Breast cancer in the West Midlands : an evaluation of screening, treatment and survival

by
Alison Mary Jennifer Teale

A thesis submitted to the Faculty of Medicine and Dentistry
of the University of Birmingham
for the degree of

DOCTOR OF PHILOSOPHY

Department of Public Health and Epidemiology
Faculty of Medicine and Dentistry
The University of Birmingham
April 1997

ABSTRACT

High quality population based data are essential to evaluate breast cancer. This study has improved data quality at the West Midlands Cancer Intelligence Unit through the development of a staging algorithm which significantly increased the proportion of cases with stage recorded. It has also demonstrated the benefits of combining cancer registry data with those from the **NHS** Breast Screening Programme (NHSBSP). This study identified stage as the most significant prognostic factor for breast cancer patients, and the favourable staging characteristics of screen detected cancers thus suggest that improved survival should result. The study also found that only ***3*** 1.6% of cancers in eligible women were screen detected in the prevalent round. **As *3*** 1% of breast cancers arose in women in the screening age range, only an estimated 10% of breast cancer patients benefit directly from screening. High interval cancer rates were identified, along with the need for clarification of the definitions used to identify interval cancers. This investigation identified changes in treatment over time across the region towards the King's Fund guidelines. However, treatment varied across the region and clear divergences from the guidelines were apparent. Furthermore, no association was identified between surgical caseload and survival, suggesting that the use of caseload as a proxy for specialist care may be questionable.

ACKNOWLEDGEMENTS

I would like to thank the following people for their contribution to this investigation

The Research and Development Department of the West Midlands **NHS** Executive Regional Office for their generous award of a New Blood Fellowship to fund this work

Dr Gill Lawrence, Dr Heather Winter and Professor Rod Griffiths for setting up and supervising the project

Dr Matthew Wallis, Mrs Margot Wheaton and staff at the Warwickshire, Solihull and Coventry Breast Screening Unit, and to all staff working in the **NHS** Breast Screening Programme in the West Midlands for contributing data for this project

Staff at the Department of Public Health and Epidemiology, University of Birmingham for their advice and support

Staff at the West Midlands Cancer Intelligence Unit, without whose work none of this would have been possible

Ms Dawn Richardson for her valuable assistance, in particular the collection of screening histories and assistance with the investigation of interval breast cancers

Mr Amreek Dhillon and Miss Rachel Oakes for assistance with breast screening data

Mr Ralph Smith and Dr Lillian Sommervaille for providing access to, and assistance with the Geographical Information System at the West Midlands Cancer Intelligence Unit

Dr Cathy Lloyd and **Mr** Ian Calvert for statistical advice

To my parents,

Steve,

and Holly

CONTENTS

APPENDICES

TABLES AND FIGURES

APPENDIX TABLES AND FIGURES

LIST OF ABBREVIATIONS

BASO	British Association of Surgical Oncologists
BBG	British Breast Group
CI	Confidence Interval
CMF	Cyclophosphamide, Methatrexate, and Fluorouracil
CO	Conservative surgery
COG	Clinical Outcomes Group
CT	Chemotherapy
DCIS	Ductal Carcinoma *In Situ*
DCO	a cancer registered from a Death Certificate Only
df	degrees of freedom
IDC	Invasive Ductal Carcinoma
DHA	District Health Authority
DOH	Department of Health
EBCTCG	Early Breast Cancer Triallists Collaborative Group
FHSA	Family Health Services Authority
FNA	Fine Needle Aspiration Cytology
GIS	Geographical Information System
GP	General Practitioner
GRACE	Generic Registration And Cancer Epidemiology System
HMSO	Her Majesties Stationery Office
HoN	Health of the Nation
HQA	Histology Quality Assurance report
HT	Hormone Therapy
HRT	Hormone Replacement Therapy
IARC	International Agency for Research on Cancer
ICDR	Invasive Cancer Detection Rate
IN	Interval breast cancer
KC62	Korner statistical return detailing breast screening activity
LCIS	Lobular Carcinoma *In **Situ***
LL	Log Likelihood Ratio
M:B	Malignant to Benign open biopsy ratio
MA	Mastectomy
MDSCR	Minimum Data Set for Cancer Registration

MH	Mantel Haenzel (Chi square test)
NA	Breast cancer in a screening non-attender
NBSS	National Breast Screening System (computer record system)
NHSBSP	National Health Service Breast Screening Programme
NPI	Nottingham Prognostic Index
NS	Non-surgical oncology
NST	No Special Type
ONS	Office for National Statistics (includes the former OPCS merged with the central Statistical Office in 1996)
OPCS	Office of Population, Censuses and Surveys
PDP11	The computer used to store the WMCIU database up to 1994
PNL	Prior Notification List
PPV	Positive Predictive Value
QA	Quality Assurance
QASS	Quality Assurance Surgical Summary report
RHA	Regional Health Authority
RR	Relative **Risk**
RT	Radiotherapy
SD	Screen detected breast cancer
SDR	Standardised Detection Ratio
SIR	Standardised Incidence Ratio
SMR	Standardised Mortality Ratio
SNOMED	Systematised Nomenclature of Medicine
SPSS	Statistical Package for the Social Sciences
TCR	Thames Cancer Registry
TNM	Tumour, Nodes, Metastases cancer staging system
UICC	Union Internationale Contre le Cancer
UK	United Kingdom
UKACR	United Kingdom Association of Cancer Registries
UKCCC	United Kingdom Consultative Committee on Cancer Research
WHO	World Health Organisation
WMCIU	West Midlands Cancer Intelligence Unit
WMRHA	West Midlands Regional Health Authority
WSC	Warwickshire, Solihull and Coventry (screening service)

WEST MIDLANDS DISTRICT HEALTH AUTHORITIES
(Current April 1994 - March 1996)

M02	Hereford
M04	Worcester
M05	Shropshire
M07	North Staffordshire
M17	Coventry
M18	Dudley
M19	Sandwell
M20	Solihull
M21	Walsall
M22	Wolverhampton
M25	South Birmingham
M26	North Worcestershire
M27	South Staffordshire
M28	Warwickshire
M29	North Birmingham

WEST MIDLANDS NHS BREAST SCREENING PROGRAMME UNIT CODES

MBD	West, North and East Birmingham
MBS	South Birmingham
MBU	South East Staffordshire
MCO	Warwickshire, Solihull and Coventry
MDU	Dudley and Wolverhampton
MHW	Hereford and Worcester
MSF	Mid Staffordshire
MST	North Staffordshire
MSH	Shropshire
MWA	Walsall and Sandwell

CHAPTER ONE

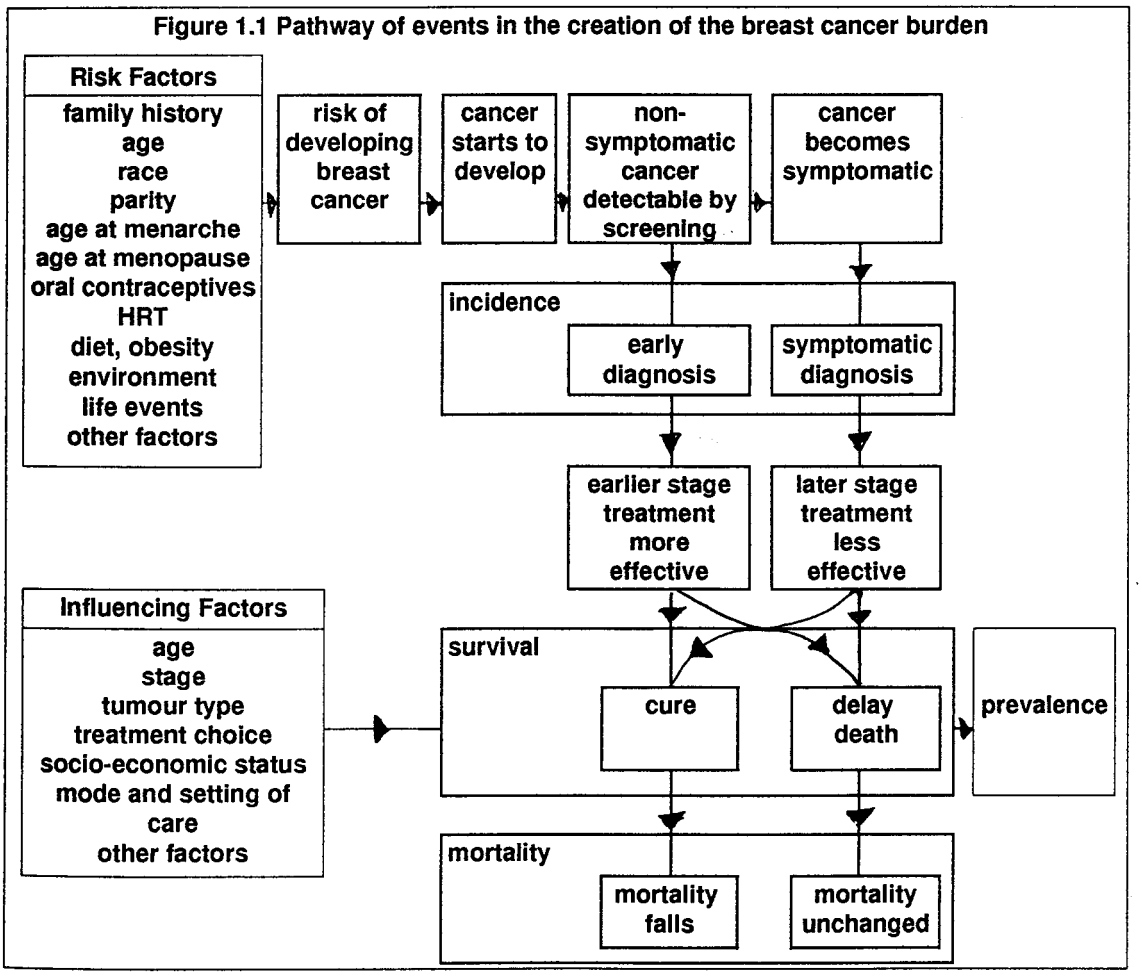
INTRODUCTION

1.1 FOREWORD

Breast cancer is the most common female neoplasm in the developed world **(IARC, 1992).** In 1988 over 26,700 cases were reported in England and Wales (Office of Population, Censuses and Surveys (OPCS), 1994), and in 1989 over 14,000 women died from the disease (OPCS, 1991). This thesis examines female breast cancer in the West Midlands and the impact of the NHS Breast Screening Programme (NHSBSP) in this region.

The burden created by breast cancer can be depicted by the processes shown in Figure 1.1. Incidence is determined by a number of **risk** factors, and reflected in mortality which in turn reflects curabilty, which is influenced by factors such as stage at diagnosis.

Figure 1.1 Pathway of events in the creation of the breast cancer burden



With breast cancer, deaths attributable to the disease can continue over a long period and the definition of cure is not simple. The impact of treatment is often determined by the stage of disease, and the earlier breast cancer is diagnosed, the more likely a patient is to enjoy prolonged remission. However, some "early" cancers are aggressive, and respond poorly to treatment, and occasionally patients with advanced cancers respond well to treatment and enjoy prolonged remission.

This chapter provides an overview of the incidence of breast cancer, prognostic factors, and a brief description of the introduction and organisation of NHSBSP.

1.2 THE INCIDENCE OF BREAST CANCER

The highest incidence rates of breast cancer are found in more economically developed countries including North America and Northern Europe, and lowest rates in less developed areas such as Asia. Japan has exceptionally low rates, although these have been increasing (Coleman *et al.,* 1993). These international differences could reflect genetic predisposition to breast cancer, or differences in exposure to risk factors associated with lifestyle or environmental exposures in different countries. Migration studies have provided evidence that risk is, in part, influenced by external factors. Japanese women living in the US experience higher rates than those in Japan, and rates in second-generation immigrants are higher still (Buell 1973). Furthermore, as lifestyle in Japan has become more Westernised, so breast cancer risk has risen. Thus the incidence in Miyagi, Japan for 35-74 year olds in 1970 was 27.8/100,000 rising to 61.2/100,000 in 1985 (Coleman *et al.,* 1993). The risk of breast cancer rises as women approach middle age, approximately doubling with each decade of life until the menopause after which the rise slows down (McPherson *et al.,* 1994). The highest risk approximately co-incides with the menopause, an observation suggestive of a hormonal influence in pathogenesis.

A full review of the evidence for putative aetiological factors is inappropriate here but the data presented in recent reviews (McPherson *et al.,* 1994, Torenson, 1994) are summarised in Table 1.1.

Table 1.1 Established and probable risk factors for breast cancer		
Factor	Relative risk	High risk group
Age	>10	Elderly
Geographical location	5	Developed country
Age at menarche	3	Menarche before age 11
Age at menopause	2	Menopause after age 54
Age at first full term pregnancy	3	First child in early 40s
Family history	>=2	Breast cancer in first degree relative when young
Previous benign disease	4 - 5	Atypical hyperplasia
Socio-economic group	2	Groups I and II
Diet	1.5	High intake of saturated fat
Body weight - premenopausal	0.7	Body mass index >35
Body weight - postmenopausal	2	Body mass index >35
Alcohol consumption	1.3	Excessive intake
Exposure to ionising radiation	3	Abnormal exposure in females after age 10
Oral contraceptives	2	Use for >=4 years when young
Hormone replacement therapy	1.5	Use for >=10 years

1.3 PROGNOSTIC FACTORS FOR BREAST CANCER

Some characteristics of breast tumours have been identified as having prognostic significance. Such prognostic factors provide a means of estimating likely overall survival, **aid** selection of appropriate treatment, allow comparison between groups of patients and improve understanding of the biology of breast cancer. In recent years many putative prognostic factors have been identified and these have been recently reviewed in a **UICC** publication (Hermaneck *et al.*, 1995). **As** the authors of this review point out, the proliferation of prognostic factors, although providing the potential to increase predictive accuracy, suffers from both poor reproducibility of results, and the inability of a large number of factors to be incorporated into a predictive system. For this reason, only those factors that are reasonably well established are presented here.

Prognosis can be described in terms of two time intervals - the time to relapse (the disease-free interval), and the overall survival. Although local control is advantageous to the patient, there is no firm evidence that it always translates into improved overall survival (Dixon 1995). Tumour factors are sometimes classified into two groups, those that are "chronological", and those that are "biological". Chronological factors indicate the length of time that the tumour has been growing. Biological factors relate to innate

histological, biochemical or genetic properties of tumours which influence malignant behaviour and are assumed by some to remain constant throughout the lifetime of the tumour and are thus phenotypic properties of the tumour. The use of such classifications may be misleading as some factors may reflect both biology and chronology. Opinion is divided over whether biological factors change with time in a process termed phenotypic progression. **A** Finnish study concluded that such progression did not occur (Hakama *et al.,* 1995), but doubts over the validity of these conclusions have been raised due to small sample size and low sensitivity of the study (Day *et al.,* 1995). It seems likely that factors such as nodal status are both chronological and biological (Pater *et al.,* 1981). The term "anatomical" is used in this thesis to describe factors which **can** be observed clinically at diagnosis (although pathological confirmation is always desirable). The term "biological" is used to describe factors for which laboratory investigation is required.

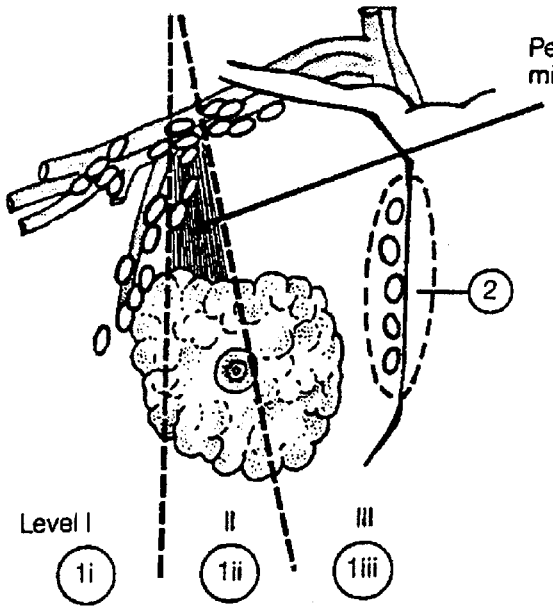
Breast cancers arise from epithelial cells lining the terminal lactiferous duct units in the breast. Cancer cells are initially confined to the duct or lobule, and are termed *in situ.* Dissemination though the basement membrane of the duct indicates invasive cancer. Tumours usually grow initially as a single mass within the breast, although some tumours are diffuse, or multi-focal, and a few primary breast tumours are sited in the axilla. The incidence of tumour recurrence post-operatively suggests that, although the majority of the tumour bulk is sited in a single mass, in many cases micro-metastatic disease is present even with apparently localised disease (Sacks *et al.,* 1993).

The breast is well supplied by lymphatic vessels which drain into nodes in the axilla and the internal nodes. The regional nodes include the ipsilateral axillary nodes, the interpectoral (Rotter's) mammary nodes and ipsilateral mammary lymph nodes. The axillary lymph nodes are divided into three levels according to their position in relation to the pectorals minor muscle, shown in Figure 1.2.

On average there are 20 nodes in each axilla, most being at level I (Bundred *et al.,* 1994). **As** tumour size increases risk of metastasis to the axilla rises, with spread to level I followed by level II and III. Occasionally "skip" metastasis occurs to level II or III via the internal mammary nodes without level I involvement (Sacks *et al.,* 1993). Involved

nodes are initially mobile, but become fixed to surrounding tissue if the cancer breaches the capsule. Fixed nodes are associated with poorer prognosis than mobile nodes.

Figure 1.2 The regional lymph nodes (Reproduced from TNM Atlas, 1992, UICC)



1. Axillary nodes

i. Level I low axilla: lymph nodes lateral to the lateral border of the pectoralis major muscle

ii. Level II mid axilla: lymph nodes between the medial and lateral borders of the pectoralis minor muscle and the interpectoral (Rotter) nodes

iii. Level III low axilla: lymph nodes medial to the medial margin of the pectoralis minor muscle

2. Internal mammary nodes

The degree of progression in relation to the site of origin is termed the disease extent. **As** disease progresses, the tumour extends beyond the breast, leading to ulceration of the skin, or fixation to the chest wall. If lymphatic vessels in the skin are blocked, oedematous swelling can give the appearance of orange peel ("peau d'orange"). The final stages of disease are reached when distant lymph nodes including the cervical, supra-clavicular, contralateral axillary or contralateral internal mammary nodes become involved or metastases appear in sites including bone, brain, liver and lungs.

1.3.1 Anatomical prognostic factors

Tumour size correlates directly with overall survival, increased tumour size being associated with decreased survival (Carter ***et*** *al.,* 1989). The pathological size of tumours is more valuable in prognosis than the clinical size (Hermaneck ***et*** *al.,* 1995), but cannot be used where the tumour was removed in several pieces, if the excision margins are not free of tumour, or where pre-operative tumour-shrinking therapy has been given. The status of the local nodes indicates the tendency of tumours to

metastasise, and again, pathological data are more prognostically valuable than clinical data (Hermaneck *et al.,* 1995). Patients with positive nodes show poorer survival than those with negative nodes, and prognosis worsens as the number of involved nodes rises. The influence of the status of the local nodes is additive with, and independent from, that of tumour size (Carter *et al.,* 1989). Metastases in distant sites are indicative of poor survival, although patients with metastasis to the supra-clavicular nodes have slightly better prognosis than those with metastases elsewhere (Miller *et al.,* 1994). The chronological nature of these factors means that status of a patient with respect to tumour size, nodal status and distant metastases is partly a reflection of the degree of delay prior to diagnosis, or the lead time. Patients presenting with small tumours, for example, will often have sought medical help earlier than those with larger tumours.

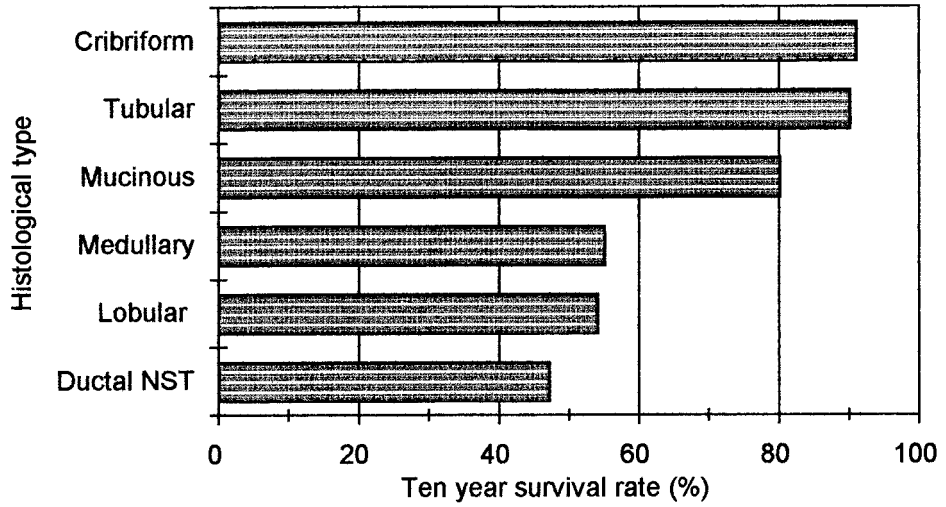
1.3.2 Biological factors

1. Tumour histology

Most breast cancers are invasive ductal adenocarcinomas. Non-invasive tumours make up a small proportion of symptomatic cancers and are subdivided into ductal carcinoma *in situ* (DCIS), and the rarer lobular variant (LCIS), depending on microscopic appearance. The risk of invasive cancer following *in situ* disease has been estimated at 25-50% at 20 years (Page *et al.,* 1995). *In situ* tumours are often associated with excellent prognosis. Concerns have been raised over inconsistency in the classification of *in situ* cancers, lack of consensus on the risk of subsequent invasive disease, and lack of knowledge of the underlying incidence of *in situ* cancer (Foucar, 1996; Fricker, 1996).

The majority *(>70%)* of ductal adenocarcinomas are of no special type (NST), having no distinguishing features on microscopic examination. **A** further 10-15% are lobular adenocarcinomas (Gallager, 1984). The remainder are "special types". These include tubular, cribriform, mucinous, papillary, medullary and adenoid cystic carcinoma. Some special types have favourable prognostic characteristics, for example mucinous tumours tend to be well differentiated (Gallagher, 1984; Ellis *et al.,* 1992). The favourable survival rates for some "special type" tumours can be seen in Figure 1.3, with ten year survival rates varying from **47%** for ductal carcinomas of no special type to 91% for cribriform carcinoma (Ellis *et al.,* 1992).

Figure 1.3 **Ten year survival rate for types of invasive breast adenocarcinoma**



Source: Ellis et al, 1992

ii. Tumour grade

Tumour grading is the classification of tumours based on morphological appearance, and is predictive of tumour behaviour. **A** number of histological features have been assessed in different grading systems including the degree of glandular formation, epithelial secretion, the size of cells, the size of nuclei, variation in the size of cells and nuclei, nuclear hyperchromatism and mitotic activity. The system currently recommended by the **NHS** Breast Screening Programme (NHSBSP, 1995a), and by the British Association of Surgical Oncologists (BASO, 1995) is that devised by Bloom and Richardson (Bloom ***et al.,*** 1957) and modified by Elston and his colleagues (Elston ***et al.,*** 1991). This system combines three features (tubule formation, nuclear pleomorphism, mitotic activity) for each of which a score of 1 to **3** is allocated based on the whether the feature is apparent to a slight (score i), moderate (score **2)** or marked degree (score **3).** Higher scores indicate a higher degree of malignancy. Tubule formation is lost in aggressive tumours and replaced by solid tumour growth. Assessment of the second feature, nuclear pleomorphism, quantifies the degree of variability of both size and shape of tumour nuclei. Normal tissue has cells with generally small nuclei, and little variation in nuclear size and shape. Classification of the third feature, mitotic activity, has been modified by Elston and his colleagues ***so*** that scoring is based on the number of mitoses per 10 high

power fields assessed at the tumour periphery, with a high mitotic count indicating a high score. The three scores are added, and banded into three grades. Grade I tumours display the best survival and grade III the poorest. Although in the past it was thought unnecessary to grade "special type" tumours, the NHSBSP currently recommends that grading is carried out for all histological subtypes. An obstacle to the widespread use of grade for prognostic purposes is lack of inter-observer reproducibility (Gilchrist *et al.,* 1985; Sloane *et al.,* 1994), although provision of clear guidelines has been found to improve consistency (Robbins *et al.,* 1995). Although grading has been shown to be prognostically useful, it is not always performed, and if recorded, it is difficult to be sure whether the recommended criteria (Elston *et al.,* 1991) were used.

Grading is also applied to *in situ* carcinomas in order to identify cancers that are likely to progress to invasive disease. Grading is based on nuclear appearance (Page *et al.,* 1995), and may be combined with the degree of comedo-type necrosis to form a prognostic classification (Silverstein *et al.,* 1995). *In situ* cancers most likely to progress are high grade. Extensive *in situ* tumour necrosis is also a sign of progressive potential.

iii. Lymphatic and vascular invasion

Lymphatic or vascular invasion is found in up to 25% of breast cancer patients, and is associated with increased risk of local recurrence after surgery (Miller *et al.,* 1994; Hermanek *et al.,* 1995).

iv. Other factors

A high proliferation rate increases the risk of poor survival and the ploidy of malignant cells is often abnormally high, with aneuploid tumours of worse prognosis than diploid tumours (Miller *et al.,* 1994).

The presence of oestrogen receptors in breast cancers has been used to predict likely response to hormonal manipulation, with receptor positive tumours producing a better response and a lower recurrence rate than receptor-negative tumours (Donegan, 1992).

Receptors for epidermal growth factor oncogene erbB2 are associated with poorer survival (Miller *et al.,* 1994). Alterations in the p53 tumour suppresser gene are the most

frequent genetic changes found in breast cancer, especially in familial breast cancer, and are associated with poor prognosis (Elledge *et al.,* 1993). Aggressive tumour behaviour has also been linked to reduced expression of the Nm23 gene (Steeg *et al.,* 1993).

1.3.3 Combining factors to produce prognostic indices

To improve prognostic accuracy systems have been produced which combine prognostic factors. The first system to be used in the UK was the clinical Manchester staging system (Paterson, 1948), which used tumour size, extent, and nodal status to assign stage. **A** number of different systems have been developed since then, based on the Manchester system. The commonest staging system used currently is the TNM system.

i. The TNM staging system

The TNM (Tumour, Nodes, Metastases) staging system was recommended in 1959 by the International Union Against Cancer (UICC) and the American Joint Committee on Cancer and has undergone revisions in 1974,1987, and 1992 (UICC, 1992). The criteria on which the TNM system is based are

- the size of the primary tumour (the T stage)
- the extent of invasion beyond the breast (also T stage)
- the malignant status and malignancy of axillary (local) lymph nodes (the N stage)
- the presence or absence of distant metastases (the M stage)

Appendix One provides a description of the classification of T, N, and M stages for breast cancer. Clinical or pathological data may be used, although clinical assessment of the lymph nodes may be unreliable and detracts from the usefulness of the staging system (Barr *et al.,* 1992), and exclusion of biological factors has also provoked criticism (Barr *et al.,* 1992, Sacks *et al.,* 1993). ***An*** important drawback of the system is that it takes no account of some factors which are critical in clinical decision-making, such as tumour grade, multi-focality, the presence of *in situ* in combination with invasive cancer, and the position of the tumour in relation to the nipple or chest wall. Despite these limitations TNM staging is useful for predicting survival (Hermanek *et al.,* 1995), and is in common use. Five year relative survival for breast cancers diagnosed in the UK from 1975-80 were reported as 84% for stage '1', 71% for stage '2', 48% for stage **'3',** and 18% for stage '4' cancers (Cancer Research Campaign, 1988). The 1992 TNM

revision is used at the WMCIU for the purposes of the national Minimum Data Set for Cancer Registration (MDSCR). However, lack of recorded stage at cancer registries is a recognised problem. For example, stage was available for only 0.89% of breast cancers diagnosed in 1982-86 in the South West region (Bristol Cancer Epidemiology Unit, 1995). That this lack of staging data is, at least in part, a reflection of the absence of staging at diagnosis is supported by a study in the South East in which, even in 1990, only **24%** of the 334 breast cancers had stage recorded in the hospital case notes (Chouillet *et al.,* 1994). The problem of absence of routinely recorded stage at cancer registries, and indeed the absence of staging at diagnosis, is a potential limitation to population-based assessments of prognosis, treatment and survival for breast cancer.

ii. The Nottingham Prognostic Index

The Nottingham Prognostic Index (NPI) (Haybittle *et al.,* 1982) combines grade, lymph node stage, and tumour size to produce three prognostic groups - good, moderate, and poor using the following formula

$$NPI = (0.2 \times \text{size in cm}) + \text{lymph node stage} + \text{grade}$$

Lymph node stage is based on pathological examination of three nodes. If all three nodes are disease-free the tumour is classified as stage **A.** If only the lowest (level I) node is involved, the tumour is stage B, and if the apical or internal mammary node is involved, stage *C.* Grading follows the modified Bloom and Richardson regimen (Elston *et al.,* 1991). The utility of this index in predicting outcome and for stratifying groups of patients has been confirmed (Todd *et al.,* 1987, Brown *et al.,* 1993).

1.3.4 Treatment factors

The aims of treatment are eradication of disease, achievement of local control, prevention of recurrence, and minimisation of side-effects. Treatment falls into three types - surgery, locoregional radiotherapy, and systemic cytotoxic or endocrine therapy.

i. Surgery and radiotherapy

surgical removal of the tumour often forms the basis of treatment for breast cancer, and many patients also receive radiotherapy to the chest wall or axilla. The modified radical (Patey) mastectomy involves removal of the breast and the axillary nodes, but the

pectoralis major and often the pectoralis minor muscles are preserved and internal lymph nodes are left *in situ*.

The physical and psychological impact of mastectomy can be severe (Sacks *et al.*, 1993), and surgical techniques which aim to preserve the breast have been developed. Conservative surgery encompasses a number of techniques. Lumpectomy is a technique in which the palpable tumour is excised and an attempt to obtain tumour-free margins is routinely made. More extensive techniques include wide excision and segmentectomy and involve the loss of more breast tissue in an attempt to clear the tumour margins. Radiotherapy is used to sterilise microscopic tumour deposits at the site of surgery and axilla, often after conservative surgery (Sacks *et al.*, 1993). Advanced cancers may be treated with radiotherapy alone for palliative reasons (Rodger *et al.*, 1994).

Early trials of conservative surgery plus radiotherapy found a survival disadvantage compared to mastectomy, but the radiotherapy doses used were inadequate by modern standards (Rubens, 1992). More recent trials have identified no survival difference between the two types of surgery, but found increased risks of recurrence after conservative surgery (EBCTCG, 1994, Fisher *et al.*, 1995, Riley *et al.*, 1995).

The cosmetic advantage of breast preservation can be offset by fears of recurrence, and psychiatric morbidity does not appear to be influenced by the type of surgery performed (Maguire, 1994). Risk of recurrence has been linked to incomplete tumour excision, the presence of an *in situ* component, vascular or lymphatic invasion, and high tumour grade (Dixon, 1995). Concern over recurrence rates following conservation surgery have been raised (Benson, 1996), with the suggestion that conservative techniques should be restricted to small, uni-focal tumours (<2cm), with complete clearance of the tumour margins. The perception that conservative surgery is better for patients was demonstrated by the quality standard set by the NHSBSP in 1992 that *>50%* of women with a tumour <=15mm should receive conservative surgery, but the need for complete tumour excision was stressed (NHSBSP, 1992). The need for surgical stringency has been addressed in guidelines from the BASO group (BASO, 1995) for surgeons treating symptomatic breast cancer. Use of reconstructive surgery has risen by more than 5-fold since 1992 (Watson *et al.*, 1995), although such techniques are not universally practised.

The recent BASO guidelines recommend that a minimum of four axillary nodes should be sampled at diagnosis (BASO, 1995), while the TNM staging system recommends that at least 6 axillary nodes are sampled (UICC, 1992). Various methods of management of the axilla are still in use, ranging from a "watch" policy" to dissection, clearance and/or radiotherapy. Axillary staging is still not always carried out. Axillary dissection may be performed to level I, II or III of the axilla. Evidence that axillary irradiation reduces recurrence rates in women undergoing conservation surgery has been found repeatedly (Bundred *et al.,* 1994). A recent meta-analysis comparing axillary clearance with radiotherapy identified no survival difference between groups, but fewer recurrences in those receiving radiotherapy (EBCTCG, 1995).

Side effects from radiotherapy include desquamatisation, nausea, and fatigue. Complications may result from intervention at the axilla, particularly if a level II or III clearance and irradiation are used. Nerve and muscle damage can be caused by surgery or irradiation, giving rise to conditions including brachial plexopathy (Spittle, 1995). It is anticipated that the peak incidence of radiation induced plexopathy has been reached due to changes in techniques in since 1970 (Maher, 1995). Lymphoedema of the arm is most severe when resulting from axillary recurrence, but can result from combined high-level dissection and radiotherapy. Around 30% of those receiving radiotherapy after a level II clearance experience lymphoedema, and radiation in addition to surgery should be avoided after level III dissection (Bundred *et al.,* 1994), and avoided or spread out over a long period where reconstruction is planned (Watson *et al.,* 1995).

ii. Systemic therapy

Systemic therapies have been developed in attempts to eliminate micro-metastatic deposits in patients with apparently localised disease, to reduce symptoms and improve survival of patients for whom surgery is unsuitable and to shrink large tumours pre-operatively in order to facilitate effective surgery. Systemic therapy is used as both a primary intervention, and as an adjuvant to locoregional surgery or radiotherapy. Adjuvant therapy preserves prognostic data at diagnosis, but the benefit to individual patients is hard to monitor. Primary tumour-shrinking treatment does allow direct assessment of progress and may reduce the impact of surgery, but involves loss of

prognostic—data. Commonly used systemic therapy can be divided into cytotoxic and endocrine (hormone) therapy. **A** number of trials have examined use of immune therapies, but none has identified a significant benefit from these drugs in terms of recurrence-free or overall survival (EBCTCG, 1992).

a. Cytotoxic chemotherapy

Many cytotoxic agents have been used for breast cancer, including alkylating agents and anthracyclines. Combinations such as Cyclophosphamide, Methatrexate, and Fluorouracil (CMF) are more effective than single agents at delaying recurrence, especially in poor-prognosis patients (EBCTCG, 1992). In this trial the proportional reduction in annual mortality for patients >= 50 years with stage **'2'** disease was 12% compared to patients not receiving CMF, and for patients ***<50*** years, 30%. Cytotoxic therapy may be useful for rapidly progressive cancer (Rodger ***et*** *al.,* 1994), but the benefits of cytotoxic drugs must be weighed against side effects such as nausea, alopecia and induction of menopause. In patients wishing to undergo reconstructive surgery cytotoxic chemotherapy should only be given after the wound has healed and only if there are no signs of infection (Watson ***et*** *al.,* 1995). Oophorectomy, which is only of benefit to younger women, produces similar survival benefit to polychemotherapy, can be achieved by surgery, irradiation or hormonal manipulation, but is also associated with unpleasant side-effects including induced menopause and osteoporosis (Richards, 1994).

b. Hormone therapy

The most commonly used hormonal agent for early disease is Tamoxifen, which although predominantly an oestrogen antagonistic, has some agonist properties. **A** reduction in mortality of 20% was found for middle aged women with stage **2** tumours treated with Tamoxifen (EBCTCG, 1992). **A** recent meta-analysis concluded that, although of more benefit in oestrogen receptor-rich tumours, a smaller benefit could also be achieved in receptor-poor tumours. Tamoxifen is considered to be of benefit for all ages (Richards, 1994), and may be used as primary treatment and an adjuvant in patients with early breast cancer. Combined polychemotherapy and Tamoxifen has been found to be more effective than either regimen alone in terms of relapse-free survival (EBCTCG, 1992), but a meta-analysis of quality-adjusted survival revealed no benefit for those receiving both therapies compared to Tamoxifen alone, a finding which authors attribute

to side-effects associated with chemotherapy (Gelber *et al.,* 1996). A number of other hormonal agents have been developed including aromatase inhibitors which are used in postmenopausal women with advanced disease (Zeneca Pharma, 1995).

iii. Treatment of *in situ* disease

There is currently no consensus of opinion on the optimal treatment for *in situ* cancer. One trial comparing lumpectomy vs lumpectomy plus radiotherapy identified a reduced risk of non-invasive or invasive ipsilateral cancer in those given radiotherapy, and authors suggest conservation surgery with radiotherapy may be more appropriate than mastectomy (Fisher *et al.,* 1993). Recent recommendations on the treatment of DCIS (Page *et al.,* 1995) include excision of tumours <4cm and mastectomy if >=4cm or widespread DCIS, postoperative radiotherapy if comedo-type DCIS, and possible use of Tamoxifen. The NHSBSP is currently undertaking a trial of patients with screen detected DCIS to identify the most appropriate treatment options (NHSBSP, 1989).

