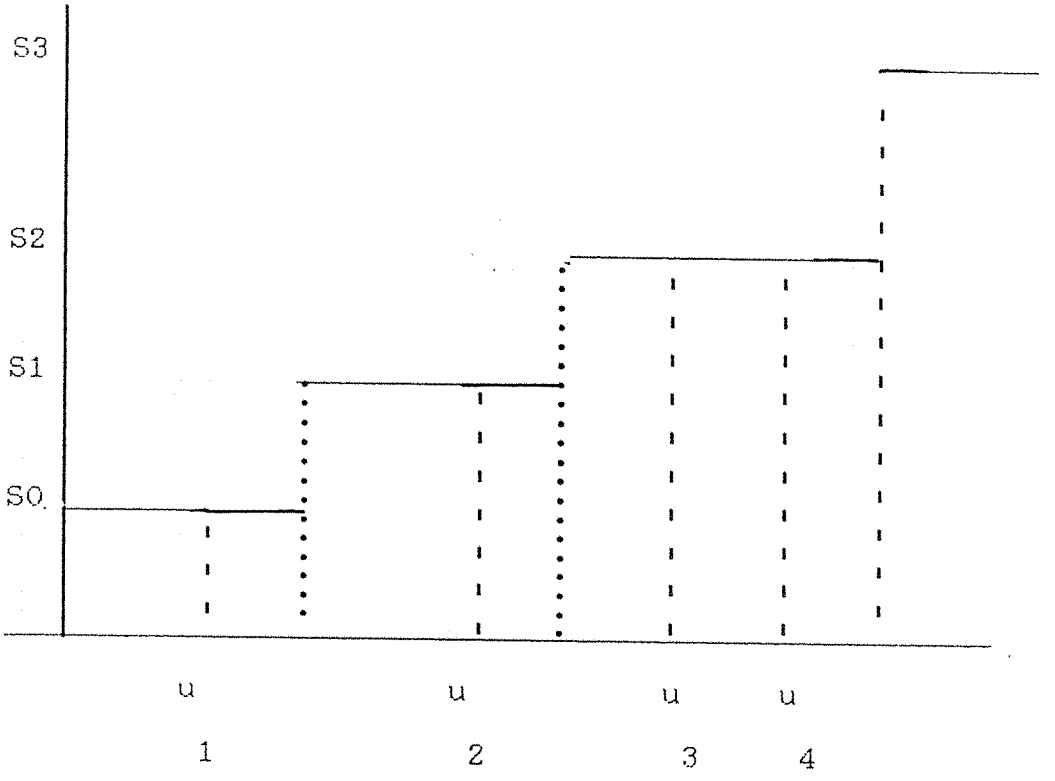
We let t_i be the time of transition from state S_{i-1} to S_i Each t_i is variable. Time t_1, the start of the disease may occur at any age but can be represented by a probability density function.

Time t_2 represents when the tumour is detectable by a screening test. The duration in S1 , t_2-t_1, depends on the growth of the tumour and the type of test used. For example mammograms can detect tumours less than 0.5cm in diameter whereas CE typically does not find tumours under 1cm. Entry into S3 does not occur at a specific time in the development of the tumour, that is when a tumour reaches a certain diameter.

If screening takes place then four classifcations emerge depending on the state that a woman is in and the sensitivity of the screening test used. See Figure 1.6.2.

If screening occurs in state S0, say at u_1, the woman is free of disease and hence screening would result in a (true) negative result. It is possible for screening to indicate the presence of disease which does not exist- this

Figure 1.6.2. The states of disease and screening



is known as a false positive result. For screening in S2, at time u_2, the disease is present but not yet detectable, because of the limitations of the test in detecting a tumour of this size. Although a woman has the disease, the inability of the test to detect it is not strictly a false negative result for the test.

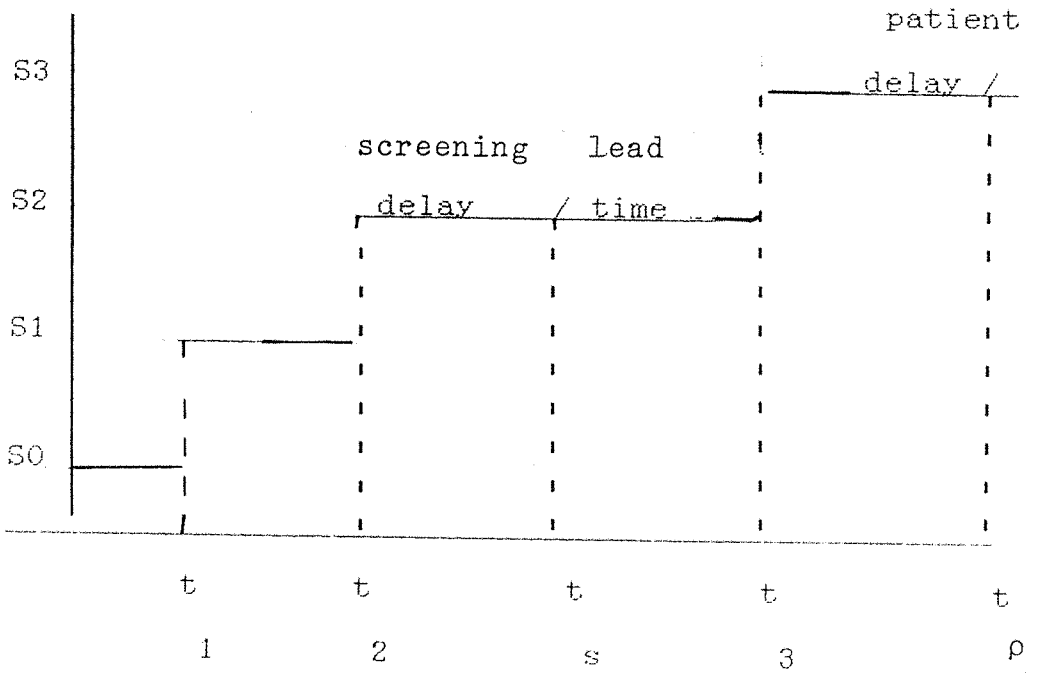
In state S2, at time u_3, the disease is detectable by the screening test. It may be detected - a true positive or because of certain features of the disease (for example

position or type) be missed - a false negative result. In the latter case further screens may occur before detection or the patient may enter S3 and hence detect disease herself. For a patient practising self-examination, signs of the disease may be apparent and hence time t_3 will be advanced.

Various delays exist, some of which are uncontrollable while others are affected by the patient screening regime and attitudes.

Figure 1.6.3. shows the delays possible. The minimum delay in detection is $t_2 - t_1$ the length of the 'silent interval'. Only advanvces in technology resulting in the detection of disease at an earlier stage and size can affect this. If we let t_s be the time that screening detects a tumour then $t_s - t_2$ can be termed the screening

Figure 1.6.3. States of the disease and delay



delay. This may be minimised depending on the time of screening in relation to t_2 and the sensitivity of screening.

Detection on screening will advance discovery of the disease by $t_3 - t_s$; this is known as the lead-time.

In the absence of screening or in its inability to detect the disease then signs or symptoms may become apparent. But a patient might not seek advice at that time but delay until time t_p. Hence the patient's delay $(t_p - t_3)$ is such that it could be eliminated without costs.

The importance of screening may be reflected in the magnitude of screening delay (or lead-time) and the duration in state S2 which may be reduced by a woman's ability to detect the disease at an early stage.

CHAPTER 2

SCREEENING

2.1 General principles of screening

Many diseases which afflict mankind can be controlled or eradicated by primary prevention. This can be done by innoculation, improved nutrition and hygiene or changes in diet. In the case of cancer primary prevention is limited. The incidence of cancer induced by occupational exposures to carcinogenics or those related to tobacco and alcohol could be reduced given political and personal backing, but little is known about the early natural history of most cancers prior to detection and hence primary prevention is not possible. However Cuzick(1986) claims that evidence strongly suggests that the amount of available oestrogen is a key factor in breast cancer and the relation of factors affecting ovarian hormones such as parity and age at first birth, menarche and menopause support this. He proposes a clinical trial using tamoxifen as an anti-oestrogen to assess its possible preventative properties.

Screening is sometimes referred to as secondary prevention since it aims not to prevent the disease but to alter the natural history of the disease which leads to morbidity and mortality.

Screening was defined by the United States Commision on Chronic Illness (1957) as "the presumptive identification of unrecognised disease or defect by the application of

tests, examinations or other procedures that can be applied rapidly".

More recently the DHSS (1976) defined screening as: "the deliberate examination of substantial segments of the population - even the entire population - in search for disease at its earliest stages is a logical extension of the role of preventative medicine".

Screening tests are not intended to be fully diagnostic, but positive findings will need to be confirmed by other procedures. It is only worthwhile screening for disease (when the concept can be applied to all diseases, the term disease will be used, otherwise cancer will be specified) if treatment to cases found on screening is more effective or cheaper than that applied to cases where symptoms have led to diagnosis. There is no point in screening for diseases which can be treated just as successfully or cheaply when symptoms appear. Nor is it reasonable to screen for untreatable diseases.

Screening can be very expensive and divert resources away from other health priorities. If the objective is a reduction in the population's mortality from disease, then it is important to realise that the life expectancy of the population may be little changed. It has been calculated that if all cancer was eradicated the effect from other competing causes is such that the life expectancy would only increase by 2 1/2 years. Screening for diseases that afflict at an earlier age would have a greater effect on

life expectancy.

At an international conference on Screening for Cancer (Miller, 1978) participants including those who had been involved with screening programmes for breast cancer agreed that the following pre-requistes should be considered before the introduction of screening as part of a public health programme.

1. The disease should be common and should cause substantial mortality and/or morbidity.

2. There should be evidence of the effectiveness of treatment of lesions discovered by screening in reducing mortality and expected level of improvements should be stated.

3. Sensitivity and specificity of the screening test to be used should be evaluated and a check made to ensure these levels are reached.

4. Any adverse effects of the test, subsequent diagnosis and treatment should be weighed against benefit.

5. There should be sufficient resources for not only screening but resultant treatment.

6. The population to be screened - the target population should be clearly defined and selected so that the predictive value of the test will be acceptable.

7. The target population should be limited to those at particular risk of the disease where such identification is possible.

8. The target population should be reachable and likely to

adhere to the programme for screening and further tests and treatment.

9. There should be an established policy for early recall for suspicious findings and routine recall in the event of negative findings.

The effectiveness of screening is often difficult to assess because of biases. It is not appropriate to compare the number of deaths among breast cancer cases compared to controls because of lead-time bias. That is, survival appears better because we are detecting the cancer at an earlier stage. The length bias is where the slower growing tumours are more likely to be detected by screening whereas the fast growing tumours will be discovered by the patients themselves when symptoms manifest themselves. Those that present for screening are self-selecting and may be at greater risk because of a family history of breast cancer and are also more likely to seek advice quickly after symptoms appear. Women may also present for screening for diagnosis of an existing complaint. Lastly there is an over diagnosis bias when lesions may be detected which will never manifest themselves during the patients lifetime. The last 2 biases will increase the yield of cancers.

The safety and acceptability of the screening test is important. Any fears over the safety or the risk of getting the disease or other diseases by screening will reduce attendance. If the test is painful or further tests are debilitating the acceptability will be low.

The reliability of a test can be measured in various ways:

1. The sensitivity of a test measures its ability to detect all diseased people in the screened population. It is expressed as the percentage of diseased for whom the test gives a positive result.

2. The specificity of a test measures its ability to identify non-diseased people. It is expressed as a percentage of non-diseased people for whom the test is negative.

3. The validity of the test is defined as Sensitivity + Specificity - 100.

4. The false positive and false negative rates are defined as 1- specificity and 1- sensitivity respectively.

In general the sensitivity and specificity of a test have to be traded off against one another. Increase in sensitivity results in decrease in specificity and vice-versa. The importance of poor sensitivity depends of the importance of delay in the outcome of the disease and the frequency of screening. Poor specificity will result in increased cost of further diagnostic tests plus a reduced response to attend further screening sessions if further tests are painful or debilitating and unnecessary. Sensitivity and specificity are not absolute attributes of the test but refer to it as it was performed at that time. The test should be reproducable but many tests suffer from observer variability.

The cost of a screening test can be measured in various ways:

1. the cost per test.

2. The cost per person screened. This depends on specificity and is estimated by the weighted costs of each of the four classifications on screening less the discounted cost of treatment in the absence of screening.

3. the cost per true positive which depends on sensitivity and specificity.

4. the cost per year of life saved. This is measured by comparing survival of screened and unscreened and depends on lead time gained.

Costs may be reduced by directing programmes at high risk groups possibly using initial test with high sensitivity and low specificity and then subsequent tests with high specificity. This will reduce unnecessary further tests. Such a secondary test may be expensive for general use.

A particular screening test may have high validity, but when used to screen a large population, it may be potentially harmful. It becomes necessary to consider alternative or compromise screening strategies.

The compromise is made if:

i. asymptomatic people with the disease could clearly benefit from the test.

ii. many asymptomatic people would not benefit

iii. most people in the population do not have the disease.

iv. using the test is expensive, time consuming or inconvenient.

v. there is potential risk that the test can damage health.

A compromise strategy has two requirements, namely that it should expose as many asymptomatic diseasedpeople as possible and expose as few people who do not have the disease as possible.

2.2 Screening for breast cancer.

For screening to be worthwhile, cancers diagnosed and treated at an early stage must stand a better chance of being cured than those treated at a later stage whether still localised or when metastatic spread has occurred. Several studies indicate that survival is better when the tumours are found at an early stage. Berkson (1952) gives the survival rates for patients after 5 years without and with metastases as .78 and .34 respectively. Cutler (1967) related certain characteristics to high mortality. These are distant metastases, satellite nodules in the skin of the breast, fixed axillary nodes, skin edema and fixation to underlying tissue. He recorded survival according to tumour size, localised and distant spread. Table 2.1 and 2.2 shows that mortality was higher for patients with larger tumours than those with smaller ones and higher for

nodal involvement than none. The relative increase in mortality as the tumour enlarged was greater for patients with negative nodes than those with positive nodes.

Table 2.1 The relationship between tumour size and survival.

(Source: Cutler, 1967)

Size	% survival after	
	5 years	10 years
T1: no mass or less than 2cm.	85-95	75-90
T2: 2-4cm	70-75	58-60
T3: 4-7cm	55	40

Table 2.2 The relationship between tumour spread and survival

(Source: Cutler, 1967)

Nodal involvement	% survival after	
	5 years	10 years
negative	75	60
positive-movable	65	54
positive-fixed	35	0

Little is known about the natural history of breast cancer before it becomes detectable. In addition it is uncertain how the cancer progresses through the different

stages. The definition of abnormality is difficult. Cancer results in recognisable characteristics in symptomatic individuals, but asymptomatic patients may have cancers at an early stage without recognisable clinical characteristics. Screening will identify a number of conditions defined as pre-clinical or pre-cancerous lesions, for example pre-invasive intraductal neoplasia. Uncertainty as to the progression of borderline cases will cast doubt on the ability of a screening programme to detect early cancers. In addition, early disease may be detected that will not develop, or not clinically surface within the patient's lifetime. These patients will suffer unnecessary mastectomies and/or radiation treatment which would detract from the value of life. One must consider the question of the maximum age for participation in a screening programme in the light of this as well as the slowing down of increases in incidence after 70. Randomised clinical trials should answer some of these questions.

As discussed in 1.3. the definition of a high risk group is extremely difficult for breast cancer. Although a group may be defined as having a risk greater than normal, this is not so marked as say the risk to smokers of lung or throat cancers. There may be identification problems, for example are those women who have taken the pill at higher risk because of the pill or because they have their first pregnancy at an older age.

Breast cancer is an important cause of mortality and

morbidity especially since it begins to take its toll at the peak of family and occupational responsibility. The incidence is low below 30 but increases between 30 and 45 to 50. Considering that the years of life lost due to breast cancer in the under 50s contribute 40% of the whole, screening in the under 50s is important. But there are major reservations in offering screening to the under 50s. Low incidence rates mean greater expenditure per case found and the greater incidence of benign disease lowers specificity. In addition, mammography is less sensitive in pre-menopausal women and radiation risk greater.

The effectiveness of screening has been under a lot of discussion. Improved survival may be illusory and due to the advancement of diagnosis not the postponment of death. The HIP results, which suggested a third reduction in mortality for those screened, did not take into account lead-time. Since lead-time is a measurement related to the rate of progression of the disease, we must assume that:

i. the disease progresses at some rate, that is there are no chronic habitues in the pre-clinical state.

ii. incidence of the disease is at the same rate as entry to the detectable state. If there is a cohort effect, this is not true.

Lead-time can be estimated by prevalence/incidence. From the HIP data this varies from 1.37 years in the 40-44 age group to 2.17 years in the 50-54s and then drops to 1.84 and 1.95 in the over 55s. The increase in the 50-54 group

could reflect the slowing down of growth rate or the plateau in incidence.

The acceptance rate for screening has varied in different trials from 37% to 82%. Individual acceptance depends on the woman's attitude to her health, the perception of vulnerability, the realisation of the severity of the disease and the belief in the effectiveness of treatment. Those who accept tend to have a higher personal history of breast diseases and symptoms. Those who refuse tend to be older, less educated, often single and from a low socio-economic group. Hence screening often did not reach the target group but became a diagnostic facility for those with symptoms.

Because of poor specificity, screening for breast cancer increases the demand on medical facilities. On initial screening, many benign diseases prevalent in the population are found which require further tests to confirm diagnosis. Although the benign disease may surface later and require treatment, screening will concentrate this demand in the present.

George(1980) found that the referal and biopsy rates fell from 7.9% and 2.5% respectively in the first year of screening to 4.3% and 1.1% in the second year and 2.7% and 1.4% in the third year. The greatest workload was generated in the 40-49 age group. George concluded that once screening was established, an average district hospital would have a work load of 4 to 7 consultations and 2 to 4

biopsies per week depending on the age groups screened.

2.3. Using risk factors for selecting women for screening

There has been little success in deciding on a woman's risk of developing cancer according to the factors discussed in 1.3. None of the variables of marital or fertility status concentrate a large enough proportion of all cases into a high risk segment. Dunn(1969) cites variables which concentrate 75-80% of cases into 60% of the population. Shapiro(1973) found 33% of cases had 3 or more risk factors, but so did 21% of patients free from the disease. Stark & Way(1974) estimated that 77% of all breast cancer patients had high risk factors associated with fertility, heriditary and benign conditions. Fournier(1977) discovered that 80% of cases were classified as high risk, but so were 50% of controls.

Soini(1978) failed to select a high risk group. One-fifth of breast cancer cases were in a low risk group, while the high risk group contained 60% of the population. He classed 5 variables; age at first birth, number of abortions, age at menopause, use of hormones and socio-economic status as statistically significant.

Wolfe(1974) and Krook(1978) looked at parenchymal patterns. They divided the mammogram patterns into four classes:

N1 = parenchymal of mainly fat with small amounts of

dysplasia.

P1 = mainly fat plus prominent ducts involving one-quarter of the breast.

P2 = prominent ducts in over one-quarter of the breast.

DY = severe involvement with dyplasia often obscuring duct patterns.

They found a higher rate of breast cancer in the P2 and DY classes. Egan and Mosteller(1977) sceptically related the relationship to higher false negative rates in these classes due to the density of the tissue.

The incidence in the DY class increased from 4 per 1000 to 40 per 1000 for age groups 35-39 to 60+. In the N1,P1,P2 and DY groups, the incidence per 1000 were 2.1, 6.9, 8.9, and 16 respectively. There appears to be a relationship between the different patterns and age at risk. Incidence is high in the DY class after 40, in the P2 class after 45, in the P1 class after 50 and in contrast the N1 class produce few cancers.

Krook suggests these patterns can be used for a screening programme whereby following an initial screen at 40, those with DY classification are screened frequently, whereas those deemed P2 and P1 only enter screening at ages 45 and 50 respectively.

Farewell(1977) considered high risk groups using family history, log-etiocholanole level, age at menarche and first birth to work out a risk ratio. One risk factor implies double, two 4-5 times, three 8-10 times and four 17 times

the risk of those without any risk indicators. Table 2.3. shows the distribution of patients according to the number of risk indicators.

Table 2.3. % distribution of women according to their risk status and disease status

(Source:Farewell,1977)

No of risk factors	Normal women	Cancer cases
0	10	0
1	35	18
2	38	35
3	15	40
4	2	7

Chamberlain(1982) specifies that it must be possible to concentrate 80% of the cases in 20% of the population if the aim is to only screen high risk women. But as the table indicates this is far from possible. We can only concentrate 82% of cases in 55% of the population or 47% of cases in 17%. If the hormone test was discarded due to difficulties in utilisation, then only 71% could be concentrated in 43% of the population.

The most important contribution to finding a high risk group appears to be from the work of Bulbrook and others for the Imperial Cancer Research Fund.(ICFR,1982 and 1983). They have followed three lines of work:

1. measurement of hormones in blood and urine samples

obtained from patients with breast cancer which were then compared with normal women.

2. hormone measurements in women without breast cancer but who are known to be at enhanced risk and comparison with those not deemed at risk.

3. measurement of hormones in a large population and then following these women to see whether a particular pattern is associated with eventual diagnosis of breast cancer.

There is indirect evidence that oestrogens are prime factors in controlling the number of breast cancer cells upon which carcinogens act and in controlling growth rates of tumour. Results so far indicate significant abnormalities in hormone status preceeding diagnosis:

i. a subnormal excretion of urinary androgen metabolites is related to enhanced risk of breast cancer in pre-menopausal women.

ii there is some evidence that luteal phase progesterone values are low in pre-menopausal cases.

iii. levels of prolactin above the 70th percentile of the normal range are commonly found in post-menopausal women who develop breast cancer.

iv. there appears to be a higher amount of biologically avaliable oestrogen in cases than in controls.

v. there is confirmation that women with high risk parenchymal patterns appear to have four times the risk of those with low risk patterns. Furthermore prolactin levels appear to be raised in women with high risk patterns.

Research is continuing to ascertain whether or not abnormal binding of oestradiol is present before the disease is diagnosed. A new approach is being made using multiple assays and parenchymal patterns with prolactin levels to identify high risk group. Such tests are however both expensive and time consuming to run.

2.4. Screening tests for breast cancer

There are four types of test for the detection of breast cancer - clinical examination(CE), mammography, thermography and ultrasound.

2.4.1. Clinical examination

CE can only detect tumours that have reached a certain size, usually 1-2 cm or more, although careful inspection and palpation have found tumours less than 1cm, especially when reactive changes in the surrounding tissue has occured. Surrounding odema may make the clinical size 30% larger than the mammographic size.

CE is cheap and safe and can be performed by doctors or nurses but is subject to considerable examiner variablility. Using non-medical staff may reduce the cost but result in low sensitivity. (Chamberlain, 1982) The effectiveness is also affected by breast characterisitics. Small tumours are harder to detect in dense breasts found in younger women. Specificity is poor in younger women as well.

2.4.2. Mammography

Mammography is a highly accurate and reproducible procedure given experienced staff, and it is capable of detecting cancers unidentifiable by other means. It involves taking two or three views of the breast. Indication of early malignancy are opacity with a clear halo, microcalcifications and localised increased vascularity.

Mammography can detect small tumours some less than 0.5cm, many represented by only a small cluster of microcalcifications. (Hermann,1982) Occasionally only a small distortion of the tissue indicates the prescence of an abnormality. Dodd(1977) found that 45.6% of the cancers found by mammography were clinically occult.

The disadvantage of the technique are the high initial and ongoing costs and the risk of radiation. These can be reduced by taking one view only which can be just as successful as 2 or 3 views. (Jakobsson & Lungren,1976) To increase sensitivity it may be necessary to re-read films although this will increase costs. Some tumours may be missed on mammography because of location on the edge of the film or of a type that doesn't show clearly by this method. They may be detected by CE though and so mammography should not be used alone.

2.4.3. Thermography

Thermography measures the temperature of the breast. Abnormal features include localised areas of increased heat emmisions, localised increased vascularity and increased

heat in the areolar area. Stark(1976) used thermography as an index of suspicion. Each woman has a thermal pattern which remains stable during her reproductive life. Hence, it can indicate malignant changes even when the lesion is too small to show up on mammography. But there is often a large false positive rate and sensitivity is low. Dodd(1977) suggests thermography could be used to establish a high risk group and as an adjuct to other tests but it should not be used on its own. When combined with CE, sensitivity of over 85% can be obtained. For the under 50s, the stability of the thermal pattern may prove an adequate substitute for more frequent mammogram.

2.4.4.Ultrasound

Ultrasound is able to distinguish between cystic, solid and complex masses. The false negative rate is high when used on asymptomatic women and therefore it is not suitable as a screening test. But when the location is known, the accuracy for tumours over 1cm is 70 to 100% and for those less than 1cm, 55 to 60%. It is possible to locate tumours less than 0.5cm but about 20 to 30 images are required.

2.5. The role of breast self examination(BSE)

The Department of Health has no formal policy on screening for breast cancer. There is an eight year research project to evaluate the feasibility of a national screening programme. (UK Trials of Early Detection of Breast Cancer, 1981) The remainder of Britain's breast

screening services are either in the private sector or run by local initiative. In the private sector BUPA and Private Patient Plan run full-scale body check-ups for £210 including breast and pelvic examinations. The availability of breast examination by the NHS depends not only on the manpower and resources allocated to it by the local health authority, but also on a person's geographical and financial placing. Thus the value of BSE should be considered as an addition or an alternative to other methods of detection.

There is some controversy over the value of BSE. Alcoe(1979) says:

"There is strong circumstantial evidence that at the moment the most practical way to improve the prognosis for those who will develop breast cancer is for all women to examine their own breasts regularly"

Smith(1980) is more reticient;

"Although BSE is held to be valuable and life saving there are problems in evaluating its usefulness. It is not clear that the practise of BSE does lead to a favourable diagnosis"

The University if Southampton and Wessex Regional Cancer Organisation conducted a survey of aspects of detection of breast cancer including doctors attitudes to screening, delays by patients in seeking help and referal delays by doctors.(University of Southampton,1983)

In the survey of 102 Southampton GPs, 55 were strongly

in favour, 43 in favour and 4 neutral for women to examine their own breasts. The same GPs were on the whole against offering women regular screening by mammography and CE. Only 12 were for regular mammography with 72 against. BSE has many advantages: it is free, convenient, needs no special equipment and takes little time. But some doctors are worried about the psychological affects on women including increased anxiety or about the danger of false reassurance. The Southampon study showed that women practising BSE were less reluctant to see the doctor, and did not delay in seeking advice. The study revealed long delays after the onset of symptoms before seeking medical advice. For the 560 women participating in the survey, the median delay when a lump was involved was 12-16 days; with pain but no lump 24 days; with changes in the breast shape 94 days and with puckering of the skin 695 days. The main reason for such delays were 'too worried to go to the doctor' and the assumption that the symptoms were not serious.

Three studies have investigated BSE practise and its affect on tumour size and stage at diagnosis: Huguley(1981), Foster(1978) and Feldman(1981).

Huguley found 67% of 2092 women practised BSE of whom 51% did so monthly and 11% every 2 months. But only 57% were judged to do it correctly.

Foster found that of 246 patients, 25% practised monthly, with a further 28% less often. There was a

significant relationship between age and practise. The correlation was negative with 50% of the over 70s versus 16% of the under 50s never having examined their breasts. Also 5% of the oldest versus one-third of the youngest practised monthly.

Those that practised BSE presented with smaller tumours and an earlier stage of development. Foster found a significant relationship between BSE and lower clinical stage. About half the patients reporting monthly BSE had stage 0 or I diseases as opposed to one-third and one-fifth of those practising less often or never. The international system of staging is: Stage 0 -in-situ; Stage I- tumours less than 2cm; Stage II- tumours 2-5cm but localised; Stage III- regional spread only and Stage IV- distant metastases.

Table 2.4. gives the relationship between BSE practise, stage and method of detection. Mammography gave the greatest percentage of Stage 0 to II tumours. Those who practised BSE had more Stage 0 to II tumours than those detected by CE even when the tumours were found accidently. Of those not practising BSE who found their tumours accidently over one-third had Stage III or IV tumours.

Patients reporting more frequent BSE had fewer positive lymph nodes. Whereas 27% of those never practising had four or more positive nodes those practising monthly or less often had 9% and 17% respectively.(Foster,1981) Huguley(1981) cites 57% of BSE practicioners with negative

Table 2.4. Relationship between BSE, method of detection and stage at diagnosis.

(Source:Foster,1978)

Method of detection	% in stage		
	0 or I	II	III or IV
BSE	27.4	56.8	15.7
BSE group but accidently	25.4	56.2	18.4
CE	30.4	43.9	25.7
Mammogramphy	61.2	31.7	7.1
Accidently(non BSE)	17.5	46.4	36.2

nodal involvement and 19.3% with four or more positive nodes. This compares with 50.2% with negative nodes and 24.1% with four or more positive nodes for those never doing BSE.

Feldman(1981) gives the tumour size at diagnosis according to BSE practise. (See Table 2.5) The mean size for those practising BSE monthly or less often was 2.5cm and for those rarely or never practising was 3.3cm. Huguley gave the size at diagnosis of 944 patients practising BSE as 2.81cm $\pm$ 0.07(SEM) and for 459 not practising BSE as 3.54cm $\pm$ 0.11(SEM) which is significantly different at 95%CI. Foster divided his patients into three groups and gave the tumour size at diagnosis as;

1.97cm $\pm$ 0.22(SEM) for those practising monthy

2.47cm $\pm$ 0.20(SEM) for those practising less often

3.59cm $\pm$ 0.15(SEM) for those never doing BSE.

Table 2.5. % cumulative frequency of maximum tumour diameter and BSE practise.

(Source:Feldman, 1981)

diameter(cm)	monthly or less often	rarely or never
less than 2	56.4	38.6
3	66.4	52.9
4	82.1	69.5
5	90.7	81.2
6	93.6	86.6
7	95.7	91.1
8	97.8	94.7
	(n=140)	(n=223)

The difference between those never and those doing BSE is signifcant at 95%CI. The sample sizes were 52, 59 and 107 respectively.

Table 2.6. gives the cumulative frequency distribution of size at diagnosis using Foster's data and assuming a lognormal distribtion. This seems appropriate since the long tail reflects tumours greater than 9 or 10cm at diagnosis. The standard deviation is calculated as $\sqrt{n} \times SEM$.

These figures suggest that 4/5ths of tumours could be found before they reach 3cm if BSE was practised monthly and this proportion could be detected by 4cm if BSE was practised less often. Opposed to this only 3/5ths of tumours in women not practising BSE are detected by 4cm.

Table 2.6. % cumulative frequency distribution of tumour size at diagnosis according to BSE practise.

(using Foster's data)

diameter	BSE practise		
	monthly	less often	never
less than 1	7	1.7	0.2
2	51.5	31.2	9.2
3	81.8	67.4	34.2
4	93.7	86.9	59.6
5	97.8	95.0	77.3
6	99.2	99.7	87.9
8	99.9	99.7	96.6
10	100	99.9	99.0

Foster does not give details of the accuracy of the techniques used and therefore the mean size of the tumours found by BSE could possibly be reduced by more effective methods. Similarly for those not practising BSE, no details are given of the patient's delay after symptoms have appeared and hence the mean of 3.59cm may not represent the size of the tumour at onset of symptoms.

Schwartz(1978) attempted to evalaute alternative screening strategies for women performing BSE. The reader is refered to the article for comments on the limitations of his model. He concludes:

i. monthly BSE would realise over 50% of the possible improvement in life expectancy that would be realised if

mortality from breast cancer was eliminated. This compares with 57% for mammography and CE yearly and 26% for them every two years.

ii.monthly BSE would detect breast cancer early enough so that 80% of cases would have no metastases developing after surgery in the life-time of the patient. This compares with 72% if yearly mammogram and CE but no BSE.

Foster calls for a need to compare screening programmes involving structured BSE and the more conventional ones. He concludes:

"It appears from our data that each woman in whom breast cancer is destined to develop has in her own hands the possibilty of increasing her chance of survival through performing BSE."

There has been several attempts in the USA to educate women by TV, publications and personal instruction(PI) in BSE. One such project using films reported 7.8% using BSE before the film, but 6 months later 80.7% reported doing BSE at least once. Another project used PI and found that 79% had practised BSE afterwards with 62% reporting doing BSE regularly every two months or less.(Foster,1977) Boyle(1981) initiated a community programme to promote early breast cancer detection. This involved teaching nurses BSE and encouraging them to teach lay women. 97% of nurses and 82% of lay women practised BSE after the programme.

Bone(1978) gives details of women seeking family planning(FPS). Of those married in 1971-75, 78% used FPS with a maximum in the non-manual social classes where the incidence of breast cancer is higher. It should be possible to reach 80-90% of women this way if it was directed at all ages but concentrating on the under 30s. The survey showed an increase in women using the service from 57% in 1970 to 74% in 1975, hence if the trend continues most women could be reached at some stage of their lifetime. It may be argued that since the incidence of breast cancer is low in the under 40s, it is not worthwhile teaching BSE to them. Also such practise may lead to an increase in the number of unnecessary biopsies. But younger women are more ready to accept BSE and it could become an established routine. In addition since the incidence of benign tumours is greater in the under 50s, such a program would allay fears in this age group and encourage women to seek advice promptly.

Miller(1985) in a review of BSE, states that BSE could be applied in a population of women without substantial increase in health resources. He points out that at present there is insufficient evidence that BSE is effective in reducing mortality from breast cancer and he calls for research into the best methods of promotion and evaluation of its effectiveness. Since survival or case fatality comparisions are biased then the only measure is the comparison of mortality from breast cancer in the BSE group and the control group.

2.6. Errors in screening tests

The efficiency and accuracy of a screening test can be measured in several ways. Table 2.7. shows the possible situations which exist after screening. Disease present may or may not be found, while the screening test could incorrectly indicate the presence of disease. The

Table 2.7. Persons classified according to disease status and the results of screening.

		Disease		
		Present	Absent	Total
Screening	Positive	x 11	x 12	M
	Negative	x 21	x 22	M'
	Total	N	N'	T

sensitivity of a screening test measures the true positive rate, that is the degree of success of a test to discover those with the disease. The specificity measures the true negative rate, that is the ability of the test to recognise the absence of diseases. Using the notation in the table gives the sensitivity as x_{11}/N and the specificty as x_{22}/N'

These measures are clearly important in a screening

programme.

The false positive and negative rates may be given as $a = x_{12}/N'$ and $b = x_{21}/N$ respectively. When comparing different tests the predictive value(PV) which is the ratio of true positives to all positives (x_{11}/M) may be used.

We need to define a false negative result for a particular test or group of tests. A person is screened negative but is later found to have the disease. This may occur when the test did not find the disease or the disease developed after the test. In section 1.6. we defined four states: S0 where there is no disease; S1 where the disease is present but not detectable; S2 where the disease is only detectable by screening and S3 where the disease is clinically apparent. A person with the disease which is at too early a stage to be detectable by screening will be classified as negative. Strictly this will be a false negative but for the purposes of this study we will define a false negative result only when the disease has reached S2.

The S1 and S2 states do not coincide for different screening tests. Unless there are clinical changes tumours are not usually palpable less than 1cm in diameter. According to their position they may be missed even when much larger than this. Mammography can detect non-palpable tumours less than 0.5cm. The minimum is usually taken as 0.2cm.(Fournier,1978). But certain types of tumours eg scirrhous may be missed by mammography.(Davey,1976)

A person with a negative result may later be found to have breast cancer. This may be the result of a false negative or that the tumour has grown quickly. The person may have entered S1 and preceeded to S2 since the last screen or have been in S1 at the time of screening. A person with a tumour less than 1cm will usually be negative for CE. The tumour will grow and may clinically surface quite quickly after screening. The false negative result is not a false negative for that test but exists because of the limitations of that test. Thus a person having a negative screen by one of the methods can not be classifed free of breast cancer, but only (most probably) free from breast cancer greater than a certain diameter.

Many studies eg Chamberlain(1979) and Stark(1974) have calculated the false negative rate for screening programmes using the number of interval cancers, that is those found prior to the next screen. This gives a useful estimate but suffers from the inability to classify interval cases as wrongly diagnosed at a previous screen or as new cases with fast doubling time.

A method of estimating the false negative rate and the total number with the disease is given by Wittes & Coulton(1974) and used by Goldberg & Wittes(1978) to estimate the false negative rates for breast cancer screening. The theory is based on a capture-recapture model which estimates the total population by two or more independent samplings. Hence if n_1 cancers are found

by screening test 1 and n_2 by screening test 2, an estimate can be made of the number of cancers and the sensitivities of the two tests.

The presence of disease may be recognised by only one or both of the tests.

If N is the number of persons with breast cancer and k is the number of tests involved, then an approxiamtion for N is $\hat{N}$ where:

$$\prod_{i=1}^{k} (\hat{N} - n_i) = \hat{N}^{(k-1)} (\hat{N} - n) \qquad 2.6.1$$

where n is the number found in total and n_i is the number found by test i.

The sensitivity of the ith test is $p_i = n_i/N$ and the combined sensitivity of all the tests is:

$$P = 1 - \prod_{i=1}^{k} (1 - p_i)$$

$$= n/N \qquad 2.6.2$$

This gives a false negative rate b = 1-P.

For k=2 equation 2.6.1 becomes $N = n_1 n_2 / y_{11}$ 2.6.3

where y are the number positive on the test.

The test sensitivity is $p_i = y_{11}/n_i$

and the combined sensitivity is $P = p_1 + p_2 - p_1 p_2$

We will only discuss mammography and CE screening in future since thermography is not suitable for general screening purposes, but such ideas could be extended to thermography or other tests.

These equations provide a simple estimate of the sensitivities but in the case of breast cancer the method is not viable for all sizes of tumour since mammography and CE are not measuring the same population. The equations can only be used when considering tumours greater than a certain size eg 1.5 or 2cm. They are however useful in ascertaining the sensitivity of two tests in a particular screening programme, that is as a method of quality control. Low sensitivity may indicate poor use of or faulty equipment, or lack of examiner experience. An estimate of the false negative rate for mammography for tumours greater than 1cm, does not indicate a similar rate for all sizes.

Let us look at those in the screened population who have breast cancer and whose tumour is of a size that is detectable by a screening test. We can represent S2 by a simple figure. Let z be the diameter of the tumour when

Figure 2.6.1. The state S2

symptoms appear. This may vary considerably and a simple model would take z as a random variable. We assume the number of prevalent cancers are evenly distributed in the size range (0.2,z). We are not concerned with tumours less than 0.2cm. Then we let:

P = probability that the tumour is in the size range (0.2,1)

1-P = probability it is greater than 1cm

p_1 = probability that a cancer in (0.2,1) is not detected by mammography

p_2 = probability that a cancer greater than 1cm is not detected by mammography

p_3 = probability that a cancer greater than 1cm is not detected by CE.

Two situations exist. We can assume that all cancers greater than 1cm will be found by one of the screening tests. Or as is more likely, some will be missed by both tests. Under the first assumption the probability of diagnosis by one or more of the tests are:

Mammography only	$P(1-p_1) + (1-P)p_3$	2.6.5
CE only	$(1-P)p_2$	2.6.6 2
Both tests	$(1-P)(1-p_2-p_3)$	2.6.7

and the probability of missing the tumour is

$$Pp_1 \qquad 2.6.8$$

Under the second assumption the probabilities of detection are:

Mammography only $P(1-p_1) = (1-P)(1-p_2)p_3$ 2.6.9

CE only	$(1-P)p_2(1-p_3)$	2.6.10
Both tests	$(1-P)(1-p_2-p_3+p_2p_3)$	2.6.11

and the probability of missing the tumour is

$$Pp_1 + (1-P)p_2p_3 \qquad 2.6.12$$

Let us illustrate this with an example. We assume that 30% of tumours not yet found are less than 1cm. We give the probabilities of detecting tumours less than 1cm and more than 1cm by mammography and more than 1cm by CE the values of 0.8, 0.95 and 0.8 respectively. This gives the values P =0.3, p_1 =0.2, p_2 =0.05 and p_3 =0.2. The overall sensitivity for mammography would be 0.905 and for CE 0.56.

Table 2.8. shows the sensitvity and false negative rates under both these assumptions. The false negative rate is similar under both assumptions, giving a combined sensitivty of 0.93. Although the contribution to detection given by

Table 2.8. The probability of detection by mammogram and CE

Detected by	Assumption 1	Assumption 2
Mamm only	.38	.373
CE only	.035	.028
Both	.525	.532
Missed	.06	.067
	1.0	1.0

CE for tumours over 1cm is small, it does add to the overall sensitivity, and should not be neglected when using mammography.

The probability P varies according to the population's history of screening. Assume the size of tumours are evenly distributed up to 1cm in diameter. This would occur given a constant growth rate for each patient and assuming that tumours less than 1cm are unlikely to display symptoms and hence be detectable by the patient. On the first screen the probability of detecting these tumours depend on the probability of detecting each size by the screening test. After the first screen, the size of tumours less than 1cm will be unevenly distributed since large tumours in this interval will more likely have been detected. The value of P at the next screen will depend on the distribution of sizes after the last screen, the growth rate of the tumours, the interval between screens and the number of new tumours which were less than 0.2cm at the last screen.

If a number of years have elapsed since the last screen, then those tumours less than 1cm not detected by the screen will have grown and hence there will be an even distribution, that is an unscreened situation will occur.

The value of P after the screen depends on the type of screening test used. If CE only is used, few cancers under 1cm will be detected. But when mammography is used, then some cancers less than 1cm will be found. hence the value

of P must be re-evaluated accordingly. In a group of refered or self-referred women, some symptoms may be apparent and so these populations will contain more tumours over 1cm than expected, therefore fewer tumours will be discovered by mammography alone. See George(1976) for a comparison of the percentage detected by both screening tests involving referred and invited women. Another consideration is the practise of BSE in the screened population. These women will detect their cancers at an earlier size than others, hence the distribution of z will be different and the ratio of P to 1-P will not be the same. Fewer cancers will be found by screening tests and the number of interval cancers will be higher. Hence the sensitivity of the screening tests will appear poor if inappropriate methods are used.