1.3.5 The mode and setting of treatment delivery

i. Consensus treatment guidelines

In 1983 a survey of 766 consultant surgeons in England revealed variations in clinical practice (Gazet *et al.,* 1985). Use of surgery varied; over 70% of patients underwent mastectomy, and over 10% of surgeons never performed conservative surgery. Use of radiotherapy and chemotherapy differed and although use of endocrine therapy was common, criteria used to select women for treatment varied. In 1986 the Kings Fund convened a panel of experts to set guidelines on treatment and the organisation of breast cancer services (Kings Fund 1986). A summary of the guidelines is shown in Table 1.2.

A survey in 1991 examined the impact of the King's Fund guidelines (Morris *et al.,* 1992). The authors found that use of conservative surgery had increased, that surgeons discussed treatment options with patients more often, and that access to counselling had improved since the original 1983 study. However, an investigation of patients treated in 1986 at two London teaching hospitals found that clinical practice still differed from the guidelines (McCarthy *et al.,* 1991). For example, only 43% of patients had axillary node sampling or clearance and few women under 50 years received cytotoxic therapy. A

recent study from the South East found that treatment varied widely (Choiullet *et al.*. 1994). Axillary surgery was carried out in only 46% of cases. The authors concluded that, five years after the King's Fund conference, some of the recommendations had not been absorbed into common practice. This situation is similar to that in the United States where, although rates of conservative surgery are rising, there are large differences in rates between areas (Farrow *et al.*, 1992), and in a study of the effect of hospital type on survival from breast cancer, use of recommended breast conserving surgery was greater amongst teaching rather than non-teaching hospitals (Lee-Feldstein *et al.*, 1994).

Table 1.2 Kings Fund Consensus Guidelines: treatment of primary breast cancer

Investigation of breast lumps and diagnostic procedures

- prompt referral for diagnosis by GP to be seen by consultant at first hospital visit
- pathological diagnostic testing should be carried out on suspicious lesions using open surgery *only* if fine needle aspiration cytology or trucut biopsy not possible
- frozen section biopsy followed by immediate mastectomy rarely justified

Treatment decisions

- age, menopausal status and tumour stage must be taken into consideration
- nodal involvement to be assessed clinically, and sampling performed at operation
- prognostic indices including grade and receptor status can be useful

Optimal surgical treatment for subgroups of patients

- no evidence that mastectomy is better than lumpectomy in terms of survival
- reconstructive surgery should be discussed with all women undergoing surgery
- mastectomy recommended for multi-focal or extensive disease
- axillary clearance recommended for extensive nodal involvement

Additional treatment

- use of medical therapies is appropriate for extensive disease
- risk of recurrence after local surgery is reduced with radiotherapy
- chemotherapy is most effective under 50 years but short courses only
- hormone therapy is most effective in over 50s, but reduces relapse rates in all cases

Organisation of breast cancer services

- services should be grouped in specialist clinics with multidisciplinary teams
- auditing of services at regional level is recommended
- mammography and histopathological diagnosis with cytology should be available

ii. The organisation of cancer care

The organisation of cancer care has been examined in relation to survival by a number of studies (Stiller, 1994). Finnish women living near hospitals with radiotherapy facilities enjoyed better survival after adjustment for age and stage compared to those living elsewhere (Karjalainen, 1990). Survival for women treated at UK teaching hospitals was

better than for those at non-teaching hospitals after adjustment for age, stage and type of operation (Basnett *et al,* 1992) and a study from the United States identified a survival advantage for patients treated at large community hospitals compared to smaller hospitals (Lee-Feldstein *et al.,* 1994). It has been hypothesised that this survival advantage results from the better treatment provided by specialists in multi-disciplinary teams which treat a large number of patients and have access to new or better treatment regimens. One study identified a reduction in **risk** of dying for those treated by specialists after adjustment for age, tumour size, socio-economic status, and nodal involvement (Gillis and Hole, 1996). No treatment data were available in this study, but the authors suggest that the observed survival advantage may reflect differences in treatment.

A recent study identified a survival benefit for those treated by clinicians with an annual workload of >30 new cases/year (Sainsbury *et al.,* 1995b), and caseload was found to have a significant influence on survival after adjustment for age, nodal status, grade, social class and treatment regimen. One Italian study, however, failed to identify any benefit in terms of mortality rates according to caseload (Bofetta *et al.,* 1993). It is thus questionable whether caseload alone is adequate as an indicator of multi-disciplinary care provision (Baum 1996, Sikora, 1996), although in one study surgeons with a high caseload were more likely have a "special interest" in breast cancer, and were more likely to use adjuvant treatment (Sainsbury *et* ***ai.,*** 1995a).

The validity of surgical caseload as an indicator of **high** quality multi-disciplinary care is currently a controversial issue. The Calman-Hine report (HMSO, 1995) set out a strategy for reorganising cancer care in the UK, focusing on the establishment of Cancer Centres and Cancer Units in which multi-disciplinary teams with site specialisation will operate. In response to the Calman-Hine report a Clinical Outcomes Group (COG) was set up to focus on the diagnosis and management of breast cancer, including recommendations on diagnostic facilities, treatment, follow-up and palliative care, and stressing the importance of the establishment of specialist breast teams (NHS Executive, 1996a). The COG Report recommended that all breast referrals should be to a specialist breast team which deals with at least 100 new cases per year. The British Breast Group (BBG), however, although advocating the management of breast cancer by specialist

teams, recommended a higher caseload of 150 cases/year as a realistic minimum for a specialist breast team, with a level of 50 cases/year being set for unusual "highlands and islands" areas (BBG, 1994), a view endorsed by BASO (BASO, 1996). On the basis of the limited evidence available, higher patient throughputs are believed by the COG to be associated with greater diagnostic accuracy, better quality surgical and non-surgical treatment, and better survival rates (NHS Executive, 1996a, 1996b).

1.4 SCREENING FOR BREAST CANCER

Concerns over high breast cancer mortality in the UK raised the possibility of introducing a population-based screening programme. A variety of screening techniques were studied in the 1960s and 1970s, and results suggested that detection of pre-symptomatic cancer could be achieved by mammography. To assess the feasibility of a programme the Department of Health appointed a committee to consider the available evidence. The committee, chaired by Sir Patrick Forrest, produced a report of findings (the "Forrest" report, HMSO, 1986), which was updated in 1991 (NHSBSP, 1991).

1.4.1 Evidence from studies of breast cancer screening

Randomised controlled trials assessed the impact of screening on breast cancer mortality by comparing mortality in those invited with that in uninvited populations. Data from the most important studies of screening by mammography are summarised in Table 1.3.

Table 1.3 Results of randomised controlled trials of breast cancer screening

Study	Start date	Method*	Interval (months)	Age (yrs)	Follow up (yrs)	RR** (95% CI) age 40-49	RR (95% CI) age 50-74	Reference***
HIP	1963	CM2	12	40-64	10	0.78 (0.52-1.18)	0.69 (0.49-0.97)	Shapiro *et al.*, 1982
Malmo	1976	M2	18-24	45-69	12	0.51 (0.22-1.19)	0.86 (0.64-1.16)	Andersson *et al.*, 1988
Swedish 2-Counties	1977	M1	24-33[1]	40-74	12	1.04 (0.68-1.59)	0.72 (0.59-0.88)	Tabar *et al.*, 1992
Stockholm	1981	M1	28	40-65	7	1.00 (0.49-2.04)	0.65 (0.40-1.08)	Nystrom *et al.*, 1993
Gothen-burg	1982	M1	18	40-59	5	0.72 (0.26-1.24)	0.91 (0.53-1.55)	Nystrom *et al.*, 1993
Edinburgh	1979	CM2[2]	24	45-64	10	0.78 (0.41-1.48)	0.84 (0.62-1.14)	Roberts *et al.*, 1990

* C clinical examination; M1 single view mammography; M2 two view mammography
** *Relative risk of death from breast cancer*
*** Main source: Wald *et al.*, 1994. References are for descriptions of individual trials
1 Round length 24 months for 40-49 year-olds, 33 months for 50-69 year olds
2 First round two view mammography, subsequent rounds single view mammography

For women aged 50-74 years, the overall reduction in mortality in those invited was **24%** (RR 0.76 95% CI 0.67 - 0.87). Assuming that the underlying risk of breast cancer death was equal in attenders and refusers, the mortality reduction in those screened was estimated at 31% (Wald ***et al.***, 1994). For women aged 40-49 years the mortality drop was 16% (RR 0.84 95% CI 0.67 - 1.06) and was not statistically significant.

Data from population-based studies were also considered. In a number of studies screening was offered to defined populations and the impact of screening assessed by comparing the odds of screened women dying from breast cancer with those of unscreened women. Data from these studies are summarised in Table 1.4. The reduced **risk** of death found in population-based studies suggest that screening can reduce mortality, but may be confounded by the "healthy screenee" effect, in which women least likely to die from breast cancer selectively attend for screening.

Table 1.4 Results of population based studies of breast cancer screening

Study	Start date	Method*	Interval (months)	Age (years)	Follow up (years)	RR** (95% CI) (all ages)	Reference***
Utrecht	1974	CM2	12-24	50-64	7	0.30 (0.13-0.70)	Collette *et al.*, 1984
Njimegen	1975	CM1	24	35-65	7	0.48 (0.23-1.00)	Verbeek *et al.*, 1985
Florence	1977	M2	24-60	40-70	8	0.53 (0.29-0.95)	Palli *et al.*, 1986
UK[1]	1979	CM1	24	45-64	10	0.80 (0.69-0.94)	EDBCTG, 1993

* C clinical breast examination; M1 single view mammography; M2 two view mammography
** Relative risk of death from breast cancer
*** Main source: Wald *et al.*, 1994. References are for descriptions of individual trials
1 Case control study based on geographical controls

1.4.2 The NHS Breast Screening Programme

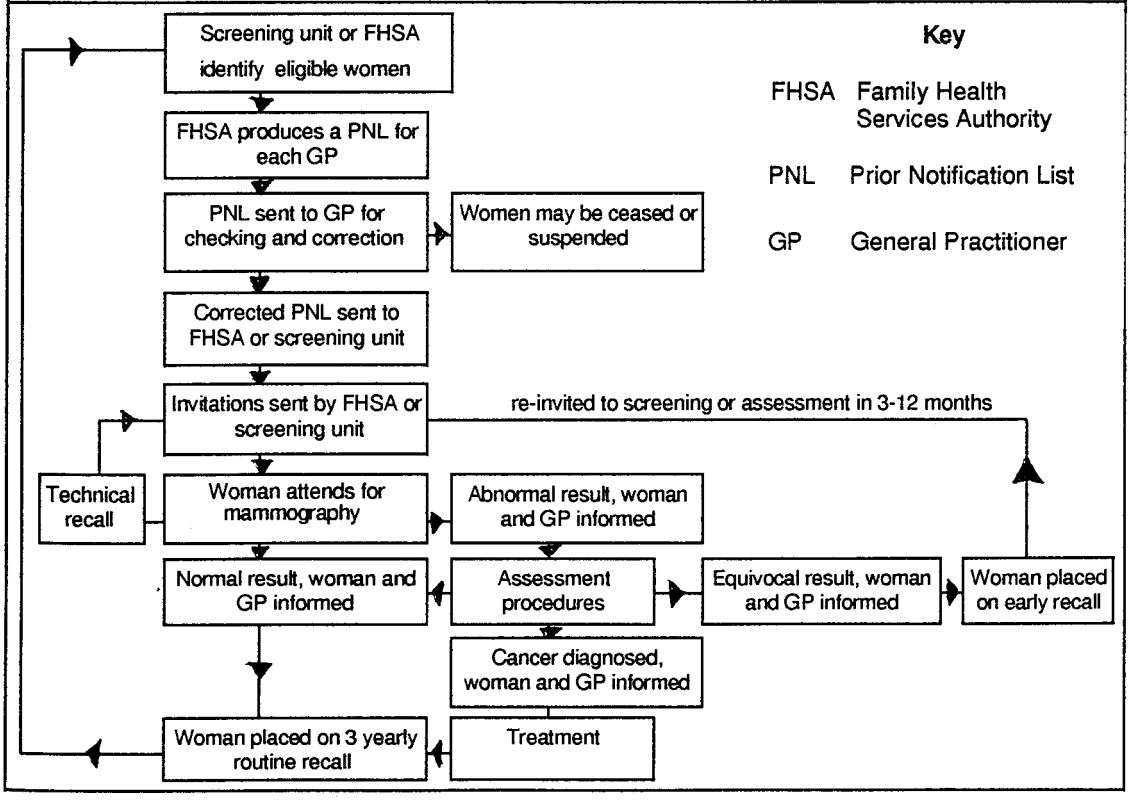
The Forrest Committee concluded that criteria set out by Wilson and Jungner in 1968 were met sufficiently by breast cancer screening by mammography and recommended that there should be a breast screening programme in the UK, and that

- the screening test should be single view, medio-lateral oblique mammography
- women aged 50-64 years should be routinely invited, women of 65 or more should be screened on demand, but those <50 years should not be eligible for screening
- the screening interval should be 3 years but kept under review
- specialist, multidisciplinary assessment teams should be used to carry out diagnosis

- a-screening record system should be developed to identify, invite and recall women, to record attendance and to monitor the screening process
- quality control arrangements should be made between and within centres so that an acceptable standard of mammography is maintained

The NHSBSP became operational in 1988, with services launched across the UK in a staggered introductory period lasting until 1992. To reflect the aims of the NHSBSP, mortality in invited women was chosen as a Health of the Nation (HoN) target of a 25% mortality reduction in women aged 50 - 69 years (HMSO, 1992, Department of Health (DOH), 1993), although monitoring changes in mortality in women aged 55 - 69 years is preferred by the NHSBSP (Quinn and Allen, 1995). The HoN target mortality rate has recently been lowered to account for a background fall in breast cancer mortality which began prior to 1988 (Quinn and Allen, 1995). The mortality rate in the baseline year 1990 was estimated at 95.1/100,000, giving rise to the original target of 71.3/100,000 (HMSO, 1992). The 1991 census revealed that the actual mortality rate in 1990 was lower at 93.4/100,000, and thus the current HoN target is 70.1/100,000 (DOH, in press). The operational procedures used by the NHSBSP are summarised in Figure 1.4.

Figure 1.4 Procedures used by the NHSBSP (adapted from Austoker, 1990)



i. Identifying eligible women and issuing invitations

The Forrest Committee concluded that women aged 50 to 64 years should be eligible for routine screening and that older women may receive screening on demand. Women under 50 years are not eligible, with evidence concerning the effectiveness of screening in this age group being contradictory (Margolese, 1996). An overview of five Swedish trials identified a non-significant 13% mortality reduction in women aged 40-49 years after 8 years (Tabar *et al.,* 1993). A meta-analysis of seven studies identified a non-significant mortality reduction of 10 - 15% in the screened groups for 40-49 year olds (Smart *et al.,* 1995), but analysis of false positive results from Stockholm concludes that the potential negative aspects of screening were substantial for younger women (Lidbrink *et al.,* 1996). In order to explore this further, the NHSBSP "age" trial began in 1991, recruiting 65,000 women aged 40 and 41 years for annual screening. An age-matched control group is being monitored to identify differences in breast cancer incidence between the groups. The trial will run for 15 years and women will be followed up to identify differences in mortality, Currently women aged from 65 years upwards may be screened on demand, but are not routinely invited. The charity Age Concern is, in conjunction with the NHSBSP, raising awareness of breast screening in this age group (NHSBSP and Age Concern, 1996). The possibility of extending the routine screening age up to 69 years is being explored at two pilot sites in the UK, which were set up in 1996 and are due to run for three years before any results are available.

The population registers used to identify eligible women are held by District Health Authorities (DHAs) which have incorporated the functions performed by the former Family Health Services Authorities (FHSAs). The small degree of inaccuracy of these lists (Bowling *et al.,* 1989; Bickler *et al.,* 1993) means that a small proportion of eligible women are not invited for screening. A basic "Forrest" unit screens around 12,000 woman a year, staggering invitations over the three year round. Women are invited in batches, usually according to GP practice. The process is initiated by a request from the screening unit to the FHSA for a Prior Notification List (PNL) providing details of eligible women. The PNL is sent to GPs for checking so that inaccuracies are corrected and women for whom breast screening is contraindicated can be temporarily or permanently excluded from the programme. The checked PNL is returned to the FHSA

or unit and used to produce invitations specifying the date, time and location of a screening appointment. Most units operate from a static site based at a hospital, and also have access to a mobile unit which visits a number of sites in the catchment area to reduce the travelling distances for women attending.

ii. Basic screening method

The Forrest Committee recommended that single view, medio-lateral oblique mammography should be used. However, almost half the 83 units operational in 1993 used two view mammography, in which a second cranio-caudal view is taken, throughout the prevalent round. Analysis of results from programmes using single view compared to those using two view mammography revealed a significantly higher cancer detection rate in those using two view screening (Moss *et al.,* 1995). A randomised trial of one versus two view screening confirmed that two view screening is more effective for identifying breast cancers (Wald *et al.,* 1995). All units have been required to offer two view mammography to women attending for their first screen from August 1995. Women returning for subsequent screens are offered single view rnammography. Independent double reading of screening mammograms has been performed by some units but is not possible at many sites due to lack of staff. The benefits of double reading are disputed, with one study concluding that double reading produces only a modest increase in cancer detection rates (Ciatto *et al.,* 1995), but another that it not only improves the cancer detection rate but is also less costly than a single reading policy (Brown *et al.,* 1996).

iii. The screening result

Results of the screening mammogram are normally sent to the woman and her GP within two weeks. Over 90% of women are informed that the screening mammogram was normal. A few women are asked to return for a repeat mammogram due to a technical fault. A quality standard of <3% of all films requiring technical recall was set in 1992 (Muir Gray, 1991). Less than 10% of women are recalled for assessment due to an abnormality on the screening film. Although not diagnostic of cancer, women are often distressed by the request for reattendance and clear information must be provided explaining the implications of recall (Austoker *et al.,* 1994; Ong *et al,* 1996).

iv. Assessment procedures

Assessment may involve a number of different investigations, usually from a specialist multi-disciplinary team which should include a radiologist, radiographer, pathologist, surgeon and breast care nurse. Repeat X-ray films are taken, and a physical examination performed for palpable lesions. Ultrasound may be employed if a cyst is suspected. In some units fine needle aspiration (FNA) is used to investigate possible cysts, and cytology of specimens obtained using FNA can be diagnostic. Suspicious lesions may have a fine-needle ("core", "trucut" or "wide bore needle") biopsy from which a histological diagnosis can be made. It is sometimes necessary to perform a surgical open biopsy, and X-ray guided localisation may be needed to locate impalpable lesions. In 1993 only 5.5% of women screened were referred to assessment, and 0.8% underwent open biopsy (NHSBSP 1995e), rates which are lower than the estimates of 10% and 1.5% respectively made before the introduction of screening (Austoker *et al.,* 1988). These low rates are largely due to use of non-invasive diagnostic techniques. Over 80% of women are informed that the assessment proved negative and are placed on routine recall. For some women the result of assessment will be a diagnosis of breast cancer. **A** small number of women have unequivocal results after assessment, and are placed on short-term "early" recall, being re-invited for screening after **3** to 12 months.

v. Treatment

Referral for treatment may be undertaken by the woman's GP, but usually the screening unit refers women to a specialist breast cancer treatment team. The NHSBSP has set quality standards to ensure that over-treatment is avoided and that conservative surgery is used when appropriate. **A** mixture of radiotherapy, chemotherapy and hormone therapy is often used in addition to surgery. All women should receive information on treatment options and where appropriate should be offered a choice of treatment.

vi. The call and recall system

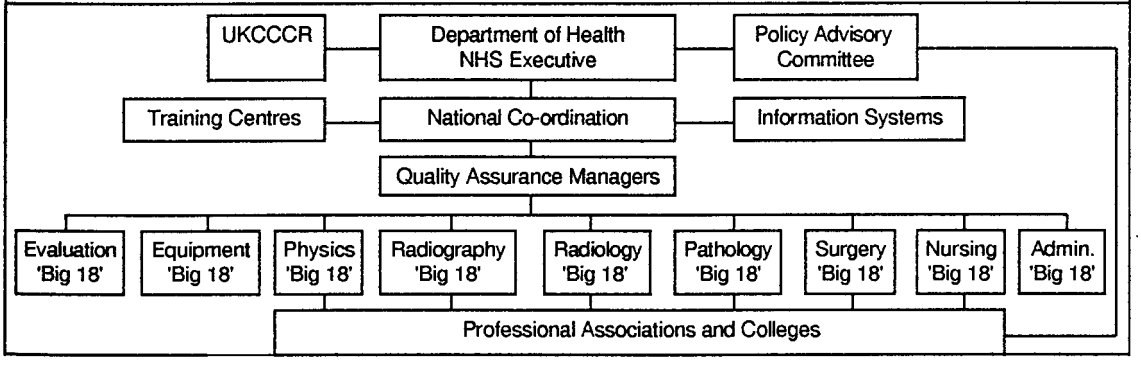
The Forrest Report recommended a **3** year screening round but specified that further research in this area was required. The NHSBSP is currently undertaking a trial to compare the effect of annual screening with routine 3 year screening. Eligible women are allocated a unique identifying number by the unit which is maintained throughout the

period of eligibility for routine screening. The call and recall system operates by identifying eligible women every 3 years based on age specifications set by the screening unit. The unique identifier allows previous screening history to be identified. It is essential that the invitation batch order is maintained so that women are re-invited on a 3-yearly basis. The need to adhere to a 3-yearly round is demonstrated by a recently suggested radiological quality standard of *>90%* of women to be re-invited within **2** months of the 3-year anniversary (Radiological Big '18',personal communication).

1.4.3 The Quality Assurance system

The Forrest Report stated that consistent high quality results will require a quality assurance (QA) programme. Following **NHS** reforms in 1996 QA services are purchased by lead purchasers representing Health Authorities in each of the eight regions in the UK. Specific aims of QA include provision of information about the NHSBSP, participation in evaluation and development, collection and review of data, organisation of multidisciplinary visits to units to assess performance, identification of problems, and monitoring remedial action. The QA system embraces surgery, radiology, pathology, medical physics, radiography, breast care nursing, health promotion, screening office administration and evaluation. The communication structure of the QA system outlined in Figure 1.5 is reproduced from QA Guidelines published in 1995 (NHSBSP, 1995c).

Figure 1.5 Communication structure of Quality Assurance in the NHSBSP



Local QA services are focused at quality assurance reference centres (QARCs) led by a QA manager responsible for co-ordinating the activities of a QA Team. The QA Team consists of representatives from each of the professional groups who represent the interests of that discipline in the QA Team, advise the QA Manager on matters relevant

to their profession and participate in the National Co-ordinating Groups ("Big 18s") shown in Figure 1.5.

Regional QA Team meetings provide a forum for discussion of QA issues and may be attended by invitation by unit directors, purchasers or other interested parties. Members of the QA Team also attend multidisciplinary visits to units at which discussions with representatives of each professional group allow assessment of quality, identification of problems and an opportunity to acknowledge good working practices.

1.4.4 Evaluation of the NHSBSP

1.4.4.1 Monitoring progress

Breast screening is a multi-stage, multi-disciplinary process and outcome measures which assess the performance of each process are required. The progress of the NHSBSP can be assessed at three levels:-

- level 1 - changes in mortality due to screening
- level 2 - calculation of values for programme sensitivity and specificity
- level 3 - calculation of surrogate outcome measures

Level I The aim of screening is to cause a fall in breast cancer mortality, but monitoring mortality alone does not provide a simple means of evaluating the success of screening. Breast cancer mortality has begun to fall in the UK (Quinn and Allen, 1995; Beral *et al.,* 1995), but the degree to which screening is responsible for this is disputed (Wright *et al.,* 1995), and there are other factors, including increased use of adjuvant hormone therapy, which are likely to have contributed to such a change. Identification of a significant mortality drop resulting from population based screening is complicated by the absence of a control population and confounded by an underlying fall in mortality (Quinn and Allen, 1995; Wright *et al.,* 1995). Evidence from randomised controlled trials suggests that any reduction in mortality due to screening is unlikely to be apparent for 7 - 10 years after the introduction of screening and quantifying.the contribution of the NHSBSP to any drop in mortality will involve complex adjustments to allow for confounding (Day *et al.,* 1989). Calculations such as these require aggregated data and thus should be undertaken at national and regional, rather than unit level (Knox, 1988).

Level 2 Calculations of sensitivity and specificity provide useful information but require large sets of standardised data (Greenstein ***et al.,*** 1995). Data are difficult to collate due to the staggered start of screening and variation in round length at different units. The calculation of sensitivity requires knowledge of interval cancer rates. It has been assumed that interval cancers arising symptomatically in the 12 months after a negative screen can be used to represent false negative test results (NHSBSP, 1993a). It **is,** however, possible that such cancers began to develop after the negative screen, and that some genuine "missed" cancers become symptomatic after 12 months has elapsed.

Level 3 The evaluation methods currently applied at national, regional and unit level involve the production of outcome measures which act as surrogate endpoints. **A** range of outcome measures and quality Standards has been developed by the **QA** system. Outcome measures produced at unit and regional level and provide information about the progress of the NHSBSP to those working in the programme, to purchasers and host trusts. Table 1.5 summarises the outcome measures used to assess whether the requirements of a successful programme are being met.

Table 1.5 Parameters used for assessing overall programme performance

Requirement	Outcome measures	Desired outcome
Sufficient women attend	Uptake rate	High
Sufficient cancers are identified*1	Cancer detection rate Standardised Detection Ratio	High High
Reduced rate of advanced disease	Small cancer detection rate	High
Few false positives *2	Referral rate Benign open biopsy rate Benign:malignant open biopsy ratio PPV of open biopsy	Low Low Low High
Few false negatives *1	Rate of interval cancers	Low

*1 measures related to programme sensitivity *2 measure related to programme specificity

For outcome measures to be useful they should be defined unambiguously and calculated consistently. High quality data must be available, and where data are available from multiple sources they must be consistent. Everyone using outcome measures must interpret them correctly, applying up to date and relevant standards, and making year on year comparisons to identify trends.

1.4.4.2 The cost effectiveness of the NHSBSP

The annual cost of running a screening programme was estimated in 1986 to be £18 million. The Forrest Report used evidence from a cost-utility analysis in which the cost per quality-adjusted life year was estimated at £3,400, and concluded that the benefits of screening justified the costs. More recently there has been disagreement over the cost-benefit balance of breast screening (Baum, 1995; Wright *et al.,* 1995).

In any screening programme there is trade off between risks and benefits. The benefits, including mortality reduction, depend on accurate detection of early cancers, identification of normal women and use of effective treatment. The risks of surgery can be minimised by avoiding unnecessary open biopsies for benign tumours, and by pre-operative diagnosis. These aims are reflected in quality standards set by the NHSBSP, and the monitoring undertaken by the **QA** service aims to ensure that unnecessary procedures are avoided.

Mammography requires compression of the breast, and can be painful, although this was not reflected in the high attendance achieved in screening trials (Hurley *et al.,* 1992). Possible causes of psychological morbidity include anxiety if recalled and increased length of time with knowledge of diagnosis (Hurley *et al.,* 1992). **A** degree of anxiety following positive screens is inevitable, but can be reduced if the information needs of women are respected (Austoker *et al.,* 1995).

Risks resulting from breast cancer screening include overdiagnosis and the detection of length bias cancers. To what extent screening detects "length bias" cancers which would never become life-threatening is difficult to ascertain. The selective detection of such cancers leads to over-treatment and unnecessary anxiety, and reduces the ability of screening to reduce mortality. During the prevalent screen it is expected that more cancers are detected than would arise symptomatically in this time, and that some of these will be large, slow growing tumours (Day *et al.,* 1989). This has been verified, but after subsequent screens the excess of cancers disappears suggesting that overdiagnosis is confined to the prevalent screen (Peeters *et al.,* 1989; Tabar *et al.,* 1992).

One issue yet to be resolved is the suitability of the three year screening round, which some believe to be too long, resulting in calls for a shorter two year round (Woodman *et al.*, 1995; Day *et al.*, 1995a, Asbury *et al.*, 1996). The risk of breast cancer rises with age and thus screening should be repeated at intervals. The Forrest Committee recommended a three year screening round based on evidence available at the time, but recognised the need for further research in this area. Symptomatic cancers presenting after a negative screen (interval cancers) may indicate low sensitivity, or may suggest that a shorter interval between screens may be required. Although the incidence of some interval cancers is inevitable, minimising their occurrence is vital to the success of screening. The current definition of an interval cancer used in the collation of national statistics is a histologically confirmed, invasive breast cancer diagnosed within 36 months of a normal screen. **As** interval cancer rates are used as a quality assurance standard for the NHSBSP, it is essential that interval cancers are uniformly identified and recorded and that radiological classification is accurate and complete. Current definitions do not account for a variety of factors which may influence interval cancer rates, and radiological classification is frequently unavailable for analytical purposes.

1.4.4.3 The cancer registration system and screening evaluation

In the **UK** a national network of cancer registries collect data on the incidence and characteristics of cancers occurring in the population, and submit a standard data set to ONS for collation of national cancer statistics. Although the current cancer registration system was established nation-wide in 1970, the data collected by different registries differed (UKACR, 1994). The WMCIU collects data from multiple sources, including pathology reports, and records details of treatment received. **In** 1993 a Minimum Data Set for Cancer Registration (MDSCR) was introduced (EL(92)95 Annex B) setting out requirements for the data to be collected by cancer registries and sent to ONS. In anticipation of the divestment of cancer registries **from** the old Regional Health Authorities to Lead Purchasers in 1996, a core contract for cancer registration (EL(96)7) was issued to ensure that the requirements of the MDSCR are maintained. Lack of TNM staging for breast cancers recorded at cancer registries has already been mentioned. The MDSCR covers many aspects of cancer registration, including timeliness and reliability, and recognises the importance of prognostic data by setting

standards for the proportion of certain cancers with staging data available. For breast cancer a minimum standard of 30% was set, with a desirable target of 80%. Whether these targets are met depends not only on efficient data supply to cancer registries, but on whether staging data are collected and stored in an assessable form within Trusts. Nonetheless, cancer registries in the UK are currently a unique source of routinely collected, population based data for evaluation of breast cancer.

Cancer registries have an essential role to play in evaluating the NHSBSP (Bull *et al.,* 1989). The NHSBSP produced guidelines on the utilisation of breast cancer data, stressing the need for communication and data exchange between the NHSBSP and cancer registries (NHSBSP, 1993b). Categories were provided by the NHSBSP into which breast cancers arising in women eligible for screening would fall - screen detected cancers, interval cancers, cancers in lapsed attenders, cancers in non-attenders and cancers in the uninvited. It is only possible to categorise breast cancers in this way with the co-operation of both the NHSBSP and cancer registries. The incidence and characteristics of cancers in these groups provides insight into the success of screening.

Examination of the prognostic characteristics of breast cancers in relation to their screening history can provide early indicators of success for the NHSBSP (Day *et al.,* 1989; Bull *et al.,* 1991). For screening to succeed the prognosis of screen detected cancers must be favourable, as a fall in mortality due to screening is likely to be preceded by a fall in the rate of advanced disease in screen detected cancers. Studies have shown that screen detected cancers are often small with negative axillary nodes (Duffy *et al.,* 1991; Tabar *et al.,* 1992; Moss *et al.,* 1994), and have more favourable biological features than symptomatic cancers (Klemi *et al.,* 1992; Crisp *et al.,* 1993; Hakama *et al.,* 1995; Rajakariar *et al.,* 1995). By combining data from the NHSBSP and cancer registries it is not only possible to examine the prognosis of screen detected cancers, but by comparing these cancers with those found outside the NHSBSP, it is possible to assess the likely impact of screening on mortality. However, missing and unreliable staging data can pose a serious limitation to this type of population-based assessment the NHSBSP.

For screening to succeed, the majority of breast cancers in the eligible population should be detected by screening. However, in a recent study of breast screening in the South East Thames area, only 1331 of the 4202 (48%) cancers in eligible women examined were screen detected (Garvican *et al.,* 1996). If the high proportion of cancers diagnosed outside the screening programme in the South East were to be combined with a higher degree of aggressiveness in these cancers compared to those found by screening, the impact of screening on breast cancer mortality will be smaller than expected. The incidence of breast cancers outside the NHSBSP in women eligible for screening provides important information about the programme. In particular, the high rates of interval cancers observed in the NHSBSP suggest low programme sensitivity, and/or that the screening round length is too long (Woodman *et al.,* 1995; Day *et al.,* 1995a, Asbury *et al.,* 1996). Other breast cancers diagnosed outside the programme could signal other significant problems. High incidence rates in lapsed attenders, who have previously attended but defaulted an appointment prior to diagnosis, would suggest that the screening experience was unacceptable, while high incidence in the uninvited would indicate inadequacy in population coverage.

1.5 AIMS OF THIS INVESTIGATION

There is much activity currently being directed towards improving the outcome for cancer patients in the UK, including women with breast cancer. The impact of the NHS Breast Screening Programme (NHSBSP), the reorganisation of cancer services following the Calman-Hine report, and improvements in treatment in response to guidelines issued by BASO and the NHSBSP must be evaluated. Cancer registries are currently the only source of the routinely collected population-based data that such assessments require. The NHSBSP routinely records data describing screen detected cancers and thus can act as additional data source. However, the validity of routinely collected breast cancer data for such assessments must be established. The first objective of this thesis was to audit prognostic data including stage recorded for breast cancers at the WMCIU and by the NHSBSP, in order to assess whether these data are of sufficiently **high** quality to be useful. Additional aims of the audit were to identify possible causes of any inadequacies in the data, and to suggest solutions. Tumour stage is of particular importance in evaluating breast cancer as it is recognised as a highly significant prognostic factor. The reported incompleteness of tumour stage recorded at cancer registries thus imposes a serious constraint on the utility of recorded data. In an attempt to resolve this difficulty the development of a staging algorithm as a possible means of allocating stages to breast cancers at the WMCIU was explored.