Nine factors must be considered which will affect the false negative rate for a screening programme. They are: age, type of breast tissue, menopausal status, risk, type of screening test, size of tumour, interpretation by the examiner, time since last screen, and type of cancer and extent of nodal involvement. Menopausal status and tissue type correlates closely with age. Friedman(1966) gives the tissue type according to age. Table 2.9. gives the false negative rate according to clinical signs, menopausal status and type of tissue.

Table 2.9. False negative rates according to certain characterisitics

(adapted from Friedman, 1966)

Factor	Rate
Clincial signs:	
Obvious signs	.17
Dominent masses	.48
Breast complaint	.53
Menopausal status:	
Pre-	.51
Menopausal	.53
Post-	.25
Type of tissue:	
Fibrous	.45
Glandular	.40
Fatty	.20

Table 2.10 uses HIP data to show how sensitivity can vary with age. The small percentage found by mammography reflects the less sophisiticated equipment in use then. But some age difference will reflect the type of tissue and menopausal status.

Table 2.10.% of cancers found according to age

(adapted from Shapiro,1977)

Age	mamm	CE	both	no
40-49	38.7	80.6	19.3	31
50-59	60.0	58.5	18.5	65
60+	61.1	69.4	30.5	36
total	55.3	66.7	22.0	132

Table 2.11. gives the percentage found according to nodal status. these figures reflect the correlation between nodal involvement and size.

Table 2.11. % found by each method according to nodal invovlement.

(adapted from Shapiro,1977)

nodal status	mamm	CE	both	no
positive	63.3	76.7	40.0	93
negative	51.6	63.4	15.0	30

Chamberlain(1979) also indicates the greater success of CE over mammography in the younger age groups. Ommiting CE in the under 50s would miss 46.2% of the cancers as opposed to 23.1% if mamography was omitted. But in the over 60s, the figures were reversed (16.7% and 50% respectively). In younger age groups the incidence of fibrous and glandular tissue is high and hence mammography sensitivity is lower than for older ages.

Both the success of mammography and CE at detecting cancer is subject to exaimer vairability. Boyd(1978) quotes a Candian experience where 9 radiologists were given 100 X-rays to consider. In only 4 cases did all 9 agree rising to 73 cases where at least 5 agreed. George(1976) also gives figures on percentage of cancers missed by medical teams consisting of surgeons and radiologists and non-medical teams consisiting of nurses and radiographers using CE and mammography. Table 2.12. gives the range of results.

Table 2.12. % of cancers missed by different examiners

(adapted from George, 1976)

Examiner	All women	Invited women
Surgeon(CE)	2	12.5
Nurses(CE)	6.5	37.5
Radiologists(mamm)	23.4	18.8
Radiographers(mamm)	18.4	18.8
Medical team(CE/mamm)	1	0
Non-medical team(CE/mamm)	2.5	12.5

The percentage cancers missed was much higher in the invited group, that is where the population was asymptomatic rather than where women attended of their own accord.

With general screening, that is with an asymptomatic population many cancers can be missed by inexperienced staff. False negative rates can be reduced by reducing

examiner variability by using more than one examiner for CE and re-reading X-rays. Venet(1971) quotes the results of clinical re-examination after positive diagnosis by mammography. 67.8% of the masses and 16.8% of the microcalcifications giving a total of 50% were found positive, of which 6 had been positive on CE but one was negative on both.

The sensitivity of a test can only be calculated when considering the size distribution of tumours prior to the test and the ability of that test to detect tumours of a certain size.

2.7.Factors affecting the frequency of screening

In most programmes screening takes place annually perhaps with early recall. This interval is chosen arbitarily and has little relation to the natural history of the disease or other factors. The frequency of screening is influenced by the following:

i. the natural history of the disease involving growth rate and spread of tumour.

ii.the cost of delaying diagnosis in terms of patient survival.

iii. the cost of screening both financially and to the patient.

iv. the risk of screening eg radiation and danger of further tests.

2.7.1. The natural history of screening

A negative result on screening does not mean that a person is free forever, or indeed at present. The optimum screening interval depends on how fast the tumour grows and disseminates. The critical issue is how long it takes for those cancers that will metastise after reaching threshold size to grow to a size where they will spread and hence the probability of survival will decrease. Tumour growth and the time of dissemination varies considerably between patients. Metastases have been found when the tumour is still less than 1cm, whereas some tumours grow large and do not seem to metastise. In the case of fast growing tumours, annual screening may not be effective. The screening interval should not be the same for all patients but adjusted to risk factors and mammographic appearance.

2.7.2. The cost of delaying diagnosis

In general the larger the tumour and the greater the extent of nodal involvement, the poorer the prognosis. The survival rate is greater when there is no nodal involvement, less when three or less nodes are affected and poor if four or more nodes are involved. Survival rates varies in each category according to tumour size. See Table 2.13.

If the tumour is slow growing and not prone to metastases the cost of delay is low. But if we have a fast growing tumour very prone to spread then the cost of delay

Table 2.13. % of breast cancer 5 year survival in relation to size and nodal involvement.

(Source: Seidman, 1972)

nodal invovlement	size (cm)		
	2	4	6
negative	82	70	75
1 to 3	68	62	56
4+	58	26	22

is very high. For tumours that metastasise before the tumour is detectable, screening has no value at present.

The cost of delay must be considered in relation to the success of treatment for various stages of breast cancer. A new treatment may increase the prognosis even when spread has occured to 3 nodes, thus spread to more than three nodes becomes the critical stage. Against this one must consider the cost of treatment and debiliating effects which will increase proportionally to tumour size and spread.

2.7.3. The cost of screening

Apart from the initial setting up costs and the cost of each screen, the increase in workload must be considered. Each screening test takes time to administer and to analyse the results. Where the tests are not highly specific, the number of referals will be high and the added costs of further tests will be involved. When the sensitivity is

high most cancers will be found, and hence if rescreening takes place too quickly, the yield will be very low since small cancers will not have developed sufficiently to be found. On the other hand, if sensitivity is low a rescreen would be necessary to pick up the missed cases. The demand on women of frequent tests must also be considered.

2.7.4. Risks of screening

Although some screening tests have no risk, radiation from mammography is considered by some to have a cancer inducing effect. Further procedures such as biopsies and operations for benign conditions also carry a risk.

The assumption that radiation can induce cancer is based on the increased incidence of breast cancer in Japanese women exposed to the atomic bomb, those treated for benign disease and those subjected to repeated floroscopy for TB.

Estimates of the number of cases induced vary according to dose-response relationship. A National Academy of Science Report in 1972 estimated 6 extra cases induced per million per year per rad tissue dose after a 10 year latent period.

Some believe that there is no safe dose but that cases induced are proportional to dosage.(Breslow,1977). The amount of radiation is related to the medium and also the number of views taken. George(1976) records radiation levels as high as 4.9 for standard industrial film dropping to 0.18 rads for Trimax XD with a rare earth phosphorus

screen. Dodd(1977) gives an absorbed dose per exposure of .15 to .19 rads. He estimates that 13 annual mammograms could be performed before the natural lifetime risk increased from 7 to 8% given an exposure of 1 rad.

Strax(1978) makes the following points in relation to radiation hazard:

i. there has been no evidence to suggest that mammography has produced cancer or led to an increase in incidence.

ii. anxiety refers to accumulated radiation from repeated exposure in special circumstances. There is no evidence that the occassional use of mammography would be hazardous.

iii. extrapolation may imply doses of more than 5 rads could be hazardous over many years, but whether there is a minimum safe dose or whether low doses accumulate is controversial.

Baal(1978) observed the ratio of cancers in irradiated breast to the expected number to be 5 to 1. The mean latent period was 23.6 years and was influenced by radiation levels. Since doses of 252 to 2250 rads were involved and the latent period was long the implications are perhaps irrelevant.

Finally, a cancer induced is not the same as a death; if screening continues it should discover the tumour at an early stage. Lives saved at sixty should be balanced against cancers induced at 70 plus.

CHAPTER 3

THE AGE OF ONSET DISTRIBUTION OF BREAST CANCER

3.1. Introduction

Chapters 1 and 2 have endeavoured to provide a background to breast cancer and the problems of screening for it. In a consideration of models for screening relationships between age, size of the tumour and duration in state S2, when a tumour is detectable by a screening test but has not manifested itself clinically are needed. There could be a considerable delay between the actual onset of the disease and diagnosis. In order to evaluate a screening policy it is necessary to investigate these relationships as well as the error in screening tests related to the size of tumour and its age.

3.2. The age distribution of breast cancer patients at diagnosis of the disease.

The OPCS publish data on the incidence of breast cancer as the registered number of cases in a population per age group for a specific year. Using the methods of Elandt-Johnson(1977) it is possible to construct a table giving the age distribution of registered cases of those who will develop breast cancer.

In Table 3.1. I have used the breast cancer incidence data for 1978(OPCS,1983) and the abridged life tables for 1977-79 for England and Wales. (Annual Abstract of

Statistics, No 118, 1982)

This gives the number of women at risk in each age group. The following notation is used:

$I_{n\,x}$ is the incidence rate in the age group x to x+n

$L_{n\,x}$ is the expected total number of years lived between ages x to x+n. These values were calculated using $L_{5\,x} = 5(l_x + l_{x+5})/2$ where l_x are the number surviving to age x.

$Q_{n\,x}$ is the probability of the disease diagnosed between age x to x+n over lifetime risk.

This is given by

$$\frac{I_{n\,x} \cdot L_{n\,x}}{I_x \cdot L_{n\,x}} \qquad 3.2.1$$

and F(y) is the probability of diagnosis before the age x+n.

Table 3.1. Age distribution of those who will develop breast cancer

Age group	I 5 x	L 5 x	I . L 5 x 5 x	Q 5 x	F(y)
0-4	0.1	493005	49300	0	0
5-9	-	492502.5	0	0	0
10-14	-	491850	0	0	0
15-19	0.2	490960	98192	0.0002	0
20-24	1.8	489910	881838	0.0015	0.002
25-29	6.4	488555	3126752	0.0055	0.007
30-34	20.3	486510	9876153	0.0173	0.025
35-39	52.7	483127.5	25460819	0.0446	0.069
40-44	94.9	477365	45301939	0.0794	0.149
45-49	146.0	467782.5	68296245	0.1198	0.268
50-54	149.2	453067.5	67597671	0.1185	0.387
55-59	160.8	431172.5	69332538	0.1216	0.509
60-64	176.6	398837.5	70434703	0.1235	0.632
65-69	201.3	351307.5	70718200	0.1240	0.756
70-74	212.0	283612.5	60125850	0.1054	0.861
75-79	225.3	196742.5	44326085	0.0777	0.939
80-84	237.9	97500*	23195250	0.0407	0.980
85+	306.6	37500*	11497500	0.0202	1.0
Total	1992.1	7111308	570319040	1.0	

* estimated using English Life Tables No 13 1970-72

Initially I fitted both a Weibull and a lognormal distribution to the whole age range using regression methods on F(y). A lognormal function gave a poor fit to the data and was rejected. The Weibull function W1(x) with scale parameter 64.88 and the shape parameter 6.05 is shown in Table 3.2. This also gave a poor fit especially in the age range 40 to 60.

Since evidence suggests that pre- and post- menopausal cancer may be two different diseases, I considered using two Weibull functions to model the data with an age split around about the menopause. The function W2(x) and W3(x) were fitted by regressing from ages 25 to 50 and 50 to 85 respectively. The parameters of 57.788 and 7.425 for W2(x) gave a good fit to the data, with a maximum difference between observed and expected values of 0.006. The Kolmogrov-Smirnoff test gives a D value at the 0.05 level of 0.08. The parameters of 64.426 and 4.696 for W3(x) gave a good fit to the data with a maximum difference between observed and expected values of 0.018. The K-S test gives a D value at the 0.05 level of 0.03. Therefore we can say that the data is consistent with the two Weibull distributions given above.

Table 3.2. Weibull distributions for breast cancer rates.

Age less than	F(y)	W1(x)	W2(x)	W3(x)
15	0	0	0	
20	0	0	0	
25	.002	.003	.002	
30	.007	.009	.008	
35	.025	.024	.024	
40	.069	.052	.063	
45	.149	.103	.145	
50	.268	.187		.262
55	.387	.308		.379
60	.509	.464		.511
65	.632	.636		.647
70	.756	.795		.772
75	.861	.910		.870
80	.939	.971		.937
85	.980	.994		.975

The two functions meet for age 47.93. Hence I have represented the age at diagnosis by a composite Weibull distribution:

$$F(x) = \begin{cases} 1 - \exp(-x/57.788)^{7.425} & x < 47.93 \\ 1 - \exp(-x/64.426)^{4.696} & x \geq 47.93 \end{cases}$$

3.3. Tumour size at discovery

The incidence data on breast cancer does not take into account the size of the tumour. Hence for any investigation into the duration of S2, tumour size must be taken into account.

I have used breast cancer data from Edinburgh.(Kerr, 1982). The sizes at diagnosis for 2116 cases registered between 1974 and 1978 are given in Table 3.3.

The distribution of tumour sizes can be represented by a lognormal distribution. The parameters found by graphical methods are a log-mean of 1.197 and a log-standard deviation of 0.6152.

There appears to be a higher than expected number of cases of tumours less than 1cm on diagnosis. The maximum difference between the observed and the expected difference values using the lognormal distribution is 0.027 (or 0.018 for tumours with diameter 1cm or more.) The Kolmogrov-Smirnoff test gives a D value at the 0.05 level

Table 3.3. The distribution of tumour sizes on detection

(data: Edinburgh 1974-78)

Tumour size (cm) less than	No of cases cum. freq.	F(x)	Lognormal distribution
1	113	.053	.026
2	434	.205	.206
3	904	.427	.436
4	1310	.619	.622
5	1613	.762	.749
6	1801	.851	.833
7	1900	.898	.888
8	1983	.937	.924
9	2007	.948	.948
10	2056	.972	.964

of $1.36/\sqrt{n}$ which for n = 2116 is 0.030. Therefore we can say that the data is consistent with a lognormal distribution with parameters of 1.197 and 0.6152.

Foster(1978) gives the tumour sizes at discovery for various BSE practises. For those who practised monthly, the fitted distribution had a mean of 1.97 and standard deviation of 1.586. The corresponding figures for those practising less often or never were 2.47 and 1.536, and 3.59 and 1.552 respectively. Foster does not differentiate between pre- and post- menopausal patients. An anaylsis using the Edinburgh data showed no significant difference

between menopausal status and size at diagnosis.

Figure 3.3.1. shows the percentage of tumours found for a particular size using the Edinburgh data and Foster's data. This clearly emphasises the importance of BSE, even if practised less regularly, in detecting tumours at a small size.

3.4. Tumour doubling time

In 1.4 I discussed the various theories on the growth of tumours. The most detailed data relating to the tumour growth is Kusuma(1972), although his mean doubling time is low in comparision to that given by others.

Table 3.4. gives the distribution of doubling times and a fitted Weibull distribution with parameters 4.927 and 0.896. This gives a reasonable fit to the data and is not rejected by the Kolmorgov-Smirnoff (K-S) test. The maximum difference between the actual cumulative frequency and the fitted one is .031. The 5% critical point for the K-S statistic is .096.

Kusuma also gives doubling time according to age. I investigated whether doubling time is significantly different before and after menopause. Table 3.5. shows the distribution of doubling time according to age less than 50 or 50 and over. The split at 50 is consistent with a median age at menopause of 49.75 (Bone, 1978).

Figure 3.3.1. Percentage of tumours found by a certain diameter.

(Source: Kerr, 1982; Foster, 1978)

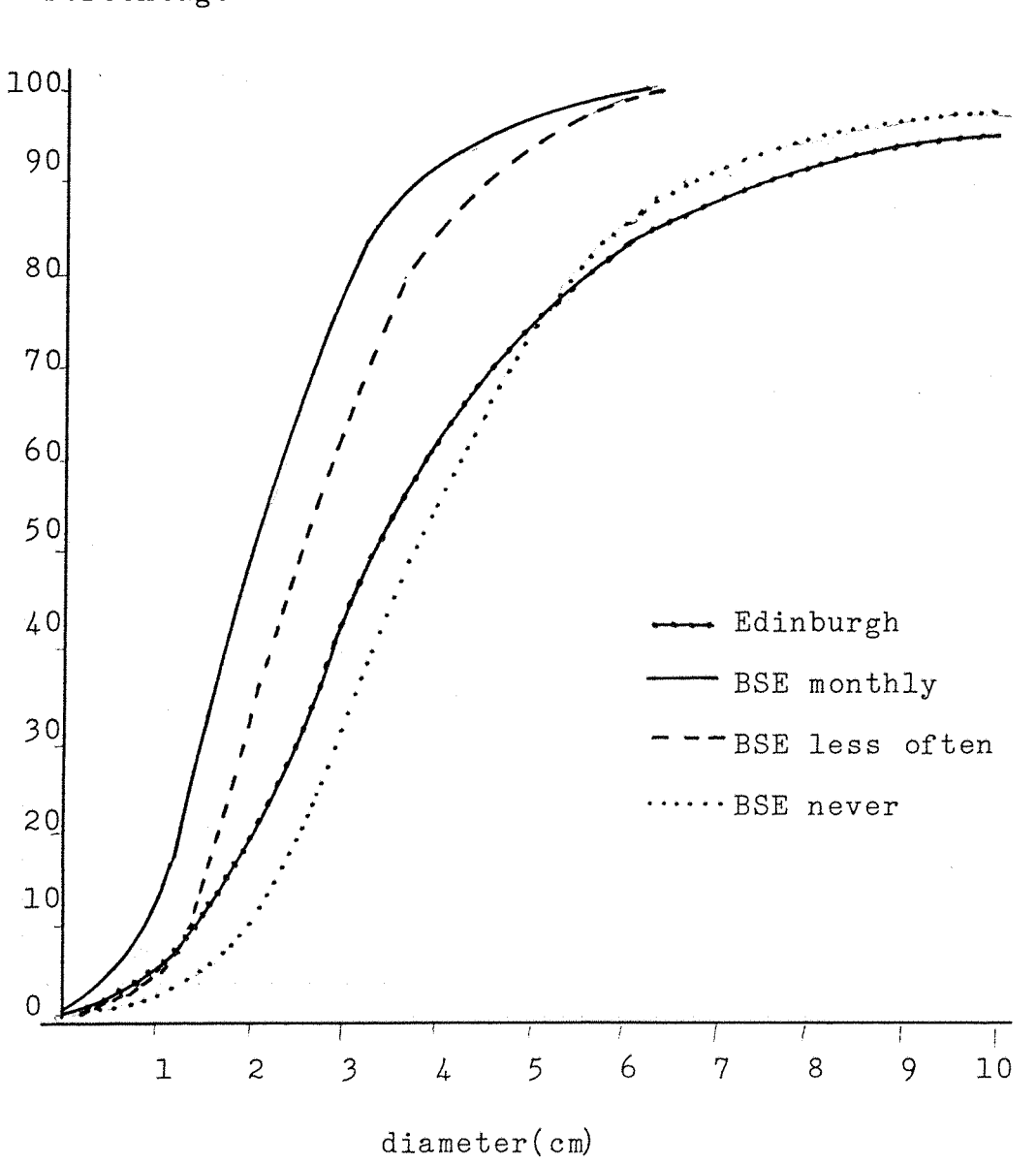


Table 3.4. Doubling time of tumours.

(Source:Kusuma, 1972)

Doubling time (months) less than	Cum. Freq. (n=199)	Weibull cum freq.
1	.186	.213
2	.387	.360
3	.493	.473
4	.593	.564
5	.658	.637
6	.714	.697
7	.759	.746
8	.779	.786
9	.789	.820
10	.819	.848

Table 3.5. Doubling time of tumours according to age.

(Source: Kusuma, 1972)

Doubling time (months) less than	Cum. distribution according to age			
	under 50s		50 and over	
	Obs.	Exp. A=3.45	Obs.	Exp. A=5.18
1	.242	.252	.158	.175
2	.500	.440	.331	.320
4	.652	.686	.564	.538
8	.818	.902	.759	.786

Initially I fitted a Weibull distribution to the data and the parameters 3.781 and 0.846, and 5.227 and 1.020 for the two groups gave a good fit. Since the shape parameters were close to 1, I fitted the data equating the second parameter to 1 and making the cumulative distribution function an exponential:

$$F(x) = 1 - \exp -(x/A) \qquad 3.4.1$$

A value of A = 3.45 gave a good fit for ages less than 50, with a maximum difference between observed values and fitted distribution of .084. The K-S statistic gave a value of .167 at the 5% significance level for a sample of 66. For ages over 50, A = 5.18 gave a maximum difference of .027. The K-S statisitc for n = 133 was .118. These means are significantly different reflecting perhaps the effects of different hormone levels on tumour growth.

3.5. The duration in state S2

The duration of a tumour in S2, or the preclinical stage has important implications for a screening programme. If the duration is short, it is unlikely that a screening programme will detect the tumour. If the duration is long then most tumours will be detected by screening and the lead time in detecting the tumour will be substantial.

Monte- Carlo methods were used to estimate the duration in state S2. The start of S2 - when a tumour is first

detectable by any screening test is taken as 0.2cm since there are references to tumours less than 0.5cm being detected and as small as 0.2cm. (Fournier,1978. Spratt, 1979). Improvements in screening tests should make the discovery of smaller tumours more common. The size at diagnosis can be given by the Edinburgh data's parameters given in section 3.3. Or if we want to consider the length of S2 when women practise BSE or not, then Foster's data (see section 2.5.) can be used for size at diagnosis. Independance of size at diagnosis and tumour doubling time is assumed. The duration of S2 was considered using two doubling times to portray pre- and post-menopausal diseases.

The random variables for the size at diagnosis distribution were generated using the Box-Muller method as described in Tocher(1963). Exponentiating produced the log-normal distribution. The Weibull cumulative distribution was generated using its inverse function. In addition to reduce variability, antithetic variables, that is for f(x) also use 1 -f(x), were incorporated.

Tables 3.6. and 3.7 show the duration in S2 for pre- and post-menopausal women using the size at detection from the Edinburgh data and the doubling time distributions with means of 3.45 and 5.18 months. The calculations were also repeated using the size on detection for various BSE practises.

Table 3.6. Duration in S2 for pre-menopausal women.

Time	Edinburgh	Cumulative distribution for BSE		
(years)	data	monthly	less often	never
1	.129	.288	.194	.108
2	.348	.442	.397	.326
3	.519	.603	.564	.501
4	.636	.714	.675	.625
5	.723	.802	.763	.711
6	.795	.864	.836	.781
7	.849	.901	.880	.842
8	.885	.927	.910	.878
9	.909	.947	.932	.906
10	.931	.959	.948	.927

Table 3.7. Duration in S2 for post-menopausal women

Time	Edinburgh	Cumulative distribution for BSE		
(years)	data	monthly	less often	never
1	.089	.158	.133	.074
2	.253	.321	.288	.234
3	.389	.466	.427	.370
4	.497	.567	.532	.484
5	.577	.653	.619	.562
6	.647	.727	.688	.636
7	.707	.787	.747	.694
8	.762	.832	.799	.746
9	.805	.869	.843	.793
10	.840	.894	.871	.830

This measure is important in ascertaining the likelihood of a screening test discovering the disease and the likely lead time for different screening intervals. If the screening interval is too short, then screening is less likely to detect the cancer than when the screening interval is long. Also when S2 is short, screening will not advance diagnosis by much and hence costs may outweigh benefits. But when S2 is long, then screening detects the cancer much earlier than the patient and hence it could be more effective at changing prognosis.

From the tables the data suggests the following:

1. For pre-menopausal patients 13% have a screening period less than one year and a half over 3 years.
2. For post-menopausal patients 8.9% have a screening period less than one year but a half over 4 years.
3. For those practising BSE monthly or less often in the pre-menopausal group 23% and 19% respectively had a screening period less than one year. Whereas half had a screening period of 2.4 and 2.7 years respectively.
4. In the post-menopausal group, 16% of the BSE monthly patients and 13% of those practising less often were in state S2 less than a year. Half of each group had a duration of over 3.3 and 3.8 years respectively.
5. Of those not performing BSE, who were pre-menopause, 11% had a screening period less than one year and 50% more than 3 years .

6. Of those not performing BSE, who were post-menopause, 7% had a screening period less than one year and 50% more than 4.6 years.

1.6. The age at menopause

Since pre- and post-menopausal breast cancer display significantly different doubling times it is important to take into account the menopausal status of patients according to age.

Table 3.8. shows the percentage of women in the pre-menopausal state at a certain age. A Weibull distribution with parameters 51.28 and 18.54 has been fitted to the data.

Table 3.8. Percentage of women pre-menopausal at a certain age

(Source: Bone, 1978)

Age	observed	expected
41	.98	.98
43	.97	.96
45	.91	.92
47	.84	.82
49	.57	.65
51	.37	.40
53	.19	.16
55	.03	.03

(n=1266)

3.7. The age at onset distribution of breast cancer

The age of onset distribution is defined for 0.2cm tumours since before this size, the natural history of the disease is not known and the cancer is not detectable. Calculation of this distribution involves the age at diagnosis distribution. According to the menopausal status different doubling times were considered. A similar calculation was made by de Senna(1983) with tumour size of 0.5cm and one doubling time distribution for all ages. Monte-Carlo methods were used as described in section 3.5. The results are shown in Table 3.9.

A Weibull distribution W1(y) was fitted over the whole range but did not give a good fit for ages 40 to 60, and the maximum difference is greater than the value of the K-S statistic at the 5% significance level for a sample of 8000.

The Weibull distributions W2(y) and W3(y) were fitted by regressing over the ages 25 to 40 and 45 to 90 respectively. These gave good fits over the two age ranges. By the K-S test the data is consistent with Weibull distributions with parameters of 53.93 and 5.64, and 58.944 and 4.189.

Table 3.9. The age of onset distribution for those who will get breast cancer

Age less than	F(y)	W1(y)	W2(y)	W3(y)
15	.002	.002	.001	
20	.006	.006	.004	
25	.013	.017	.013	
30	.035	.039	.036	
35	.084	.079	.084	
40	.171	.142	.169	
45	.285	.234		.276
50	.387	.352		.395
55	.521	.492		.527
60	.654	.637		.659
65	.771	.771		.778
70	.871	.875		.872
75	.938	.943		.936
80	.974	.979		.973
85	.991	.994		.990
90	.997	.999		.997
Weibull		a: 59.801	53.93	58.944
parameters		b: 4.657	5.64	4.189

The two functions are equal for age 41.72. Therefore I have taken a composite function to represent the age at onset distribution:

$$F(y) = 1 - \exp -(y/53.93)^{5.64} \qquad y<41.72$$

$$= 1 - \exp -(y/58.944)^{4.189} \qquad y >=41.72$$

The figure of 41.72 does not represent the age at menopause but possibly the age at which hormone changes which lead to menopause start to occur. The age at diagnosis distribution given in 3.2 was also a composite with a change of parameters at 47.93. The difference of 6.21 years between these two figures may be consisitent with the time of transition from a pre- to a post-menopausal state.

3.8. The false negative rate of a screening test.

Some authors have quoted false negative rates for screening tests by using the number of interval cases. See Table 3.10. Two points must be considered in conjunction with these figures:

i. interval cases may consist of false negatives and fast growing tumours. it is not possible to differentiate between the two.

ii. slow growing tumours missed on one screen may not be detected until the next screen For example, a large tumour

Table 3.10.FNR of screening test using interval cases.

Source	Cancers found	FNR	Method of estimation on screening *
Venet(1977)	132	.26	47 interval cancers
Dodd(1977)	836	.13	123 interval cases
		.08	if 55 of 123 cases which only showed on mamm were defined as new cases giving 68 interval cases.
Chamberlain	29	.15	5 interval cases
		.29	if 5 cases detected on 6 month, 1 year and 2 year rescreen are classed as false negative.

*by mammography and CE

detected on a subsequent screen could either be fast growing or have existed at the last screen. Thus interval cases may not represent all the false negatives at previous screens.

Chamberlain considered these two points. Of the 5 interval cases, abnormalities at the previous screen could be detected in 4 cases when films and medical notes of CE were examined. Since rescreens occured at 6 monthly intervals on two occasions and then after a 12 month

interval, it is possible to classify all cases detected after the first screen as false negatives. No details are given to the actual sizes of these 5 tumours so no comment can be made as to the likelihood of these having been missed previously. An alternative calculation of the relative false negative rate can be made by using the total number of cancers detected by either screening test. This rate could be regarded as the minimum FNR. Table 3.11 shows the wide variation of FNR for mammography and CE in different screening programmes.

Table 3.11. Relative false negative rate using the total number of cancers detected by both screening tests

Source	no of cancers detected	FNR	
		CE	mamm
Shapiro(1977)	132	.33	.45
Chamberlain(1979)	29	.35	.30
George(1976)	16	.12	.18
Dodd(1977)	836	.47	.07
deWaard(1978)	111	.58	.01
Bears(1979)	683	.52	.09

Burns(1979) reports on patients diagnosed with breast cancer between 1971-77. 80 out of 613 breast cancer cases were classifed as negative by mammography, 30 of these had CE performed with mammography. Of these 18 had obvious clinical signs. The rest were either classified as

suspicious or had a family history of breast cancer. Hence biopsies were performed which confirmed the disease.

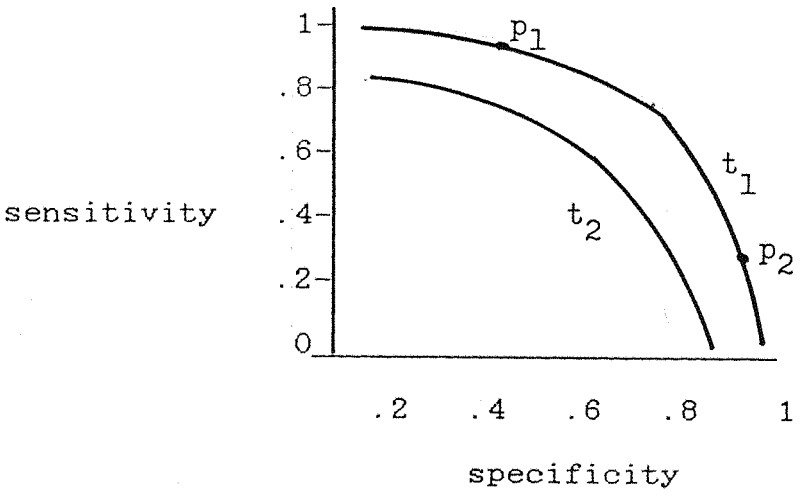
In the other 50 cases, a lump was found by the patient but the mammography showed negative. Delays occured; often a 3- month recall was ignored by the patient or the doctor delayed because of false reassurance from a negative mammogram. Delays especially occured in younger patients when the radiologists were less likely to recommend biopsies. Burns concludes that mammography is likely to have a false negative rate of at least 5-7%.

3.8.1. False negative rate versus false positive rate

A decrease in false negative rate of a screening test may be reflected in an increase in the false postive rate. This is because to avoid missing cases of the disease, a decision may be made to investigate further a case where there is some measure of suspicion of disease. Thus there will be an increase in the number of biopsies but few more cancers found. This leads to an increase in the costs of screening per cancer detected.

Figure 3.4.1. shows the curve relating sensitivity to specificity. Sensitivity is 1- false negative rate and specificity is 1- false positive rate. The t_1 and t_2 curves represent different screening tests eg mammogragaphy and CE. At point p_1 sensitivity is high but specificity is low. At point p_2 specificty is high but sensitivity is low. Decisions must be made on which point along the curve

Figure 3.4.1.Sensitivity versus specificity



gives a satisfactory balance between the two measures. Section 2.6. mentions which factors affecting the sensitivity or false negative rate can be controlled. Decisions like taking more than 1 view on mammography, re-reading mammograms by more than one examiner, further investigation of suspicious cases will increase sensitivity. Each decision will result in further costs The last two will result in more cases being classified as positive and going forward for more tests. This will probably result in more false positives, that is a lower specificity. The optimum trade off between the two will depend on the cost of further tests opposed to the future costs of not detecting the tumour at an earlier stage.

An example of whether to investigate suspicious results further is given by Rombach(1980) who gives 6 classfications of mammography according to an index of

suspicion. Recommendation for biopsy for cancers in different categories influences not only the number of detected cancers but also the number of false positives. Table 3.12 shows the 6 classifications and the distribution of 14695 mammograms. 100 out of 108 cancers were classified

Table 3.12 Classification of mammograms

(Source:Rombach, 1980)

grade	classification	no of cancers	not cancers	total
5	malignant	50	6	56
4	suspicious	50	157	207
3	1/2 year recall	5	196	201
2	benign & palpable	1	195	196
1	benign & not palpable	1	4396	4397
0	no abnormalities	1	9637	9638
	total	108	14587	14695

either as obvious or suspicious, with only 163 out of 14587 disease-free patients deemed in these categories. The sensitivity and specificity of recommending further categories are given in Table 3.13

Attempts to increase the sensitivity in a population based screening programme will involve lowering the diagnostic threshold leading to a higer biopsy rate unless additional tests eg hormone assays could be used to help confirm or reject suspicion.

From Table 3.13 it can be seen that to increase

Table 3.13. Grades of suspicion, sensitvity and specificity.

grade	cum. no of cancers	sensitivity	biopsy benign	specificity
5	50	.46	6	.999
4,5	100	.93	163	.898
3,4,5,	105	.97	359	.975
2,3,4,5	106	.98	554	.962
1,2,3,4,5	107	.99	4950	.661

sensitivity from .93 to .97 decreases specificity from .989 to .975 and to increase sensitivity to .98 decreases specificity to .962. The implications of these figures are best considered by looking at the extra number of biopsises to find one more cancer. Including classification 3 for a biopsy involves only finding 5 more cancers in the 201 extra cases. Moving up to the next group and biopsing those classified in number 2 only finds one more cancer at the expense of 196 biopses.

3.8.2. False negative rate and tumour size

One of the chief factors affecting the FNR of a test is the size of the tumour. There is little data available on the detection probabilities for different sizes of tumours. Feig(1978) presents data on the number of tumours of different sizes found by mammograms and/or CE. See Table 3.14.

Table 3.14. No of tumours found according to size.

(Source:Fieg,1978)

found by	diameter(cm)				
	less than 0.5	0.5-1	1-2	2-3	over 3
mamm	8	22	27	14	5
CE	1	15	26	10	3
both	1	5	13	9	3
total	8	32	40	15	5

Using Wittes and Cotten's formula (equation 2.6.4) we can estimate the number in the screened population for each size group and hence the sensitivites of each test. Table 3.15 shows the results for sizes over 1cm, since the method is only valid when both tests are able to detect the size range. Size over 3cm is omitted since the sample is small.

The sensitivities for mammography are based on techniques used in the 1970s which have since been improved

Table 3.15. Sensitivity of different test according to size of tumour

test	diameter(cm)	
	1 to 2	2 to 3
mammography	.50	.90
CE	.48	.64
both	.74	.96
no. of cancers	54	15.6

and so the figures may be lower than expected today. The CE sensitivities for over 1cm seem low especially when compared with the size at diagnosis of tumours of those practising BSE. One explaination of this is that women who are aware how to perform BSE correctly are better at detecting changes or lumps in their own breasts than medical persons.

When comparing sensitivities of tests according to size of tumour it is important to realise that size on CE or mammogram may not be the actual size of the tumour. This is especially true for CE where surrounding odema may increase the size by up to 30% Tables 3.16 and 3.17 record the size estimated at the test versus the actual size on biopsy.

3.16. Size on clinical examination versus actual size

(Source:Rombach,1980)

		actual size(cm)						
		-1	-2	-3	-4	-5	5+	total
	-1		1					1
	-2	1	7	1				9
CE	-3	1	8	4				13
size	-4		2	2				4
	-5		7	1				8
	5+		3	2		2	1	8
	total	2	28	10		2	1	43

3.17. Size on mammography versus actual size

(Source:Rombach,1980)

		actual size(cm)							
		-.5	-1	-2	-3	-4	-5	5+	total
	-1	2	5	2	1				10
	-2		8	24	6			1	39
Mamm.	-3		1	11	7				19
size	-4			2			1		3
	-5		1	1	1		1		4
	5								+
total		2	15	40	15		2	1	75

For mammography 49% appeared the correct size on film and 35% appeared larger. For CE only 28% were estimated correctly with 67% feeling larger. In addition 37% of CE detected tumours were 2 sizes or more larger as opposed to 8% of mammographic detected ones. Calculating the Chi-squared test for CE and mammography gave values of 1.97 (.75< p <.90) and 0.623 (.50< p <.75) respectively. The above results must be considered bearing in mind that no analysis has been made to see if estimations of size are influenced by the examiner's experience of errors in estimation between feeling or viewing and actual size. Secondly, there is a delay between mammography or CE and biopsy. Such delay means that the tumour will have grown and so size on biopsy will be greater than size at the screening time. This should be reflected in larger tumour

sizes on biopsy than estimated by screening. Clearly, this is not displayed by the tables.

The sensitivity of mammography at detecting certain sizes of tumours may be more consisitent for all tumours of that size, but the sensitivity of CE to detect a particular size of tumour will vary considerably depending on changes in surrounding tissue.

3.8.3 Modelling false negative rates

For very small tumours the probability of detection is 0 and for large tumours almost 1. While tumours are small the FNR will decrease slowly but as they increase in size then the FNR may decrease more rapidly. See Figure 3.8.1.

One possible representation of this is:

$$b(x) = \exp - (x/c)^{d} \qquad 3.8.1$$

where x is the size of the tumour and c and d are parameters.

Figure 3.8.1.

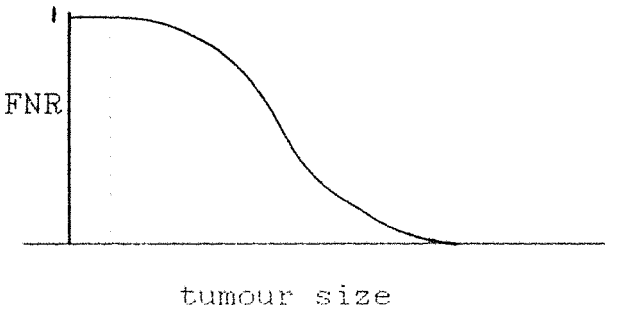


Table 3.18 shows the FNR for different sizes of tumours using various parameters.

Table 3.18. FNR according to tumour size for differing parameters

parameters	diameter(cm)						
	0.5	1	1.5	2	2.5	3	4
.8 , 2	.68	.21	.03	.002	0	0	0
1 , 1.5	.70	.37	.16	.06	.02	.01	0
1.2 , 1.3	.73	.46	.26	.14	.07	.04	.01
2 , 1.5	.88	.70	.52	.37	.25	.16	.06
2.5 , 2	.96	.85	.70	.53	.37	.24	.08

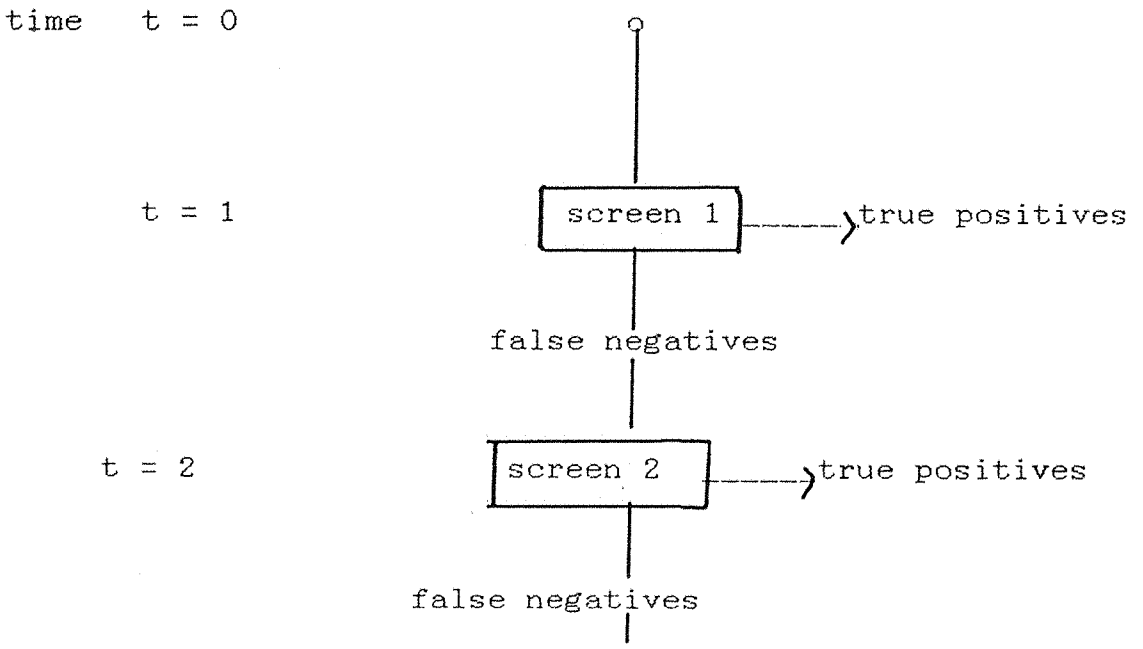
Using parameters .8,2 gives a highly sensitive test especially for tumours over 1cm. With parameters 1,1.5, the FNR declines more slowly with increasing size but with 1.2,1.3 the FNR remains quite high even for tumours over 2cm. With parameters 2,1.5 and 2.5,2, the FNR is high for tumours up to 1cm. The former FNR decreases more rapidly than the latter but still has a FNR of about 7% even when tumours are over 3cm. The second and third examples may be consistent with screening by mammogram, whereas the last two examples may represent screening by clinical examination.