The West Midlands represents approximately one tenth of the UK population and thus provides a large study population for ecological studies. The recognition of patterns of disease within a population is essential if resources aimed at reducing mortality are to be effectively targeted. The second objective of the thesis was to produce an overview of temporal and geographical patterns of disease in the West Midlands, and to focus in detail on the tumour and patient characteristics of breast cancers in the region. The major prognostic factors for breast cancer have been established by a number of studies, including tumour stage, grade, and type of treatment. This study attempted to confirm the significance of established prognostic factors in the study group in order to identify those factors which, by manipulation, could bring about improvements in survival for breast cancer patients in the West Midlands.

Use of optimal treatment is crucial if mortality from breast cancer is to be reduced, and current guidelines from BASO and the NHSBSP **aim** to improve treatment. Monitoring the degree and speed with which such guidelines change clinical practice will be essential, but it is too soon to tell whether they have had any effect. The first guidelines on the management of breast cancer in the UK were issued by the King's Fund in 1986. Although now superseded, examination of treatment pre- and post-dating the King's Fund guidelines allows assessment of the ability of treatment guidelines to alter clinical practice. Evidence from elsewhere in the UK suggests that treatment for breast cancer improved following the availability of the King's Fund guidelines, although sub-optimal treatment was still common after 1986. ***An*** aim of this study was to examine the treatment recorded for breast cancers diagnosed in the West Midlands in 1983 and 1987, spanning the publication of the King's Fund guidelines, in order to examine treatment patterns across the region and identify any changes in clinical practice. The study also aimed to establish an audit process to facilitate the routine performance of future studies and provide baseline data against which any improvements in treatment following recent guidelines could be measured.

The multi-disciplinary Clinical Outcomes Group (COG) has recently issued a recommendation that breast cancer care should be delivered within high caseload surgical teams on the assumption that treatment from surgeons with a high caseload translates into better patient outcomes. However, there is only limited evidence for a positive relationship between high patient throughput, better treatment and improved survival. **A** further aim of this study was thus to assess whether any differences were apparent in recorded treatment according to surgical caseload, and to examine whether any survival differences according to surgical caseload could be identified.

The evaluation of the NHSBSP is hampered by the lack of a control population, and analysis of the survival for women with screen detected cancer compared to cancer detected symptomatically will always be confounded by lead time bias, It is essential that other means of predicting the impact of screening on the population are identified and routinely undertaken. Although screening may bring about a significant fall in mortality for women with screen detected cancer, unless these cancers make up a large proportion

of breast cancers in women eligible for screening, the ability of the NHSBSP to bring about a significant mortality drop in the population will be reduced. Despite the importance of this type of evaluation in assessing the likely impact of the NHSBSP, currently very little research in this area has been published. One objective of the investigation was to develop a means by which breast cancers diagnosed in women eligible for screening could be categorised according to their screening history. This allows the proportions of cancers detected by screening and outside the programme to be identified, and in addition, allows the prognostic characteristics of cancers according to screening history to be examined, providing vital insight into the likely impact of the NHSBSP in the region.

In recent years attention has been focused on the occurrence of interval breast cancers in the NHSBSP, which may indicate poor programme sensitivity and/or that the screening round is too long. Although important decisions about the future of the NHSBSP, such as a potential shortening of the round length, utilise interval breast cancer data, the definition and interpretation of these cancers remains unclear. In order to explore this issue an examination of these cancers was included in the study.

The NHSBSP represents a major commitment of NHS resources and it essential that the beneficial effects of screening are maximised. The NHSBSP Quality Assurance (QA) system was established to ensure a high quality of service for the programme, and has devised a number of outcome measures to act as performance indicators and markers of likely success for the programme. It is essential that these outcome measures provide a valid indication of the performance of the programme. A further aim of the thesis was to assess these outcome measures and their associated quality standards, using data from the West Midlands to identify ways in which the validity of these outcome measures could be improved.

CHAPTER TWO

DATA SOURCES AND METHODS

2.1 DATA SOURCES

Most of the data presented in this thesis were derived from records held at the WMCIU. Screening histories were obtained from the West Midlands NHSBSP in the form of data relating to individuals, and from standard statistical reports collated by the NHSBSP computer record system. Published data were obtained from the Office for National Statistics (ONS) formerly the Office of Population, Censuses and Surveys (OPCS), the Department of Health (DoH), and from other European and UK cancer registries.

2.1.1 Data from ONS, DoH and other cancer registries

This thesis is primarily concerned with female invasive breast cancer, defined by the International Classification of Diseases (9th revision) as 1740.0 - 174.9 **(WHO,** 1977). Baseline Health of the Nation data were obtained from the Public Health Common Data set ***(HMSO,*** 1993). Survival data from other UK cancer registries were obtained from published reports (Yorkshire Cancer Organisation, 1995; Scottish Cancer Intelligence Unit ,1993, West Midlands RHA and WMCIU, 1996; Centre for Cancer Epidemiology **,** Manchester, 1992; Cancer Epidemiology Unit, University of Bristol, 1995; Thames Cancer Registry, 1994). Mortality data for the West Midlands for 1981-92 were obtained from ONS and used to derive mortality rates. Some of the data presented in the overview of breast cancer in the West Midlands in Chapter Four were calculated for the 1995 Annual Report of the West Midlands Director of Public Health (WMRHA, 1996).

The incidence of female invasive breast cancer was derived from registrations at the WMCIU. The incidence of female ***in situ*** breast cancer, defined by ICD9 233.0, and of male breast cancers defined by ICD9 175, were also determined for comparison. Numbers of deaths from breast cancer in the West Midlands by age, year and DHA were obtained from ONS. Estimates of the resident mid-year female population in each West Midlands DHA by five year age group and year were obtained from the Information Department of the WMRHA.

2.1.2 Data recorded by the WMCIU

The WMCIU has been population based since 1957, covering a fixed geographical area defined by the former West Midlands **RHA,** with a current population of over 5.3 million. Data are derived from multiple sources including pathology reports, hospital discharge summaries, notifications from medical records departments, consultants and **GPs,** post mortem reports and death certificates. Upon receipt of details of a cancer, identifyingdetails are used to ascertain whether the cancer is already registered. If a previous registration exists, tumour details are compared to decide whether the new tumour is a recurrence of a previously registered primary tumour, or is a new multiple primary tumour. Details of recurrences are added to the primary registration. If a previous registration exists, or if the cancer is a multiple primary tumour, a new registration is created, and personal identifiers, the anniversary date and the site of the cancer recorded. Use of multiple sources enables the WMCIU to record prognostic information using rules which differ between cancer sites. Up until 1994 data were held on the PDPl 1 computer database. In 1994 a new system - the Generic Registration And Cancer Epidemiology database (GRACE) - was introduced.

Prognostic data recorded at the WMCIU for breast cancers include the TNM stage, tumour size, the status of the axilary nodes, the extent of disease, the presence or absence of metastases, tumour grade and histology. Patient related information including personal identifiers and date of birth are also recorded along with address details including the postcode which may be used to allocate indicators of material deprivation. Treatment data describing surgery, radiotherapy, hormone therapy and chemotherapy are supplied to the WMCIU by hospital medical record departments and radiotherapy clinics and are recorded with dates regardless of the time since diagnosis.

2.1.3 Calculation of incidence and mortality rates

Directly age standardised incidence and mortality rates were calculated using the Standard European Population as a reference, and 95% confidence intervals (CIs) were constructed using the methods described by Boyle and Parkin (Boyle and Parkin, 1991). Directly standardised rates for each DHA were compared to regional rates to produce Standardised Incidence Ratios **(SIRs)** and Standardised Mortality Ratios **(SMRs).** 95%

CIs were constructed, and DHAs with an SIR or **SMR** with a 95% CI not including 100 were considered to differ significantly from the region. Maps demonstrating variations in incidence, mortality and survival across DHAs and NHSBSP breast screening unit catchment areas of West Midlands were produced with the assistance of Geographical Information System Project staff at the WMCIU.

2.1.4 Data recorded by the NHSBSP

Details of screening episodes are entered onto the computer record system at each screening unit. For nine of the ten units in the West Midlands this record system is the Oxford National Breast Screening System **(NBSS).** One unit uses Kodak software, but the data items collected are identical to those on the **NBSS.** Diagnostic details including histology, tumour size, grade and nodal status are recorded, but stage is not recorded by NHSBSP. Details of treatment and follow-up are also recorded and used in the calculation of outcome measures. Screening units are required to submit reports detailing screening activity over each 12 month period to the DoH. These "Körner" data are collated onto "KC62" reports by software supplied to the units. In addition specialised reports are available which detail histological and surgical data. All data supplied by West Midlands units for this study were anonymised prior to presentation.

2.1.4.1 KC62 reports

Data for the period 1 April 1994 to 31 March 1995 were provided by all ten units using the KC62 version 2.8 report. The format of a KC62 Version 2.8 report is shown in Appendix Two. Version 2.8 operates on a cohort basis so that women in each table may be followed from screening to the final outcome, allowing data quality checks to be calculated by cross-reference within and between tables. The increasing use of diagnostic cytology was recognised in this version by detailing cytological procedures separately from diagnostic histology. Unfortunately all histology procedures are grouped so that it is impossible to separate open biopsies from non-surgical procedures, Unlike previous versions of the report, outcome measures are calculated by the KC62 report.

2.1.4.2 Histological QA (HQA) reports

The histology QA (HQA) report was developed by the NHSBSP to provide pathological data for screen-detected and interval breast cancers. The format of an HQA report is

shown in Appendix Two. Prognostic data such as invasive status, tumour histology, size, grade, nodal status and presence of vascular invasion are reported. The method used to identify women for inclusion in HQA reports differs from that used by KC62 reports version 2.8. Data for 1992-95 were supplied by the 9 West Midlands units which support the HQA software. The unit with Kodak software was unable to produce these additional reports. HQA reports provide data describing screen detected cancers and interval cancers but as many units in the West Midlands were not routinely recording the characteristics of interval cancers on their computer system at the time of investigation it was not possible to investigate interval cancers from HQA reports.

2.1.4.3 Surgical QA (QASS) reports

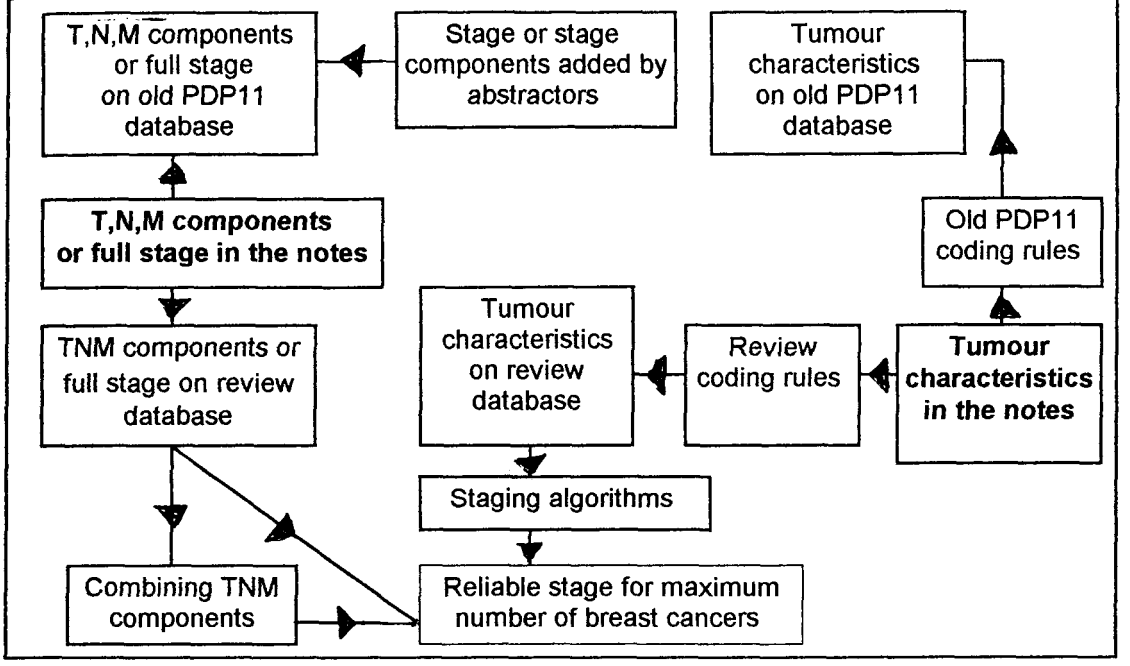
The surgical QA summary (QASS) report was developed by the NHSBSP to provide surgical data with which to assess the programme. The format of a QASS report is shown in Appendix Two. Details of diagnostic biopsies, treatment stratified according to choice of surgical treatment (conservative surgery or mastectomy) and size are provided. Surgical outcome measures (NHSBSP 1992, NHSBSP, 1996b) focus on the need to minimise the weight of benign biopsies, promotion of conservative surgery for small tumours, and minimising the use of open surgical biopsy. Data for 1992-95 were supplied by the 9 West Midlands units which support the QASS software. The unit with Kodak software was unable to produce these additional reports.

2.2 METHOD ONE: INVESTIGATION OF DATA AVAILABILITY AND VALIDITY AT THE WMCIU

2.2.1 An audit of staging and tumour characteristic data at the WMCIU

An audit of staging and tumour characteristics data at the WMCIU was undertaken to assess the quality of routinely recorded prognostic data. Data held on the old PDP11 database, and in case notes kept at the WMCIU, were reviewed in relation to TNM stage and stage components, tumour size, the status of axillary nodes, disease extent, and the presence or absence of metastases. Figure 2.1 describes the pathways of data transfer under investigation.

Figure 2.1 Pathways for recording stage and tumour characteristic data at the WMCIU



The 2511 female breast cancers registered at the WMCIU in 1983 were selected as the study population. The prognostic data under review had originally been abstracted onto the PDP11 database. The coding instructions available at the time of original abstraction onto the PDP11 were suggested to be insufficiently detailed and sometimes not adhered to. Using the available documentation, and after discussions with staff at the WMCIU, rules which applied to the use of codes on the PDP11 were compiled and are provided in Appendix Three. The suspected inconsistencies in interpretation of data and in the use of codes on the PDPI 1 were confirmed in these discussions, Some fields on the PDP11 were used to record several types of data. In addition, ambiguities in the data on the PDP11 resulted from non-adherance to or misunderstanding of the coding rules so that a single code was used to record more than one thing e.g., one code was used to record either the presence of negative axillary nodes or that nodal status was unknown. The introduction of a new WMCIU database, GRACE, in 1994 allowed problems in the structure of the PDPl 1 to be eliminated, and the abstraction rules used to record breast cancer data on the PDP11 to be reviewed. ***An*** independent database and coding system were developed for the review. The review database also held data imported from the PDP11, which were translated for analysis using the rules in Appendix Three.

Where TNM stage, or IT', Tu' or ᵀᴺᴹ components were recorded in the case notes they were transcribed onto the review database. Tumour characteristic data in the notes were recorded using rules described in Appendix Four.

The review database was designed with three fields for size, one for each of three sources - pathological, radiological, and clinical. ***An*** additional field was provided to record the size considered the most accurate. This "best" size was selected so that pathological size was selected preferentially, followed by radiological and then clinical sie. The "best" size should have matched that on the PDP11 if data transfer was complete and accurate at the original abstraction.

For the review, one field was provided to record nodal status and a second to record mobility. Pathological data were used in preference to clinical, and an additional code was included to identify cases for which the status of the nodes was not reported despite pathological sampling. The mobility of nodes was recorded as mobile, fixed or unknown. To allow comparison of data from the PDP11 and the review, clinical and pathological data held on the PDP11 were combined to give a PDP1i "best" nodal status, with pathological data taking precedence over clinical data. Nodes considered suspicious "at operation" but with no investigations reported were considered clinically positive.

The following codes were available on the review database to record disease extent

i. ***in situ***
ii. limited to the site of origin
iii. invading neighbouring structures
iii. probably limited to site of origin
v. possibly invading neighbouring structures
vi. insufficient information

Two new fields were created, one to record the presence of multiple lesions, and a field to record why a tumour was not treated including "found at ***post mortem***".

For. the review a single metastases field was provided and data recorded so that any positive site produced a positive code even if other sites were negative. The need for

comparison between clinical and "at operation" fields was eliminated by use of a hierarchical selection process, with "within site" variation defined differently to "between sites" variation.

within site — Pathological data were selected in preference to clinical data for a single site. For example, a clinical skin metastasis subsequently found on biopsy to be benign would be recorded as a pathoiogically negative result

between sites — If more than one site was mentioned, positive data took precedence For example, a cancer with clinical skin metastasis that proved to be pathologically negative, but with clinical bone metastasis without pathology would be recorded as clinicallypositive

This avoided misclassification resulting from a negative pathology at one site in the presence of clinically positive metastases at another site. For the review suspicion of metastases was also recorded, and was retained in the clinical metastases field on GRACE. This code was appropriate for cases where symptoms associated with metastases were mentioned, but no investigations were carried out, e.g. patients presenting in poor health with suspicious symptoms such as back ache, and allowed an assessment of how abstractors had interpreted suspicious symptoms. Similarly to local nodes data, "best" metastases data were identified from the PDPI 1. This process selected "at operation" data in preference to clinical data within site, and positive data in preference to negative data between sites. *An* important addition was that cancers without data referring to metastases were assumed to be free of metastases.

2.2.2 Quality assurance measures assessing data transfer

Quality assurance measures were calculated in relation to transfer of data from the notes to the PDP 11. The formulae used to calculate measures are provided in Appendix Five.

1. Completeness — Completeness of transfer describes the proportion of cancers with data in the notes for which data were present on the PDP11, regardless of the accuracy of transfer

ii.	Accuracy	Accuracy describes the proportion of cases with data in the notes and the PDP11 for which the data were consistent
iii.	Mismatched data	Calculated for tumour size only. Mismatched data describes the proportion of cases with a size in the notes which was inconsistent with the size on the PDP11
iv.	Added data	The proportion of added data indicates the proportion of cases for which data on the PDP11 did not originate in the notes

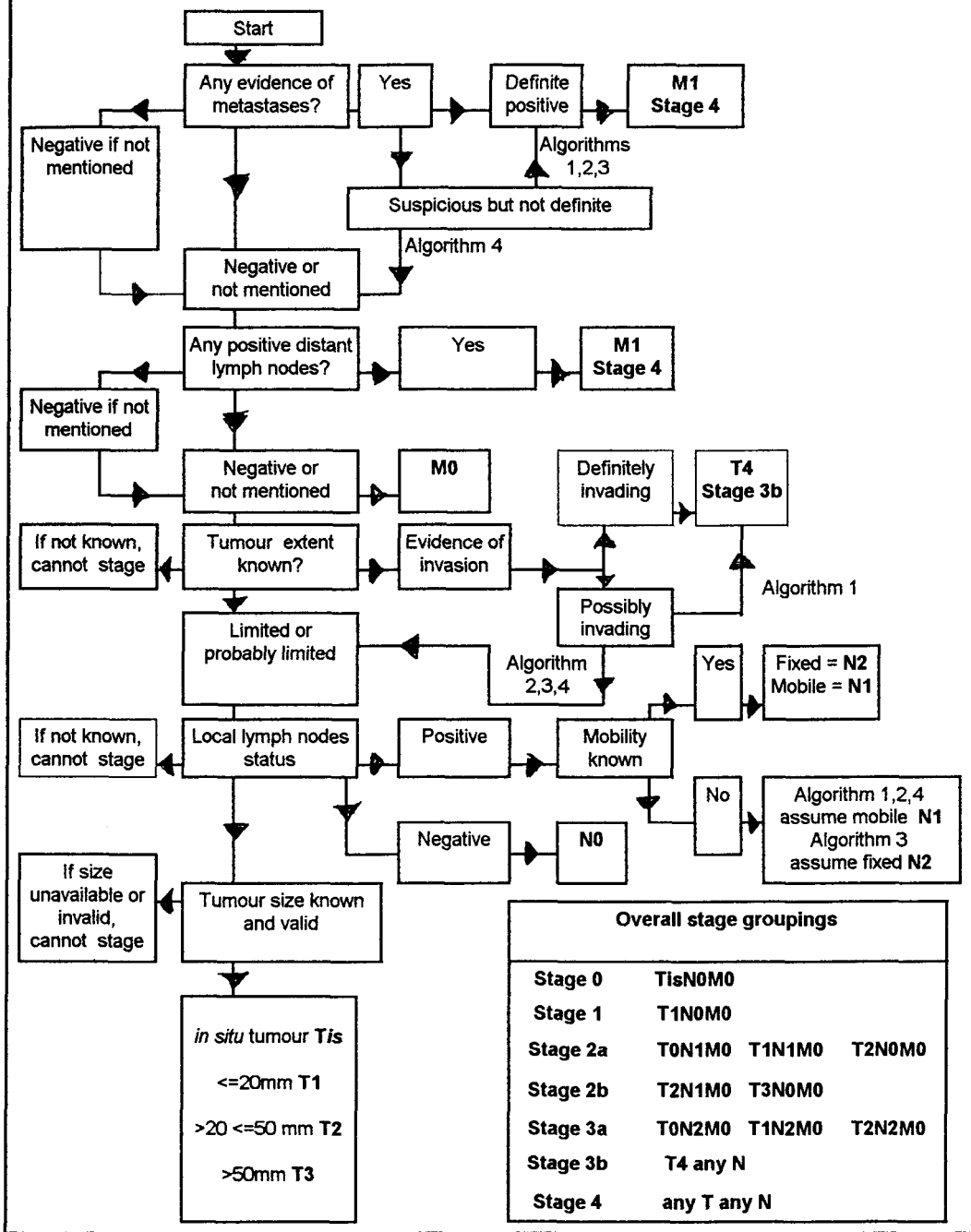
2.2.3 Design of a staging algorithm

A possible means of allocating stage at the WMCIU was to combine the separate tumour characteristics incorporated in the TNM system. The study group defined in section **2.2.1** for an audit of staging data was also used to assess the feasibility of a staging algorithm. The construction of the TNM staging system is described in Appendix One. Four staging algorithms, shown in Appendix ***Six***, were written for the package Epi **Info** (Dean et al., 1990) combining tumour characteristics using a protocol based on the TNM system. Figure ***2.2*** describes the steps used to allocate stage.

Each algorithm contained a different combination of assumptions about the data, and their validity was examined by analysis of survival in order to identify the most valid algorithm. No attempt was made to stage the **16** cancers identified at ***post mortem,*** and the 33 registered from death certificate only. ***6*** confirmed sarcomas were excluded. 413 tumours were clinically diagnosed as breast cancer but had no histology recorded. Only 6 of 2049 (0.3%) histologically confirmed cancers were non-carcinomas, and it was thus considered unlikely that many of these 413 cases were not carcinomas. All cancers without histological confirmation were assumed be carcinomas and included in the study. Tumours with ***in situ*** histology or Paget's disease without underlying carcinoma were allocated ***Tis.*** The data available were not sufficient to allow categorisation beyond the main TNM components, and no attempt to allocate sub-classifications of 'T' or 'N' components was made. The "best" tumour sizes identified during the review were used

to allocate hybrid clinical/pathologicalT/pT components. Although *6* 'N3' components were recorded in the notes, no data describing internai mammary lymph nodes, which indicate an 'N3', were found during the review, and N3 was not incorporated into any of the algorithms. Cases were allocated MO if evidence that metastases were absent was reported, or if no mention of metastases was made. These assumptions are similar to those used elsewhere (Chouillet *et al.*, 1994).

Figure 2.2 Flow diagram of information required to allocate stage to breast cancers



2.2.3.1 Assumptions used differently by the four staging algorithms

Limitations imposed by data completeness and quality dictated that three further assumptions be incorporated into the staging algorithms. Each algorithm incorporated a different combination of assumptions. The different combinations and the predicted influence on survival by stage of incorrect assumptions are shown in Table 2.1.

i. Disease extent

The review system provided two codes to indicate invasion beyond the site of origin. The first code indicated definite tumour extension. The second code indicated equivocal evidence of tumour extension. Two versions of the algorithm were created, one using both codes as criteria for a 'T4' (Algorithm 1), and the other only using the "definite" code (Algorithm **2).** Algorithms 1 and 2 were identical in all other respects.

ii. Mobility of local nodes

Of 909 cancers with positive local nodes, ***36*** had mobility specified. Algorithm **2** was designed so that local nodes of unknown mobility were assumed mobile, and assigned 'N1'. Algorithm **3** assumed that local nodes of unknown mobility were fixed, and assigned **'N2'.** Algorithms **2** and **3** were identical in all other respects.

iii. Clinical suspicion of distant metastases

46 cancers had evidence suspicious of metastases. In Algorithm 2 such evidence was assumed to be positive and assigned 'Ml'. Algorithm **4** assumed that such cases were not metastatic and assigned 'MO'. Algorithms 2 and **4** were identical in all other respects.

Table 2.1 Assumptions used in the four staging algorithms and the predicted effects of invalid assumptions

Algorithm	Disease extent	Nodal mobility	Metastatic status	Likely effect of assumption being invalid	Stages affected	Direction of effect of invalid assumptions
1	Both "definitely" and "possibly" invading beyond the site of origin codes included as a T4	Assumed mobile unless stated as fixed	Assumed positive if suspicious	Underestimation of cases with fixed nodes resulting in contamination of N1 category	2A	Survival worse than expected
					2B	Survival worse than expected
				Overestimation of cases invading beyond the site of origin (T4) by inclusion of "possibly" invading cases, resulting in contamination of T4 category	3A	Very few cases
					3B	Survival better than expected
				Inclusion on non-metastatic cases in M1	4	Survival better than expected
2	"Definitely" invading beyond the site of origin codes included as T4	Assumed mobile unless stated as fixed	Assumed positive if suspicious	Underestimation of cases with fixed nodes resulting in contamination of N1 category	1 - 3A	Survival worse than expected as T1 to T3 mask real T4 cases
				Underestimation of cases invading beyond the site of origin by exclusion of "possibly" invading cases, resulting in contamination of T1-T3		
				Inclusion on non-metastatic cases in M1	4	Survival better than expected
3	"Definitely" invading beyond the site of origin codes included as T4	Assumed fixed unless stated as mobile	Assumed positive if suspicious	Overestimation of cases with fixed nodes resulting in contamination of N2 category	2A	Survival better than expected
				Underestimation of cases invading beyond the site of origin by exclusion of "possibly" invading cases, resulting in contamination of T1-T3	2B	Survival better than expected
					3A	Survival better than expected
					3B	Fewer cases
				Inclusion on non-metastatic cases in M1	4	Survival better than expected
4	"Definitely" invading beyond the site of origin codes included as T4	Assumed mobile unless stated as fixed	Assumed negative unless stated positive	Underestimation of cases with fixed nodes resulting in contamination of N1 category	1-3B	Survival poorer than expected
				Underestimation of cases invading beyond the of origin by exclusion of "possibly" invading cases, resulting in contamination of T1-T3		
				Underestimation of metastatic cases in M1	4	Survival better than expected

2.2.3.2 Selection of the most valid staging algorithm

Al cases except those diagnosed at post mortem or registered from a death certificate only ***(DCO*** cases) were allocated a stage using each of the four staging algorithms. Cause specific survival was calculated using the Kaplan Meier function provided by **SPSS** for Windows **(SPSS** Release 6.0). Survival was calculated in months from diagnosis and cases were followed from diagnosis to the end of the study, defined as 3 1st December 1994, at which point those still alive were censored. Deaths fiom breast cancer were identified using the cause of death recorded on the death certificate. Deaths for which breast cancer was the primary cause (1 **153** deaths), and those with breast cancer present at the time of death but not the primary cause (146 deaths) were counted as events. Deaths from other causes were censored on the date of death.

Since stage has a natural order, tests were used to assess whether a significant trend in survival existed across stages. **Chi** square tests for equality of survival by stage were performed. The influence of each algorithm on survival by stage was assessed. These analyses examined whether, for example, stage '2a' from Algorithm 1 had a significantly better or worse survival experience than stage '2a' from Algorithm **2.** Rather than examining survival stratified by stage, an analysis was designed that examined survival for each stage stratified by algorithm. Survival analyses were performed based on the instances of each stage according to the algorithm by which it was produced. The following criteria were used to identifythe algorithm that was most valid and useful

- stages must fall into the correct order according to survival
- stages must display a statistically significant trend in survival
- sufficient cases should be staged for the algorithm to be of practical use

The selection process was based on pair-wise comparisons summarised in Table **2.2.**

Table 2.2 Pairs of algorithms compared to identify the most valid staging algorithm

Assumption under assessment	Algorithm supporting assumption	Algorithm not supporting assumption
Disease extent "possibly invading" can be interpreted as "definitely invading"	Algorithm 1	Algorithm 2
Positive local nodes of unknown mobility can be assumed to be fixed	Algorithm 3	Algorithm 2
Suspected metastatic sites can be assumed to be involved	Algorithm 2	Algorithm 4

2.3 METHOD TWO: AN EXAMINATION OF BREAST CANCERS DIAGNOSED IN 1983 AND 1987

Female breast cancers registered at the WMCIU in West Midlands residents which were diagnosed in 1983 and 1987 were used as the study populations in investigations of prognostic characteristics, treatment and survival. The 2.7% of cases diagnosed in 1983 and 8.4% of cases diagnosed in 1987 for which the case notes could not be located at the WMCIU, for which no age or date of birth was available, or for which no **DHA** of residence or postcode was available were excluded.

2.3.1 Definition of patient and tumour characteristics

Breast cancers in the study populations were characterised in terms of age at diagnosis, DHA of residence, level of material deprivation, stage at diagnosis, tumour grade, tumour histology, treatment type and the annual caseload of the treating surgeon.

Age at diagnosis was grouped into three bands **<50** years, 50 - 69 years, and 70 years or older. The **<50** year age group was assumed to be made up of mainly pre-menopausal women, whereas the oldest group was assumed to be mainly post-menopausal. The middle group was assumed to be a mixture of pre-, peri- and post-menopausal women.

The District Health Authority (DHA) of residence was used to partition the data into geographical areas within the West Midlands. The boundaries of the 15 DHAs in existence in 1994-95, the period in which much of this work was undertaken, were used, although mergers have since taken place. Similarly the **14** Regional Health Authorities ***(RHAs)*** in existence in 1995 have been used to describe inter-regional differences although these were reorganised in 1996.

Townsend scores (Townsend ***et al.,*** 1982) were used as an indicator of material deprivation. Townsend scores were derived from four data items collected during the 1991 census; access to a car, employment status, degree of overcrowding, and type of housing tenure. Each enumeration district was allocated a score based to these variables according to the protocol devised by Townsend (Townsend ***et al.,*** 1982) to provide a quantitative measure of material deprivation. In this study level of deprivation for cases in the study group was estimated from that of the enumeration district of residence,

which were in turn identified using the postcode of residence. Townsend scores allocated to cases diagnosed in 1983 were banded into quintiles to provide five deprivation levels. The cut-off scores identified in this process were used to allocate deprivation levels to cases diagnosed in 1987.

The staging algorithm selected by the method described in section 2.2.3.2 was used to assign stage at diagnosis. For some purposes stages were grouped into three levels - "early" (stages '0', '1' and '2a'), "locally advanced" (stages '2b' and **'3',** and "metastatic" (stage **'4').** Tumour histology was assigned according using the International Classification of Diseases Morphology codes (WHO, 1977) recorded at the WMCIU. The histology groups allocated for the purposes of the investigation are shown in Table **2.3.** Special types of invasive carcinoma were defined as medullary, tubular, mucoid, cribriform, papillaryand adenoid cystic carcinomas.

Table 2.3 Definition of histological groups

Histological group	Definition
Ductal carcinoma, no special type (NST)	Infiltrating duct carcinoma excluding special types
In situ carcinoma	In situ carcinoma including DCIS and LCIS
Lobular carcinoma	Invasive lobular carcinoma
Sarcoma	Breast sarcomas (mostly cystosarcoma phyllodes)
Special types	Breast carcinoma of "special type"
Histology not recorded	No pathology report available

Tumour grade was obtained direct fiom pathology reports during the audit of data for cases diagnosed in 1983. For cases diagnosed in 1987, grade was obtained fiom data recorded on GRACE. Some of the data on GRACE had been imported directly from the PDP11, and the method previously used to store grade on the PDP11 necessitated grouping grade into two groups - "good" (grade I or II), and "poor" (grade III, and tumours described as anaplastic or undifferentiated).

2.3.2 Treatment data

2.3.2.1 Definition of treatment types

The treatment data included in analyses were restricted to therapy given within **six** months of diagnosis. The term "conservative surgery" was used to describe all variants

including lumpectomy and wide local excision. "Mastectomy" is used to describe all variants of mastectomy. The term non-surgical oncology encompasses radiotherapy, hormone therapy and cytotoxic chemotherapy. Non-surgical therapy data simply described whether or not treatment was recorded. The treatment recorded for each patient was described using two treatment variables described in Table 2.4.

Table 2.4 Criteria used to assign treatment variables

Treatment variable	Options
Surgical	No surgical or non-surgical treatment recorded
	No surgical treatment recorded
	Conservative surgery with no mastectomy
	Mastectomy with or without previous conservative surgery
Non-surgical	No non-surgical treatment recorded
	Radiotherapy only
	Hormone therapy only
	Both radiotherapy and hormone therapy
	Cytotoxic chemotherapy +/- other non-surgical treatment

The data did not allow for determination of axillary node clearance, whether oophorectomy had been performed or the agents and dosages used in chemotherapy.

2.3.2.2 Identification of surgical caseload

A method similar to that used by Sainsbury (Sainsbury ***et al.,*** 1995b) was used to identify surgical caseload. Data for two three-year periods were examined. Caseload of relevant surgeons was estimated from the number of new breast cancers treated from 1982 to 1984. The number treated within 6 months of diagnosis by each surgeon was counted, and an annual rate calculated. The process was repeated for cases diagnosed from 1986 and 1988 to calculate the caseload for surgeons treating new cases in 1987. Patients having surgery were classified according to the caseload of the surgeon, and if both conservative surgery and mastectomy were performed, on the basis of the surgeon performing mastectomy.

2.3.2.3 Comparison of treatment with the King's Fund guidelines

The Kings Fund guidelines provided guidance on use of diagnostic and staging techniques, provision of advice and counselling, and discussing treatment choices with

patients. Changes in treatment patterns from 1983 to 1987 were investigated in order to assess whether the following guidance was acted upon

- age, menopausal status and stage should be considered when allocating treatment
- mastectomy and conservative surgery are of equal survival benefit, but mastectomy should be used for extensive disease
- patients with extensive disease should be offered adjuvant therapy
- conservative surgery should be followed by radiotherapy to reduce the risk of recurrence
- cytotoxic chemotherapy is most effective for women aged *<50* years
- hormone therapy is most effective in over women over 50 years, but may be of benefit in reducing recurrence rates in all ages

A further recommendation was that care should be delivered by multi-disciplinary specialist teams. Whether or not cases were treated by a multi-disciplinary team could not be ascertained from data held at the WMCIU, but the surgical caseload was examined to identify changes which could be related to multi-disciplinary working, and also to allow examination of whether any survival benefit with increased caseload was apparent as has been found elsewhere (Sainsbury *et al.*, 1995b; Gillis and Hole., 1996).

2.3.3 Survival analysis

Breast cancers diagnosed at post mortem or registered from a death certificate only were excluded from analyses of survival.

2.3.3.1 Univariate analysis of survival

Software designed by Hakulinen and Abeywickrama (Hakulinen *et al.*, 1985) was used to calculate five year relative survival rates using the 1981 England and Wales life table as the comparison group. Cases were stratified according to year of diagnosis, and by age group, DHA, level of deprivation, stage, grade, histological group, treatment type and surgical caseload. All causes of death were included in analyses, and live cases were censored on 31st December 1994. Ascertainment of deaths was assumed to be complete at the censor date, and no cases were withdrawn from the data set. Statistical testing of variation in survival was performed using the log rank test supplied in the software.

2.3.3.2 Multivariate analysis of survival

me influence of multiple factors on survival was assessed using Cox proportional hazards modelling (Cox, 1972). Tests of statistical significance were set at the 5% level. SPSS software (SPSS release 6.0) was used to examine the 1041 invasive breast cancers in the combined study populations for which stage, grade and Townsend score were available. All factors were entered as categorical covariates. Plots of log[-log survival] for cases stratified according to each covariate were constructed and did not suggest that the assumption of proportionality was contravened. The categories used for all variables except treatment were the same as for univariate analysis. In addition, year of diagnosis was entered as a covariate.

To simplify analyses treatment type was described in one covariate combining surgical and non-surgical therapy. Conservative surgery and mastectomy were grouped as "surgery". Those with non-surgical treatment only or without any treatment were combined into a category representing those with least intervention. Surgical cases were grouped into those with surgery alone, and those that had recorded adjuvant therapy. Cases with adjuvant therapies were further classified into four groups - those with hormone therapy only, with radiotherapy only, with radiotherapy and hormone therapy, or with chemotherapy either alone or in combination with other adjuvant therapies.

Models were created for all cases, and for each stage separately, using two methods. Models in which each covariate was entered singly were produced to assess the significance of each covariate. To assess the impact of adjusting for stage, this process was repeated with a baseline model including stage. Overall models were produced using forward stepwise conditional selection. Selection for entry depended on the global Chi square, with the significance level for entry set at <0.05. Covariates in the model with a probability of the likelihood ratio statistic of >= 0.1 were rejected. Relative risks for each covariate were described by comparison with a reference category.

2.4 METHOD THREE: DERIVATION AND INTERPRETATION OF SCREENING HISTORIES FOR BREAST CANCER

2.4.1 Identification of screening histories

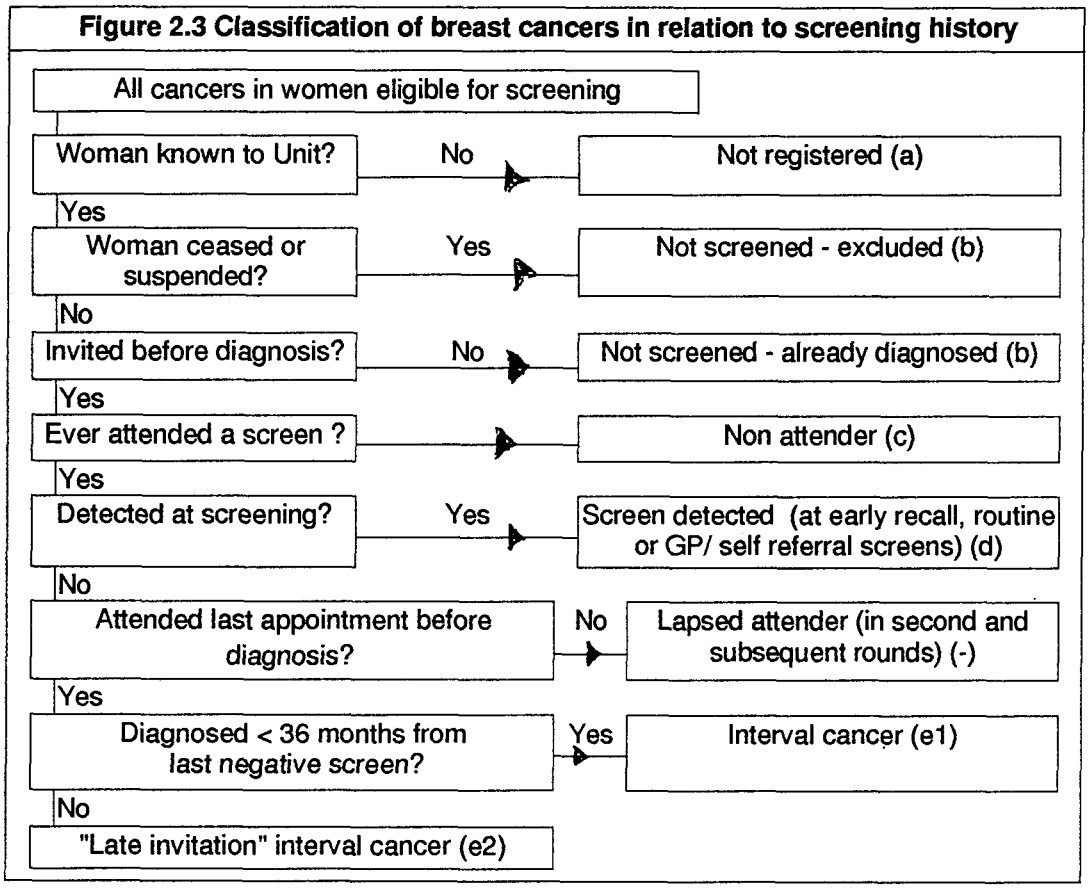
A test population was identified as all female breast cancers registered at the WMCIU resident in the catchment area of the Warwickshire, Solihull and Coventry breast screening unit. The catchment area of this Screening Service includes about 20% of the female population in the age group eligible for screening, and was thus considered to constitute a sample large enough to be reasonably representative of the region. Screening activity in the prevalent screening round (17th April 1989 to 31st March 1992) was considered. Over this period more than 80,000 women were invited, and over 62,000 attended for screening.

A retrospective analysis was devised to identify and categorise breast cancers which might have been detected by screening in the prevalent round. It was necessary to define a time limit, starting at the time of screening, beyond which a cancer arising symptomatically could not have been detected at the original screen. This time depends on the average sojourn time of breast cancer, i.e., the length of time that a cancer is detectable pre-clinically. The mean sojourn time for breast cancers in women aged 50-69 years was estimated from the Swedish Two Counties Study to be 4.12 years (Paci *et al.*, 1991). This means that, on average, a breast cancer is detectable pre-clinically for 4.12 years, and that a cancer detected symptomatically could have been found by screening for the previous 4.12 years. In this investigation the limit was set at five years in accordance with the NHSBSP statistical KC62 version 2.8 report in which women who have been screened before are divided into those screened less than, or more than five years previously. Cancers arising more than 5 years from the last screen were excluded.

Woman were designated as eligible for having their cancer detected by screening if the cancer was diagnosed when aged 49-68 years, and if invited, diagnosis was within five years of the last invitation. Screening histories were located using surname, forename, date of birth and postcode to identify registration and screening details at the unit.

Cancers diagnosed before the unit was operational, in women registered at neighbouring units, or that were detected following a screen or invitation in the second or third round were excluded. Cancers were classified using the scheme in Figure 2.3. A cancer in a lapsed attender cannot arise until a woman has attended and been invited again, and thus could not be identified immediately after the prevalent round. The term "lapsed attender" **is** also used by the **NHSBSP** to describe cancers arising symptomatically at >=36 months from the last negative screen in women no longer eligible for a routine invitation. In this investigation such cancers were termed "late invitation" interval cancers and comprised those cancers arising at >=36 months in women no longer eligible for a routine invitation, and cancers arising in women re-invited at >=36 months from the last screen but in whom the cancer arose prior to re-invitation.

Figure 2.3 Classification of breast cancers in relation to screening history



a. Cancers in women not registered at the unit

Cancers diagnosed from 17th April **1989** to 31st March **1992** in women aged 50-64 years, resident in the catchment area, of whom no details were identified at the unit or other units in the West Midlands.

b. Cancers in those not screened

Cancers in women registered at the Unit, who did not attend an invitation between 17th April 1989 and 31st March 1992 because they had been excluded by their GP, or because the invitation post-dated a diagnosis of breast cancer between 17th April 1989 and 31st March 1992. In both cases diagnosis occurred in women aged 50-64 years.

c. Cancers in non attenders

Cancers in women registered at the unit, that arose after a non-attendance, in women that received their last invitation between 17th April 1989 and 31st March 1992 aged **50-64** years. This category was sub-divided into cancers diagnosed at <36 months and >=36 months from the last invitation. Cancers diagnosed at >60 months from the last invitation were excluded.

d. Screen detected cancers

Cancers detected at screening between 17th April 1989 and 31st March 1992 in women aged 50-64 years. This category was sub-divided into cancers detected at routine screens and those detected at screens resulting from self or GP referral or early recall within the first round.

e. Interval and "late invitation" cancers

Cancers arising after a negative screen, in women attending their last negative screen between 17th April 1989 and 31st March 1992 aged 50-64 years who had either not received their next appointment or had not had the opportunity to attend the appointment. This category was sub-divided into cancers diagnosed <36 months (e1 intervals cancers) and >=36 months (e2 "late invitation" interval cancers) from the last negative screen. Cancers diagnosed at >60 months from the last screen were excluded.

2.4.2 Examination of the definition and classification of interval cancers

Interval and "late invitation" interval cancers identified by the method described were selected for further examination. The interval time was calculated as the time from the last negative screening mammogram to the histological confirmation of the cancer. Histological classification is vital as breast lumps may be lymphomas or metastatic disease. Paget's disease without underlying invasive disease was classified as *in situ*. Cancers were classified as "core" or "non-core" according to the criteria in Table 2.5.

The criteria include cancers arising in women at higher than normal risk and those presenting symptomatically in women on early recall but do not include cancers detected at early recall. Cancers diagnosed more than 36 months from the last screen ("late invitation" intervals) were excluded.

Table 2.5 Criteria defining "core" interval cancers and cancers in "high risk" women
Core intervals
A cancer that has been histologically confirmed as breast cancer
A cancer that is of invasive histology
A cancer in a woman whose last screen was between the ages 50 and 64
A cancer that occurred less that 36 months from the last negative screen date
Cancers in high risk women
A symptomatic cancer arising in a woman placed on early recall after her last screen
A cancer arising in a woman last screened as a self or GP referral
A cancer arising in a woman with a history of breast cancer

Radiological classification of the 213 "core" intervals was undertaken by Consultant Radiologists working in the West Midlands programme. Screening films were loaded onto a multi-viewer with negative films in a ratio of interval : normal of 1:2. The viewer was read by the visiting radiologists and by home radiologists. Screening films were identified as "positive" if two radiologists would have recalled the case, and if the area of suspicion corresponded to the position of the cancer, based on diagnostic mammograms. Classifications were allocated using the rules in Table **2.6.**

Table 2.6 Radiological classification of interval cancers			
Classification	**Diagnostic mammogram available?**	**Result of diagnostic mammogram**	**Review of screening mammogram**
True interval	Yes	Positive	Negative
False negative	Yes	Positive	Positive
Occult	Yes	Negative	Negative
False negative	No	No films available	Positive
Not false negative	No	No films available	Negative
Unclassifiable	No	No films available	Insufficient data

Based on the Northern region scheme (Simpson *et al.*, 1995) cancers without mammograms from the time of diagnosis were classified as "false negative" if the screening films were identified for recall when reviewed, and "not false negative" if the

screening films were passed as negative. A small group of "unclassifiables" still remained for which insufficient surgical or pathological data were available. A method similar to that used by Peeters (Peeters *at al*, 1989) was employed to estimate the effect of screening at 2 years. It was assumed that false negative cancers presenting in the third year would be detectable at 2 years and that occult cancers could not be screen detected at any point. The "not false negative" group were assumed to be made up proportionately of true and occult cancers. No assumptions were made concerning unclassifiable cancers, which were retained in the denominator of calculations of the proportion detectable. The rate of true intervals becoming detectable in each 12 month period after a negative screen was assumed to remain constant and was estimated as the rate arising in the first 12 months. The true intervals becoming detectable in the third year after screening was assumed to be made up of some cancers which had been detectable at less than 24 months (which could have been found if screening had occurred at two years) and some that only became detectable after 24 months. The annual rate of true intervals was used to estimate the proportion of the true intervals detected in the third year that were detectable at two years. The number of "detectable" cancers arising in the third year (the sum of the false negative and detectable true intervals) was used to estimate the proportion detectable by screening at two years.

Routine screening activity data for 1989 - 1995 were obtained from KC62 reports and used to calculate Standardised Detection Ratios (SDRs) for 50 - 64 year olds. This age standardisation technique calculates the ratio of the observed invasive cancers to the number expected based on the results of the Swedish Two-County Study, an SDR of 1 indicating parity with the Swedish study (Blanks *et al.,* 1996a).

Diagnostic and screening details of 155 women with a cancer detected at their incident second round screen, who also attended in the first round, were used to calculate individual round lengths for women. All 155 women had been found to be screen negative at their first screen, placed on routine recall, and arrived at their second screen at which cancer was detected as a result of routine invitation. The proportion of women re-invited at 36 months or more from the last screen was assessed.

The expected underlying incidence of invasive breast cancer in 1992 in the population covered by the screening unit was estimated for 50-64 year olds using cancer registration data for 1981 to 1987. The pre-screening trend in incidence rates in this age group was extrapolated by linear regression to calculate the expected rate in 1992.

2.4.3 Assessment of prognostic data in relation to screening history

Prognostic data with which to assess the impact of screening were investigated using five groups of cancers identified by the method described in section 2.4.1. Histology, tumour size, nodal status, and grade were examined. Data were derived from pathology reports held at the WMCIU if available, but if absent, data recorded by the unit for screen detected cancers and some of the interval cancers were used.

The characteristics of prevalent round, screen detected breast cancers, interval breast cancers, cancers in non-attenders, and second round incident screen detected cancers were examined. Breast cancers registered at the WMCIU which were diagnosed in 1987 in women aged 50 - 64 years provided a pre-screening symptomatic control group. Prevalent screen detected cancers were included only if found at routine screens. Cancers detected at early recall, or after GP or self referral were excluded as there were too few in these groups to permit detailed analysis. Cancers in non-attenders were included if diagnosis occurred within 36 months of the last invitation. Interval cancers identified as "core" intervals by the method described in section 2.4.2 were examined. In addition, 156 incident (re-screen) screen detected cancers from the second screening round were examined.

Pathological tumour size and nodal status were used in preference to clinical data which were used in the small number of cases for which pathological data were unavailable. Tumour sizes were excluded if the excision margins were not clear, if the tumour was not excised in one piece, or if pre-operative therapy was recorded. Nodal status was classified as negative or positive, and a third category indicated the presence of distant metastases including distant nodes. The proportion of *in situ* breast cancers, lobular carcinomas and special types in each post-screening group was compared with that in the control group. To examine the significance of the differences between proportions Chi square tests with Yates correction were performed. The distribution of cases by size,

nodal andmetastasis status in each group was compared with that of controls, and the significance of differences in distributions was tested using the Mantel-Haenzel Chi square test for linear association.

2.5 METHOD FOUR: ASSESSMENT OF THE OUTCOME MEASURES USED BY THE NHSBSP

The NHSBSP has devised a number of outcome measures with which to assess the performance of the Programme. Table 2.7 describes some outcome measures calculated from KC62 reports. Quality standards have been set in relation to these outcomes. As the NHSBSP has developed new measures have been introduced, some measures have been withdrawn and changes in the definitions of measures and in the standards set have occurred. The rationale on which these standards are set is that, to achieve results comparable to the Swedish Two Counties Study, the NHSBSP must achieve equivalent or better interim screening characteristics. The measures defined in 1988 (NHSBSP, 1988) set a single standard for each outcome, as **all** units were in the prevalent round. Standards were set at a level that should have been met by **all** units. The 1993 Objectives for the Breast Screening Programme (NHSBSP, 1993b) introduced standards split into "acceptable" and "achievable" levels. An acceptable standard was set as the minimum to be attained. Achievable standards were set at a level already achieved by 25% of units and provided a target to be aimed at. The 1996 Quality Standards (Radiological Big '18', personal communication) have changed the terminology. A "minimum" standard is set a level at which, if not met, remedial action must be taken. "Expected standards" are set at a level which should be attained at **all** units.

Data were obtained from KC62 reports supplied by the ten screening units in the region for the period 1 April 1993 - 31st March 1994. Outcome measures are presented in relation to the current quality standards in order to identify and discuss the difficulties associated with using the currently recommended outcome measures. In addition, treatment data for 1992 to 1995 provided by QASS reports supplied by nine screening units are used to assess some of the treatment outcome measures used by the NHSBSP.

Table 2.7 Outcome measures used by the NHSBSP 1988 -1996					
Measure	**NHSBSP (1988)[1]**	**Objectives for the NHSBSP (1993b)[2]**		**1996 Radiological Standards[3]**	
	Acceptable	**Acceptable**	**Achievable**	**Minimum**	**Expected**
Uptake rate	70%	70%	75%	70%	-
Cancer detection rate	>50/10,000 (P)	>50/10,000 (P, ninv) >35/10,000 (I, ninv)	60/10,000 (P, ninv) No (I) standard	>2.7/1000 (P, inv) >3.0/1000 (I, inv)	>3.6/1000 (P, inv) >4.0/1000 (I, inv)
SDR	-	-	-	0.75	1.00
Small cancer detection rate	>15/10,000 (P, noninv, <=10mm)	15/10,000 (P, inv, <=10mm)	-	>0.7/1000 (P, inv, <10mm) >1.35/1000 (P, inv, <15mm) >0.75/1000 (I, inv, <10mm) >1.5/1000 (I, inv, <15mm)	>0.9/1000 (P, inv, <10mm) >1.8/1000 (P, inv, <15mm) >1.0/1000 (I, inv, <10mm) >2.0/1000 (I, inv, <15mm)
DCIS rate	-	-	-	>10% <20%	
Referral rate for assessment	<10%	<7% (P) <7% (I)	<5% (P) <3% (I)	<10% (P) <7% (I)	<7% (P) <5% (I)
Benign open biopsy rate	-	<0.5% (P) <0.35% (I)	<0.4% (P) No (I) standard	<2.7/1000 (P) <2.0/1000 (I)	<1.8/1000 (P) <1.0/1000 (I)
Malignant: benign ratio	1:3	1:1 (P) 1.5:1 (I)	-	-	-
PPV (open biopsy)	-	>50%	>60%	-	-
Early recall rate	-	-	-	<1%	-
Pre-operative diagnosis by cytology rate	-	-	-	>70%[4]	-

Key To Abbreviations

P Standard for prevalent (first time) screens
I Standard for incident (second or subsequent) screens
ninv Standard includes non-invasive cancers
inv Standard includes invasive cancers only
PPV Positive predictive value
SDR Standardised Detection Ratio for invasive cancers
DCIS Ductal carcinoma *in situ*

Sources

[1] Quality Assurance Guidelines for Mammography (NHSBSP, 1988)
[2] Objectives for the Breast Screening Programme (NHSBSP, 1993b)
[3] Radiological Standards (Radiological Big '18', personal communication
[4] Target for palpable lesions identified by cytology and needle histology

CHAPTER THREE

RESULTS: THE AVAILABILITY AND VALIDITY OF ROUTINELY COLLECTED PROGNOSTIC DATA

3.1 INTRODUCTION

The validity of data is defined by IARC as the proportion of cases with a given characteristic recorded which truly has the attribute (Parkin *et al.*, 1994). It is essential that data recorded at the WMCIU and by the NHSBSP are both complete and valid if they are to be used in the assessment of breast cancer. In this chapter the results of studies are presented which address the need for high quality prognostic data by assessing breast cancer data routinely recorded at the WMCIU and by the NHSBSP in the West Midlands.

To undertake studies of survival and to identify stage shift resulting from early diagnosis, complete and accurate staging data are required. Ideally TNM stage would be recorded for all breast cancers at diagnosis and supplied to the WMCIU. Complete and accurate transcription would then result in a fully staged database. If valid data describing tumour characteristics were available it would also be possible to confirm the validity of stages by comparison with tumour characteristics. **A** review of prognostic data held at the WMCIU **is** described, focusing on TNM staging and those tumour characteristics incorporated into the **TNM** system. The review contributed to the development of new breast cancer coding rules and an improved coding system for incorporation into the WMCIU's new computer system GRACE. The design and impact of a staging algorithm as a possible means of increasing the proportion of cases staged for analytical purposes is also described.

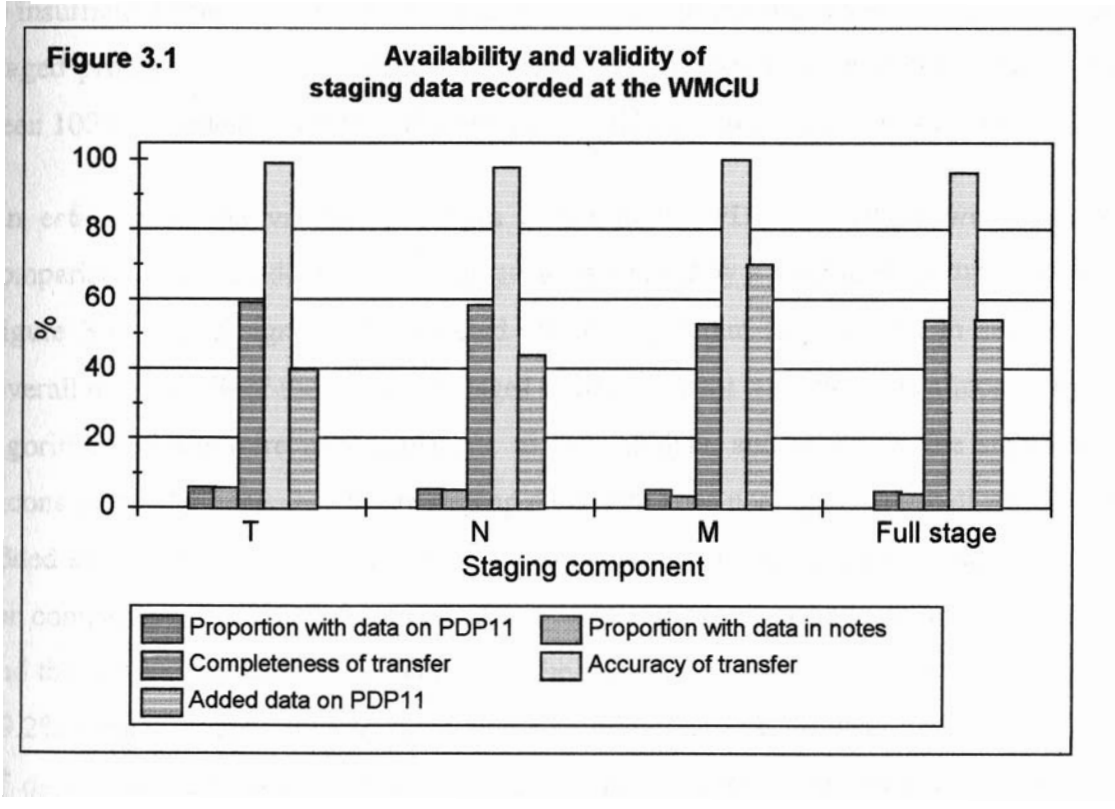
High quality prognostic data are essential to the evaluation of the NHSBSP, and must be available for breast cancers detected both inside and outside the screening programme. The prognostic data collated in HQA reports provided by West Midlands Breast Screening Units was assessed. The availability of prognostic data for breast cancers arising within the catchment area of the Warwickshire, Solihull and Coventry Screening Unit was examined according to screening history and compared with data recorded for cases diagnosed in 1987 prior to the start of screening.

3.2 AN AUDIT OF STAGING AND TUMOUR CHARACTERISTIC DATA HELD AT THE WMCIU

2444 (97.3%) of the 2511 breast cancers diagnosed in the West Midlands in 1983 were eligible for analysis. A total of 67 cases were ineligible. The notes for 34 cases were not located, and notes for 25 cases contained insufficient clinical detail to be of use. Neither DHA nor postcode of residence was available for 7 cases. One case was excluded because neither age nor date of birth was available.

3.2.1 Staging data

Figure 3.1 summarises quality assurance (QA) measures for staging data which describe the transfer of data from case notes onto the old PDP11 database at the WMCIU. The QA measures shown in Figure 3.1 are reported in detail in Table A5.1 of Appendix Five which also provides definitions of all the QA measures presented in Chapter Three.

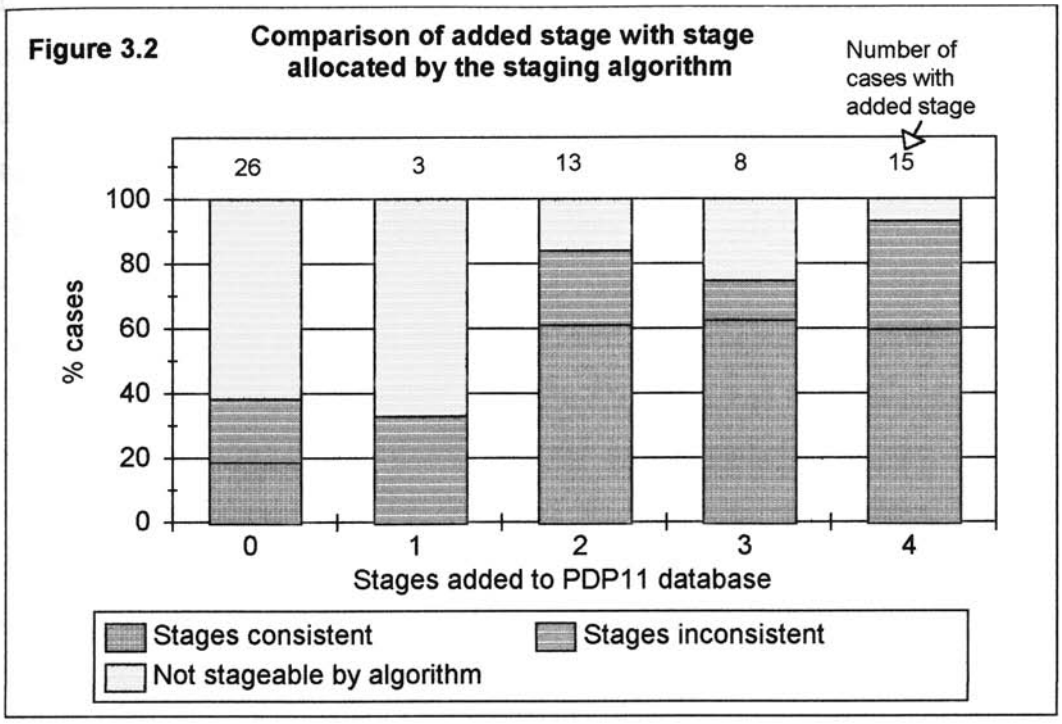


Only 4.8% of cancers had stage recorded on the old WMCIU PDP11 database, and 4.1% had stage recorded in the notes. One or more stage components were recorded on the database for just 5% of cases, with a similar proportion having components recorded in

the notes. The availability of both full stage and stage components was thus extremely poor. Completeness of transfer was highest for advanced stages (stage '3' 61.8%, stage '4' 63.2%), and lower for earlier stages. No stage '0's, indicating *in situ* disease, were found in the case notes, so that the 26 stage '0's recorded on the PDP11 database had been added by abstractors in the absence of data in the notes, and were termed "added stages". Accuracy of transfer was high for all stages, being 96.3% overall. Proportions of added stages on the PDP11 database were lowest for stage '1' and highest for stage '4', suggesting that addition of stage by abstractors was biased towards later stages.

Poor completeness, high accuracy and high proportions of added data were also seen for **'T', 'N',** and 'M' components. Staging components for the 115 cancers without recorded stage but with staging components were combined using the TNM protocol to allocate stage. If no 'M' component was present, 'M0' was assumed. 8 cancers were not staged due to insufficient data. This process increased the overall proportion staged by 113%, but the staged proportion was still under 10%. Thus, even if data transfer at the WMCIU had been 100% complete, less than 10% of cases would have had staging data on the PDP11,

An estimate of the validity of stages added to the PDP11 database was made by comparison of these added stages with stages allocated by a staging algorithm, shown in Figure 3.2. The design of the selected staging algorithm is presented in section 3.3. Overall only 41.5% of the 65 added stages were consistent with the stage allocated by the algorithm, 35.4% were unstagable due to lack of data, and in 23.1% the stages were inconsistent. Agreement with the staging algorithm does not necessarily indicate that the added stage was correct, and no stage allocated by a clinician at diagnosis was available for comparison, but lack of agreement indicates inconsistency between the added stage and the tumour characteristics recorded in the case notes. Of cancers with added stage '0', 19.2% were allocated a higher stage by the algorithm and 61.5% were unstageable. Lack of data describing status of the local nodes was responsible for most of the unstaged cases. For cancers with an added stage '4', 60% were in agreement with the algorithm, but 33.3% were allocated a lower stage by the algorithm as no evidence of distant metastases was recorded in the notes.



3.2.2 Tumour characteristic data

Summary QA measures for tumour characteristic data recorded for the **2444** cases eligible for analysis are shown in Table 3.1. The definitions of QA measures are provided in Chapter Two and in more detail in Appendix Five. Detailed descriptions of the coding systems used on the PDP11 database, and that developed as part of this audit, are provided in Appendix Three and Appendix Four respectively.

Table 3.1 Availability and validity of tumour characteristic data

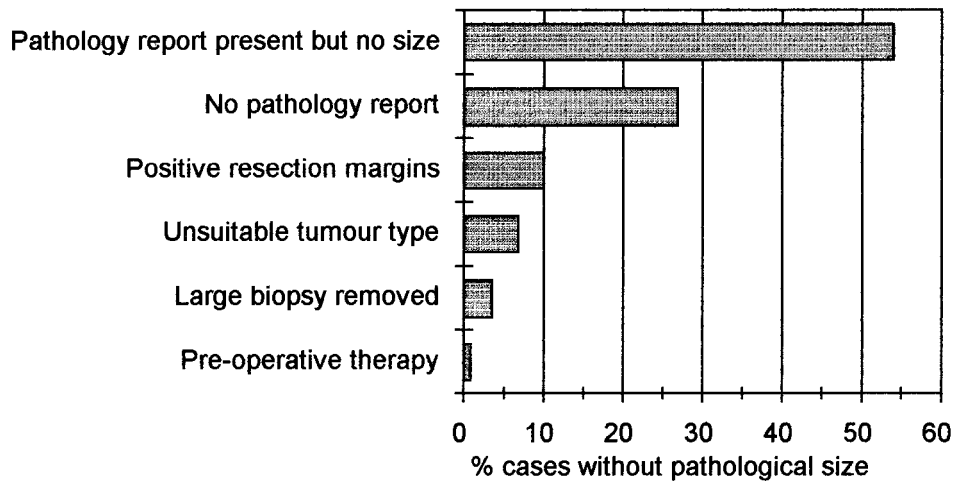
Tumour characteristic	% cases with data on PDP11	% cases with data in case notes	Completeness (%)	Accuracy (%)	Added data (%)
Size	63.2	65.1	86.5	84.6	10.8
Local node status	96.2	71.4	99.0	88.4	22.2
Local node mobility	5.1	1.5	69.4	100.0	79.8
Disease extent	83.4	36.5[1] or 85.0[2]	88.7	80.0[3] or 85.5[4]	7.8
Positive metastases[5]	9.0	9.3	69.2	91.4	18.7
Distant node status	94.4	6.7	97.5	61.0	93.2

[1] excluding "possible" and "probable" disease extent data
[2] including "possible" and "probable" disease extent data
[3] assuming that "possibly invading" is equivalent to "invading" disease extent
[4] assuming that "possibly invading" is equivalent to "limited" disease extent
[5] QA measures for positive data only

3.2.2.1 Tumour *size*

Tumour size was recorded on the **PDPI** 1 for **1544** of the **2444** cases examined **(63.2%),** and valid sizes were found in the notes for **1591 (65.1%).**Of "best" sizes derived from the notes (the definition of "best" size is provided in Chapter Two), **66.4%**were pathological, **33.2%** clinical, and **0.3%** radiological. No "best" size was identified for **853** cases **(34.9%),** of which **19.6%** had a size incorrectly recorded on the **PDP11** database, including **90** cases with an invalid pathological size. For prognostic purposes pathological size is considered most accurate, but **1387 (56.8%)** cases had no valid pathological size recorded in the notes, of which **275** with an invalid pathological size. Reasons for invalidity of pathological size are described in Appendix Four. Reasons for lack of pathological size in the sample examined are shown in Figure **3.3,** with the most common reason for invalidity being the presence of positive resection margins. In **367** cases **(26.5%)** without pathological size, no pathology report was available. However, in a further **745** cases **(53.7%),**a pathology report was present but size had not been recorded.

Figure 3.3 **Reasons for lack of pathological size on review database**



Overall completeness and accuracy of transfer of size was high at **86.5%** and **84.6%** respectively. However, **10.8%** of sizes had been wrongly added by abstractors. The completeness and accuracy of pathological size **(89.4%** and **89.7%** respectively) was

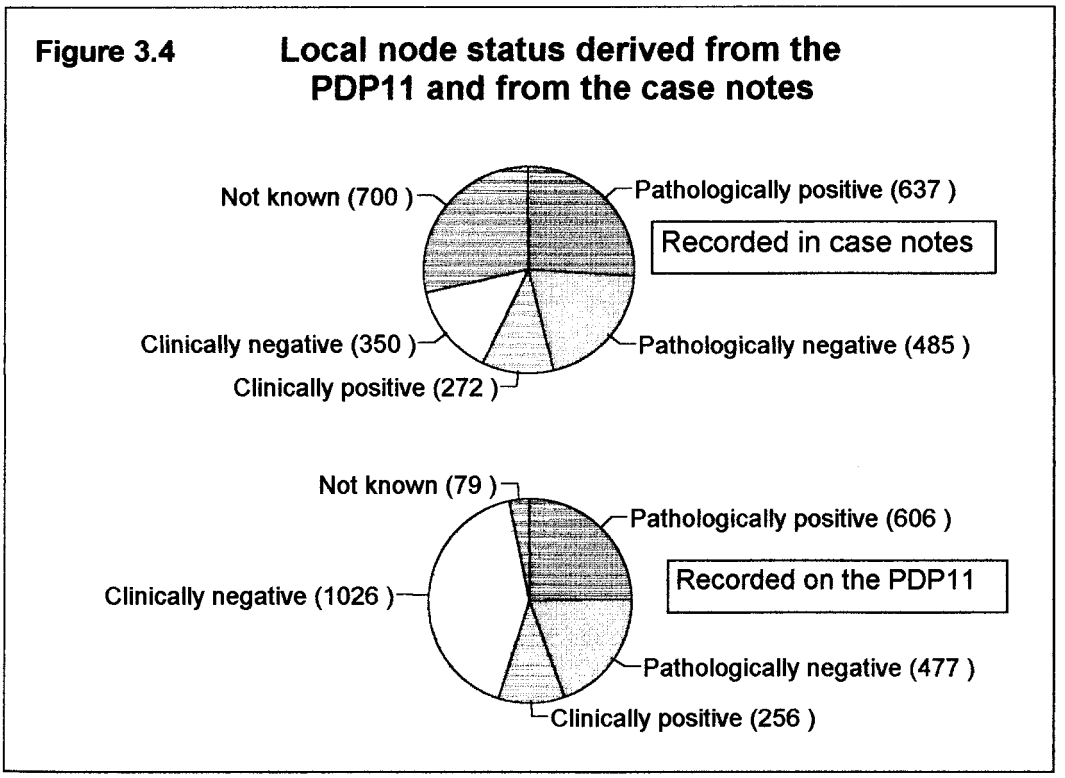
higher than for clinical size (81.1% and **73.9%).** Completeness and accuracy of transfer are not independent of coding changes and of the consistency with which selection of size **for** entry operated. For example, if a clinical size of 15mm and a pathological size of 10mmwere both available and valid, pathological size should have been recorded. If clinical size had been selected incorrectly, even if transferred accurately it would not match the "best" size of 10mm,giving the appearance of inaccurate transfer.