3.8.4. False negative rate and tumour age.

The false negative rate for a particular screening test at time t, measured from the time the tumour is first detectable by a screening test, depends on the distribution of tumour sizes at time t plus the ability of a screening test to detect tumours of different sizes.

Figure 3.8.2 shows the situation at time t_1 and t_2 assuming that tumours are only detected at screening. The distribution of tumour sizes at screen i is $s_i(t)$. Some tumours will be found, the others will continue to grow and

Figure 3.8.2.



a new distribution of tumour sizes will exist before the next screen.

The percentage of tumours less than diameter d at time

t years after initiation is given by:

$$F(d,t) = \exp - DT/A \qquad 3.8.4$$

where

$$DT = \frac{t}{4.33(\mathrm{Ln}\ d - \mathrm{Ln}0.2)} \qquad 3.8.5$$

and A is the mean doubling time of the tumours.
Table 3.19 and 3.20 give the distribution of tumour sizes after t years assuming no self-detection has occured.

These tables are included here to show at a glance what percentage of tumours are below 1 cm, t years after initiation. Thus in the first two years just over one third of the faster growing ones and over a half of the slower

Table 3.19. % frequency of tumour sizes with doubling time of 3.45 years

	size(cm)				
no of years	0.2-1	1-2	2-3	3-4	4+
1	60.7	9.8	3.8	2.2	23.8
2	36.8	12.9	5.5	3.2	41.5
3	22.4	12.7	6.0	3.6	55.3
4	13.6	11.2	5.8	3.7	65.7
5	8.2	9.3	5.2	3.5	73.9
6	5.0	7.3	4.5	3.1	80.0

Table 3.20. % frequency of tumour sizes with doubling time of 5.18 years

	size(cm)				
No of years	0.2-1	1-2	2-3	3-4	4+
1	71.7	7.6	2.8	1.6	16.4
2	51.4	11.4	4.5	2.6	30.1
3	36.8	12.9	5.5	3.2	41.5
4	26.8	13.0	5.9	3.6	51.0
5	19.0	12.3	6.0	3.7	59.0
6	13.6	11.2	5.8	3.7	65.7

growing ones may not be easily detected by clinical examination. If we assumes self detection is very likely in tumours over 4cm then over half the tumours should be detected within 2 1/2 years if fast growing and 4 years if slow growing.

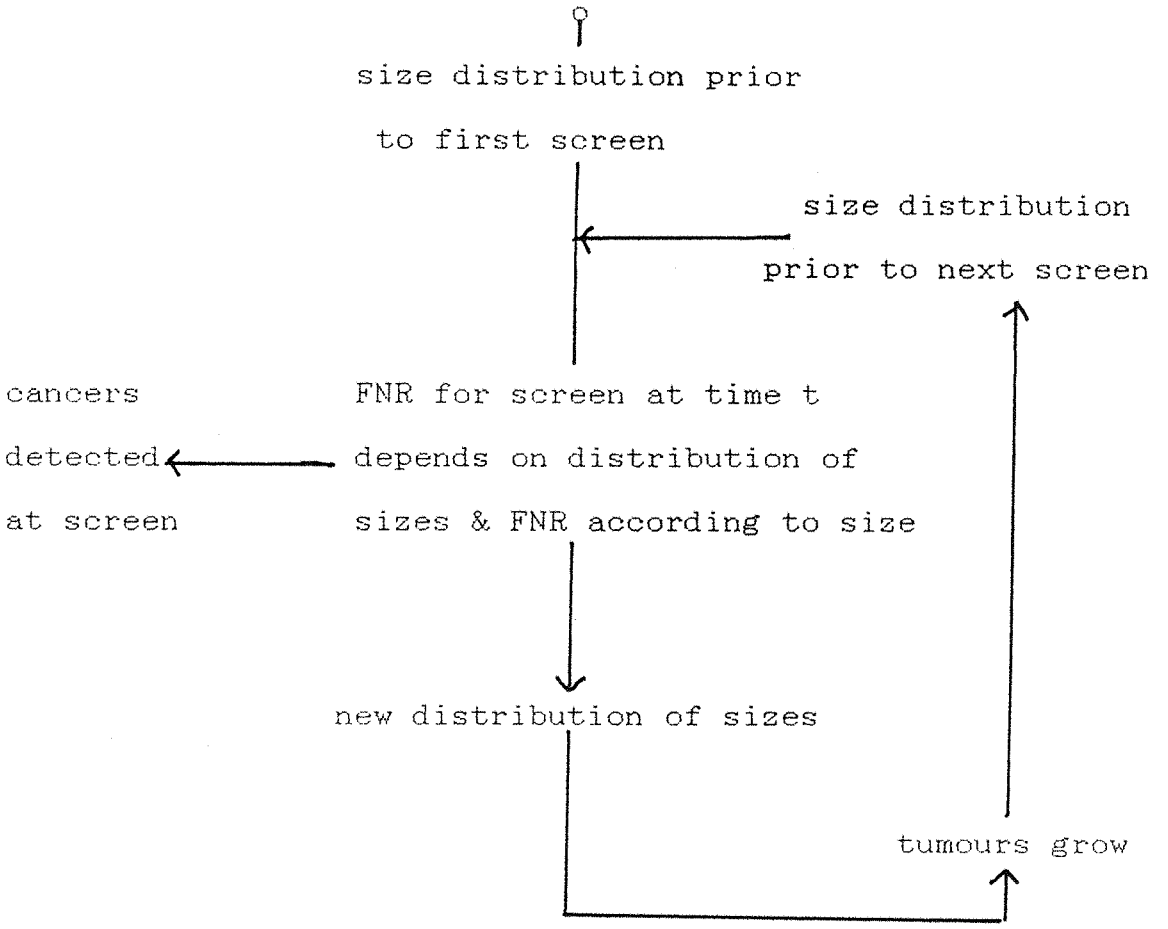
The false negative rate for a particular screen according to tumour age depends on the distribution of tumour sizes. This depends on two main factors:

i. the population's screening history; that is if they have been screened previously, how long ago and by how sensitive a test.

ii. the growth of tumours since the last screen; which depends on the growth rate and the duration.

Calculation of the FNR according to tumour age involves an iterative process which is represented in Figure 3.8.3.

Figure 3.8.3. Schematic diagram for the calculation of FNR for a series of screens



For the first screen t years after tumour initiation, the distribution of sizes can be given by equation 3.8.4. The FNR at the screen is given by:

$$\int_{0.2}^{\infty} b(x)\, f(x)\, dx \qquad 3.8.6$$

where $f(x)$ is the frequency of tumour size x, $b(x)$ is the FNR for size x.

After screening those tumours not detected form a new distribution. Between screens, the tumours will continue to grow. Using the distribution after the screen, the time interval between screens and a growth rate, the distribution of sizes at the next screen can be calculated. These stages can then be repeated to give the FNR on successive screens according to the age of the tumour.

CHAPTER 4:

A SCREENING MODEL

4.1 Introduction

Evaluation of randomised controlled trials to measure the effectiveness of screening programmes is difficult (Provok, 1981). Such trials are expensive to carry out, few women who develop the disease may be involved and the results inconclusive. The value of mammography alone or as an adjunct to clinical examination has not been evaluated in properly designed trials. Nor has BSE been considered as an alternative or an addition to other screening tests. The present UK trials being carried out may provide some answers (UK Trials for Early Detection of Breast Cancer, 1981) but these will be restrictive:

i. the trials involve only women age 45 to 64 and will not evaluate the effect of screening pre-menopausal women.

ii. mammography is performed bi-annually with CE in between. No variation of this is being considered.

The American Cancer Society report in 1980 (see Gaśtrin, 1981) recommends the following guidelines:

1. All women over 20 should perform BSE monthly and have CE every 3 years before 40 and yearly thereafter.

2. Between 34 and 40 each woman should have a mammogram which can be used as a point of reference for further mammograms. After 40 and below 50, mammograms should only be taken if deemed necessary because of family history or indications on the first mammogram.

3. All women over 50 should have a yearly mammogram.

Trials to evaluate the effects of such recommendations on the mortality and morbidity of breast cancer are expensive to run and such frequency of screening may not represent the best choice in terms of altering prognosis, ensuring good attendance for tests and the demands on the health service in terms of costs and manpower. A screening model provides a cheap and flexible way to consider the likely effects of various screening policies on breast cancer mortality and morbidity.

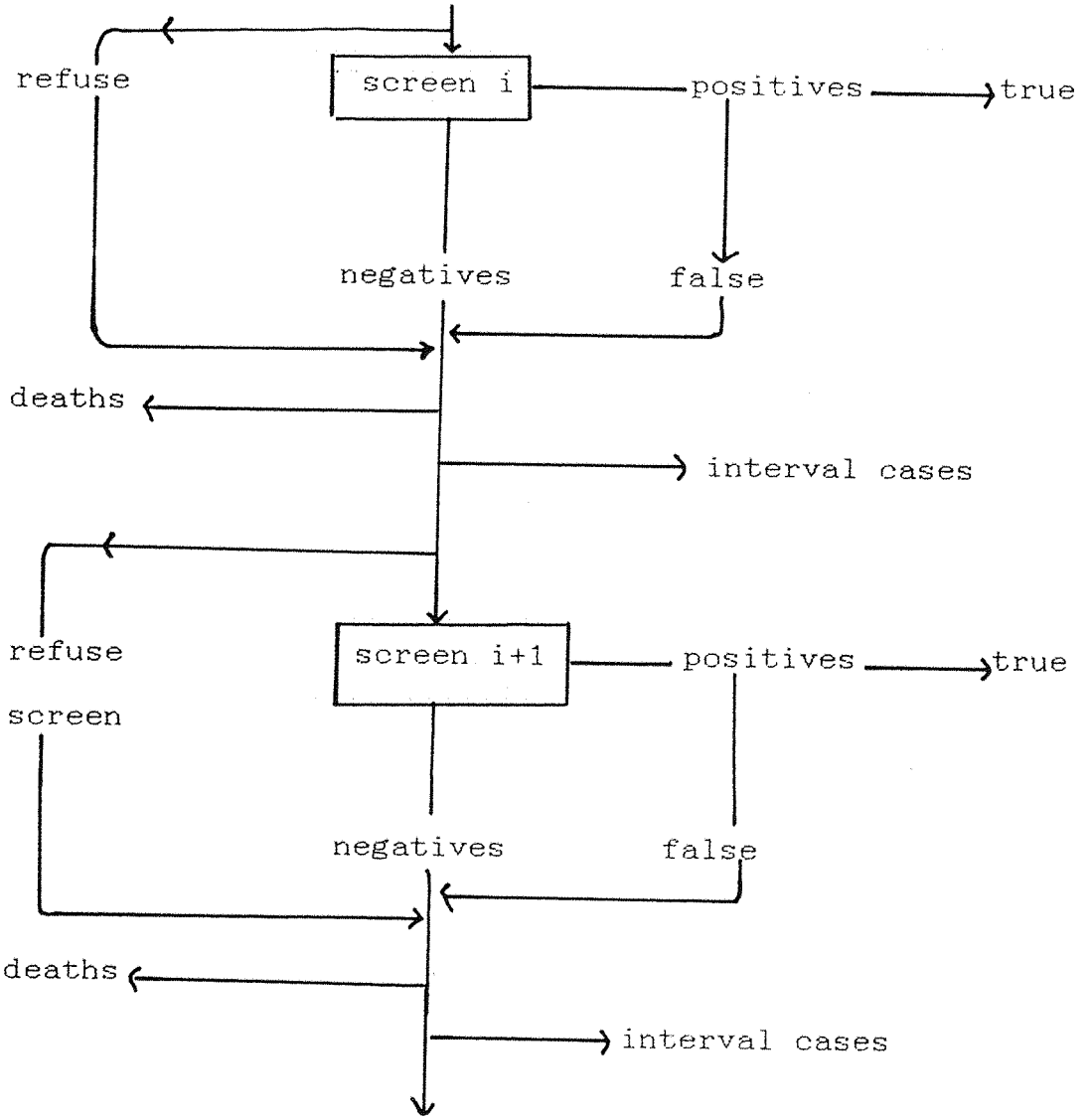
4.2. The screening population

A screening population consists of women who will never get the disease (or they will die before its detection) and those who will get the disease and it will be discovered before death from other causes.

Figure 4.2.1. depicts screening a population on two successive screens. Of those screened any person with a positive test will undergo further tests eg a biopsy to confirm the results. If disease is confirmed the person is classified as a true positive by the test and leaves the programme. If presence of disease is not confirmed, these false positives will return to the screening programme.

Those women classified as negative on the test will continue in the screening programme. Any person with the disease who is classified negative will continue in the

Figure 4.2.1 Screening a population on two successive screens



screening programme until discovery of the disease at a later date either by screening or self-detection during the screening interval.

New developing cases may also become apparent during the screening interval. Some women may refuse the first screen and future screens or accept some but not all future

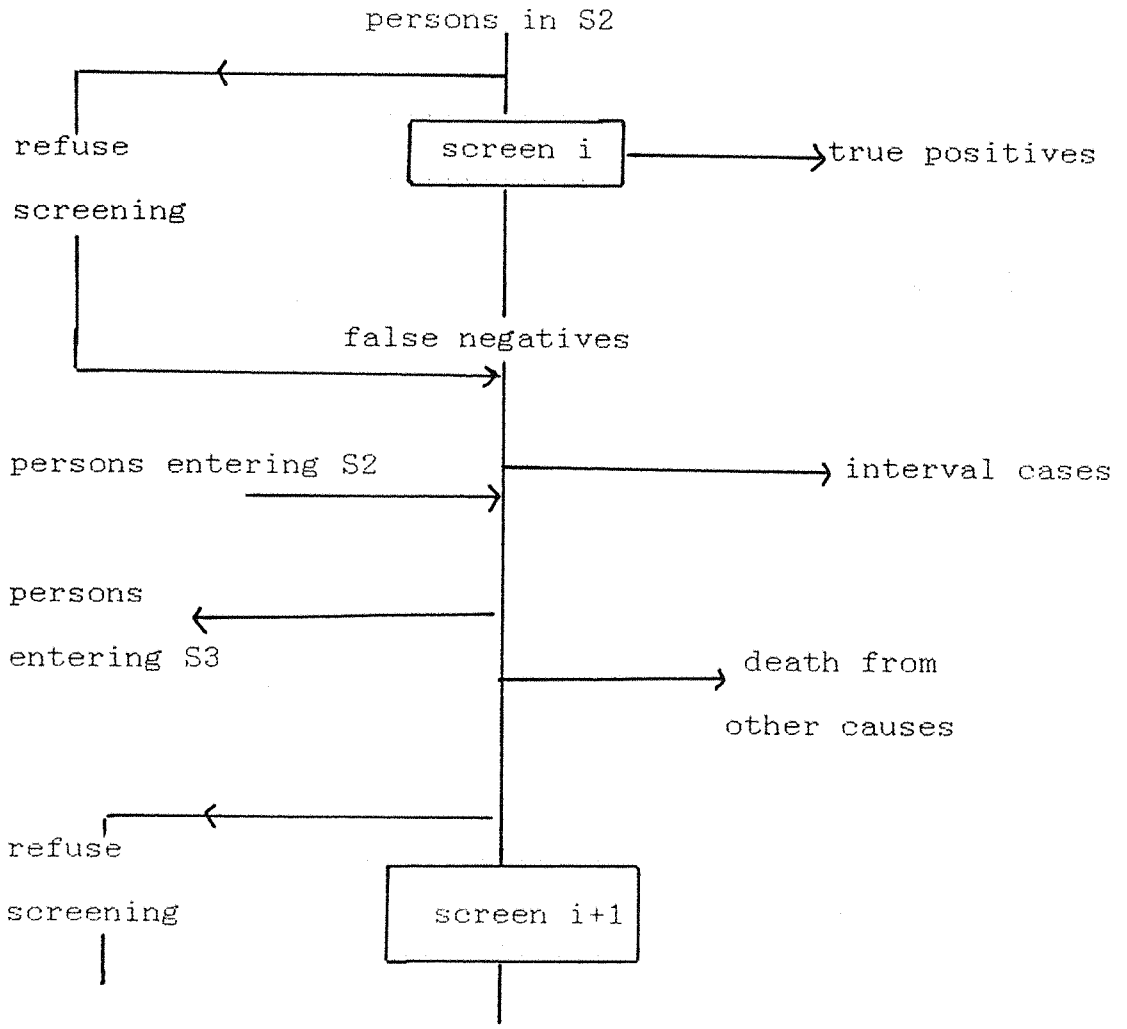
screens. During the screening interval death from other causes will decrease the screening population. This diagram of the screening population is the first stage in developing a model which incorporates the number of screens performed, the number of persons in the four categories on testing and hence the yield and cost per cancer found.

In this study, I have chosen to consider only those women who will get breast cancer in their lifetime and consider the effect of screening on the detection of the disease. In Figure 4.2.2 we consider what happens on screening to those who have the disease, but it has not been discovered although it is of a size to be detectable by a screening test.

The prevalence of disease are those women in state S2 and the incidence is those women entering S3, that is where they become aware of their disease and seek medical advice. For those with the disease, there are only two outcomes at the test: i. a true positive result or ii. a false negative result. Those with the disease not detected will either enter S3 before the next screen and be classed as an interval cases or present for the next screen which may or may not detected the disease.

Between screens a new group of women who have developed detectable cancers will enter this 'diseased population'. They will present for the next screen unless they detect their cancers prior to this time. Some women will leave the population because of death.

Figure 4.2.2. Prevalence, incidence and detection of those afflicted by the disease



Some mathematical relationships may be derived:

Let u_i = the time of screen i

a = the age group

m = screening method or test

P = prevelance of detectable disease for age group a

i,a at screen i

$I_{i,a}$ = incidence of new cases entering S2 between the ith and i+1 screen for age a

$D_{i,a}$ = death in age group a

$b_i(z_1)$ = false negative rate at ith screen where z_1 depends on tumour size, doubling time and sensitivity of test

$r_i(z_2,t)$ = probability false negative at ith screen will present for next screen at time t, where z_2 depends on tumour doubling time and size.

$s_i(z_3)$ = probability new cases will present for next screen where z_3 depends on the doubling time of tumour, length of screening interval.

v = screening interval

define $r_{i+1}(z_2, t)$. $r_i(z_2,t) = r_i(z_2,t)$ 4.2.1

Section 3.7. gives the incidence of new cases entering S2 for 5 year age groups. Hence for time interval v between screens, the incidence for age a is $v.I_{i,a}/5$.

$P_{i,a}$ can be obtained by using the incidence of cases entering S2 and the distribution of duration in S2 given in section 3.5.

Hence:

$$P_{i,a} = \sum_{j=1}^{n} l(j) \; I_{(a-j)} \qquad 4.2.2$$

where $l(j)$ is the frequency of tumours whose duration in S2 is j years and $I_{(a-j)}$ is the incidence of 0.2cm tumours for age (a-j).

n can be taken as the longest duration in S2.

In using the incidence in 3.7. for any calculations we assume that the screening population produces the exact incidence and is unbiased. Those attending screening are self-selecting and could be at higher risk of getting bre ast cancer. If we were directing screening to those women deemed say with twice the risk then the value of I must be adjusted accordingly. Alternatively using values from clinical trials may provide a better estimation.

Figure 4.2.3. shows the situation for two successive screens for a particular age group. For simplicity we assume all will attend for screening. A screening test will discover

$$(1 - b(z_i)) P_{i,a} \qquad 4.2.3$$

cancers but

$$b(z_i) P_{i,a} \qquad 4.2.4$$

will go undetected.

Before the next screen I new cases will develop of which

$$(1 - s_{i3}(z)) I_{i,a} \qquad 4.2.5$$

will clinically surface before the next screen.

$$(1-r_{i2}(z)) b_{i1}(z) P_{i,a} \qquad 4.2.6$$

of the false negatives will be found before the next screen.

The prevalence at the next screen can be given by the recursive formula:

$$P_{i+1,a} = r_{i2}(z) b_{i1}(z) P_{i,a} + s_{i3}(z) I_{i,a} - D_{i,a} \qquad 4.2.7$$

This can be written as:

$$P_{i+1,a} = r_{12}(z,t) \prod_{j=1}^{i} b_{j(z)i,a} + \sum_{k=1}^{i-1} r_{k+1\,2}(z,t) \prod_{j=k+1}^{i} b_{j1}(z) (s_{k3}(z) I_{k,a} - D_{k,a}) + s_{i3}(z) I_{i,a} - D_{i,a} \qquad 4.2.8$$

4.3. The two state model

One important facet of a screening model involves looking at the delay in detection of the disease for different screening policies. For our purposes we assume that this delay is measured from when the disease is first detectable by a screening test.

Shahani et al(1977) have developed a two-state screening model in which screening of those who will develop the disease occurs at intervals x for i = 1,2.... The probability of a false negative (that is a tumour will be missed) is given by b, and the probability of an incorrect positive result is a.

Some of the main results are:

The mean number of tests before discovery of the tumour is:

$$\sum_{i=0}^{\infty} S(x_i) + b/(1-b) \qquad 4.3.1$$

The mean number of false positives is:

$$a \sum_{i=1}^{\infty} S(x_i) \qquad 4.3.2$$

The mean delay is:

$$E(D) = \sum_{n=1}^{\infty} x_n \sum_{i=1}^{n} \left(S(x_{i-1}) - S(x_i) \right) (1-b)b^{n-i} - E(T) \qquad 4.3.3$$

where S(x) is the survival function for the women who will

get the disease.