A total **of** 90 invalid pathological sizes had been recorded on the PDP11, and furthermore, 49 cancers with a valid pathological size in the notes had a different size recorded on the PDP11 which matched the clinical size in the notes. Although possibly due to chance, a likely explanation is that abstractors overlooked valid pathological sizes. The reliability of size on the PDP11 was further reduced by the finding that 18.9% of cases had no size recorded on the PDP11 database despite identification of a clinical size during the review.

3222 Local node status and mobility

Figure **3.4** shows local node status data recorded on the PDP11 database and in the notes.

Figure 3.4 **Local node status derived from the PDP11 and from the case notes**



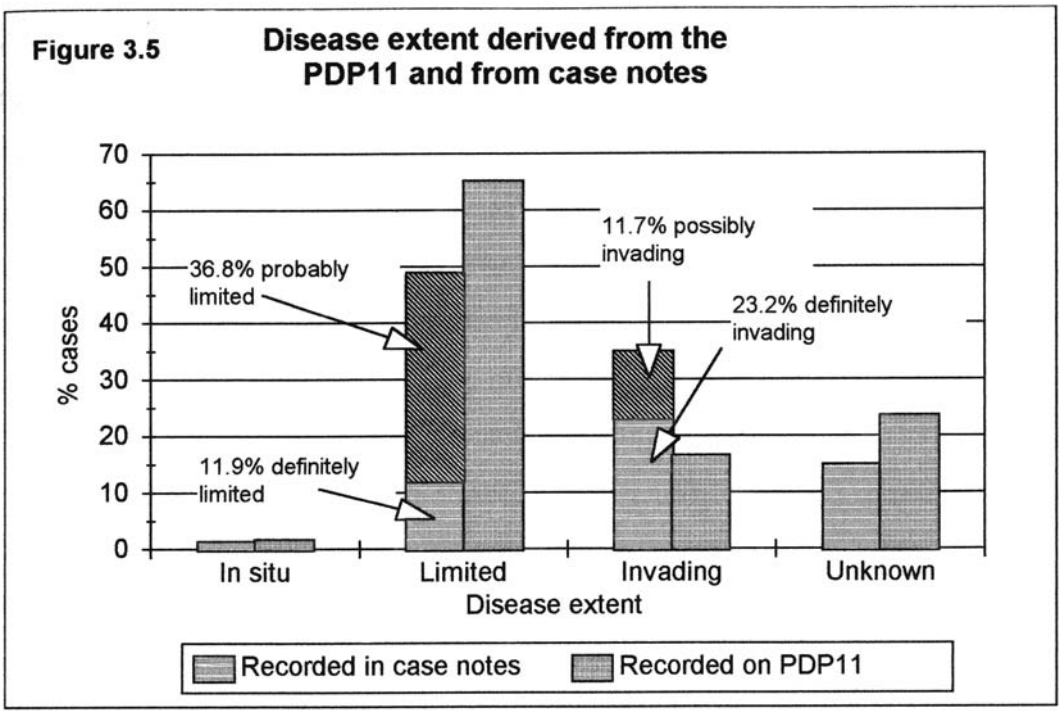
Although local node status was recorded on the PDP11 database for 2351 cases (96.2%), only 1122 (45.9%) had pathological local node status recorded in the notes, and a further 622 (25.5%) had clinical node status recorded although pathological data were absent. Of the 1322 cases without pathological node data in the notes, 476 (36.0%) had a pathology report available but nodal status was not recorded. In 41 cases node sampling was recorded in the notes, but no pathology for this procedure was recorded, indicating that the data sent to the WMCIU were incomplete. In 700 cases (28.6%) no clinical or pathological data describing the local nodes were available.

The completeness and accuracy of recording nodal status at the WMCIU was high at 99.0% and 88.4% respectively. A problem with the data on the PDP11 database is shown in Figure 3.4 by the proportion of cases recorded as "clinically negative". Data on the PDP11 database indicated that 1026 cases (42.0%) fell into this category, but data in the notes revealed that only 350 cases (14.3%) could be placed in this group. Of 1026 cases recorded as "clinically no nodes" on the PDP11 database, 65.9% had been incorrectly allocated this code to indicate lack of data instead of the rarely used "no data" code.

Local node mobility, which is required for TNM staging, was recorded on the PDP11 database for only 125 cases (5.1%). However, just 37 cases (1.5%) had mobility recorded in the notes. The completeness of transfer of mobility data was 69.4%, and although accuracy was 100%, nearly 80% of mobility data had been added in error by abstractors.

3.2.2.3 Disease extent

Figure 3.5 shows that 1593 cases (65.2%) were recorded as "limited to the site of origin" on the PDP11 database compared to 291 cases (11.9%) recorded on review as "definitely limited" based on data in the notes. The interpretation of data in relation to disease extent is complex. The introduction of new coding categories allowed definite and possible evidence of invasion beyond the breast to be recorded. Even if "probably limited" cases are assumed to be limited to the breast, the "limited" proportion on the PDP11 was significantly higher than that in the notes (65.2% vs 48.7%, $p<0.05$). Conversely, if only "definitely invading" cancers were assumed to be invading beyond the breast, significantly fewer cases were "invading" on the PDP11 database (16.6% vs 23.2%, $p<0.05$).



Overall completeness of recording disease extent was high at **88.7%,** and accuracy was over **80%,** although calculation of accuracy was complicated by the interpretation of new coding categories. For "possibly invading" cases, completeness was **82.9%.** Accuracy, if equivalence to "invading" on the **PDP11** database is assumed, was **30.0%,** but when equivalence to the "limited" code on the **PDP11** database was assumed, rose to **61.3%.** This suggested that characteristics that were suspicious of invasion had been interpreted inconsistently at the original abstraction. For "probably limited" cancers, completeness was **91.3%,** and accuracy assuming equivalence to the "limited" **PDP11** code, was over **98%.** The proportion of added data was low at **7.8%.** The disease extent field on the **PDP11** database had not been used to identify ***in situ*** cancers, demonstrated by the observation that of the **34** ***in situ*** cancers identified, **88.2%** had been coded on the **PDP11** database as "limited". The new coding rules allowed disease extent to be recorded for 52.2% of cancers that had no disease extent data on the **PDP11** database, including **66** "definitely invading" cases. In a small number of cases, disease extent data were no longer hidden, **as 45** of **61** cancers recorded as "multiple lesions" or "found at ***post mortem"*** on the **PDP11** database had data in the notes. In addition, ***in situ*** cases were identified in the new disease extent field in addition to the use of the appropriate **ICD9** code.

3.2.2.4 Distant metastases and distant nodes

Identification of positive distant metastases from the PDP11 database required comparison of clinical and "at operation" data, which was complicated as "at operation" data could not be translated directly as pathological as some codes were based on observation at surgery only. For the review all metastatic sites were coded together, and clinical and pathological data were coded in one field on a hierarchical basis, making data analysis easier. Only 220 cases had any metastases data recorded on the PDP11 database, and over 77% had no data relating to metastases recorded in the notes. 31 cases (1.3%) had pathologically confirmed metastases, and 196 (8%) clinically positive metastases. Only 3 cancers had all suspected sites confirmed as negative, and clinically negative data were available for 269 (11.0%). A further 46 (1.9%) had suspicious symptoms such as back ache which had not been investigated further. Only 157 of the 227 cases with positive data (69.2%) in the notes had positive data on the PDP11 database, although accuracy was 91.4%. **A** total of 76 cancers with positive data had been overlooked, and 18.7% of positive data on the PDP11 had been incorrectly added. Of 46 cancers with "suspicious" data in the notes, only 31.4% had been recorded as clinically positive on the PDP11 database indicating inconsistent interpretation of data.

The recording of distant node status suffered from the same problems as those identified for local nodes. Only 164 cases (6.7%) had distant nodes data recorded in the notes, compared to 2307 cases (94.4%) with distant node status recorded on the PDP11. Overall, 93% of data on the PDP11 database had been added, mostly as a result of incorrect use of the "clinically no nodes not present" code, which was used for 95% of cancers on the PDP11 database. Although overall completeness was 97.5%, accuracy was only 61%, indicating that distant node data had been omitted on the PDP11.

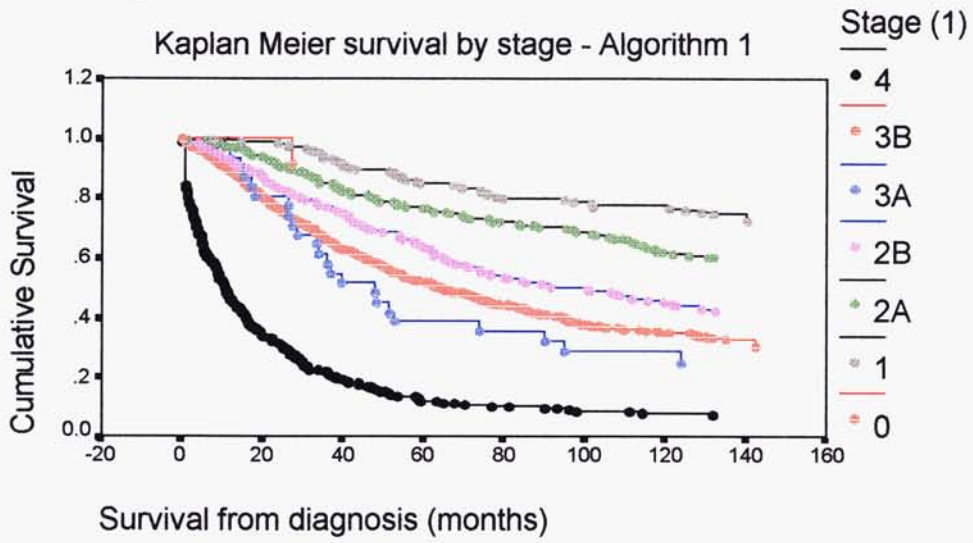
3.3 THE DESIGN AND VALIDATION OF A STAGING ALGORITHM

Data recorded for the review were used in the development of four staging algorithms from which one algorithm was selected for future use. The assumptions contained within each algorithm, and the method by which the algorithms were tested are provided in Chapter Two. Programme listings for each algorithm are reproduced in Appendix Six.

3.3.1 Survival by stage from the four algorithms

Figures 3.6 - 3.9 show survival stratified by stage calculated from the four algorithms.

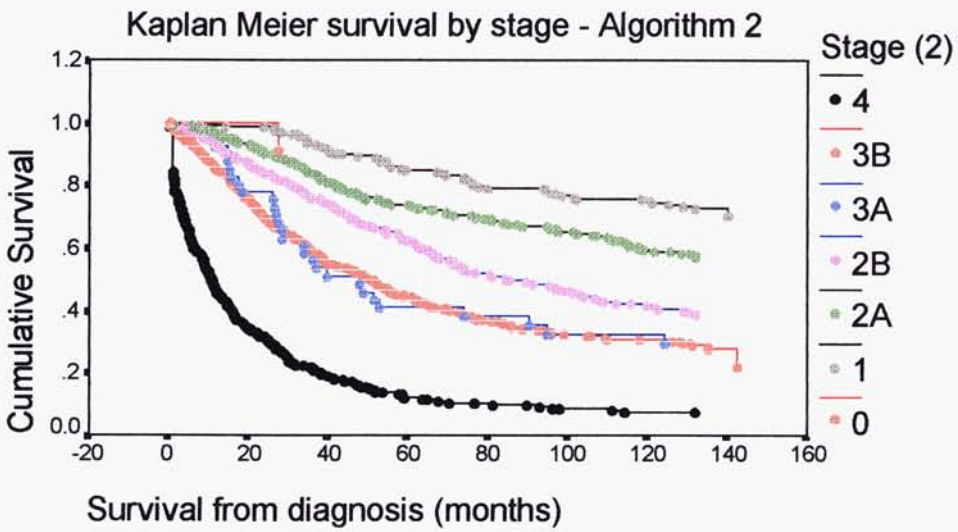
Figure 3.6



Chi-square test for linear trend of survival by stage from Algorithm 1pooled over all stages

Test	Statistic	Degrees of freedom	P value
Log rank	375	1	<0.001

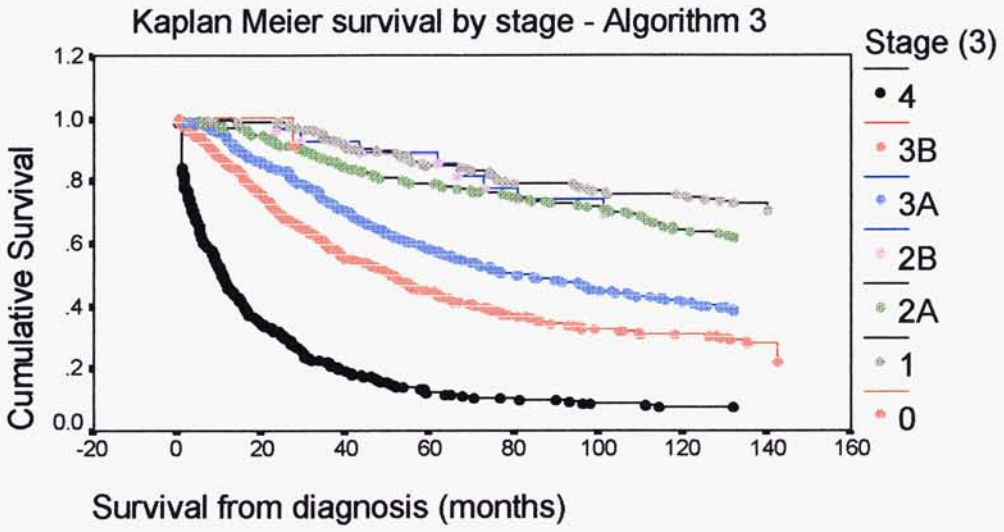
Figure 3.7



Chi-square test for linear trend of survival by stage from Algorithm 2 pooled over all stages

Test	Statistic	Degrees of freedom	P value
Log rank	486	1	<0.001

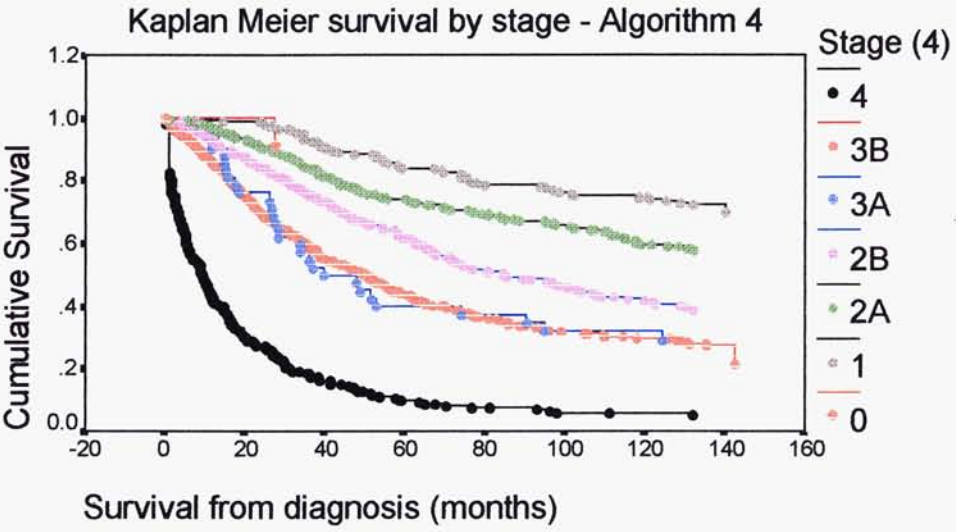
Figure 3.8



Chi-square test for linear trend of survival by stage from Algorithm 3 pooled over all stages

Test	Statistic	Degrees of freedom	P value
Log rank	444	1	<0.001

Figure 3.9



Chi-square tests for linear trend of survival by stage from Algorithm 4 pooled over all stages

Test	Statistic	Degrees of freedom	P value
Log rank	495	1	<0.001

Table 3.2 shows the number of cancers allocated to each stage by each staging algorithm, and five year survival rates calculated at 60 months or at the closest previous time point.

Table 3.2 Cumulative five year survival by stage according to four algorithms								
Stage	Algorithm 1		Algorithm 2		Algorithm 3		Algorithm 4	
	No.	5 year survival (%)	No.	5 year survival (%)	No.	5 year survival (%)	No.	5 year survival (%)
0	12	-**	12	-**	12	-**	12	-**
1	192	84.8	228	84.9	228	84.9	230	84.2
2A	285	76.8	345	74.0	229	79.6	349	74.3
2B	170	64.2	228	62.9	28	89.2	235	62.3
3A	31	38.7	41	41.0	357	58.3	42	40.0
3B	660	52.2	404	44.7	404	44.7	419	43.9
4	316	12.3	316	12.3	316	12.3	275	10.0
Unstaged#	778	70.9	870	70.0	870	70.0	882	69.2

** 1 stage '0' case died from breast cancer. The pathology report stated intraductal carcinoma.

Total unstaged includes 55 cases that were DCO, found at PM or not carcinoma; these were excluded in the calculation of mean survival for those cases that were unstaged

All four algorithms staged over 63% of cancers, with algorithm 1 staging 68% and algorithm **4** *64%.* For all algorithms, survival by stage displayed a highly significant trend, with stage '4' showing the poorest survival. Algorithm **4** displayed the most significant statistics for trend, and algorithm 1 the least. Algorithm 1 showed inversion of the survival from stages '3b' and '3a' so that stage '3a' cancers displayed poorer survival than stage '3b'. This inversion of stages 3'a' and '3b' was displayed by all algorithms except Algorithm 3, which inverted stages '2a' and '2b'. The influence of the assumptions created differences in the proportion of cases that could be staged, and in the survival calculated for each stage.

3.3.2 Survival by stage according to staging algorithm

Figures 3.10 - 3.15 show survival curves for each stage stratified by the algorithm used. For stage '1', survival according to algorithms 2, 3, and **4,** was similar, but that from Algorithm 1 was slightly higher. For stage '2a' cases, survival from algorithm 3 was the highest, followed by algorithm 1, with similar curves for algorithms 2 and **4.** For stage '2b', survival for those staged by algorithm 3 was much higher than for algorithms 1, 2, and **4,** and a similar pattern was seen for stage '3a' tumours. Survival for stage '3b' was higher for those staged by algorithm 1 than from the other three algorithms. For stage '4 tumours, survival for tumours staged by algorithms 1, **2** and 3 were identical so that only the curve for algorithm 1 is visible, but survival from algorithm **4** was slightly poorer.

Figure 3.10

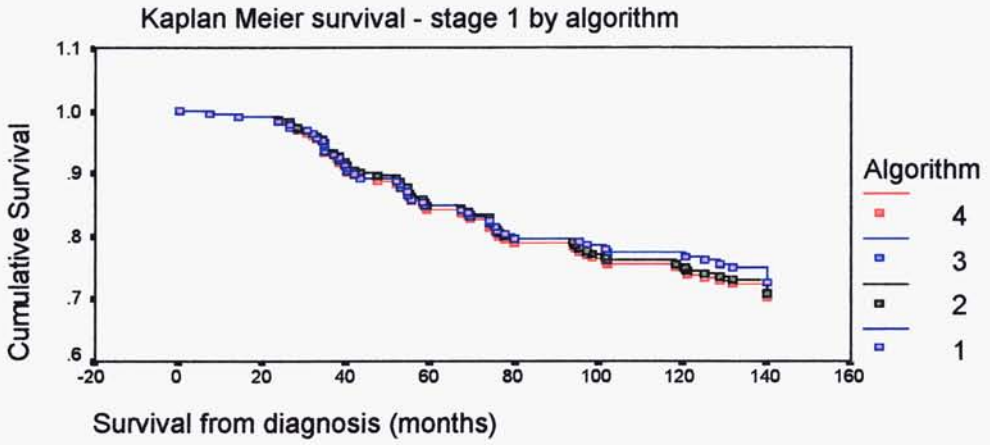


Figure 3.11

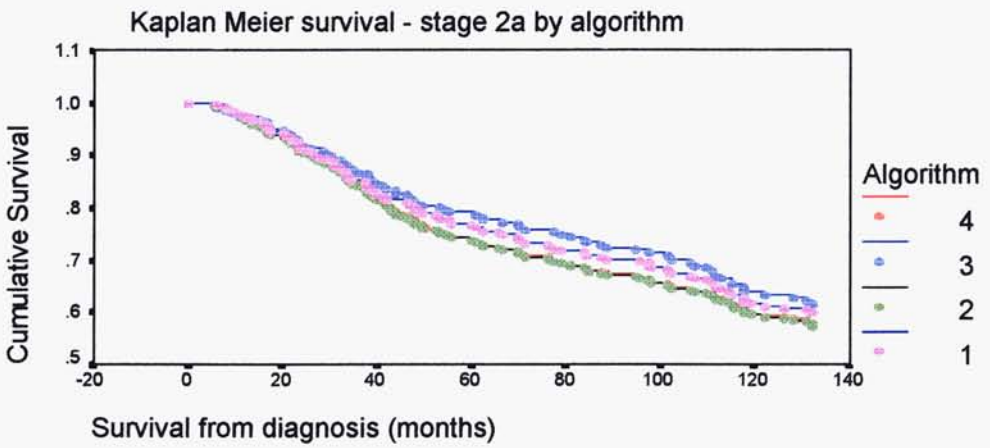


Figure 3.12

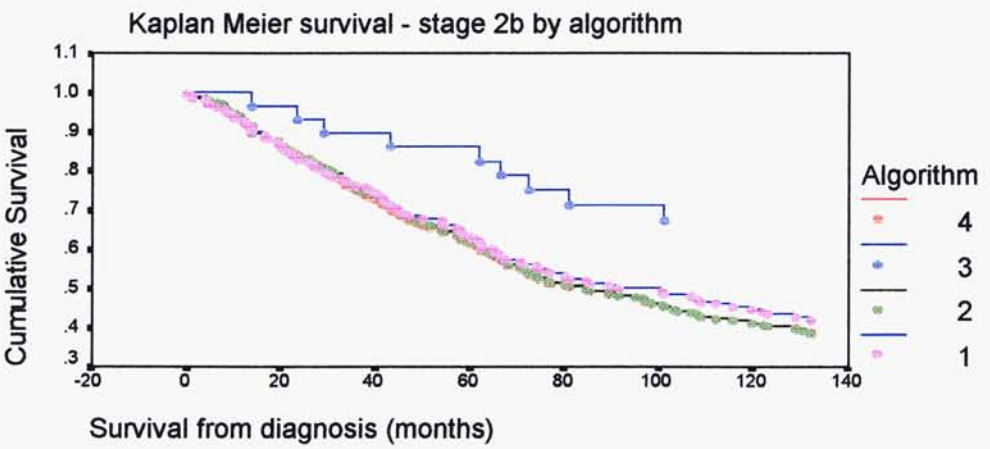


Figure 3.13

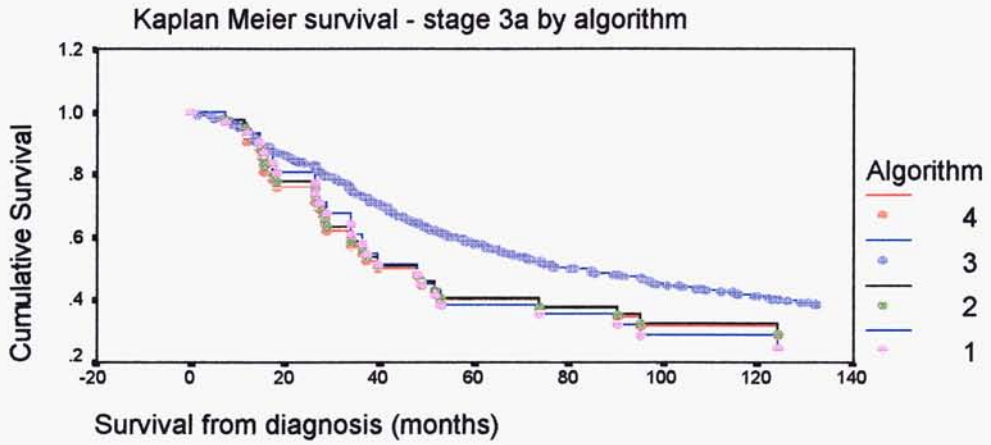


Figure 3.14

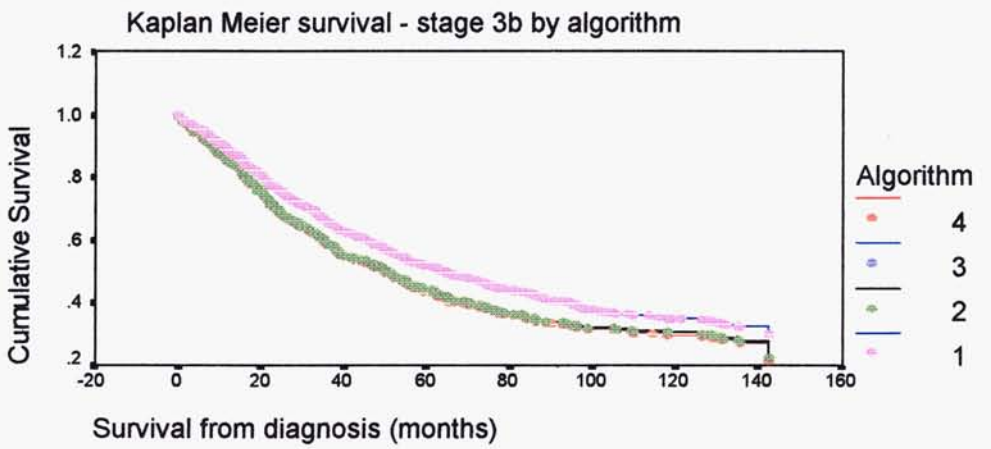
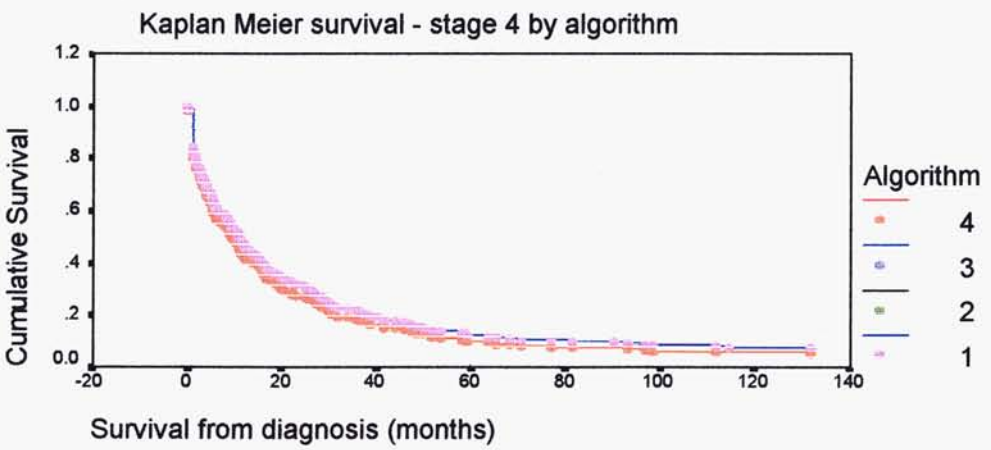


Figure 3.15



3.3.3 Selection of the best algorithm

3.3.3.1 Comparing Algorithm 1 and Algorithm 2

Algorithm 1 staged more cases than algorithm 2, with 31.8% and 35.6% unstaged respectively, but the assumption used in algorithm 1, that the two disease extent codes were equivalent was not supported. Five year survival rates for stage '3b' cases (Table 3.2) were 52.2% and 44.7% for algorithms 1 and 2 respectively, and a log rank test for difference in survival for stage '3b' from each algorithm was significant (p=0.02). Algorithm 1 inverted stages '3a' and '3b', and although the trend in survival across stages was highly significant for both algorithms it was more significant for algorithm 2. Algorithm 1 was thus rejected.

3.3.3.2 Comparing Algorithm 2 and Algorithm 3

Both algorithms staged 64.5% of cases. However, the assumption that nodes of unknown mobility were fixed, used in algorithm 3, was not supported. Presence of fixed local nodes is the criterion for 'N2', with a minimum stage of '3a'. If mobile nodes were allocated 'N2' this would increase the survival of stage '3a' cases. Five year survival rates for stage '3a' cases shown in Table 3.2 were 41% and 58% for algorithms 2 and 3 respectively, and a log rank test for difference in survival for stage '3a' cases from each algorithm gave a near significant value of 3.06 (p=0.08). In addition, the trend in survival across stages was stronger for algorithm 2 than for algorithm 3. Algorithm 3 was therefore rejected.

3.3.3.3 Comparing Algorithm 2 and Algorithm 4

Algorithm 4 staged fewer cases than algorithm 2, with 36.1% and 35.6% of cases unstaged respectively. The assumption that suspicion of metastases was indicarive of positive metastases was not supported. The five year survival rates for stage '4' cases (Table 3.2) were 12.3% and 10.0% for algorithms 2 and 4 respectively, and a log rank test for difference in survival for stage '4' cases from each algorithm was non-significant (p=0.20). Despite the lack of significance, this result suggetes that cases allocated a stage '4 by algorithm 2 due to suspected metastases, when allocated a lower stage by algorithm 4, left behind cases of a poorer survival in stage '4'. Furthermore, the trend in survival by stage for algorithm 4 was stronger than for algorithm 2. Algorithm **2** was thus rejected.

3.3.4 The selected staging algorithm

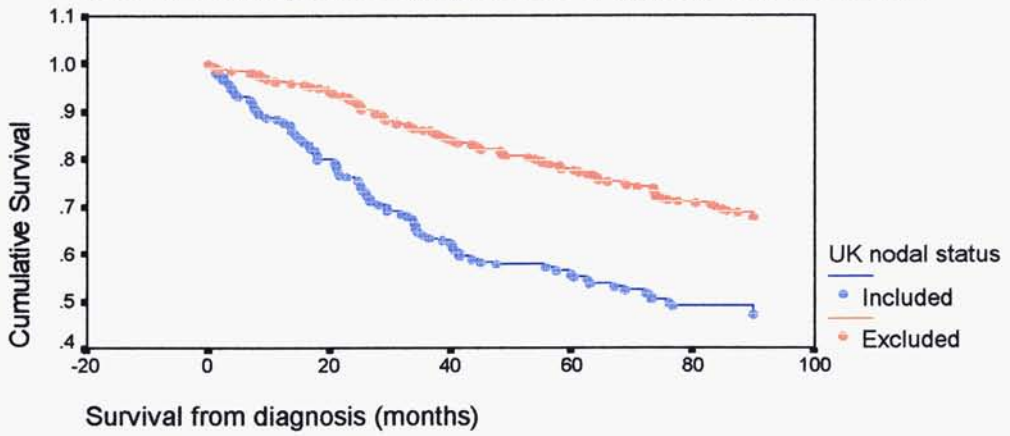
Algorithms 1, 2 and 3 were rejected leaving algorithm 4 as the selected staging algorithm. The selected algorithm used only definite evidence of disease extent beyond the breast to allocate 'T4', assumed that local nodes of unknown mobility were mobile, and assumed that suspicion of distant metastases was unfounded in the absence of definite evidence. The lack of nodal mobility data meant that very few cancers were allocated to stage '3a', making this stage category of little practical use. The observed five year survival rates for stages '3a' and '3b' were similar (40.0% based on 42 cancers and 43.9% based on 418 cancers respectively), and displayed stage inversion, and in an amendment to the selected algorithm, these two stages were amalgamated into a single stage '3'.

All four algorithms excluded cases with unknown nodal status if evidence indicating a stage '3b' or '4' was absent. To examine whether this was justified, algorithm 4 was modified to allocate stages to cancers with unknown nodal status by assuming that the nodes were negative. Figure 3.16 shows survival for stage '2a' cases including and excluding those with unknown nodal status. Including cases with unknown nodal status significantly reduced the survival for stage '2a' cases (log rank statistic 26.04 $p<0.0001$), suggesting that at least some of these cases had nodal involvement. Exclusion of cases with nodes unknown was retained in the selected algorithm, algorithm 4.

Figure 3.16

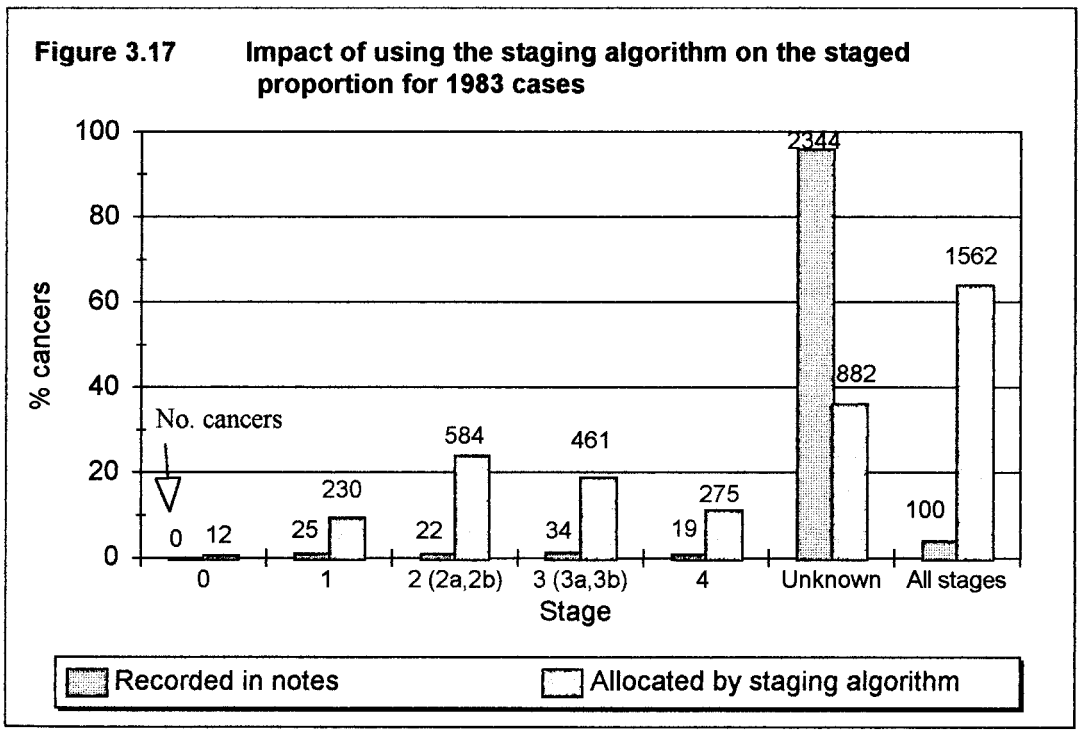
Kaplan Meier survival for stage '2a' cancers (algorithm 4)

effect of including and excluding cancers of unknown (UK) nodal status



using algorithm 4, 1562 cases (63.9%) were staged, producing an almost 13-fold improvement in the staged proportion, shown in Figure 3.17. The largest change (other than that for stage '0' cases, of which none were recorded in the notes) was seen for stage '2' cases, with a 26-fold increase, the lowest increase being for stage '1' cases for which the algorithm brought about an **8** fold increase. For advanced stages (stages '**3**' and '4', a 12 fold increase was seen, demonstrating that, even for advanced cancers, many had not been staged in the notes. However, despite the algorithm 36.1% of cases remained unstaged including 55 cancers excluded as registrations from **DCO** or post mortem report, or non-carcinoma histology. The remaining unstaged cases had insufficient data even after the review of the notes. The cumulative 5 year survival rate for unstaged cases was 69%, midway between stage '2a' (79.2%) and stage '2b' (62.3%), suggesting that this group of cases contained earlier stage cases rather than a preponderance of advanced disease.

Figure 3.17 **Impact of using the staging algorithm on the staged proportion for 1983 cases**



More data items are required to allocate stages '1' to '3a' than for stages '3b' and '4'. Only a single data item is required to allocate stage '3b' or '4' (invasion beyond the breast or positive metastases respectively), but in the absence of data indicating 'T4' or 'M1', size, nodal status and disease extent must all be recorded. In the sample examined, 28.6% had no nodal status and 34.9% had no size recorded in the notes.

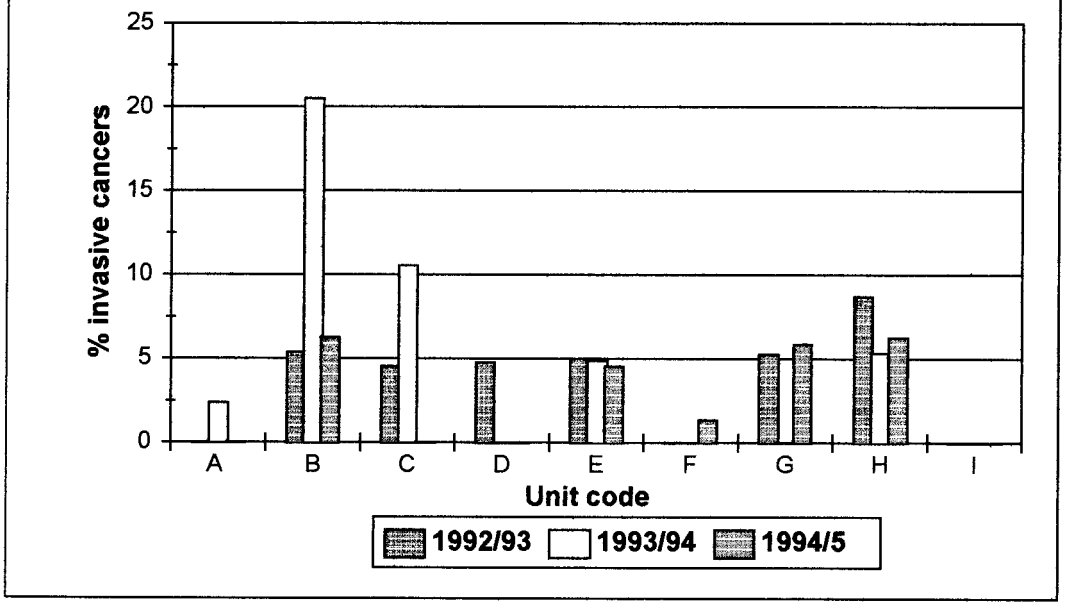
3.4 THE AVAILABILITY OF PROGNOSTIC DATA FOR EVALUATION OF THE NHSBSP

3.4.1 Data available from standard NHSBSP reports

Prognostic and diagnostic data recorded on the computer record system at nine West Midlands screening units and compiled into HQA reports were examined. The format of HQA reports is described in Appendix Two.

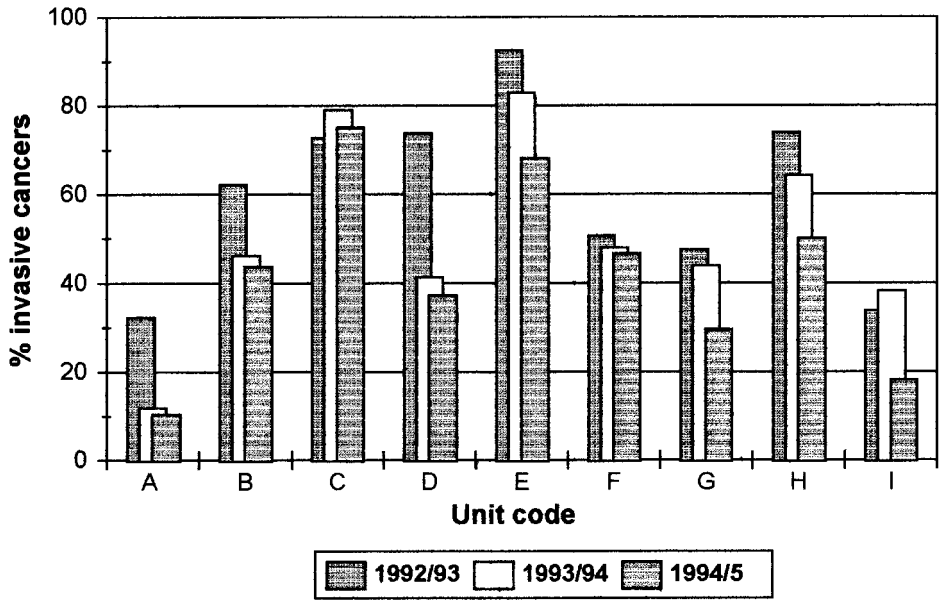
The proportion of invasive cancers of unknown size recorded at each screening unit from 1992 - 1995 are shown in Figure 3.18. There was variation across the region, and despite generally falling rates, three units failed to achieve the **<5%** minimum standard set for unknown size (Radiological Big '18', personal communication). In addition to restricting analysis of prognosis, cancers of unknown size cannot contribute to the small cancer rate, which is one of the outcome measures used by the NHSBSP discussed in Chapter Six.

Figure 3.18 **Proportion of cancers with unknown size by screening unit 1992-95**



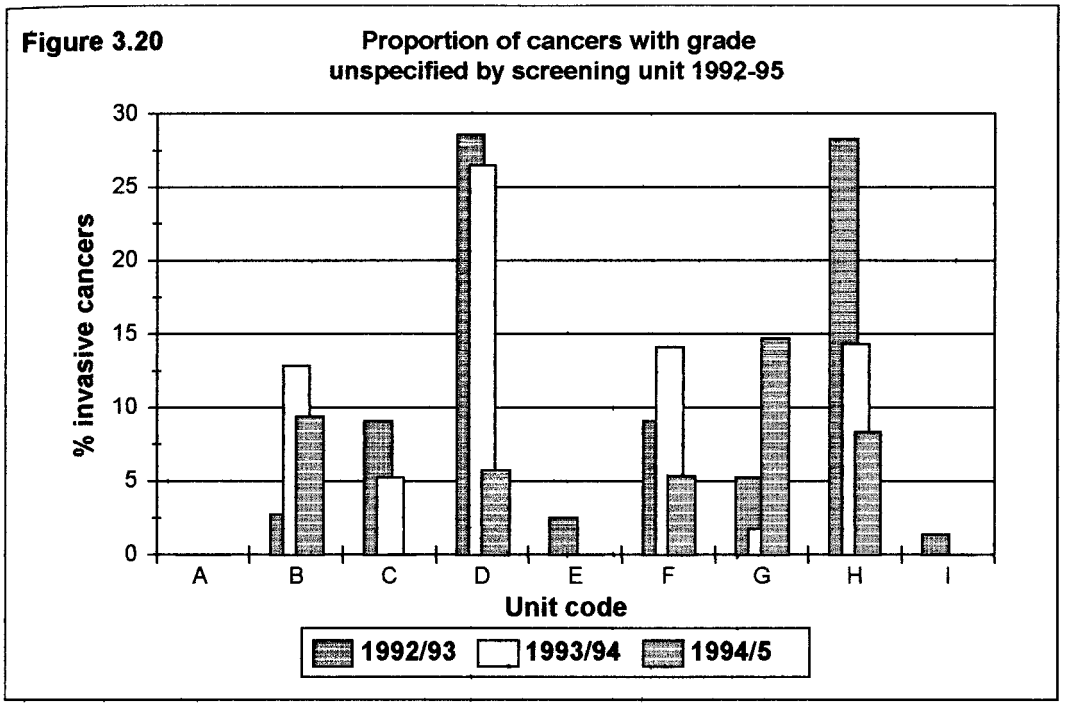
Data describing local node status should be provided for all breast cancers detected by screening (NHSBSP, 1996). However, Figure 3.19 shows that many units had high proportions of cancers with unknown nodal status. In 1992/93 rates of cancers with unknown nodal status varied from 18% to 92% across the region.

Figure 3.19 **Proportion of cancers with nodal status unknown by screening unit 1992-95**



Despite falling rates from 1992 onwards, in 1994/95 three units had nodal status recorded for less than 50% of invasive cancers. For screening to be effective it is recommended that >70% of cancers are node negative (Tabar et ***al.,*** 1992), but if data are incomplete it is impossible to tell whether this is the case. One reason for lack of data may be that node sampling is performed as a secondary surgical procedure of which details are not routinely entered onto screening unit computer systems.

Variation in the proportions of cases without grade recorded are shown in Figure 3.20. Generally falling rates from 1992 to 1995 suggest that the proportion of cases being graded and/or the supply of data to screening units has improved, especially, for example, at Units D and **H.** The increasing rate at Unit G is cause for concern and should be investigated. Tabar also recommended that at least 30% of grade III screen detected tumours should be <15mm (Tabar et ***al.,*** 1992). However, the **HQA** report current in 1996 did not strati& by both size and grade, and used the size limit <=15mm and not <15mm, making this outcome measure impossible to calculate. The surgical **(QASS)** report did strati& by grade and size but did not provide the <15mm group either.



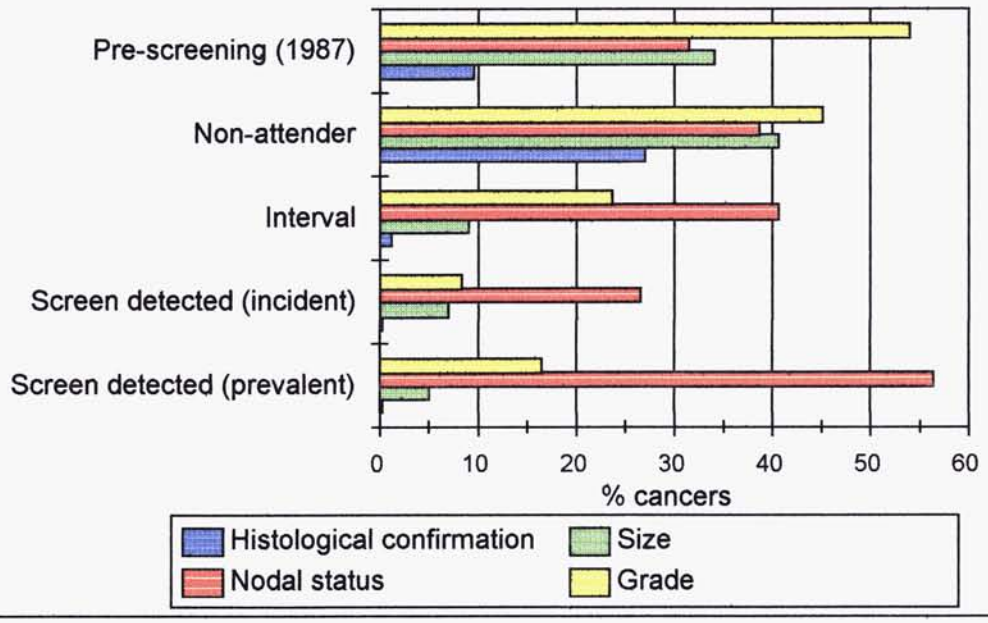
3.4.2 THE AVAILABILITY OF PROGNOSTIC DATA IN RELATION TO SCREENING HISTORY

The screening histories and the availability of prognostic data for 755 cancers in women invited to screening at the Warwickshire, Solihull and Coventry Breast Screening Unit were analysed. The allocation of screening history is described in Chapter Two and the prognosis and screening histories of this group of cancers is explored in more detail in Chapters Four and Six. In addition, data recorded for **788** breast cancers diagnosed in 1987 in women aged 50 - 64 years were examined as pre-screening controls. Data held at the WMCIU and supplied by the screening unit were combined to maximise data completeness. The proportions of cancers in each group with histological confirmation, tumour size, nodal status and grade recorded are shown in Figure 3.21.

Histological confirmation was available for all the screen detected cancers, and for all but 1.1% of interval cancers. The proportion of cancers in non attenders without histological confirmation was 26.9%, higher than the 9.5% found in pre-screening controls. Even amongst screen detected cancers size was unavailable for 4.8% and 6.8% of prevalent and incident cancers respectively, despite the quality standard set by the NHSBSP of **<5%** of cancers be without recorded size (Radiological 'Big 18', 1996). 8.9% of interval cancers

were without size, but over 40% of cancers in non-attenders had no size available, again higher than the 34.1% in pre-screening controls.

Figure 3.21 **Proportion of cases with missing prognostic data according to screening history**



The poorest availability of nodal status was observed in prevalent screen detected cancers, of which over half had no data recorded, even when data from pathology reports held at the WMCIU were included. As the WMCIU should be sent pathology reports relating to all surgical procedures including those which might be classified as secondary procedures not recorded by the NHSBSP, this suggests that the lack of data on HQA reports also reflects failure to supply data to the screening units as well as a possible lack of pathological investigation. Interval cancers and cancers in non-attenders were similar with around 40% of cases without nodal status compared to 31.3% of pre-screening controls. Encouragingly, only 26.3% of incident screen detected cancers were without nodal status suggesting that availability of this data item is improving.

Tumour grade was available for the majority of screen-detected cancers, with only 16.4% of prevalent and 8.3% of incident screen detected cancers being ungraded. Grade was most often unavailable in symptomatic cases, with 23.5% of interval cancers, and 44.9% of cancers in non-attenders, compared to 53.9% of pre-screening controls.

3.5 SUMMARY

An audit of prognostic data recorded at the WMCIU revealed that, in general, the accuracy and completeness of the data recorded were high, but identified several factors which reduced the validity of data recorded at the WMCIU.

- only 4.8% of the 2444 cases studied had stage recorded on the WMCIU database, and only 4.1% had stage recorded in the notes held at the WMCIU
- even when stage was derived from individual staging components found in the case notes, it was only possible to obtain a stage for 10% of the cancers investigated
- the accuracy of the recording of stage on the PDP11 was high at 96.3%, but only 54% of stages in the notes were recorded on the PDP11
- 45.8% of stages on the PDP11 were derived from the notes, the remainder being added by abstractors. Of added stages, only 40% were consistent with the stage allocated by the staging algorithm
- size was available for 65.1% of breast cancers, local node status for 71.1% and disease extent for 83.4%
- the completeness and accuracy of tumour characteristic data recorded on the old WMCIU PDP11 database was high at 87% and 85% respectively for size, 99% and 88% for local node status, 97.5% and 61% for distant node status, 89% and 86% for disease extent and 67% and 91% for positive metastases
- misuse of codes for local node status occurred on the PDP11 so that cases with unknown nodal status were indistinguishable from those with negative nodes. Survival analysis for stage '2a' cases including and excluding cases with unknown nodal status suggested that some of these cases were node positive
- this misuse of codes limited the usefulness of prognostic data recorded on the PDP11. Improved coding rules and a new database structure which rectified the problems identified have been implemented in the new **WMCIU** database GRACE
- in 745 of the 1387 cases (53.7%) without a pathological size, and 476 of the 1322 cases (36.0%) without pathological node status, a pathology report had been sent to the WMCIU but did not report these prognostic factors

A staging algorithm was designed to combine coded tumour characteristics. The algorithm was able to stage 63.9% of cases, a 13-fold increase in the staged proportion. The algorithm was able to improve the staged proportion sufficiently for use in examination of treatment, prognosis and survival, although its use was restricted by the completeness of the data, and some of the assumptions made may not always be valid.

Analysis of the HQA data recorded by the NHSBSP revealed similar problems to those seen with pathological data at the WMCIU. For the screening year 1994/95

- an average of 42.1% of cases had no local node data recorded, with the rate varying from 10.3% to 75.0% across the region
- three screening units had no size recorded for >5% of cancers, with rates varying from 0% to 6.3% across the region
- five out of the nine screening units had no grade recorded for >5% of cases, with rates varying from 0% to 14.7% across the region

The availability of prognostic data varied according to screening history. Histological confirmation and prognostic data were available more often for screen detected cancers detected compared cancers diagnosed outside the screening programme.

- size was available for 95.2% of prevalent, and 93.2% of incident screen detected cancers, 91.1% of interval cancers, 59.5% of cancers in non-attenders and 65.9% of pre-screening controls
- local node status was available for only 43.9% of prevalent, and 73.7% of incident screen detected cancers, 59.6% of interval cancers, 61.5% of cancers in non-attenders and 68.7% of pre-screening controls
- grade was available for 83.6% of prevalent, and 91.7% of incident screen detected cancers, 76.5% of interval cancers, 55.1% of cancers in non-attenders and 46.1% of pre-screening controls

To facilitate the evaluation of the NHSBSP, the completeness of prognostic data must be improved, particularly in relation to local node status for screen detected cancers. The completeness of prognostic data for breast cancers diagnosed outside the programme needs particular attention.

CHAPTER FOUR

RESULTS: AN EXAMINATION OF BREAST CANCER IN THE WEST MIDLANDS

4.1 INTRODUCTION

This chapter describes breast cancer in the West Midlands, utilising published data and findings from analyses undertaken at the WMCIU. An overview is presented to describe the magnitude of the problem in terms of incidence, mortality, and survival from the disease in the region. Attention is focused on differences across the West Midlands, and variation by age and time trends is reported.

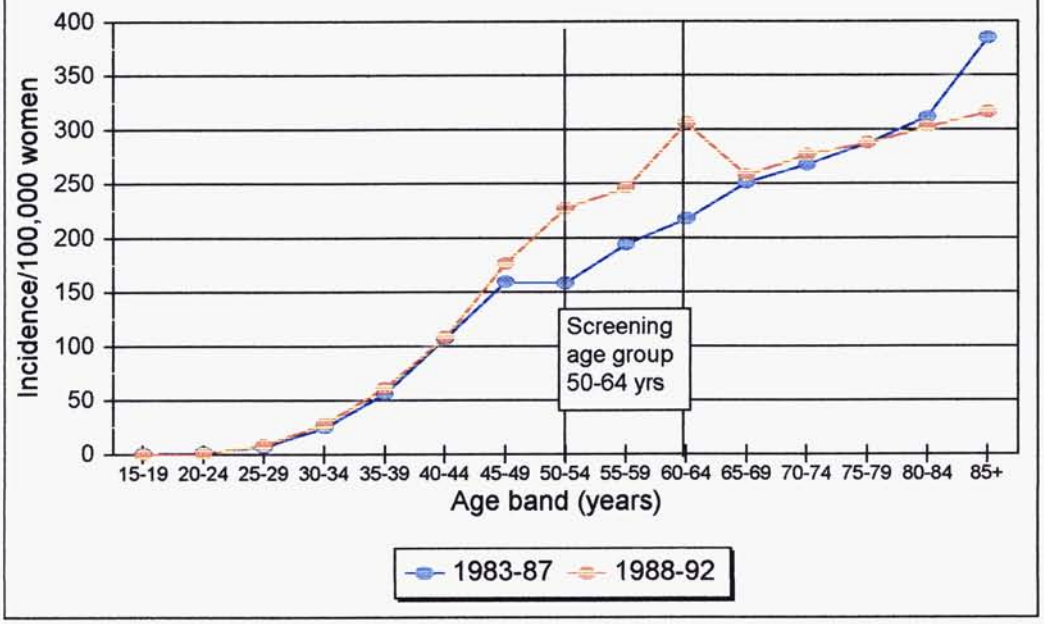
The results of two studies examining breast cancer in the West Midlands are then presented. Patient and tumour characteristics for breast cancers registered at the WMCIU which were diagnosed in 1983 and 1987 are described. The prognostic characteristics of breast cancers in relation to their screening history are then described in order to assess the likely impact of screening on breast cancer in the West Midlands.

4.2 AN OVERVIEW OF BREAST CANCER INCIDENCE, MORTALITY AND SURVIVAL IN THE WEST MIDLANDS

4.2.1 Incidence

The West Midlands had the second highest incidence of 14 Regional Health Authorities (RHAs) in 1988 with a Standardised Incidence Ratio (SIR) of 107 in comparison to the average for England and Wales. Between 1983 and 1987, in the five years before the introduction of screening, there were on average 2640 female invasive breast cancers diagnosed each year in the region. Over this period 91 male invasive breast cancers were diagnosed, an average of 18 per year. In addition to female invasive cancers, a total of 134 female *in situ* cancers were diagnosed over this period, an average of 27 per year. The number of female *in situ* cancers diagnosed each year was thus little over 1% of the female invasive total, demonstrating the low incidence of symptomatic *in situ* disease prior to the introduction of screening. The risk of breast cancer rose with age, demonstrated by the age specific incidence rates in Figure 4.1.

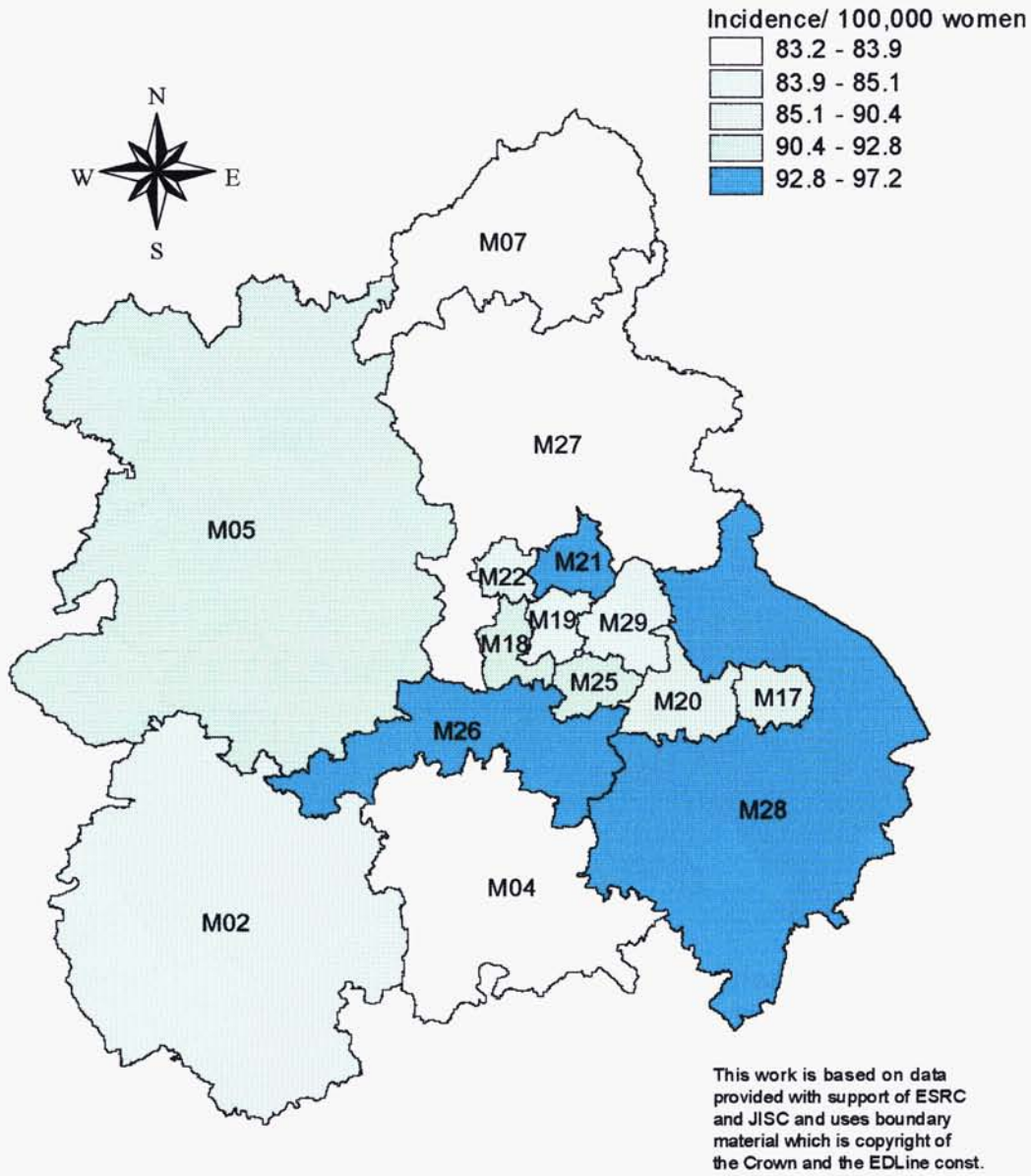
Figure 4.1 **Age specific incidence of breast cancer in the West Midlands in 1983-87 and 1988-92**



The NHSBSP is expected to produce a rise in incidence in those screened. Comparison of age specific rates in 1983-87, before the introduction of screening in the region, and 1988-92, encompassing the introductory period of 1988-91, shows a clear divergence for ages 50 to 64, but little difference for other ages. Rates also diverged for the oldest group (85+ years), but the population is small in this group, and the numbers of cancers diagnosed were similar at 163 cases/year in 1983-87 to 154 cases/year in 1988-92.

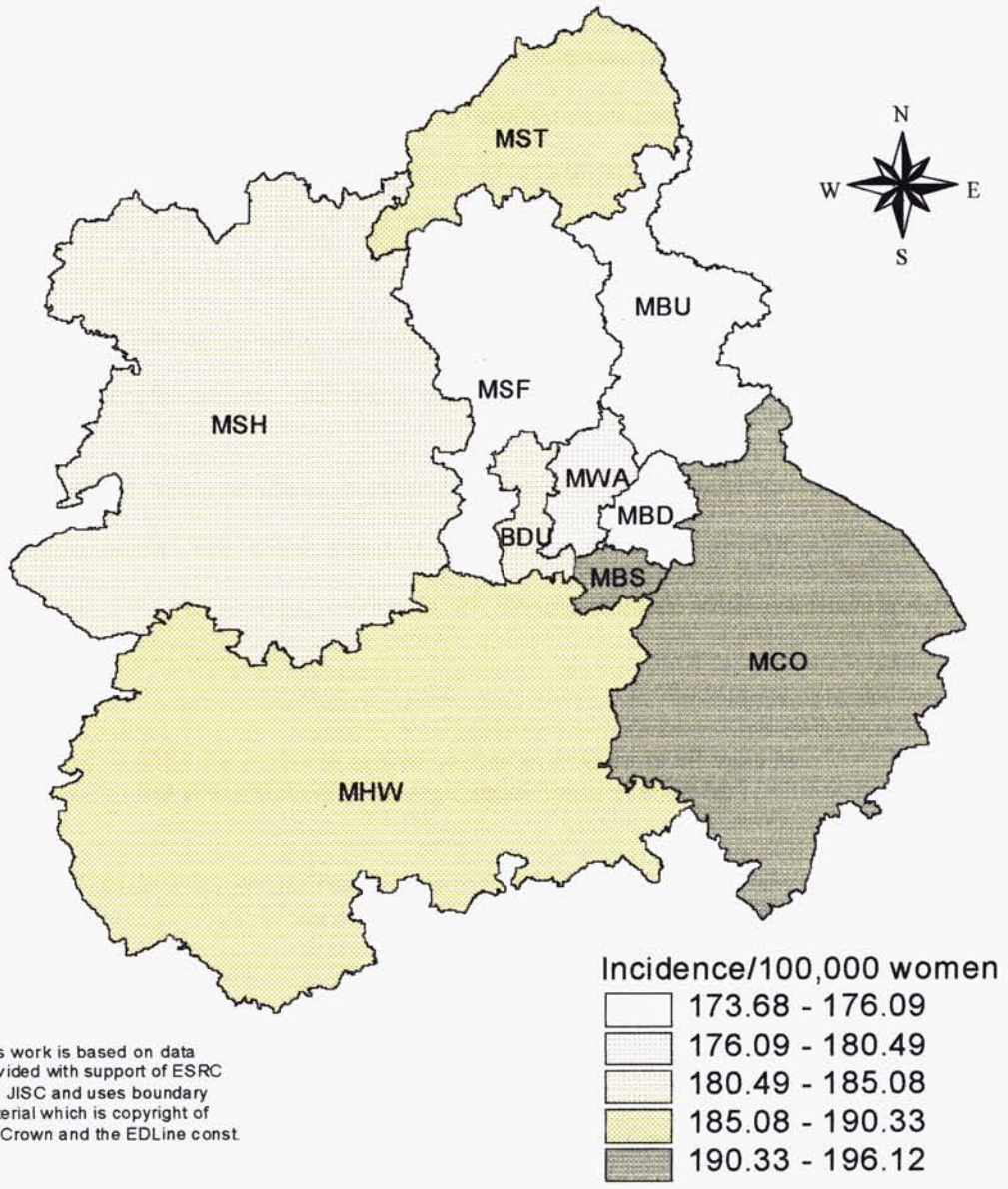
The variation in directly age standardised incidence rates across the West Midlands for 1983-87 for women of all ages is shown in Figure 4.2. There was significant variation in incidence across the region. 7 DHAs (Hereford (M02), Worcester (M04), North Staffordshire (M07), Sandwell (M19), Wolverhampton (M22), and North Birmingham (M29)), had incidence rates significantly lower than the regional average of 89.1/100,000, and 2 DHAs (North Worcestershire (M26) and Warwickshire (M28)) had rates significantly higher than the regional average.

Figure 4.2
Age standardised incidence of invasive breast cancer in women of all ages in the West Midlands in 1983-87



Prepared with the assistance of the West Midlands Cancer Intelligence Unit GIS Project

Figure 4.3
Age standardised incidence of invasive breast cancer in women aged 50-64 years in West Midlands NHSBSP catchment areas in 1983-87



This work is based on data provided with support of ESRC and JISC and uses boundary material which is copyright of the Crown and the EDLine const.

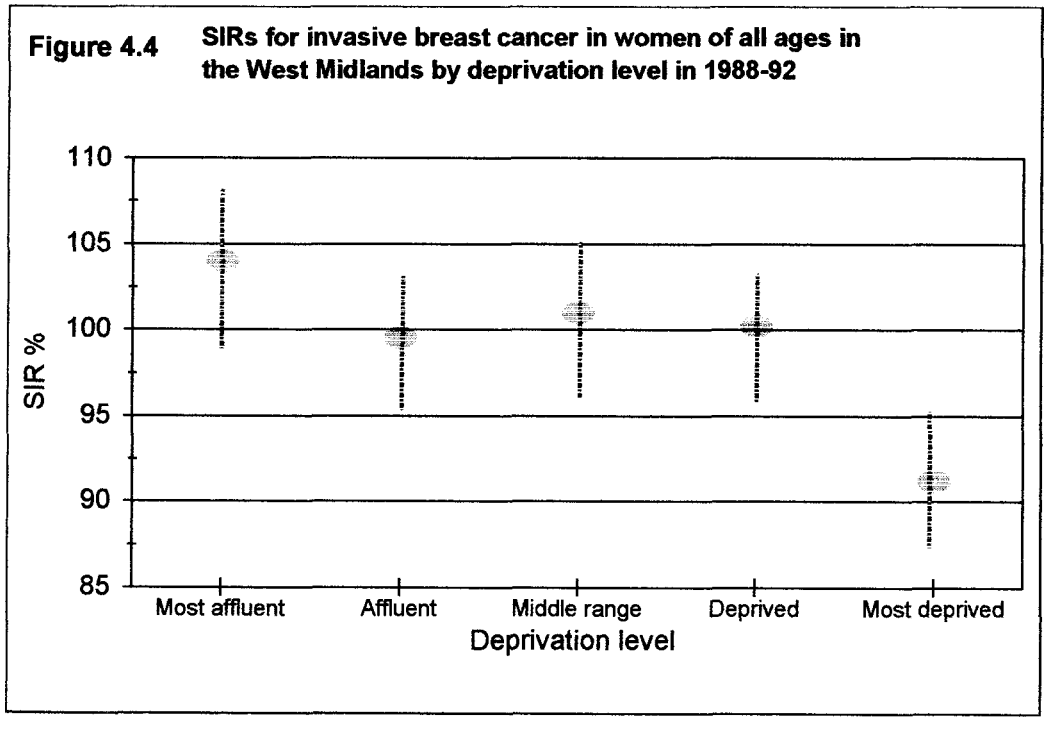
Prepared with the assistance of the West Midlands Cancer Intelligence Unit GIS Project

The NHSBSP operates in the West Midlands from 10 screening units with catchment areas defined by the boundaries of the 22 DHAs in existence in 1987. Figure 4.3 shows age standardised incidence rates in the screening age group (50-64 years) for the

catchment areas before screening began, In Figure 4.3 each catchment area is labelled with the unit code allocated by the Department of Health. Incidence rates varied from 173.7/100,000 in North Birmingham (MBD), an area with a largely inner city population, to 196.1/100,000 in South Birmingham **(MBS),** an area encompassing both deprived and affluent districts. Thus the catchment area of the North Birmingham screening service had a pre-screening incidence amongst women aged 50-64 years 7.2% below the regional average of 186.2/100,000, whereas that of the South Birmingham service was 5.1% above the regional rate. Differences in the baseline incidence rates are important because after the introduction of screening such variations could influence the cancer detection rates achieved by different screening units, with those in a low incidence areas likely to achieve lower detection rates than those in high baseline incidence areas. Differentiating between this effect and the influence of screening sensitivity, which also affects detection rates, is difficult.

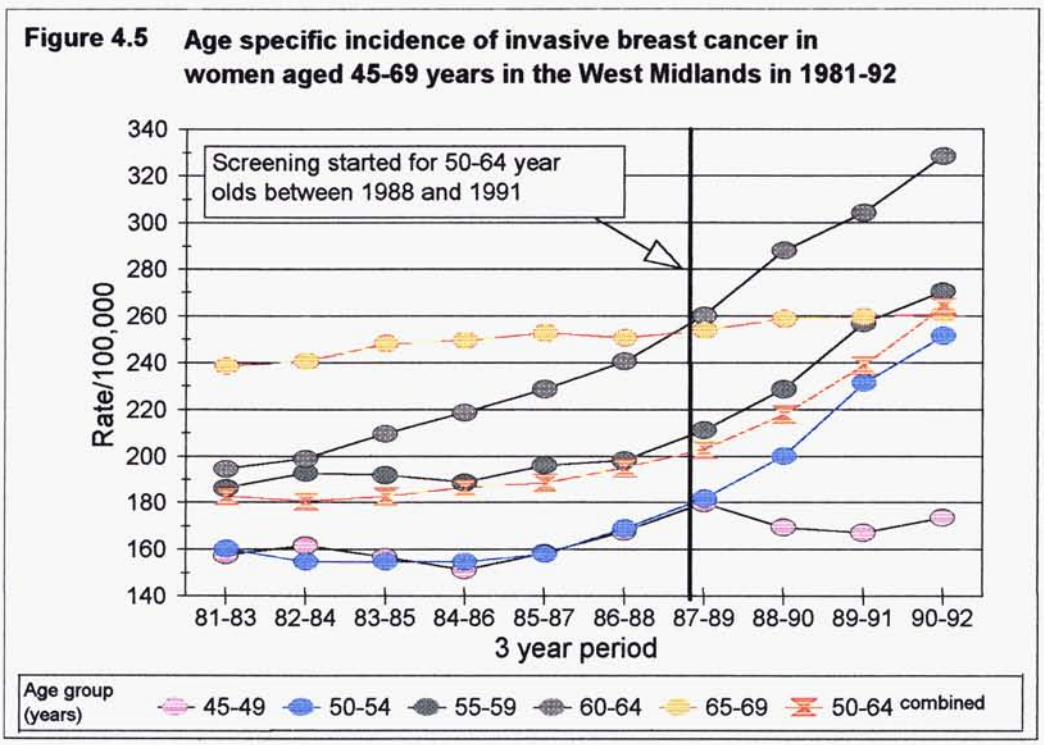
Figure **4.4** shows **SIRs** for breast cancer in the West Midlands according to level of material deprivation. The clear trend in incidence, with higher **SIRs** in more affluent areas, reflects the increasing incidence of breast cancer with increasing affluence reported elsewhere (Leon, 1988).