E(T) is the expected value of S(x)

b is the false negative rate for the test.

For periodic screening intervals x =nf where f is the screening interval.

Shahani(1977) makes three assumptions for the screening model:

1. Discovery of the tumour is only by screening.

2. The false negative rate is constant irrespective of the population's screening history.

3.Screening begins at age 0.

In addition it is usually assumed that screening is performed at constant intervals. These assumptions may be relaxed to give a more robust screening model.

4.3.1. Discovery of the tumour

Tumours are not just discovered by screening. Symptoms may develop between screening or women will find a sign, usually the presence of a lump, that will make them seek medical advice. On the other hand it is argued that a negative screening result will give women a false confidence and such signs and symptoms may be ignored.

If the screening interval is large then many false negatives from a previous screen will be detected by the patient as well as some rapidly developing new cases. For a short screening interval then a large proportion of false

negatives will go undetected until the next screen.

The probability of tumours being discovered between screens depends on their size immediately after the last screen, the growth rate of the tumour and the length of the screening interval . .

If the probability of women detecting their own tumours is II(D,v) then:

$$II(D,v) = \int_{0}^{\infty} \int_{0.2}^{\infty} D(x,t)\, h(x)\, dx\, dt \qquad 4.3.4$$

where D(x,t) is the distribution of tumour size x at time t after the last screen.

h(x) is the probability of a woman finding a tumour size x.

Equation 4.3.3 may be rewritten to allow for the likelihood of interval cases:

$$E(D) = \sum_{n=1}^{\infty} x_n \sum_{i=1}^{n} (1-II(D,v)(S(x_{i-1}) - S(x_i))b^{n-i}(1-b)$$
$$+ \sum_{n=1}^{\infty} x_{nn} \sum_{i=1}^{n} II(D,v)\,(S(x_{i-1}) - S(x_i))b^{n-i} \qquad 4.3.5$$

where $x_n = nf$ and $x_{nn} = nnf$

4.3.2. The false negative rate

The false negative rate is not constant but must be related to the age and size of the tumour. We assume that the false negative rate decreases with increasing size of tumour. In equation 4.3.2 and 4.3.3 we can replace the constant b by a function b(t) which is dependent on the age of the tumour at the time of screening. The age is measured from when a patient enters state S2.

We replace b^{n-i} by $z(b(t))$ where

$$z(b(t)) = \prod_{k=1}^{n-i} b(t) \quad \text{if } n \neq i$$

$$= 1 \quad \text{if } n = i$$

4.3.6

where t is measured from the midpoint of the screening interval and is given by $t = (k - .5)f$

$(1 - b)$ is replaced by

$$(1 - b(t_1)) \text{ where } t_1 = (n - i + .5) \qquad 4.3.7$$

4.3.3. The starting age of screening

It is important to consider the effect on delay, and also the reduction on the number of screening tests given when the starting age of a screening programme is varied.

In Shahani's equations the origin may be established at any desired age. The simplest way of allowing for a delayed start is to put

$$x_n = a + (n-1)f \qquad n = 1,2,.....$$

$$x_i = a + (i-1)f \qquad i = 2,3,4,.....$$

$$= 0 \qquad i = 1$$

(see de Senna, 1983)

I have developed a different expression in order to allow for a differing false negative rate. If we let the age at start be s, then the first screen is for n at m = Int(s/f)

Hence:

$$E(D) = \sum_{n=m}^{\infty} x_n \left(1 - S(x_{n-1}) \right) b^{n-m} + \sum_{i=m}^{n} \left(S(x_{i-1}) - S(x_i) \right) b^{n-i} \Big) (1-b) - E(T) \qquad 4.3.8$$

which can be more usefully written as

$$E(D) = \sum_{n=m}^{\infty} x_n (1-b_n) \left(\sum_{j=1}^{m-1} (S(x_{j-1}) - S(x_j)) b_{j-1} \right) + \sum_{i=m}^{n} (S(x_{i-1}) - S(x_i)) b_{i-1}^{n-i} - E(T) \qquad 4.3.9$$

Combining the equations depicting varying false negative rates and age at start then the estimated delay is:

$$E(D) = \sum_{n=m}^{\infty} x_n \left(\sum_{j=1}^{m-1} (S(x_{j-1}) - S(x_j)) (1-b_j(t_4)) z'(b_j(t_3)) \right) + \sum_{i=m}^{n} ((S(x_{i-1}) - S(x_i)) (1-b_i(t_2)) z(b_i(t_1)) - E(T) \qquad 4.3.10$$

where

$$z'(b(t)) = \prod_{k=m}^{n-1} b(t) \quad \text{if } n \neq m$$

$$= 1 \quad \text{if } n = m \qquad 4.3.11$$

$$z(b(t)) = \prod_{k=1}^{n-i} b(t) \quad \text{if } n \neq i$$

$$= 1 \quad \text{if } n = i \qquad 4.3.12$$

and $t_1 = (k-.5)f$ $t_2 = (n- i+.5)f$

$t_3 = (k-j+.5)f$ $t_4 = (n -j+.5)f$

4.3.4. The frequency of screening

Periodic screening is perhaps the most appealing: it is easy to carry out in terms of administration and for the patient. But it may be prudent to increase the frequency of screening as breast cancer incidence rises with age, especially after the menopause. For example screening every 5 years before the menopause then every two years after could be considered. Such a choice of interval varying with age must be kept reasonably simple to facilitate adherence to it.

Consider two screening intervals f_1 and f_2, where A is the age at which the patient moves from one to the other. Then

$$x_n = nf_1 \quad \text{for } n < A/f_1$$

$$= A + (n - A/f_1)f_2 \quad \text{for } n >= A/f_1$$

$$n=1,2... \qquad 4.3.13$$

4.4. Applications of a screening model

The two state model is useful for comparing the delay in detection with different screening regimes which vary the screening interval and the age at starting screening. The false negative rate function can be varied to reflect different screening tests. In addition we can look at the mean delay when the screening method is self-examination which may be performed monthly or less often.

In this section I have demonstrated the variations which can be made to the basic mean delay equation to allow for a more realistic view of breast cancer screening. A final model should make allowance for the following:

1. detection by the patient either before screening commences or in the screening intervals. This will vary according to the patient's attitude and BSE practise.
2. varying sensitivity of a particular test according to the screening history of the population.
3. varying screening intervals begining and ending at various ages.

CHAPTER 5

SCREENING, DELAY IN DETECTION AND SURVIVAL

5.1 Introduction

One important facet for screening when considering delay or lead time is the length of time a woman is in state S2. This can be estimated by using the distribution of sizes at diagnosis for Edinburgh (Kerr, 1982) and the doubling time of the tumour. Table 5.1 shows the cumulative distribution of time in S2 for various doubling times.

Table 5.1. % cumulative distribution in state S2

years	doubling time(days)				
	25	50	100	150	200
1	84.42	4.08	0.05	0	0
2	100	85.78	4.43	0.28	0.07
3		99.98	36.15	3.80	0.70
4		100	84.15	20.45	3.65
5			99.23	53.45	14.22
6			100	84.78	36.55
7				97.61	63.25
8				99.84	84.82
9				99.97	96.22
10				100	99.42
11					99.92
12					100

With rapidly growing cancers with a doubling time of 50 days or less, the duration is short. But with slower growing tumours of 150 or 200 days, then half have a duration of nearly 5 years or over 6 years respectively. The weighted mean duration for each doubling time given from 25 to 200 days is 14 months, just over 2 years, 3 3/4 years, 5 1/2 years and 7 years respectively.

5.2. Duration of different stages of breast cancer.

It is necessary to consider the progress of a woman in each stage of breast cancer because stage at detection results in different survival patterns. The main partitioning of tumours are Stage 1 - up to 2cm; and Stage 2 - from 2-5cm with localised spread only. Stage 3 and 4 involve regional spread and distant metastases respectively. Table 5.2. gives the duration for different doubling times of cancers in the two partitions over 0.2cm but less than 2cm and 2 to 5cm.

Table 5.2. Duration in years of tumours in two size partitions

size	doubling time(days)						
	25	50	100	150	200	250	300
under 2cm	0.68	1.37	2.73	4.10	5.46	6.83	8.19
2-5cm	0.27	0.54	1.09	1.63	2.17	2.71	3.26

Hence tumours with a doubling time of less than 100 days will grow from under 2cm to over 5cm in less than a year and for those with doubling time less than 200 days in less than the two year screening interval.

5.3. False negative rates according to size of tumour for different screening tests

When tumours are young in age the FNR will be high since the tumours will still be small and difficult to detect. As the tumour ages and grows in size the FNR will decrease. On the first screen in a population, the test will detect the larger tumours and will miss a large proportion of the smaller tumours. The actual percentage will depend on the type of test used and the conditions existing at the time of the test eg examiner experience. If the time to the next screen is long then these small tumours which were not detected will increase in size so that the FNR on the rescreen will be smaller. If screening is frequent the FNR will decrease slowly since the large tumours will have been detected previously and the small ones will not have grown much in the screening interval. The relative decrease in FNR with time does of course depend on the doubling time of the tumours.

Using the method on Section 3.7. it is possible to calculate the FNR according to the sensitivity of a particular test for a certain size of tumour and according to the interval between tests. Table 5.3. and 5.4. give the

FNR for various screening tests performed every one or two years allowing for tumour growth with mean exponential doubling time of 3.45 and 5.18 months. The parameters for sensitivity according to sizes are used as given in Section 3.8. The initial FNR given is for a screen one or two years after the tumour has reached 0.2cm in diameter.

Table 5.3. FNR on successive yearly screens

Screen no.	FNR according to sensitivity of test with parameters;			
	(1.2, 1.3)		(2, 1.5)	
d/time;	3.45	5.18	3.45	5.18
1	.494	.591	.611	.705
2	.342	.471	.449	.586
3	.249	.388	.334	.491
4	.192	.332	.257	.419
5	.169	.295	.206	.367
6	.140	.271	.174	.328
7	.130	.256	.154	.302
8	.125	.247	.143	.284
9	.123	.241	.137	.273
10	.122	.238	.134	.266

Table 5.4. FNR on successive yearly screens

Screen no.	FNR according to sensitivity of test with parameters			
	(1.2, 1.3)		(2, 1.5)	
d/time	3.45	5.18	3.45	5.18
1	.301	.417	.408	.532
2	.135	.250	.202	.344
3	.066	.160	.103	.226
4	.037	.111	.056	.156
5	.024	.085	.035	.114

It is clear to see that for two yearly tests the FNR decreases more rapidly than for one yearly tests reflecting that the longer duaration between screens give a higher proportion of larger tumours at screening which are easily detectable. Also in each instance, the faster doubling time results in a quicker decreasing FNR. Initially the FNR for yearly screening, for slow growing tumours is 15 to 20% greater than that for faster growing tumours but after 6 years is roughly double. For 2 year screening, initially the FNR is 30 to 38% higher for slower growing tumours with ratios of 3 to 1 after 4 years.

The effect of doubling time on FNR at screening can be displayed by using a simple function for FNR according to tumour size. For example:

1 for tumours less than 0.2cm

0.4 0.2 up to 1cm

0.1	1 up to 5cm	
0	5cm and over	5.1.1.

The effect of doubling time can be seen in Table 5.5. For a doubling time of 100 days, the FNR falls quickly to

Table 5.5.FNR for yearly screen and differing doubling times

year	doubling time(days)		
	100	200	300
1	.246	.313	.339
2	.168	.270	.313
3	.118	.233	.289
4	.089	.205	.269
5	.071	.184	.252
6	.061	.170	.239
7	.054	.160	.229
8	.051	.153	.215

under 10% in 4 years. But for slower doubling times of 200 or 300 days, 15% and 20% respectively of tumours undetected by previous screens continue to go undetected by the next screen even after 8 years.

Such calculations of FNR on repeated screens should take into account patient detection between the screens. A simple representation involves using the data in

Foster(1977) for size at diagnosis when women do not perform BSE. The percentage of tumours not detected during the screening interval can be calculated as a patient 'screen'.

Table 5.6 shows the FNR for screening every 2 years allowing for a mid-interval patient detection.

Table 5.6. FNR for 2 yearly screens allowing for for patient detection on alternate years starting at year 1.

year	parameters:	(1.2, 1.3)		(2, 1.5)	
	d/time	3.45	5.18	3.45	5.18
1		.749	.825	.749	.825
2		.288	.419	.410	.551
3		.591	.732	.550	.703
4		.119	.241	.195	.349
5		.474	.657	.400	.596
6		.065	.163	.111	.242
7		.411	.608	.316	.521
8		.049	.130	.081	.190
9		.387	.582	.280	.478
10		.045	.117	.071	.167

The FNR for patient detection will remain high since the screen will detect the larger tumours that a patient will find easy to detect. Comparing the FNR for patient detection after three years for a doubling time of 3.45 months of .591 and .550 for the more sensitive and less sensitive tests the effect is seen clearly. That is the more sensitive test will detect a larger proportion of the tumours that will be most easily be found by the patient herself, and hence the patient detection FNR will be higher than for the less sensitive test. Also for the slower doubling time of 5.18 months the FNR for patient detection will remain higher than for a doubling time of 3.45 months.

The FNR for patients practising BSE can be found in the same way using Foster's data(1977) for the size at detection by BSE as the sensitivity according to tumour size. See Tables 5.7. to 5.10.

When BSE is performed monthly the FNR is very high and remains so especially for slow growing tumours, where over 4 years it drops from .979 to .904 compared with a drop from .968 to .852 in the faster growing tumours. This is because BSE performed monthly will only detect the few tumours that have grown to a sufficent size to be felt. If BSE is performed 3 times a year then the FNR will be lower at each screen than for monthly BSE since more tumours will have grown to a detectable size over the period.

Table 5.7. FNR for BSE performed monthly with doubling time of 3.45 months

month \ years	1	2	3	4
1	.968	.906	.873	.858
2	.961	.902	.871	.857
3	.955	.898	.870	.857
4	.949	.895	.868	.856
5	.944	.892	.867	.855
6	.938	.899	.865	.855
7	.933	.886	.864	.854
8	.928	.884	.863	.854
9	.923	.881	.862	.853
10	.918	.879	.861	.853
11	.914	.877	.860	.853
12	.910	.875	.859	.852

Table 5.8. FNR for BSE performed every 4 months with doubling time of 3.45 months

month \ years	1	2	3	4
4	.896	.771	.663	.592
8	.855	.731	.635	.576
12	.813	.695	.612	.563

Table 5.9. FNR for BSE performed monthly with doubling time of 5.18 months

month \ years	1	2	3	4
1	.979	.944	.923	.911
2	.974	.942	.922	.910
3	.971	.940	.921	.909
4	.968	.938	.919	.909
5	.965	.936	.918	.908
6	.962	.934	.918	.908
7	.959	.932	.917	.907
8	.956	.931	.915	.906
9	.954	.929	.914	.906
10	.951	.927	.913	.905
11	.949	.926	.912	.905
12	.947	.924	.912	.904

Table 5.10. FNR for BSE performed every 4 months with doubling time of 5.18 months

month \ years	1	2	3	4
4	.929	.860	.794	.741
8	.907	.837	.774	.727
12	.884	.815	.757	.715

.5.4. Delay in detection.

The time between when the tumour reaches 0.2cm and when it is found by screening is called the screening delay. The formulas for calculating this mean delay which are given in Chapter 4 require an estimation of the expected age from the onset distribution. The value of E(X) can be found by evaluating

$$\int_0^{\infty} x\, f(x)\, dx \qquad 5.3.1$$

where f(x) is the probability density function for age at onset. The parameters of the function are given in Section 3.7. and since the function is a composite one involves integrating between 0 and 41.72 and 41.72 and ∞ . The value of E(x) was calculated as 53.21 years.

5.4.1. Screening delay.

The screening delay in years measured from when a tumour is 0.2cm can be calculated for screening at various intervals using equation 4.4.3. but allowing for a differing FNR according to the age of the tumour as found in this chapter.

Table 5.11 gives the results when patient detection is ignored, that is detection is only by screening. The two paramters are explained in Section 3.8. and may be consistent with mammogram and CE respectively.

The increase in delay as the screening interval

increases appears to be linear with larger increases for a slower doubling time and for less sensitive tests (that is the second set of parameters as opposed to the first).

Table 5.11. Screening delay in years for different screening intervals

parameters for FNR according to size	screening interval(years) 1	1.5	2
d/time = 3.45months			
(1.2,1.3)	1.32	1.47	1.61
(2,1.5)	1.56	1.72	1.88
d/time = 5.18			
(1.2,1.3)	1.58	1.77	1.95
(2,1.5)	1.92	2.14	2.34

In reality, it is unlikely for there to be no cancers found during the screening interval. This is especially so for screening tests every 2 years and more so for less sensitive tests like CE. Table 5.12. shows the screening delay for yearly and 2 yearly screens. Two calculations were made:

i. assuming no interval detected cases in performing a screen with FNR of 1.

ii assuming patient detected cases by a 'screen' at the mid interval.

Table 5.12. Screening delay in years allowing for patient detection(PD)

parameters	screening interval			
	1 year		2 year	
	No PD	PD	No PD	PD
d/time 3.45 months				
(1.2,1.3)	1.57	1.37	2.11	1.64
(2,1.5)	1.80	1.52	2.37	1.77
d/time 5.18 months				
(1.2,1.3)	1.82	1.64	2.44	1.99
(2,1.5)	2.17	1.89	2.82	2.22

The first calculation was made since the figures for delay calculated in the second part are not comparable to Table 5.11.

These results indicate the reduction in the mean delay allowing for patient detection. This reduction is an under estimate since patient detection was not measured continuously over the interval. The reduction in delay allowing for PD between the more sensitive and less sensitive tests is the same for 1 or 2 yearly screens for both doubling times. This reduction is over 1 1/2 months for the faster doubling time and nearly 3 months for the slower doubling time.

5.4.2 Detection delay and BSE

BSE appears to be important in diagnosing tumours at a smaller size. Table 5.12 shows the delay when BSE is practised at various intervals. For 1 and 2 monthly BSE Foster's size at detection for BSE monthly was used and for 4 and 6 monthly BSE his 'less often' size at detection was used. (Foster, 1977) Even though the FNR at each self-examination was high the mean delay in detecting the tumour is quite low and compares favourably with the delay figures in Table 5.11. But account must be taken that we have assumed the likelihood of detecting the tumour is independent of the last examination. This may be an over-simplfication since the location of the tumour or tissue density may make detection difficult even when the tumour reaches a reasonable size.

Table 5.13. Detection delay in years with BSE

BSE every(months)	doubling time	
	3.45	5.18
1	1.22	1.60
2	1.41	1.85
4	1.56	2.01
6	1.64	2.12

Table 5.14. records the estimated delay when BSE is practised twice yearly between yearly screens. These indicate a reduction of delay of 1.3 months and 2.4 months

Table 5.14. Delay for yearly screening with BSE practised twice yearly

test parameters	doubling time(months)	
	3.45	5.18
(1.2,1.3)	1.34	1.62
(2,1.5)	1.45	1.81

when a less sensitive test replaces BSE yearly for the two doubling times. For the more sensitive test, the reduction in delay is twice these figures. Such a reduction in delay by performing yearly screening may not be so marked when BSE is performed more often.

5.5.Survival according to tumour characteristics

Many studies have drawn conclusions about survival over a fixed time period eg 10 years. Such a measure is not synomonous with a cure since relapse could occur after this. A more robust method to analyse survival data is to compare the observed overall mortality of the patient group with the expected mortality of controls. An example quoted by Langlands(1979) involving 704 women suggests the two figures approach each other after 21 years and thus 30% of women with Stage I and II cancers were cured.

In a study of 3878 patients (Langlands, 1979) there was a 58% excess in mortality between 15-20 years after treatment ($p< 0.001$, 95%CI = 25%,98%). In an alternative

method of expressing excess mortality in terms of the number of years follow up for 15-20 years, the excess rate is 18.8 per 1000 woman years. Whereas the rate of observed to expected mortality can approach unity, there is often still a continuing excess mortality rate.

Various studies have pointed to the main prognostic factors in breast cancer as tumour size, nodal involvement and spread. A multiple regression anaylsis of 298 patients in Nottingham (Haybrittle, 1981) found that lymph node stage and tumour grade was highly significant ($p < 0.01$) and tumour size was less so ($p<0.05$). Langland's (1979) study showed that:

i. Survival in the first 10 years is related to tumour size with the average mortality being about twice the normal population between 10-20 years after treatment.

ii. The presence of positive lymph nodes had no effect after 10 years.

iii. After 10 years, the ratios of observed to expected death for stages I and II were similar ($\chi^2 = 0.54$, df=2)

iv. When a fixed time survival is considered, there is a significant difference in survival between different stages that persists up to 20 years.

Berkson(1952) showed the effects of metastases on survival in 6426 cases. In table 5.15, the net annual death rate and the life expectancy were based on the proportion of the population not subjected to mortality from breast cancer.