Figure 4.4 **SIRs for invasive breast cancer in women of all ages in the West Midlands by deprivation level in 1988-92**



Age specific incidence rates for the period 1981-92 are shown in Figure 4.5. Three year rolling rates were calculated to reduce the impact of fluctuations between years. As well as the clear rise in incidence with age, some time trends are apparent. In the period 1981-87, prior to the introduction of screening in the region, incidence rates in 50-64 year olds rose by a total of 14.4% from 180.1/100,000 in 1983 to 206.4/100,000 in 1987, an average rise of 3.6% per year. After the introduction of screening in 1988 the expected rise in the 50-64 year group is seen, with rates in this group diverging from the lower rates in women aged 45-49 years, however, who are not eligible for screening. The pre-screening rise in incidence amongst 50-64 year olds may have contributed to the post-screening rise, complicating evaluation of screening, which is made more difficult by the staggered start of screening in the region.

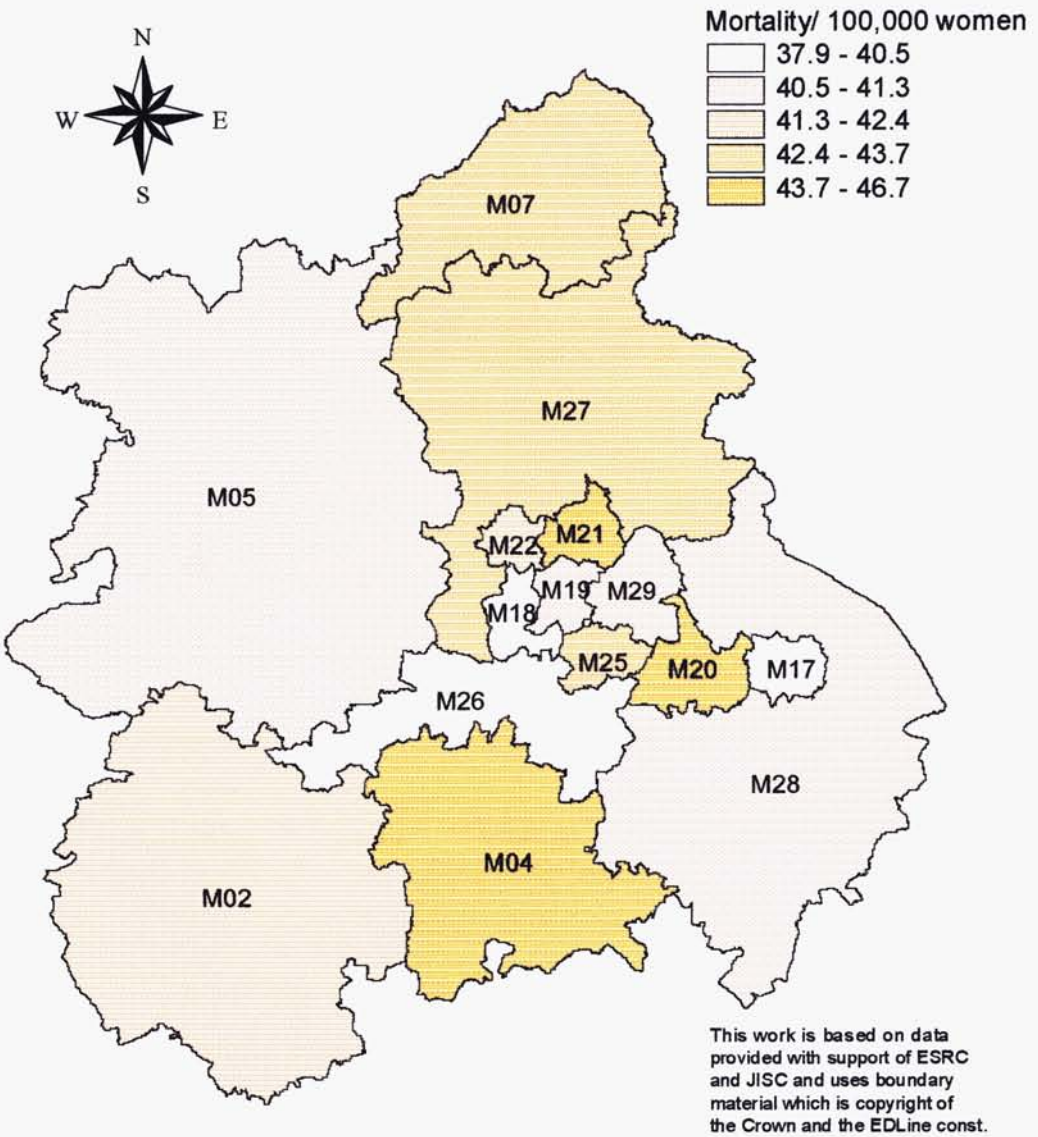
Figure 4.5 **Age specific incidence of invasive breast cancer in women aged 45-69 years in the West Midlands in 1981-92**



4.2.2 Mortality

From 1989-91 there were 1700 deaths from breast cancer each year in the West Midlands. The region had relativity high mortality compared to elsewhere, with a Standardised Mortality Ratio (SMR) of 105 in 1989 in comparison to the England and Wales average. Mortality rates for 1988-92 across the region are shown in Figure 4.6.

Figure 4.6
Age standardised mortality from invasive breast cancer in women of all ages in the West Midlands in 1988-92



Prepared with the assistance of the West Midlands Cancer Intelligence Unit GIS Project

Although mortality rates varied from 37.9/100,000 in Coventry (M17) to 46.7/100,000 in Walsall (M21), no district was significantly different from the regional average. These inter-district variations are also apparent in the 50-69 year age group and have an impact on the mortality drop required to meet the Health of the Nation (HoN) target. The percentage drop required by each DHA to achieve the target rate of 70.1/100,000 varied

from 36% in Worcester (M04) to 17% in Warwickshire *(M28)*. A drop of 36% would be impractical to achieve in the time available, while a drop of over 16% should be possible in Warwickshire. A more feasible approach to achieving the HoN target is being taken by the West Midlands NHS Executive. The 25% reduction is being applied regionally, *so* that each DHA must achieve a 25% reduction, including those with relatively low baseline mortality (Baker, personal communication). Unlike incidence, mortality showed no clear trend according to level of deprivation, shown in the *SMRs* in Figure 4.7. This lack of association suggests that, although affluent women were more likely to develop the disease, this group had better survival than deprived women, resulting in the similar mortality across deprivation levels.

Figure 4.7 **SMRs for invasive breast cancer in women of all ages in the West Midlands by deprivation level in 1998-92**

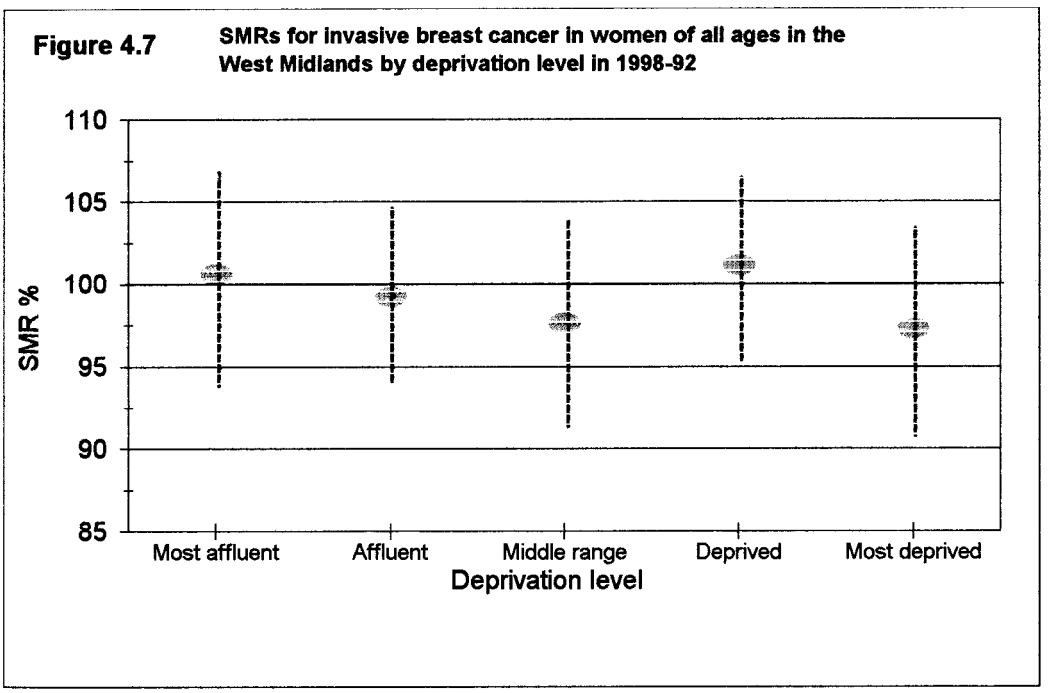
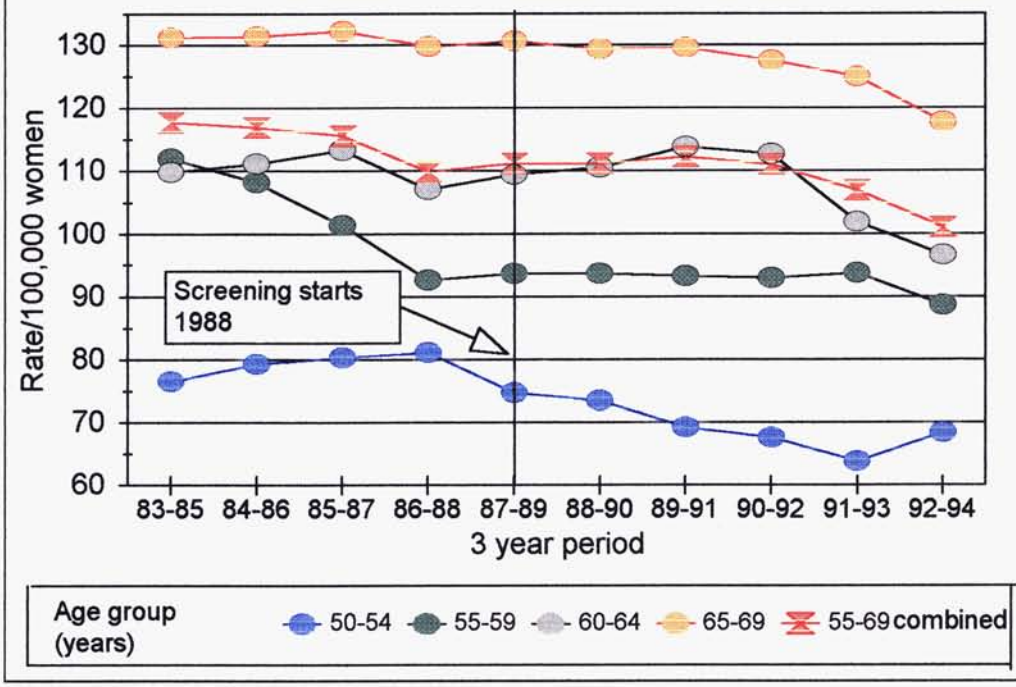


Figure 4.8 demonstrates a general fall in mortality from 1983 to 1994. From 1983-87 mortality fell, particularly in 55-59 year olds, although mortality in 50-54 year olds rose slightly. From 1990 onwards a downward trend in mortality for women aged 55-69 can be seen. The 55-69 year age group is considered suitable for monitoring changes in mortality in relation to the NHSBSP since it is in this group that any effect would be expected to be apparent (Quinn and Allen, 1995). From 1983-87, prior to the introduction of screening, mortality in this group fell by 4.9%. A further drop of 8.9% was identified between 1988, which marked the start of a staggered introduction of

screening in the region, and 1994. As the NHSBSP was not fully operational until 1991, these recent reductions in mortality are unlikely to be due to screening, being more likely to reflect improvements in treatment. It is difficult to evaluate the significance of trends from data spanning a relatively short time, and monitoring over the next decades is necessary to identify ongoing changes in mortality.

Figure 4.8 **Age specific mortality from invasive breast cancer in women aged 50-69 years in the West Midlands in 1983-94**



4.2.3 Survival

Table 4.1 presents five year relative survival rates which were calculated using similar techniques from five UK cancer registries.

Table 4.1 Five year relative survival rates for breast cancer in selected UK regions

Area	Period	Survival %	Source
Yorkshire	84 - 88	63	Yorkshire Cancer Organisation 1995
Scotland	83 - 87	64.3	Information and Statistics Division, NHS in Scotland, 1993
West Midlands	84 - 88	65.5	West Midlands RHA and WMCIU 1996
North West	84 - 88	64.9	Manchester Centre for Cancer Epidemiology, 1992
South West	84 - 88	67.8	Bristol Cancer Epidemiology Unit, 1995
South East	86 - 88	69	Thames Cancer Registry, 1994

Five year survival rates varied from **63%** in Yorkshire to **69%** in the South East, with the West Midlands taking a middle position at 65.5%. These rates are not adjusted for age, and thus differences could reflect variations in age distribution.

Figure **4.9** shows five year relative survival rates for women in the West Midlands diagnosed with invasive breast cancer between **1984** and **1988** according to age at diagnosis. The overall pattern indicates worsening survival with increasing age. The highest survival rate of **72%** was observed amongst **45 - 49** year olds, although younger women fared worse than this. Amongst women aged 50 and **74** years survival showed little change with age, with survival rates worsening after age **74** years.

Figure 4.9 **Relative 5 year survival by age for women with invasive breast cancer diagnosed in 1984-88 in the West Midlands**

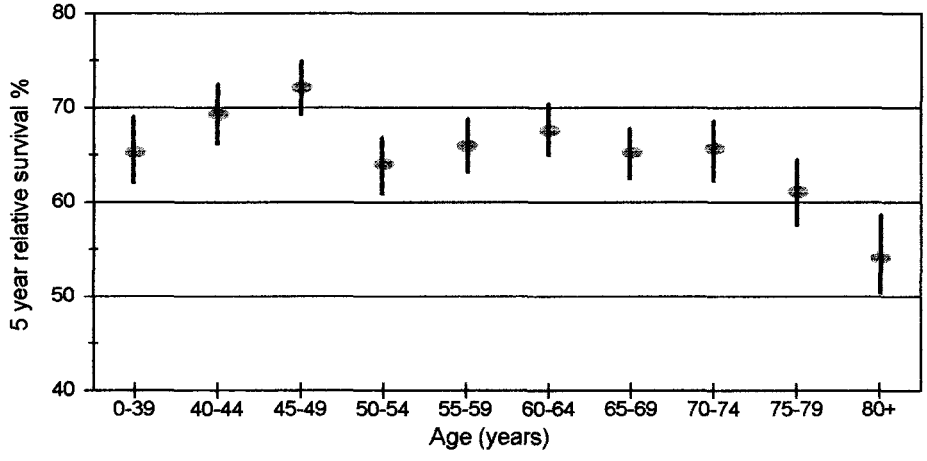
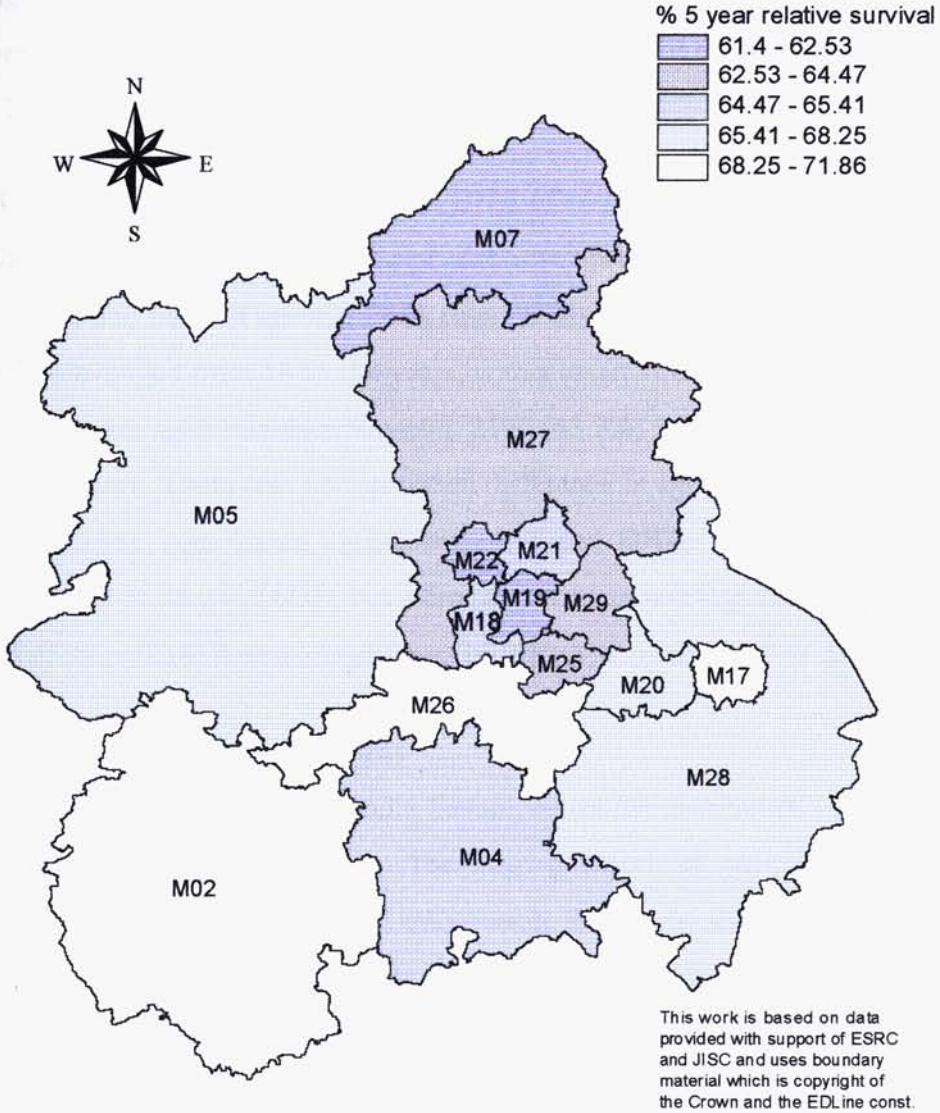


Figure **4.10** shows five year relative survival rates for women diagnosed with breast cancer in the West Midlands between **1984** and **1988.** Survival rates varied from **61.4%** in North Staffordshire **(M07)** to **71.9%** in Coventry **(M17).** North Staffordshire and Sandwell **(M19)** displayed survival rates significantly lower than the regional rate of ***65.5%,*** while Coventry, North Worcestershire *(M26)* and Hereford **(M02)** displayed showed rates significantly higher than the regional average. These inequalities in survival are likely to reflect variation in factors including stage, the biological characteristics of tumours, treatment factors, and characteristics such as diet and lifestyle. Stage is a reflection of the earliness of diagnosis, which may differ across the region, as knowledge and understanding **of** breast cancer, and availability of diagnostic facilities differ.

Figure 4.10
Five year relative survival for women of all ages with invasive breast cancer diagnosed in 1984-88 in the West Midlands



There is some evidence that survival may be improving in the UK. In the West Midlands five year relative survival rates rose from 64.5% for 1981-85 to 65.5% for 1984-88. In Yorkshire rates rose from under 60% in 1976-80 to 67% for 1986-88. Survival in Scotland has risen from 56.4% in 1968-72 to 64.3% for 1983-87. These changes could reflect changes in treatment, or earlier symptomatic diagnosis.

4.3 THE CHARACTERISTICS OF BREAST CANCER IN THE WEST MIDLANDS

Prior to analysis 2.7% of the 2511 cases diagnosed in 1983, and 8.4% of the 2907 cases diagnosed in 1987 were excluded, leaving 2444 and 2663 cases respectively for analysis from the two years. Tables detailing the characteristic of these cases are presented in full in Tables A7.1 to A7.11 in Appendix Seven.

4.3.1 Age at diagnosis and DHA of residence

Tables A7.1, A7.2 and A7.3 in Appendix Seven present the distribution of cases by age, by DHA, and by both age and DHA respectively. In the overall study population, age at diagnosis ranged from 18 to 104 years, with a mean of 62.4 years. Overall 1129 (22.2%) of cases were aged less than 50 years, 2194 (43.2%) 50-69 years, and 1754 (34.5%) 70 years or more. There was no significant difference in age between years (Chi square 3.4 df 2 p=0.18). Age varied significantly between DHAs for the years combined (Chi square 61.9 df 28 p=0.0002), but although variation was significant in 1983 (Chi square 46.4 df 28 p=0.16) it was insignificant in 1987 (Chi square 36.9 df 28 **p=0.**12).

4.3.2 Stage at diagnosis

Tables A7.4 and A7.5 in Appendix Seven present the distribution of cases by stage, and by stage and age respectively. The staging algorithm described in Chapter Three allocated stage to 2841 cases (56.0%) overall. Stage was allocated to significantly more cases diagnosed in 1983 than in 1987 (63.9% vs. 48.6%, Chi square 121.0 df 1 p<0.0001). The lower proportion staged in 1987 reflected lack of local node data, caused partly by methods used to abstract data on the PDP11 database prior to transfer onto GRACE. Lack of stage was most common in the youngest cases of which 50.8% were unstaged, with the lowest unstaged proportion in those aged 50-69 years (39.0%). The distribution of cases by stage is shown in Figure 4.11. In the total study population 27 cases (0.5%) were stage '0, 494 (9.7%) stage '1', 678 (13.4%) stage '2a', 456 (9.0%) stage '2b', 788 (15.5%) stage '3' and 398 (7.8%) stage *'4'*, with 2236 remaining unstaged.

Figure 4.11 **Distribution of breast cancers diagnosed in 1983 and 1987 by stage**

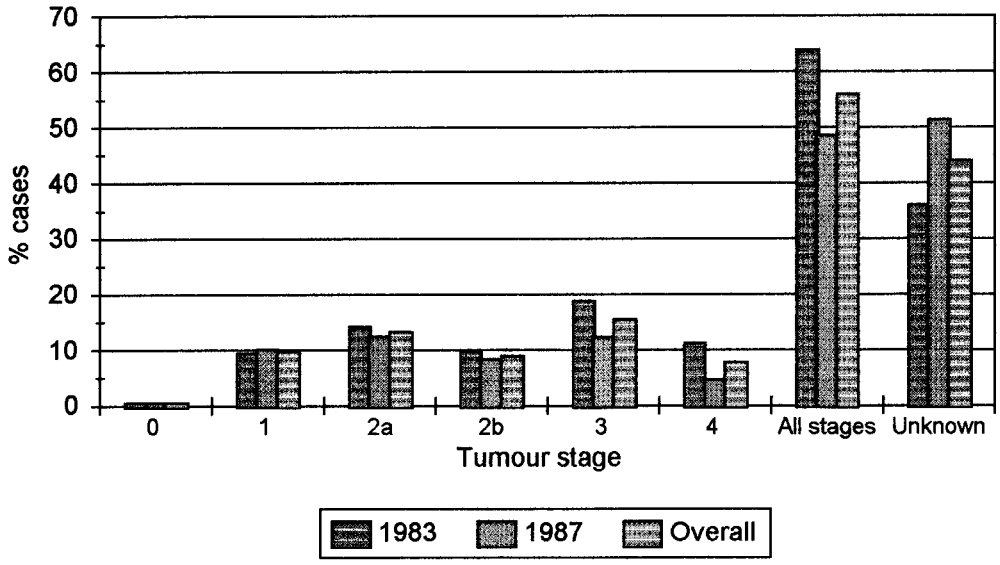
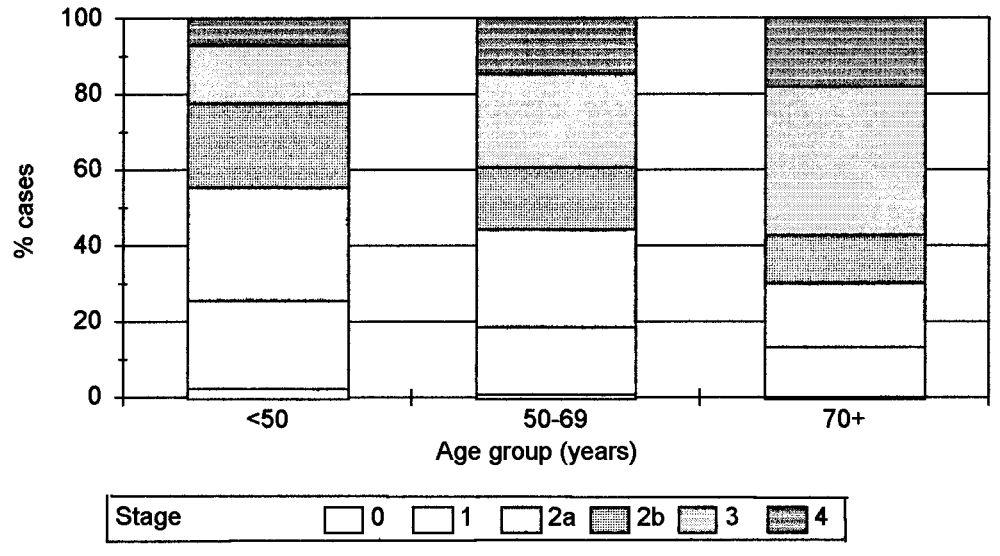


Figure 4.12 **Distribution of breast cancers diagnosed in 1983 and 1987 by stage and age**



Amongst the 2841 staged cases the stage distribution became more advanced with increasing age, shown in Figure 4.12. For the combined study population, of 947 staged cases in women aged 70 or more, 172 (18.2%) were stage '4' compared to 38 of 556 cases (6.8%) in women aged less than 50 years, with a significant association between increasing age and poorer stage (Chi square 105.8 df 4 $p<0.0001$).

4.3.3 Tumour grade

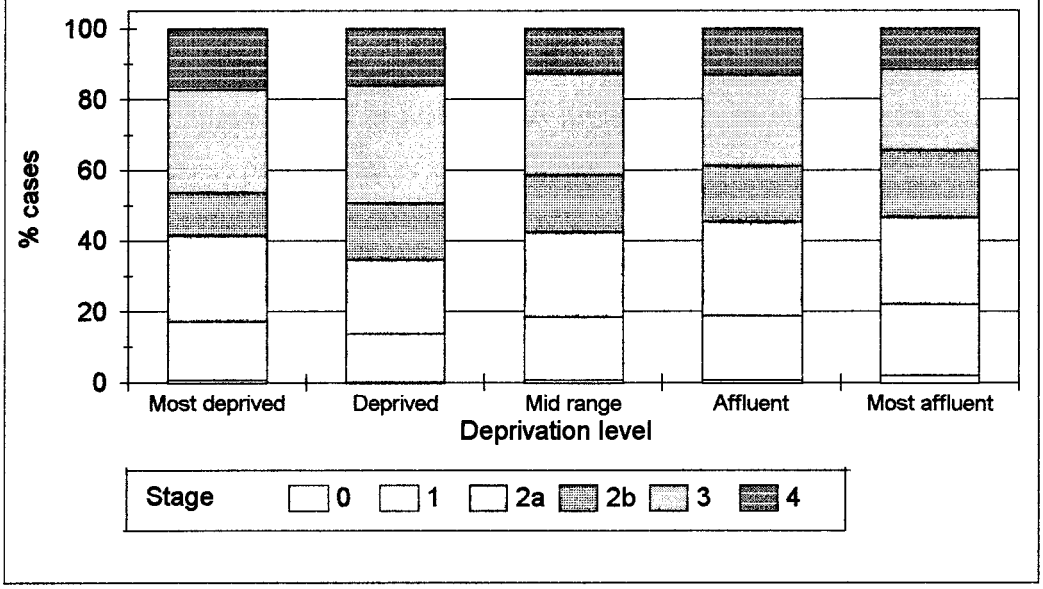
Tables A7.6 and A7.7 in Appendix Seven present the distribution of cases by grade, and by grade and age respectively. Overall only 1723 cases (33.9%) had grade recorded, with a significantly higher proportion graded in 1987 than in 1983 (30.5% vs. 37.1% Chi square 25.1 df 1 p<0.0001). Of graded cases, a higher proportion were recorded with a "good" grade (grade I or II) in 1987 (39.1%) than in 1983 (23.8%). "Good" grade cases were older than "poor" grade (Chi square 8.12 df 2 p=0.017), with 31.1% of the 559 "good" grade cases being over 70 years old compared to 25.3% of the 1164 "poor" grade cases. Whether these variations with age and year reflected genuine trends, changes in the selection of cases undergoing pathological investigations, or selective reporting of data to the WMCIU was not investigated.

4.3.4 Deprivation level

Tables A7.8, A7.9 and A7.10 in Appendix Seven present the distribution of cases by deprivation level, by deprivation level and age, and by deprivation level and stage respectively. Significantly more cases diagnosed in 1987 fell into the more affluent categories (Mantel Haenzel Chi square 11.6 df 1 p=0.0007) compared to those diagnosed in 1983. The proportion of younger women rose with increasing affluence (Chi square 63.4 df 8 p<0.0001). Of the 2052 women in the most affluent or affluent areas, 548 (26.7%) were aged less than 50 years, whereas of the 1896 women in deprived or very deprived areas, only 348 (18.4%) were in this age group (Chi square 39.2, df 1 p<0.0001).

Stage distribution worsened with increasing deprivation, shown in Figure 4.13. Of the 1153 staged cases in affluent or very affluent women, only 141 cancers (12.2%) were stage 'IV, compared to 179 of the 1088 cancers in deprived or very deprived women (Chi square 8.2, df 1 p=0.004). The reverse was true of early stage cancers, with, of the 1153 staged cases in affluent or very affluent women, 223 cancers (19.3%) being stage ‘I’, compared to only 165 (15.2%) of the 1088 cancers in deprived or very deprived women (Chi square 6.8, df 1 p=0.009).

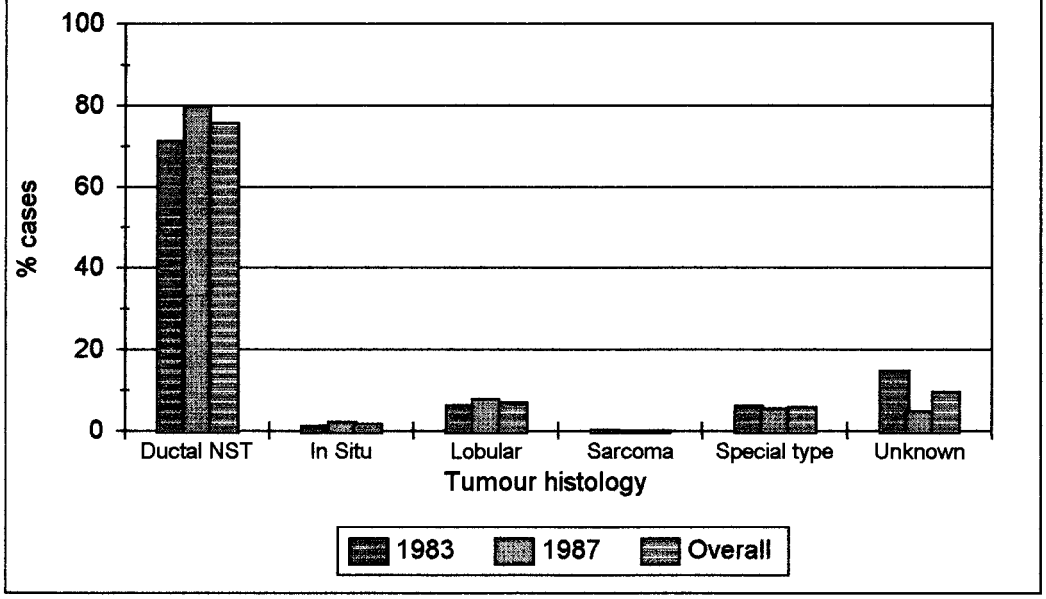
Figure 4.13 **Distribution of cases diagnosed in 1983 and 1987 by stage and deprivation level**



4.3.5 Histological group

The distribution of cases by histological group is shown in Figure 4.14 and is also described in Table A7.11 of Appendix Seven.

Figure 4.14 **Distribution of breast cancers diagnosed in 1983 and 1987 by histological group**



Significane more of the 5077 cases studied had no recorded histological confirmation recorded at the WMCIU in 1983 than 1987, with 362 cases without histology in 1983 compared to 126 cases in 1987 (14.8% vs 4.8%, Chi square 146.7 df 1 p<0.0001). This fall suggests an improvement in histological investigation or reporting to the WMCIU over this period. The proportions of cases without histology by age and deprivation level in he study group are described in Tables A7.12 and A7.13 in Appendix Seven. The proportion without histology rose with age (Chi square 271.0 df 2 p<0.0001), with 18.8% of the oldest and 2.2% of cancers in women under 50 years having no histology. Lack of histological confirmation rose with deprivation (Mantel Haenzel Chi square 36.0 df 1 p<0.0001), with 11.4% of the most deprived and 7.1% in the most affluent being without recorded histology. The overall level of histological investigation was good with 85.2% of breast cancers diagnosed in 1983 and 95.2% of cases diagnosed in 1987 having histological type recorded on the WMCIU database.

Of cases with histological confirmation a total of 86 *in situ* cancers and 9 sarcomas were recorded. Of the 4503 invasive carcinomas, 3841 (85.3%) were ductal carcinomas of no special type (NST), 358 (8.0%) lobular carcinoma and 295 (6.6%) special type carcinomas, with similar distributions between years (Chi square 4.3 df 2 p=0.16).

4.4 PROGNOSTIC CHARACTERISTICS OF CANCERS ACCORDING TO SCREENING HISTORY

A total of 755 breast cancers arising post-screening and 788 pre-screening controls diagnosed in 1987 were examined. The post-screening cancers were divided into prevalent and incident screen detected cancers, cancers in non-attenders, and interval cancers. The interval cancers examined were the "core" set identified by the method described in section 2.4.2 of Chapter Two. Of the cancers examined, 693 of the post-screening cancers and 695 of the pre-screening controls had confirmed invasive histology. The prognostic characteristics of these cases were examined. The relative incompleteness of data for cancers in non-attenders and pre-screening controls, which may have contributed to the apparent differences found, was discussed in Chapter Three.

4.4.1 Invasive status

The invasive status of the cancers analysed is shown in Table 4.2.

Table 4.2 Invasive status of breast cancers according to screening history

Invasive status	SD* (prevalent)		Core interval		NA**		SD (incident)		Control	
	No.	%	No.	%	No.	%	No.	%	No.	%
Invasive	269	93.08	213	95.09	78	90.7	133	85.26	695	88.2
*In situ***	20	6.92	11	4.91	2	2.33	22	14.1	17	2.16
Unknown	0	0	0	0	6	6.98	1	0.64	77	9.77
Total	289	100	224	100	86	100	156	100	789	100

* SD screen detected
** NA non attender
*** Including Paget's disease without underlying invasive carcinoma

Only 2.2% of the pre-screening control group were *in situ*. The *in situ* proportion in non-attenders was similar at 2.3% (Chi square 0.08 p=0.8). Interval cancers had a higher *in situ* proportion than controls (Chi square 3.7 p=0.06), but this difference was not significant. 6.9% of prevalent screen detected cancers were *in situ*, significantly more than controls (Chi square 11.9 p=0.0006). 14.1% of incident screen detected cancers were *in situ*, significantly higher than controls (Chi square 37.5 p<0.0001).