Table 5.15. Effects of metastases on survival

(Source: Berkson, 1952)

metastases	death rate(p.a.)	life expectancy	% of normal
absent	.10	16.6	75.8
present	.27	6.9	31.4

Charlson(1980,1982) and Boyd(1981) partitioned patients according to the patients history of duration of symptoms and changes in symptoms to produce a concept of rate of growth. Three partitions of slow, intermediate and fast were made to take into account the duration of symptoms and the occurance of transition events like increases in size or consistency, spread, changes in skin, shape, pain or odema of the breast or surrounding tissues. See Table 5.16

Boyd(1981) calculated the observed and expected number of deaths of each of the three categories and for other factors. This rate is a measure of the relative death rate in each group compared to the whole patient population of 756 and assesses survival for the whole follow-up period.

Table 5.16 Classification of growth rates

(Source:Boyd, 1981)

		Transition	
		Absent	Present
Duration of symptoms	under 4 months	SLOW	INTERMEDIATE
	4 month or more	INTERMEDIATE	FAST

Table 5.17, 5.18 and 5.19 show the relative death rate according to prognostic factors and taking growth with one of the other variables.

Table 5.17. Relative death rates according to prognostic factors (Source:Boyd,1981)

	No	Obs/Exp	χ^2
Growth rate			
slow	93	.59	
intermediate	488	.98	15.1
fast	168	1.34	p <.001
Clinical stage			
I	349	.71	
II	147	1.11	27.3
III	212	1.53	
Nodal status			
negative	144	.57	
1 to 3	327	.89	56.9
4 or more	183	1.81	

Table 5.18 Relative death rate according to growth rate and stage

(Source:Boyd, 1979)

Growth	Clinical stage I	II	III
slow	.41(54)	.49(17)	1.37(20)
intermediate	.65(226)	1.21(88)	1.44(145)
fast	1.89(69)	1.18(42)	1.93(47)

numbers in brackets. Significant trend of growth rate and survival

$$\chi^2 = 12.23,\ p = 0.005$$

Table 5.19. Death rates according to growth rate and nodal status

(Source:Boyd, 1979)

Growth	Nodal status negative	1 to 3	4 or more
slow	.20(17)	.58(45)	.99(22)
intermediate	.61(100)	.84(202)	1.97(116)
fast	.62(22)	1.23(80)	1.87(45)

numbers in brackets. Significant trend of growth rate and survival

$$\chi^2 = 10.1,\ p = 0.001$$

In Table 5.18 the rates observed when growth and clinical stage were considered show large distinctions according to growth rates in Stage I and II. In Stage II only small differences in death rate were seen.

In Table 5.19 the death rate doubled from slow to fast for nodal status 1 to 3 and 4 or more and trebled for negative nodes.

Charlson(1984) investigated growth categories, stage and nodal status for 191 deaths in 495 breast cancer patients. See Table 5.20. She used Cox's analysis of survival and found stages IIb(more than 1 positive node) to IV had poor prognosis and only negtive nodes had a good prognosis.

Table 5.20 Distribution of 191 deaths in 465 patients

(Charlson, 1984)

	Rate of progression			
stage	slow	intermediate	fast	total
I	0/6(0)	7/34(21)	0/1(0)	7/41(17)
II	9/38(24)	62/208(30)	15/39(38)	86/285(30)
III	13/21(62)	24/42(57)	34/45(76)	71/108(66)
IV	2/3(67)	5/7(71)	20/21(95)	27/31(87)
nodal status				
negative	5/18(28)	24/119(20)	8/17(47)	37/154(24)
1 to 3	5/20(25)	17/37(46)	13/23(56)	35/80(44)
4+	5/8(63)	31/50(62)	16/21(76)	52/79(66)

Figures given are deaths over number of patients. Percentage of deaths in brackets.

If the average annual mortality rate was considered then the overall difference between the 3 partitions was significant(χ^2 = 39.8, p<0.001) with a rate of 4.2%, 4.4% and 10.5% respectively. When patients were divided into Stage I, II and IIa(up to one node only) the three mortality rates were 2.7%, 4.0% and 7.3% with significant difference(χ^2=7.7,p< 0.05) Among the stage IIIb(over 1 node involved) and IV patients the mortality rates were 8.4%, 9.1% 2 and 15.4% respectively which were significant.(χ^2=10.0, p< 0.01)

Elwood(1980) tried to assess whether delay from first symptom to diagnosis (that is patient delay) affected prognosis. Results show that women who have a short delay between appearance of the first symptom and diagnosis have a better long-term survival rate than those with long delays, even when survival is assessed from the first symptom. This difference is not so marked in Stage I patients. See Table 5.21

Table 5.21. Mortality ratio according to stage and delay

(Source: Elmwood, 1980)

Stage		No	Mort. rate	χ^2	P value
I	S	391	.99		
	L	188	1.04	.05	.8
II	S	139	.95		
	L	40	1.17	.80	.4
III	S	41	.90		
	L	29	1.17	.82	.4
IV	S	38	.97		
	L	44	1.08	1.10	.3
Total	S	609	.89		
	L	201	1.37	15.56	.0001

S= delay 1 month or less. L= delay 1 year or more

CHAPTER 6

TOWARDS A SCREENING MODEL FOR BREAST CANCER

6.1. Introduction

In this chapter we draw together the results of the previous chapters and consider their merits in providing a basis for a screening model for breast cancer. In Chapter 4 we considered the use of using a survival model to estimate a mean screening delay. The model relies on the onset distribution of cancer, the calculation of which involves using incidence data with a life table approach and a tumour doubling time distribution and calculation of the FNR for the given size distribution at a particular test. The screening population may not be representative of the whole population. Estimation of the doubling time distribution of tumours may be unreliable since methods of calculation do not consider the very fast or very slow growing tumours. The distribution around the mean is not given and it is hard to consider the relative merit of a delay of 1.3 years compared to 1.4 years say. But such a model is necessary when the effects of starting screening at different ages is considered.

Studies that consider survival by considering stage of disease or nodal involvement involve anomalies since metastases already established may not become apparent until several years after detection of the primary tumour. Thus we may consider that the crucial consideration for the

detection of tumours is the probability of metastases according to tumour size which in turn depends on the delay in detection and the doubling time of the tumour.

6.2. screening delay: an alternative model

A simple and easy to use model to assess the effects of various screening policies can be formulated less dependent on the more detailed knowledge and where such detail is minimised to allow for sensitivity analysis.

In developing such a model we need to look at the effect of screening on a tumour with a specific doubling time where the sensitivity of the test is given in terms of tumour size.

Figure 6.1. shows the screening model for screening at time t_i, the size of tumour at screening is d_i and the probability of detection is $p(d_i)$

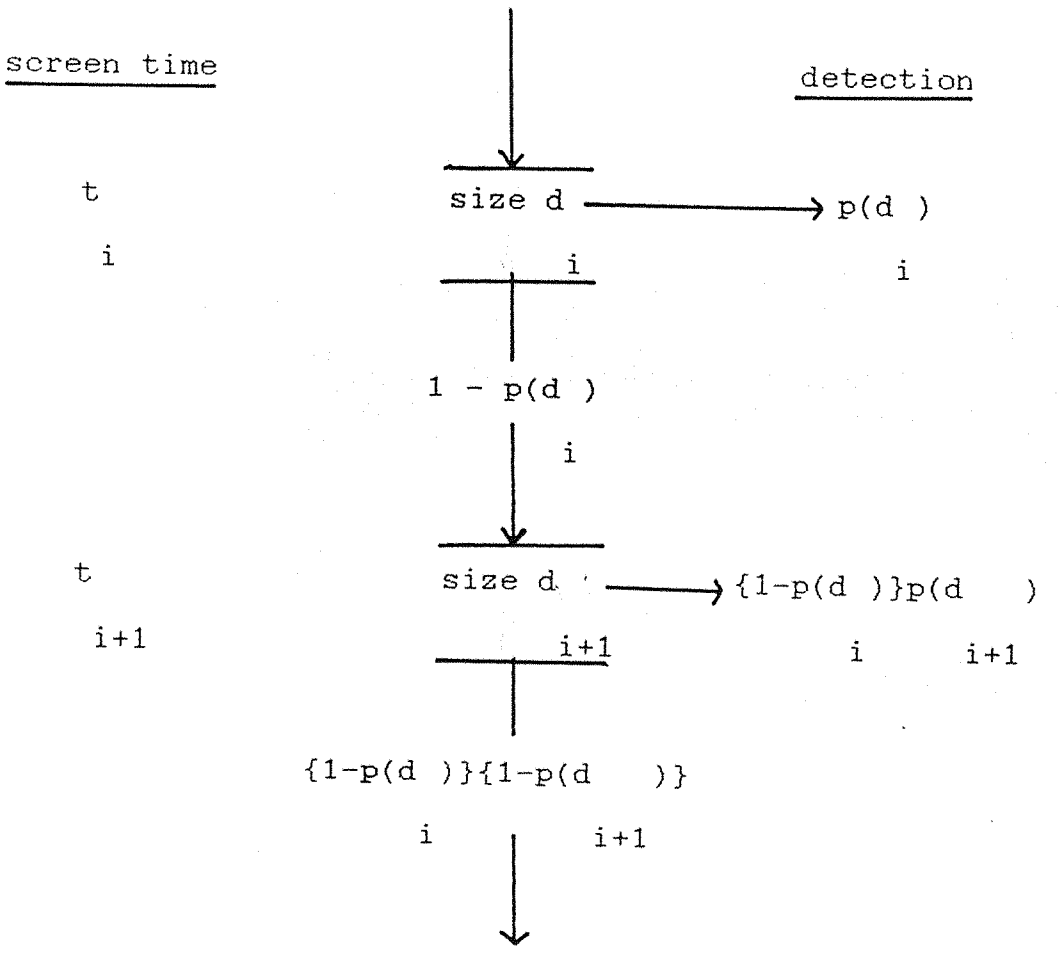
The mean delay can be given by

$$D = p(d_i)t_i + \sum_{i=2}^{i-1} \left[\prod_{j=1}^{j} (1-p(d_j)) \right] p(d_i)t_i \qquad 6.1$$

The size of the tumour at time t is

$$d_t = \exp\left(\frac{t}{4.33\ DT} + \ln 0.2 \right) \qquad 6.2$$

Figure 6.1 Screening and delay



where DT is the tumour doubling time and the false negative rate is given by

$$\exp - (d_t/a)^b \qquad 6.3$$

where a and b are parameters.

Table 6.1. and 6.2 gives the delay when considering tests with parameters of (1.2, 1.3) and (2, 1.5), which we have previously discussed as being possibly consistent with

Table 6.1. Delay on screening with FNR test parameters (1.2,1.3)

DT(days)	screening intervals(years)		
	1	1.5	2
30	1.01	1.26	1.51
50	1.33	1.60	1.85
100	1.98	2.35	2.65
150	2.50	2.97	3.35
200	2.92	3.50	3.96
300	3.59	4.37	4.98

Table 6.2. Delay on screening with FNR test parameters (2,1.5)

DT(days)	screening intervals(years)		
	1	1.5	2
30	1.19	1.44	1.69
50	1.63	1.90	2.15
100	2.60	2.96	3.26
150	3.42	3.90	4.27
200	4.14	4.74	5.20
300	5.36	6.20	6.83

mammography and CE, and screening intervals of 1, 1.5 and 2 years.

These figures were calculated by finding the mean delay allowing for when the first test took place from 1 month then in monthly intervals up to 1, 1.5 or 2 years, after

the tumour reached threshold size.

6.3. Metastases and delay

In screening for breast cancer it is hoped to save lives- if the disease is localised or regionalised, that is metastases have not formed, then the removal of the primary tumour should be successful in treating the disease. We may look at the effects of screening by considering the reduction in the probability of metastases for different screening regimes.

Using Koscielny's(1984) data for the probability of metastases according to tumour size we can construct curves relating the probability of metastases according to delay in detection for a specific doubling time. This involves the following steps:

1. Calculate the percentiles for metastases using Koscielny's lognormal distribution with mean 3.16 and SD of 2.62.(see Chapter 1.5)

2.Calculate the number of doubling times required to reach these percentile sizes and substract the number of doubling times to reach 0.2cm. (since we are interested in delay measured from when a tumour first becomes detectable.) This is given by

$$3(\mathrm{Ln}\ d - \mathrm{Ln}\ 0.2)/\mathrm{Ln}\ 2 \qquad 6.3.$$

3. The number of doubling times multiplied by the actual doubling time will then give the relationship between delay

and probability of metastases.

The results of these calculations are shown in Table 6.4 and Figure 6.2.

Table 6.4. Delay in years and probability of metastases

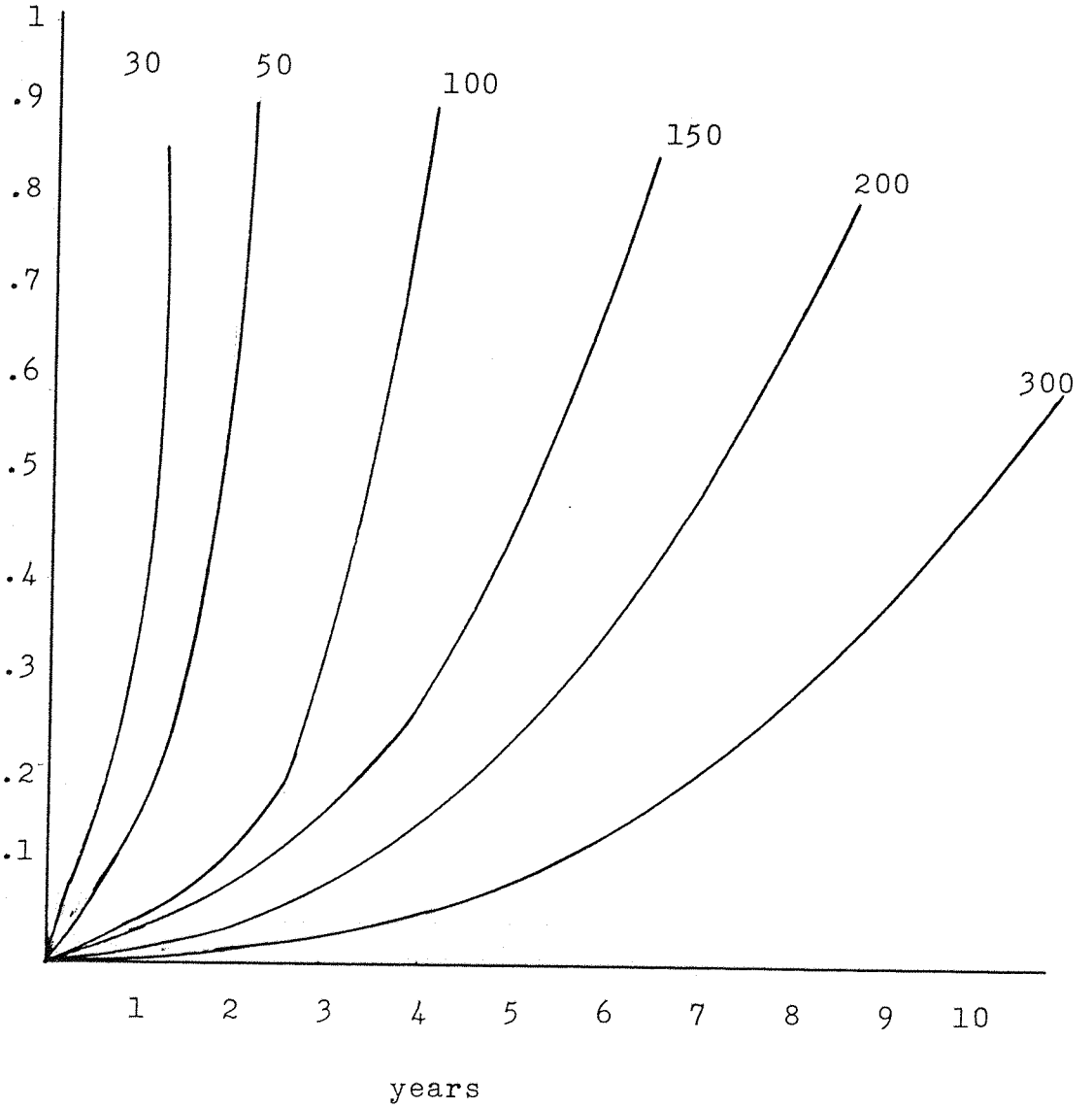
Probability	Doubling time					
	30	50	100	150	200	300
.1	.54	.91	1.81	2.72	3.62	5.43
.2	.69	1.15	2.31	3.46	4.62	6.93
.3	.80	1.34	2.68	4.02	5.35	8.03
.4	.90	1.49	2.99	4.48	5.97	8.96
.5	.98	1.64	3.28	4.92	6.56	9.84
.6	1.07	1.78	3.56	5.34	7.12	10.68
.7	1.16	1.94	3.87	5.81	7.75	11.62
.8	1.27	2.12	4.24	6.35	8.47	12.71

6.4. Screening delay and metastases

By considering the screening delay and the probability of metastases with delay we can compare different screening regimes and tests. Figure 6.3 shows the relationship between doubling time and probability of metastases for various screening policies. These were calculated using Tables 6.1 and 6.2 and Figure 6.2.

To find the probability of metastases for a screening population involves knowledge of the distribution of doubling times. For our calculations we take Kusuma's data

Figure 6.2. Probability of metastases according to delay for various doubling times.



(see Chapter 3.5.)and used Figure 6.3. Table 6.5 shows the results.

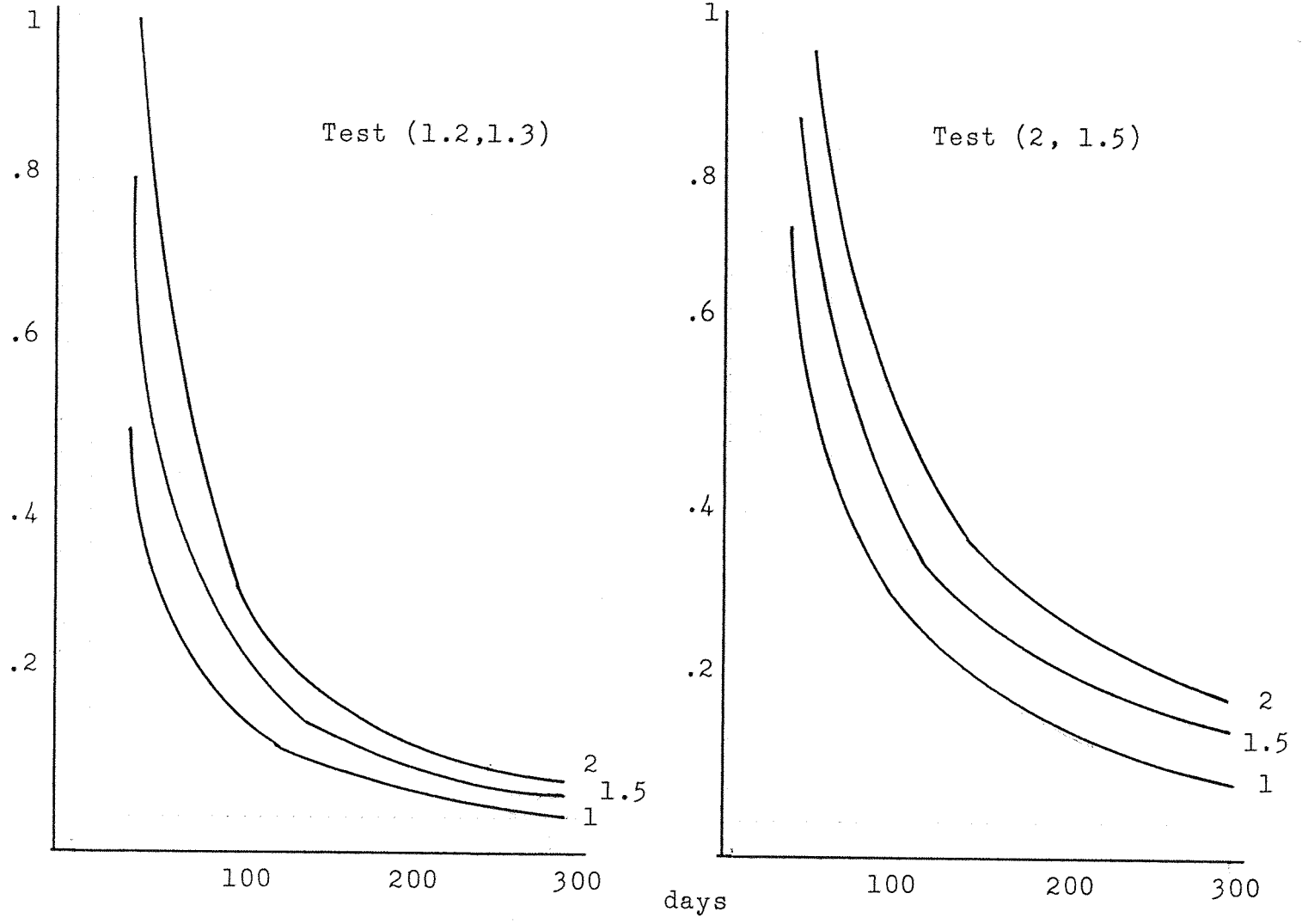
Table 6.5. Probability of metastases for different screening regimes

Test parameters	1 year	1.5 years	2 years
(1.2,1.3)	.22	.34	.45
(2,1.5)	.36	.48	.58

The figures in Table 6.5. give an indication of the reduction in metastases that may be expected by more frequent screening and by using a more sensitive test. One year screening instead of two year screening could reduce metastases by 51% in a more sensitive test and by 38% if a less sensitive test is used.

By using the Edinburgh data on size at detection(Kerr,1982) and size at detection when BSE is practised(Foster,1977) we can calculate the probability of metastases when there is no screening and when BSE is performed. The estimates of metastases are 55% using Kerr, 37% for BSE monthly and 45% for BSE less often. This suggests that regular and frequent BSE could reduce metastases by 33%.

Figure 6.3. Probability of metastases according to doubling time of tumour for test parameters (1.2,1.3) and (2,1.5) with screening times 1 to 2 years.



6.4 Benefits and costs

The subject of the benefits and costs of screening have been tackled in various ways involving lifetable and other approaches which include allowing for radiation risk. (Chiacchierini, 1978; Fox, 1978; Schweiter, 1974; Dubin, 1979). Such approaches involve an evaluation of human life or years of life lost. I have restricted this sections to a few comments on benefits and costs since such evaluation goes beyond the scope of this thesis.

Most screening treatment in the UK is done by the National Health Service which employs cost-effectiveness in the determination of its policies. This avoids the problems involved in measuring the value of life.

The cost of screening does not only include the cost of screening but also the cost of referals, biopsies and treatment. It varies considerably according to the type of test and also who carries out the tests and if double checks are made. See Table 6.6.

Savings to the NHS may occur if costs of treatment of early cancers found at screening are less than if cases were detected when symptoms showed. Estimates were made of the costs of treatment including nursing care over 20 years (or until patient death) by Simpson(1977). See Table 6.7.

Table 6.6. Cost of screening

(Source: Chamberlain, 1979)

Test	Cost(£)
CE(nurse)	3.72
CE(doctor)	4.40
Mammogram	8.58
Mamm.+ nurse	9.55
Mamm.+ doctor	10.24
Doctor+nurse+double reading on mamm.	13.62

Table 6.7. Cost of treatment according to stage

(Source: Simpson, 1977)

Stage	Primary treatment	Total cost up to 20 years*
I	762	1980
II	988}	
III	969}	2860
IV	451	3190

* discounted at 7%

The classification of stage at diagnosis involves an identification problem since metastases may be present but not detectable. We need to compare the cost of treatment without metastatic spread to one with it. Possible a Stage I cancer may be taken as representing the former since the main cost involves masectomy without the additional costs of radiation treatment or chemotherapy. Thus one facet of the costs and benefits to the health service involves equating the cost of screening, and the treatment of screen detected cancers with the cost of treatment of cancers for patients not offered screening, to find a break even point.

The cost of performing n screens is

$$s_c = nc_t + (an + y)c_b \qquad 6.4$$

where c_t = cost of each test

n = number of tests performed

a = false positive rate

c_b = cost of biopsies

y = the yield

If we assume that the probability of metastases for the screened population is m_s and without screening m_w then the cost of treating screened cancers is

$$c_s = y\{m\ t_m + (1-m)\ t_w\} \qquad 6.5$$

where t_m = treatment cost when there are metastases

and t_w = treatment cost when there are no metastases.

The cost of non-screen detected cancers is

$$c_{ns} = y\{m\ t_w - (1-m)\ t_w\} \qquad 6.6$$

If we equate the cost of screening and treating screened detected cancers with the cost of treatment for non-screened detected ones we can estimate the yield required per screen to break-even.

The following points should be noted:

1. This is a simplified version since the cancers would not be detected in the same time period as the screened cancers and hence the cost should be discounted to present value.

2. The estimation of the probability of metastases used above are an average for continuous screening. On the first screen, tumour sizes will be larger and hence the probability of metastases greater.

3. The yield for any screening programme will vary according to the population screened eg. their risk. The yield on subsequent screens will drop and depend on the screening interval.

4. The above considerations only reflect the cost to the NHS and adding other benefits that accrue from saving lives

will indicate a lower yield necessary for a break-even point.

6.6 Conclusions

The inputs into a screening model for breast cancer depends on what information is required as output. In the previous sections I have gathered together observations on various facets of the disease, and have formalised them into mathematical relationships.

A complete model for breast cancer screening would involve the following inputs:

1. Incidence especially related to women with different characteristics or risk factors.
2. Knowledge of tumour growth and distribution of growth rates and whether rates are affected by age or menopausal status.
3. The tumour size for patient detection and the duration in a detectable state for screening.
4. The spread of breast cancer according to size of tumour.
5. The ability of different tests to detect tumours, and how this varies according to tumour size, type and location.
6. Women's behaviour between screens -principally BSE.

In addition a model to evaluate the costs and benefits of screening would require:

1. The cost of screening tests and further tests eg biopsy.
2. The cost of treatment according to extent of disease at

diagnosis and continuing costs of treatment.

3. Attendence rates at screening.

In Chapters 4:3 and 6:2-4, I have investigated two sub-models. The first involves calculation of the mean screening delay postulated by Shahani(1977) which I have extended to allow for varying false negative rates according to tumour size, patient detection between screens and differing starting ages for screening. the second which I have derived is a simple model which can be used to compare the probability of metastases according to the screening interval and type of test used.

The complete picture of the effect of screening by using a screening model for breast cancer to emulate screening in the population would best be constructed by using a simulation approach. Each woman could be portrayed by an information matrix containing her age(or D.O.B.), the probability of getting cancer, the probability of attendence at screening, her BSE practise, and information conditional on her getting cancer namely:age at which the cancer becomes detectable, the growth rate, and the size of tumour when metastases will be established. Schematic diagrams for screening a population have been included in Chapter 4.

Thus for each woman, the effect of screening can be assessed; specifically the cost given no disease, or the cost when disease is detected, the spread at detection and

subsequent cost of treatment. By varying the age at start of screening, the frequency of the screening and the type of test used, it is possible to assess the likely effects and costs of screening a specific population. Lastly, it must be remembered that the actual effects of any screening regime can only be assessed by population based screening, that is by performing clinical trials.

Alcoe, S.Y. and McDermot, R. (1979). Is there a case for teaching Breast Self Examination . Can. J. Public Health 70, 418-422

Armstrong, B. (1976) Recent trends in breast cancer incidence and mortality in relation to changes in possible risk factors. Int. J. Cancer 17, 204-211.

Baal, E. et al. (1978). Occurance of breast carcinoma in women treated with X-rays for benign disease of the breast, in Prevention and Detection of Cancer, H.E. Nieburgs (ed). Marcel Dekker.

Bartosynski, R. et al. (1982). Metastatic and systemic factors in neoplastic progression, in Probability models and cancer, L. le. Cam and J. Neyman (eds), North Holland Publishing Co.

Baum, M. (1981). Breast Cancer: the facts. O.U.P.

Baum, M et al. (1982). Clinical trials in Early breast cancer. Birkhauser Verlag. Basel.

Berkson, J. and Gage, R.P. (1952). Survival curve for cancer patients following treatment. J. Am. Stat. Assoc. 47, 501-515.

Blumenson, L.E. and Bross, Irwin. (1969). A mathematical analysis of the growth and spread of breast cancer. Biometrics, 25, 95-109.

Bone , M. (1978). The family planning services: changes and effects. OPCS.

Boyd, N.F. (1978). A preliminary appraisal of the effects of length biased sampling and observer variability on screening, in Screening in Cancer, A.B. Miller (ed). UICC Technical Report Series Vol 40. Geneva.

Boyd, N.F. et al. (1981) Clinical estimation of the growth rate of breast cancer. Cancer 48, 1037-1042.

Boyle, M. et al. (1981). Effectiveness of a community program to promote early breast cancer detection. J. Surgical Oncol. 18, 183-188.

Breslow, L. et al. (1977). Report of NCI Ad hoc working gropu on the gross and net benefits of mammography in mass screening for the detection of breast cancer. J. Nat. Cancer Inst. 59, 473-478.

Brinton, L.A. et al. (1979). Breast cancer risk factors among screening program participants. J. Nat. Cancer Inst. 62, 37-43.

Chamberlain, J. et al. (1979). Error rates in screening for breast cancer by clinical examination and mammography. Clinical Oncology, 5, 135-146.

Charlson, M. (1982) Clinical indices of disease progression, in Clinical trials in early breast cancer. M. Baum (ed)

Charlson, M.E. and Feinstein, A.R. (1980). A new clinical index of growth rate in the staging of breast cancer. Am. J. Med. 69, 527-535.

Charlson, M.E. and Feinstein, A.R. (1984). Rate of disease progression in breast cancer. A clinical estimate of prognosis within nodal and anatomic stages. JNCI 72, 225-231.

Collette, H.J. et al. (19). Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1, 1224-1226.

Coombs, L.J. (1979). A prospective study of the relationship between benign breast diseases and breast carcinoma. Prev Med. 8, 40-52.

Cutler, J. and Myers, M.H. (1967). Clinical classification of extent of disease in cancer of the breast. J. National Cancer Inst. 39, 193-207.

Cuzick, J. (1986). The prevention of breast cancer. Lancet Jan. 83-86.

Davey, J.B. (1976). Thermography in the early detection of breast cancer in Health Control in Detection of Breast cancer. H. Bostrom (ed).

Day, N.E. and Wahrendorf, J. (1982). Possibilities for the reduction of breast cancer mortality, in Clinical trials in early breast cancer M. Baum (ed).

Dept of Health and Social Security. (DHSS). (1976). Prevention and Health: Everybody's Business. H.M.S.O.

Dodd, G.D. (1977). Present status of Thermography, Ultrasound and Mammography in breast cancer detection. Cancer 39, 2796-2805.

Dunn, J. (1969). Epidemiology and possible identification of high risk groups that could develop cancer of the breast. Cancer 23, 775-780.

Egan. R and Mosteller, R. (1977). Breast Cancer mammography patterns. Cancer 23, 775-780.

Elandt-Johnson, R.C. (1977). Age of onset distribution of a disease. Estimation from incidence data. Inst. of Stat. Mimeo Series. No 119. University of Carolina.

Elandt-Johnson R.C. and N.L. (1980) Survival models and data analysis. Wiley.

Elwood, K. and Moorehead, W.P. (1980). Delay in diagnosis and long-term survival in breast cancer. B.M.J. 1, 1291-1294.

Farewell, V. (1977). The combined effects of breast cancer risk factors. Cancer 40, 931-936.

Feig, S.A. et al. (1978). Detection of early breast cancer. Evaluation of 16754 women by xero-mammography, thermography and clinical examination in Prevention and Detection of Cancer. H.E. Nieburgs (ed), Marcel Dekker.

Feldman, J.G. et al. (1981). Breast self examination: Relationship to stage of breast cancer at diagnosis. Cancer 47, 2740-2745.

Forrest, A.P.M. and Kunkler, P.B. (1968)
. Prognostic Factors in Breast Cancer. Livingstone.

Foster, R.S. et al. (1978). Breast self-examination practices and breast cancer stage.
New Eng. J. of Med., 299, 265-270.

Fournier, Von D. (1978). Growth rate of 53 primary breast carcinomas with an observation time of 0.2 to 11 years, in Prevention and Detection of Cancer,
H.E. Nieburgs. Marcel Dekker.

Fournier, D.V. (1980). Growth rate of 147 mammary carcinomas. Cancer 45, 2198-2207.

Fox, S.H. et al. (1978). Benefit/Risk analysis of aggressive mammographic screening.
Radiology 128, 359-365.

Friedman, A.K. et al. (1966). A co-operative evaluation of mammography in seven teaching hospitals.
Radiology 86, 886-891.

Gastrin, G. (1981). Breast Cancer control (an early detection programme). Almquist and Wiksell, Stockholm.

George W.D. et al. (1976). Screening for breast cancer. B.M.J. Oct, 858-860.

George, W.D. et al. (1980). Hospital work load produced by breast cancer screening programms run by trained non-medical staff. B.M.J. 281, 653-654.

Gershon-Cohen, J. et al. (1963). Roentgenography of breast cancer moderating concept of biologic determinism. Cancer 16, 691.

Gold, J. and Josimouich, J. (eds) (1980). Gynecologic Endocrinology. Harper and Row.

Goldberg, J.D. and Wittes, J.J. (1978). The estimation of false negatives in medical screening. Biometrics 34, 77-86.

Grove, J. S. (1985). Wolfe's mammographic classification and breast cancer risk: The effect of misclassification on apparent risk ratios. Br. J. Radiology. 58, 15-19.

Hall, E.J. and Laing, A.H. (1968). Growth rate of tumours in Prognostic Factors in Breast Cancer. A.P.M. Forrest and P.B. Kunkler (eds). Livingstone.

Haybrittle, J. (1981). Prognostic factors in breast cancer. Review on Endocrine-related cancer. Supp. 10, 49.

Hermann, G et al. (1982). Non-palpable tumour of the breast - radiologic presentation. Br. J. Radiography 55, 623.

Heuser, L. et al. (1979). Relation between mammary cancer growth, kinetics and intervals between screening. Cancer 43, 857-862.

Heuser, L. et al. (1979). Growth rates of primary breast cancers. Cancer 43, 1888-1894.

Holliday, H.W. et al. (1983). Initial results from a programme of breast self-examination. Clinical Onconolgy. 9, 11-16.

Huguley, C.M. and Brown, R.L. (1981). The value of breast self-examination. Cancer 47, 989-995.

Hutchinson, G and Shapiro, S. (1968). Lead-time gained by diagnostic screening for breast cancer. J. Nat. Cancer Inst. 41, 655-681.

Imperial Cancer Research Fund. (1982). 1981 Scientific Report.

Imperial Cancer Research Fund. (1983).
1982 Scientific Report.

Jakobsson, S and Lungren, B. (1976). The efficiency of breast cancer screening by a simplified method of mammography, in Health Control in Detection of Cancer, H. Bostrom (ed). Almquist and Wiksell International, Stockholm.

Kerr. (1982) private communication.

Koscienlny, S. et al. (1984). Breast cancer: relationship between the size of the primary tumour and the probability of metastatic dissemination. Br. J. Cancer. 49, 709-715.

Koscienlny, S. et al. (1985). A simulation model of the natural history of human breast cancer.
Br. J. of Cancer. 52, 515-524.

Krook, P.M. (1978). Mammographyic parenchymal patterns as risk indicators for incident cancer in a screening program: an extended analysis.
Am. J. Roentgenology. 131, 1031-1035.

Kusuma, S. et al. (1972). The gross rates of growth of human mammary carcinoma. Cancer 30, 594-599.

Langlands, B. (1979). Long term survival of patientsa with breast cancer: a study of the curability of the disease. B.M.J. 2, 1247-1251.

Malaise, E.P. et al. (1973). The relationship between growth rate, labelling index and histological type of human solid tumours. Eurp. J. Cancer 9, 305.

Meyer, J.S. and Bauer, W.C. (1976). Tritiated thymidine labelling index of benign and malignant human breast epithelium. J. Surg. Oncol. 8, 165-181.

Miller, A.B. (ed) (1978). Screening in Cancer. UICC Technical Report Series, Vol 40. Geneva.

Miller, A.B. (1985). Self-examination in the early detection of breast cancer. J.Chronic Diseases. 38, 527-540.

Moolgavkar, S.H. (1980)/Two stage model for carcinogenisis epidemiology of breast cancer in females.
JNCI 65, 559-569.

Office of Population Census and Surveys. (OPCS) (1975). Medical and Population Studies No 29. HMSO.

OPCS (1978). Trends in Mortality. Series DH1 No 3.HMSO.

OPCS(1983)1978 Cancer Statistics.Series MB1 No 10. HMSO.

Petrakis,N.L.(1982).Breast in Cancer Epidemiology and Prevention. D. Schottenfeld and J.H. Fraumeni. (eds). W.B. Saunder Co. Philadelphia.

Prorok, P.C. et al. (1981). Concepts and problems in the evaluation of screening programmes. J. Chronic Disease. 34, 159-171.

Rombach, J. (1980). Breast cancer screening. Staflieus Scientific Pub. Co. Brussels.

Rutquist, L.R. (1985). Long-term survival of 485 young breast cancer patients. Cancer 55, 658-665.

Scheurlen, H, (1982) On the biology of breast cancer, in Clinical Trials in early breast cancer M. Baum (ed.)

Schmitt, E.L. and Threatt, D. (1984) Characterisitcs of breast cancer in an incident cancer population. Am. J. Radiology 143, 403-406.

Schwartz, M. (1978). A mathematical model used to analyse breast cancer screening strategies.Op.Res. 26, 937-955.

Seidman, H. (1972). Cancer of the Breast:Statistical and Epidemiological Data. Am. Cancer Soc. New York.

de Senna, V. (1983). Early detection of breast cancer: the choice of a screening policy. Phd Thesis. University of Southampton.

Shahani, A.K, and Crease, D.M. (1977). Towards models of screening for early detection of disease. Adv. Appl. Prob. 9, 605-680.

Shapiro, S. (1977). Evidence on screening for breast cancer from a randomised trial. Cancer 39, 2772-2782.

Shapiro, S. et al. (1973). Changes in 5 year breast cancer mortality in a breast cancer screening programme. 7th National Cancer Conference Proceedings. Am. Cancer Soc.

Shapiro, S et al. (1968). The search for risk factors in breast cancer. AM. J. Public Health. 58, 820-835.

Smith, E.M. et al. (1980). The effect of breast self-examination practices and physician examination on the extent of disease at diagnosis. Prev. Med. 9,409-417.

Smithers, D.W. (1968).Clinical assessment of growth rates in human tumours.Clinical Radiology. 19, 113-119.

Soini, I and Hakama, M. (1978). Failure of selective screening for breast cancer by combining risk factors. Int. J. Cancer 22, 275-281.

Soini, I. and Hakama, M. (1979). Inverse association between risk factors for benign and malignant lesions. Scand. J. Soc. Med. 7, 79-85.

Spratt, J. amd Spratt, J. (1979). Growth rates in Cancer of the Breast, W.L. Donegan and J. Spratt (eds).

Stark, A.M. (1976). Critical appraisal of new detection techniques, in Breast Cancer: Trends in Research and Treatment, J. C. Heuson (ed). Raven Press.

Stark, A. M. and Way, S. (1974). The screening of well women for the early detection of breast cancer using clinical examination with thermography and mammography. Cancer 33, 1671-1679.

Stravraky, K. and Emmon, S. (1974). Breast cancer in premenopausal and post-menopausal women. J. Nat. Cancer Inst. 53, 647-654.

Strax, P. (1981). The HIP study update. Reviews on Endocrine-Related Cancer Supp 10, 59.

Steel, G.G. (1967). Cell loss as a factor in the growth rate of human tumours. Eurp. J. Cancer 3, 381-387.

Strax, P. et al. (1978). Millirad mammography in mass screening, in Prevention and Detection of Cancer, H.E. Nieburgs. Marcel Dekker.

Stroll, B.A.(ed) (1976).Risk factors in Breast Cancer. Vol 2 in series New Aspects of Breast Cancer. W. Heinmann Medical Books.

Tabar, L. et al. (1985) Reduction in mortality from breast cancer after mass screening with mammography. Lancet April 1, 829-832

Tocher, K.D. (1963) The Art of Simulation Hodder and Stoughton.

Thomas, B. A. et al. (1981) Breast cancer population screening by single view mammography with selective clinical examination- a pilot study. Clinical Oncology 7, 201-204.

Thomas, B. A. et al. (1983) The Guildford breast screening projects. Clinical Oncology 9, 121-129.

Tulinius, H. et al. (1978). Reproductive factors and risk for breast cancer in Iceland.
Int. J. Cancer. 21, 724-730.

UK Trial for Early Detection of Breast Cancer.
Br. J. Cancer 44, 618.

University of Southampton and Wessex Regional Cancer Organisation. (1983). Southampton Breast Study.

Venet, L. et al. (1971). Adaquacies and inadaquacies of breast examinations by physicians in mass screening.
Cancer 28, 1546-1557.

Verbeek, A.L. et al. (1984) Reduction of breast cancer mortality through mass screening with modern mammography.
Lancet 1, 1222-1224.

de Waard, F. (1978). Recent trends in breast cancer incidence. Prev. Med. 7, 160-167.

de Waard, F. et al (1964). The bimodal age distribution of patients with mammary carcinoma. Cancer 17, 141-151.

Walter, S. and Day, N. (1983). Estimation of the duration of a pre-clinical disease state using screening data.
Am. J. Epidemiology. 118, 865-885.

Wittes,J. et al.(1974). Capture-recapture methods for assessing the completemess of case ascertainment when using multiple information sources.
J. Chronic Diseases 27, 25-36

Wolfe, J. N. (1976). Risk for breast cancer development determined by mammographic parencymal pattern.
Cancer 37, 2486-2492.