4.4.2 Prognostic characteristics

The histological type, size, nodal status and grade of cases is shown in Table 4.3.

4.4.2.1 Histological type

Although rates of lobular cancer varied from 6.0% in prevalent screen detected cancers to 11.7% in interval cancers, no post-screening group had a proportions of lobular carcinoma significantly different from pre-screening controls (all Chi square tests with Yates correction p>0.05). The proportion of special type cancers varied between post-screening groups. The rate amongst interval cancers (4.7%) was similar to the control group (4.8%), (Chi square 0.01 p=0.91) but prevalent and incident screen detected cancers had significantly more special type cancers, with rates of 12.6% (Chi square 17.7 p<0.0001) and 14.3% (Chi square 15.7 p<0.0001) respectively. Only 1.3% of cancers in non-attenders were special type, but this was not significantly less than the control group (Fisher's exact test 2.5 p=0.24). These variations could reflect difference in pathological expertise or data quality rather than real differences in histology.

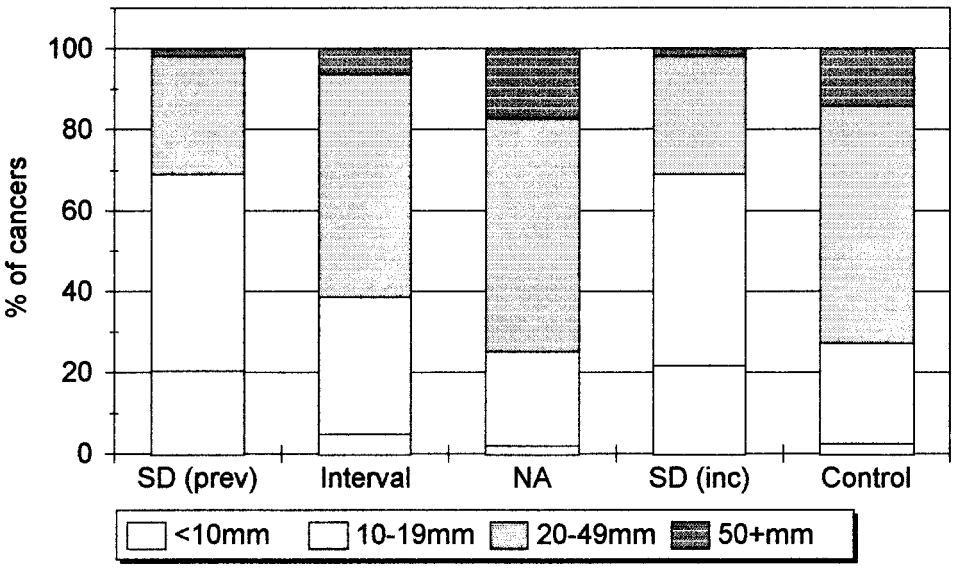
Table 4.3 Prognostic characteristics of invasive breast cancers according to screening history										
Prognostic feature	Post-screening cancers								Pre-screening control group	
	SD (prevalent)		Interval		NA		SD (incident)			
	No.	%	No.	%	No.	%	No.	%	No.	%
Tumour histology										
Ductal carcinoma NST	219	81.41	178	83.57	71	91.03	108	81.19	593	79.38
Lobular carcinoma	16	5.95	25	11.74	6	7.69	8	6.02	66	8.84
Special type carcinoma	34	12.64	10	4.69	1	1.28	19	14.29	36	4.82
Total	269	100	213	100	78	100	133	100	695	100
Size (mm)										
<10	53	19.7	10	4.69	1	1.27	27	20.3	12	1.73
10 -19.9	124	46.1	65	30.52	11	13.92	59	44.36	123	17.7
20 - 49.9	75	27.88	107	50.23	27	34.18	36	27.07	289	41.58
>50	4	1.49	12	5.63	8	10.13	2	1.50	68	9.78
Unavailable/ invalid	13	4.83	19	8.92	31	40.51	9	6.77	203	29.21
Total	269	100	213	100	78	100	133	100	695	100
Local nodes/ metastases										
Negative	77	28.62	71	33.33	18	23.08	71	53.38	258	34.54
Positive	38	14.13	52	24.41	23	29.49	27	20.3	217	29.05
Distant metastases	3	1.12	4	1.88	7	8.97	0	0.00	38	5.09
Unavailable	151	56.13	86	40.38	30	38.46	35	26.32	182	31.32
Total	269	100	213	100	78	100	133	100	695	100
Grade										
Good (I or II)	183	68.03	96	45.07	20	25.64	94	70.68	146	21.01
Poor (III)	42	15.61	67	31.46	23	29.49	28	21.05	198	28.49
Unavailable	44	16.36	50	23.47	35	44.87	11	8.27	351	50.5
Total	269	100	213	100	78	100	133	100	695	100

4.4.2.2 Tumour size

The distribution of cases by size and screening history is shown in Figure 4.15. Only 27.4% of the pre-screening control cancers were <20mm, compared to 69.1% of prevalent screen detected cancers and 69.4% of incident screen detected cancers. 38.7% of interval cancers were <20mm, and only 27.7% of cancers in non-attenders fell into this category. Compared with pre-screening controls, prevalent and incident screen detected cancers were smaller (Mantel Haenzel (MH) Chi square 144.3 $p<0.0001$, and 94.1 $p<0.0001$ respectively). Interval cancers were also significantly smaller (MH Chi

square 13.8 p<0.0001). Cancers in non attenders had a similar size distribution to the control group (MH Chi square 0.3 p=0.60).

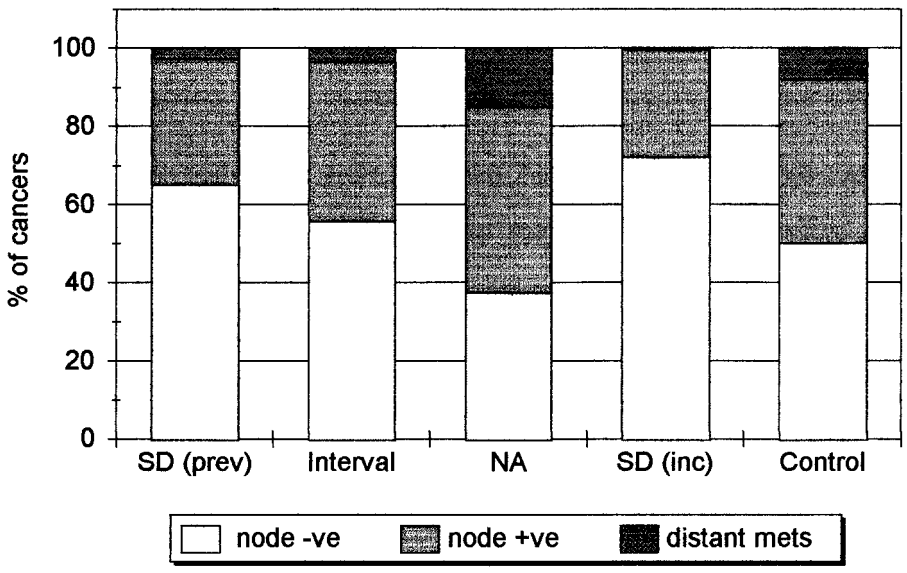
Figure 4.15 **Size of breast cancers according to screening history**



4.4.2.3 Nodal and metastatic status

The distribution of cases by nodal and metastatic status is shown in Figure 4.16.

Figure 4.16 **Nodal status of breast cancers according to screening history**

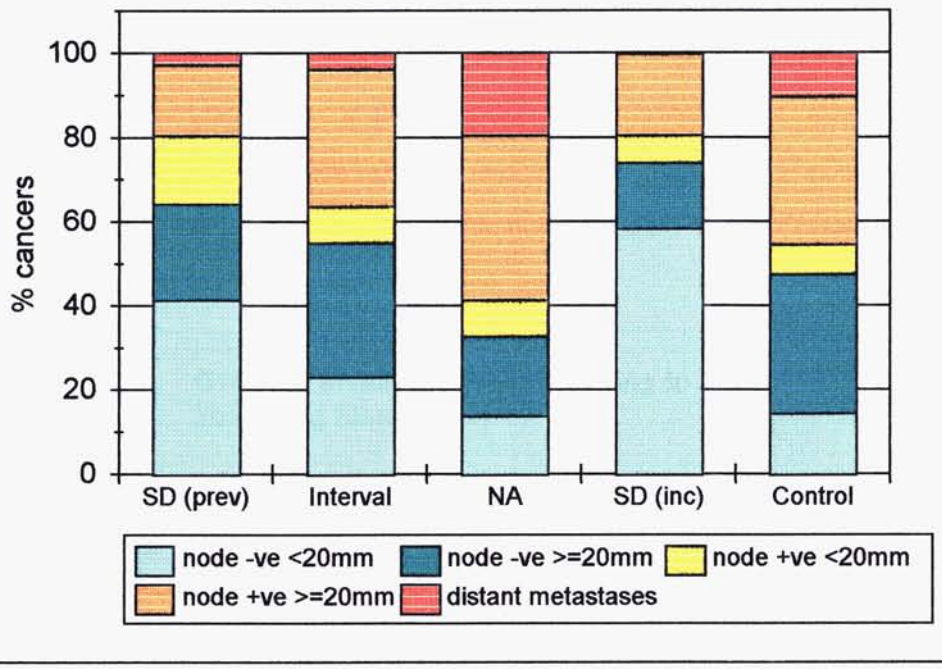


Compared to pre-screening cancers, prevalent and incident screen detected cancers had fewer involved nodes or metastases (MH Chi square 9.9 p=0.002, and 19.2 p<0.0001 respectively). Interval cancers were similar to controls (MH Chi square 2.6 p=0.12), but cancers in non-attenders had poorer characteristics (MH Chi square 4.3 p=0.04).

4.4.2.4 Combined tumour size and nodal and metastatic status

Size, nodal status and metastatic status were combined into a simple TNM-type stage (similar to TNM stage). The distribution of cases by stage and screening history is shown Figure 4.17.

Figure 4.17 **Simplified stage for breast cancers according to screening history**



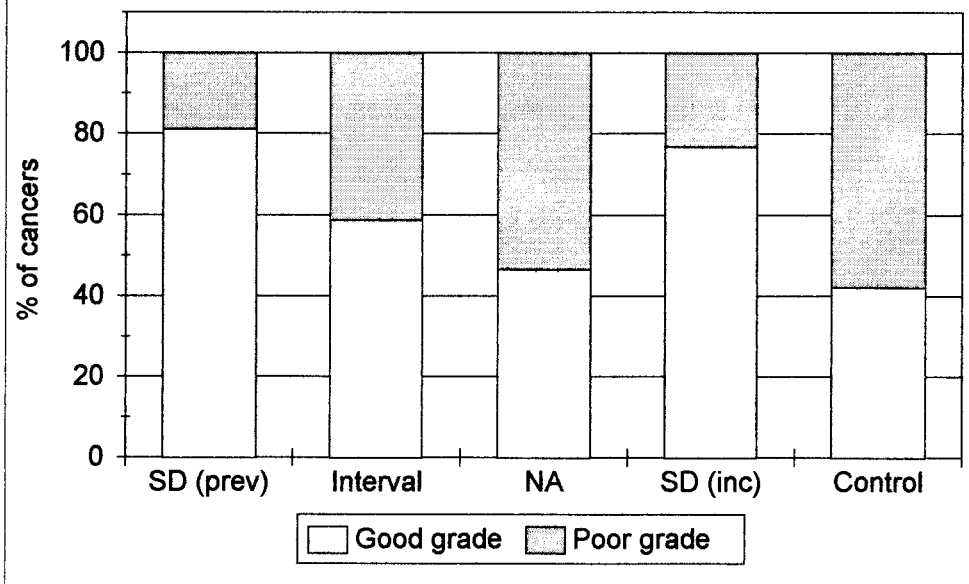
Cancers with negative nodes that were less than 20mm in diameter were classified as the earliest stage, and those with positive distant metastases were classified as the most advanced. At the first screen it is expected that some cancers will be large and slow growing. If these prevalent cancers are excluded from the post-screening group, the remaining cancers in non-attenders, intervals and incident screen detected cancers should all be newly incident, and differ from the control group only in that some of them were diagnosed earlier than others (Duffy *et al,* 1991). The stage distribution of prevalent screen detected cancers was significantly more favourable than the control group (MH

Chi square **33.98** p<0.01), with incident screen detected cancers also having a more favourable distribution (MH Chi square **50.9** p<0.01). The distribution of interval cancers was significantly more favourable than the control group (MH Chi square **4.8 p=0.03).** Cancers in non-attenders displayed a similar stage distribution to the control group (MH Chi square 0.3 p=0.56). The prognosis of all the incident cancers grouped together was also more favourable than the control group, (MH Chi square **25.8** p<0.0001), suggesting a general shift towards earlier stage which should lead to improved survival.

4.4.2.5 Tumour grade

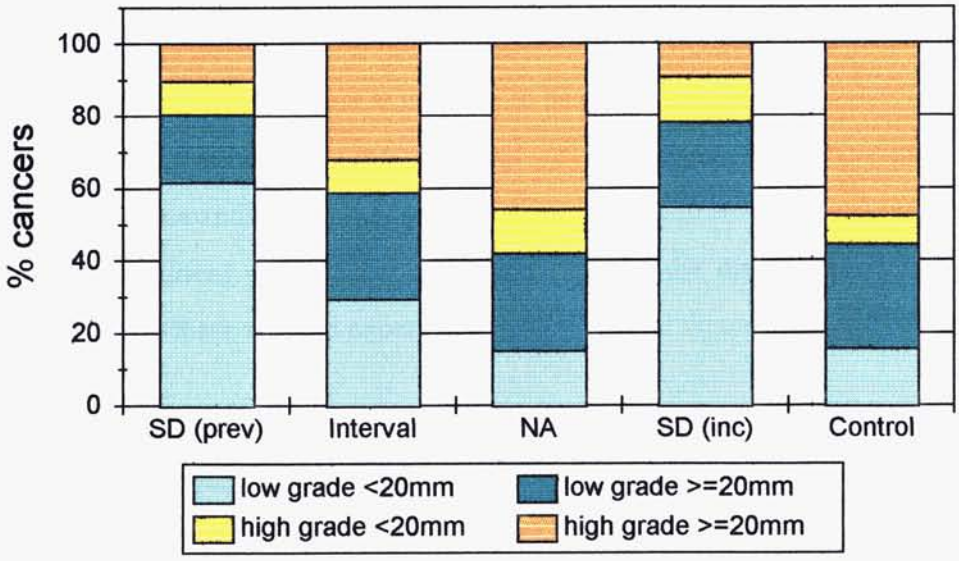
The distribution of cases according to grade and screening history is shown in Figure 4.18. Over 57% of control group cancers were of poor grade. Prevalent and incident screen detected cancers had significantly lower proportions of poor grade cancers, with 18.7% and **23.0%** being poor grade respectively. Although more interval cancers were poor grade (**4** 1.1%) than screen detected cancers, this proportion was significantly lower than the control group. **53.5%** of cancers in non-attenders were of poor grade, similar to the control group.

Figure 4.18 **Grade of breast cancers according to screening history**



Cancers with both size and grade recorded were grouped into four prognostic categories according to their grade and size (<20mm, or 20mm or more), with small, good grade cancers allocated to the best prognostic group, and large, poor grade cancers to the poorest. The resulting distributions are shown in Figure 4.19.

Figure 4.19 **Size and grade of breast cancers according to screening history**



Prevalent screen detected cancers were significantly smaller and of better grade than the control group (MH Chi square 107.9 $p<0.01$), as were incident screen detected cancers (MH Chi square 64.5 $p<0.01$). Although interval cancers were larger and of poorer grade than screen detected cancers, prognosis was more favourable than the control group (MH Chi square 12.7 $p<0.01$). Cancers amongst non-attenders were similar in size and grade distribution to control group cancers (MH Chi square 0.002 $p=0.96$).

4.5 SUMMARY

Patterns of disease across the West Midlands revealed trends and associations which could have an impact on strategies to reduce the impact of breast cancer in the region.

- incidence rates of invasive breast cancer varied significantly across the West Midlands prior to the introduction of screening in 1988, with a regional age standardised incidence rate of 89.1/100,000

- variation in baseline incidence rates between districts could influence the apparent cancer detection rates achieved by screening units, with those in low baseline incidence areas likely to achieve lower detection rates than those in high baseline incidence areas independently of the quality of screening
- incidence rates in the screening age group were rising before the introduction of screening, with rates in 50-64 years olds rising by 3.6% per year from 1983-87
- a pre-screening trend in falling mortality in 55-69 year olds continued after screening began in 1988 but is unlikely to be due to screening
- pre-screening variations in mortality and survival suggest differences in prognostic characteristics and/or treatment across the region
- although affluent women display higher rates of the disease, mortality patterns indicate that these women experience better survival than more deprived groups

Detailed examination of the 5077 cases diagnosed in 1983 and 1987 revealed that

- 56% were staged by the staging algorithm, with lack of local **nodal** status being responsible for most of those remaining unstaged
- 85.2% of cases in 1983 and 95.2% in 1987 had histological type recorded. Significantly more cases had no recorded histology in 1983 than 1987, suggesting an improvement in histological investigation or reporting over this period
- only 1.6% of cancers were *in situ*, and of invasive carcinomas, 85.5% were ductal carcinomas of no special type, 8.0% were lobular carcinomas and 6.6% carcinomas of "special type"
- 22.2% of women were aged *<50* years, 43.2% 50-69 years, and 34.5% 70 years or more, with a higher proportion of cancers in affluent women occurring below the age of 50 years
- there were significant associations between increasing age, increasing deprivation and poorer stage

Analysis of the characteristics of breast cancers diagnosed after the introduction of screening revealed differences in prognosis according to screening history.

- only 2% of pre-screening cancers were *in situ* compared with 6.9% of prevalent and 14.1% of incident screen detected cancers, indicating possible overdiagnosis

- 27.4% of pre-screening cancers were <20mm, compared to 69.1% of prevalent and 69.4% of incident screen detected cancers, 38.7% of interval cancers and 27.7% of cancer amongst non-attenders
- 34.1% of pre-screening cancers had involved local nodes or metastases, compared to 15.3% of prevalent and 20.3% of incident screen detected cancers, 26.3% of interval cancers and 38.4% of cancers in non-attenders
- 57% of pre-screening cancers were of poor (high) grade, compared to 18.7% of prevalent and 23.0% of incident screen detected cancers, 41.1% of interval cancers and 53.5% of cancers in non-attenders
- screen detected cancers had the most favourable prognosis, and cancers in non-attenders the poorest. The prognosis of interval cancers was poorer than for screen detected cancers but more favourable than that of pre-screening cancers

These results suggest that women with screen detected cancer should display better survival than those with breast cancer diagnosed outside the NHSBSP, although lead time bias must be accounted for in any analysis of survival. There is some evidence that overdiagnosis and/or length bias may be occurring in the NHSBSP due to detection of relatively high proportions of *in situ* and special type cancers. The poor prognosis of cancers in non-attenders suggests that women who do not access the NHSBSP are a particularly high risk group.

CHAPTER FIVE

RESULTS: TREATMENT OF AND SURVIVAL FROM BREAST CANCER IN THE WEST MIDLANDS

5.1 INTRODUCTION

In this chapter the findings from an examination **of** treatment data recorded at the WMCIU are presented. Data recorded for breast cancers diagnosed in 1983 and 1987 were examined, with treatment restricted to that given within six months of diagnosis. The data were examined in relation to two issues. Firstly the reliability of routinely recorded treatment data was considered. Treatment patterns were then compared with the King's Fund guidelines (King's Fund, 1986) to identify changes which may have been implemented following the publication of the guidelines.

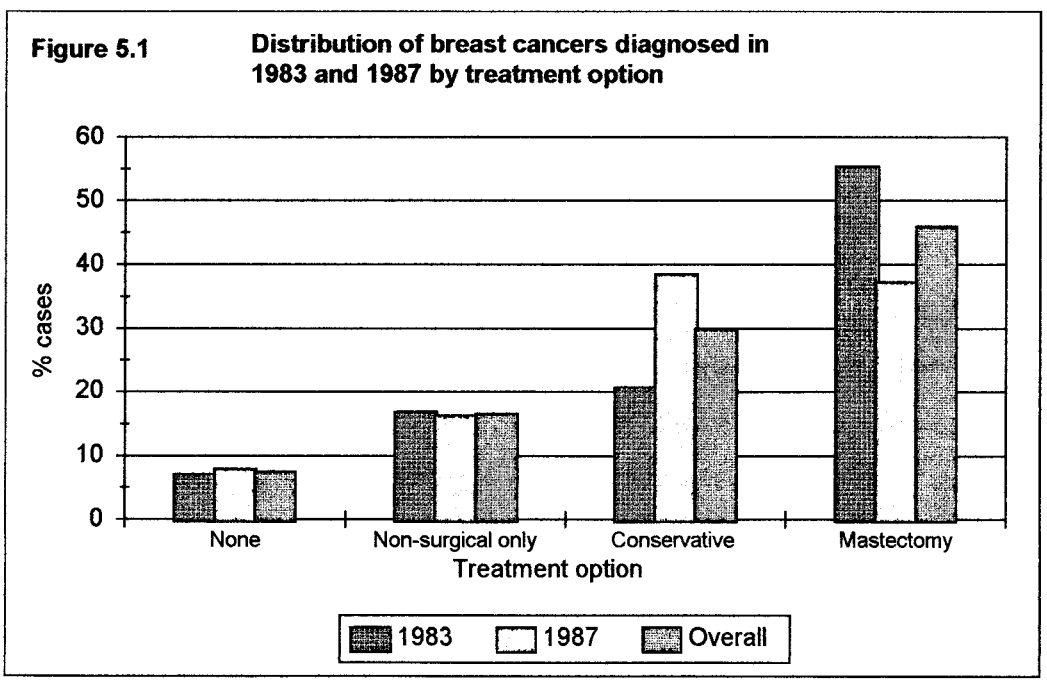
5.2 THE NATURE AND VARIATION OF RECORDED TREATMENT

Treatment data are provided in full in Tables A8.1 to A8.10 in Appendix Eight.

5.2.1 Recorded surgical treatment

Tables A8.1 and A8.2 in Appendix Eight provide the distributions of recorded treatment by year and age. The frequency of recorded treatment is shown in Figure 5.1.

Figure 5.1 **Distribution of breast cancers diagnosed in 1983 and 1987 by treatment option**

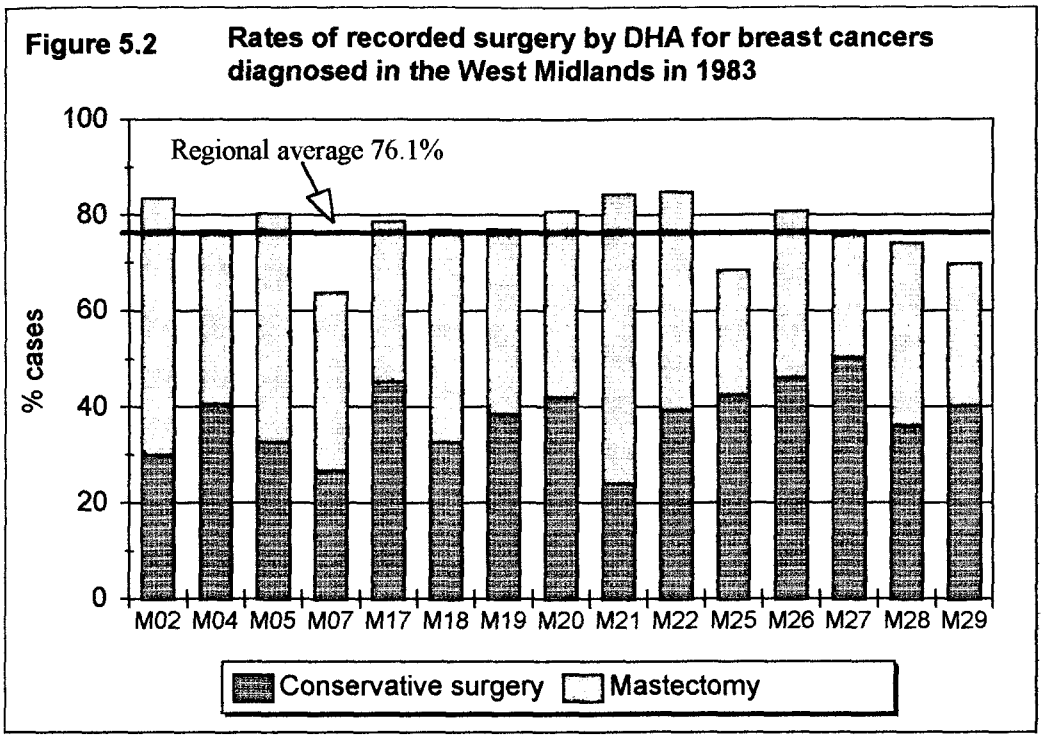


Of the 5077 cases studied only 381 (7.5%) had no treatment data recorded, 841 (16.6%) had non-surgical oncology data only, 1520 (29.9%) had conservative surgery, and 2335 (46.0%) had a mastectomy recorded. No treatment data were recorded for 172 cases (7.0%) in 1983 and 209 cases (7.9%) in 1987. Reasons for lack of treatment were known for 84.3% of the cases in 1983 and 38.8% in 1987. They included poor physical condition, refusal, and presentation with disease too advanced to treat. In the remainder of the cases it was unknown whether treatment had been absent or the data were incomplete. Overall the oldest women were more likely to have no treatment recorded, with 23 of the 1129 aged <50 years (2.0%), and 251 of the 1754 women aged 70 years or more (14.3%) having no treatment recorded.

Use of non-surgical oncology in the absence of surgery changed little over this period, being recorded for 413 cases in 1983 (16.9%) and for 428 cases in 1987 (16.3%), and was recorded most often for elderly women, with 61.8% of the 841 cases in this treatment group aged 70 years or more. Table A8.4 in Appendix Eight details recorded non-surgical oncology for patients without recorded surgery. For this group of 841 cases, hormone therapy was recorded for 758 (90.1%), and radiotherapy for 302 (35.9%), with chemotherapy recorded for only 52 cases (3.7%).

The frequency with which each type of surgery was recorded showed highly significant changes between 1983 and 1987. Mastectomy was recorded for 1353 cases in 1983 (55.4%), but for only 982 cases in 1987 (37.3%) (Chi square 165.8 df 1 $p<0.0001$). Conservative surgery rates showed a simultaneous rise, being recorded for 506 cases in 1983 (20.7%) and for 1014 cases in 1987 (38.5%) (Chi square 190.8 df 1 $p<0.0001$). These changes are in line with the King's Fund recommendation to use conservative surgery where appropriate. However, other possible reasons for differences in the rates must be considered. To estimate whether recording of surgical treatment was reasonably complete, overall surgical (mastectomy plus conservative surgery) rates across the 15 DHAs were compared. Although variations in the use of mastectomy versus conservative surgery might be expected, partly as a result of clinical preference, large differences in the overall surgical rate could signify areas with poor data completeness. Figure 5.2 shows overall surgical rates in 1983.

Figure 5.2 **Rates of recorded surgery by DHA for breast cancers diagnosed in the West Midlands in 1983**



Overall surgical rates varied significantly (Chi square **41.7,** df **14** p<0.001). However, when the DHA with the lowest rate, North Staffordshire **(M07), is** excluded, the variation ceases to be significant. North Staffordshire had the highest proportion of metastatic cases **(14%)** compared to the regional average of **11%,** which although failing to achieve significance, suggests that the apparent lack of surgery in North Staffordshire reflects, at least to some extent, clinically advanced disease which would not be appropriately treated by surgery. Similar observations were made using **1987** data, suggesting that recording of surgery was reasonably complete.

In both years the frequency of recorded surgical treatment varied with age. Figure **5.3** shows that younger women were more likely to have conservative surgery than older groups in both years. In **1983, 144** of the **526** women aged under **50** years **(27.4%)**had conservation surgery recorded, falling to **202** of the **1088 50-69** year olds **(18.6%)** and **160** of **the 830** women aged **70** or more **(19.3%)**(Mantel Haenzel Chi square **10.2** df **1** p=0.001). In **1987** the trend with age was more marked, with conservation surgery recorded for **314** of the **603** women aged less than 50 (52.**1%), 436** of the **1106 50-69** year olds **(39.4%)** and **264** of the **924** women in the oldest group **(28.6%)** (Chi square **85.5** df 1 p<0.0001).

Figure 5.3 **Surgical treatment according to age for breast cancers diagnosed in 1983 and 1987 in the West Midlands**

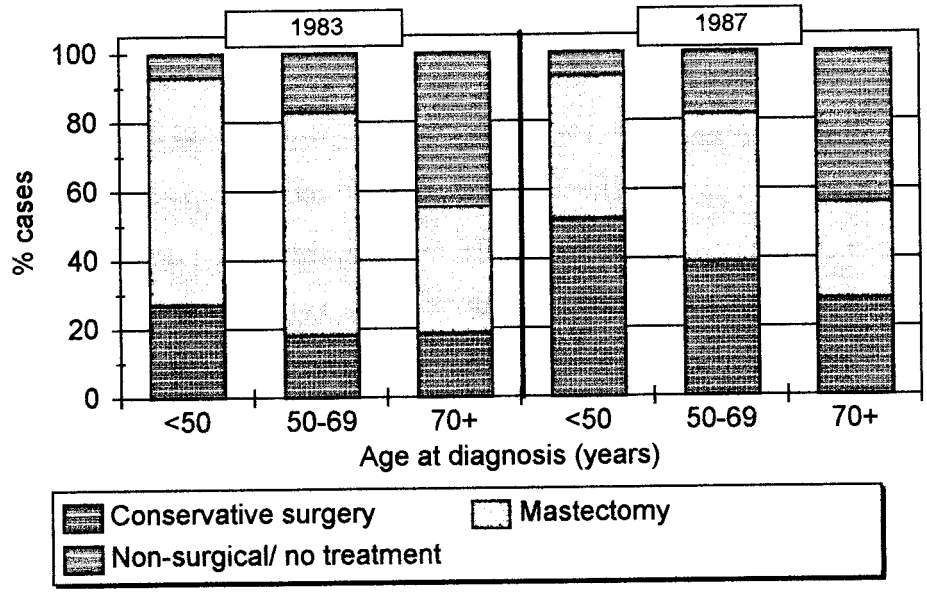
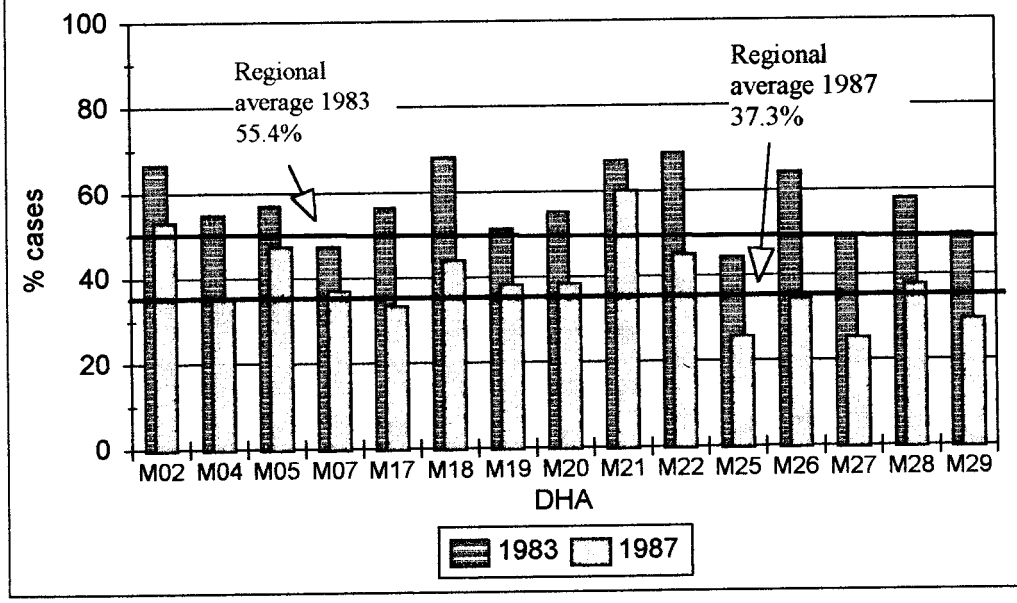


Figure 5.4 **Mastectomy rates by DHA for breast cancers diagnosed in the West Midlands in 1983 and 1987**

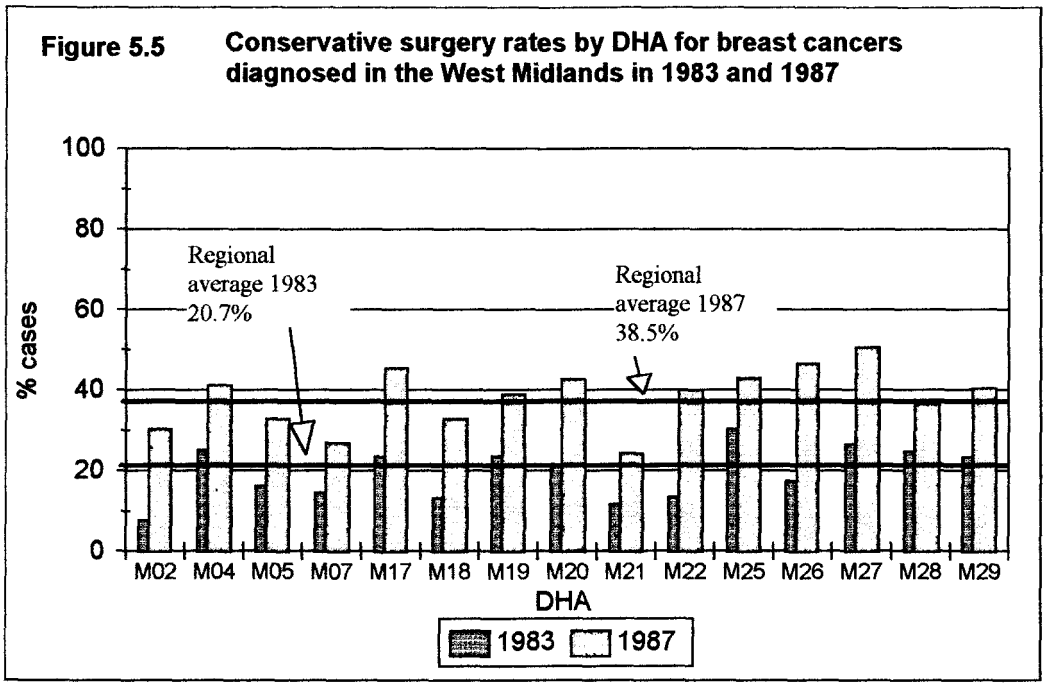


Variation in treatment between DHAs is described in Table A8.3 in Appendix Eight. Recorded mastectomy rates varied across the region, shown in Figure 5.4. Mastectomy rates fell in all DHAs from a regional average of 55.4% in 1983 to 37.3% in 1987, but

inter-district differences were significant in both years. In **1983** mastectomy was recorded for only **44.3%** in South Birmingham *(M25)* rising to **68.9%** in Wolverhampton *(M22),* with significant variation across **DHAs** (Chi square **58.9** df **14 p<0.0001).** In **1987** rates ranged from **25.3%** in South Staffordshire **(M27)** to **60.2%** in Walsall **(M21)** (Chi square **89.9** df **14** p<0.0001). Hereford **(M02)** and Walsall maintained consistently high mastectomy rates relative to other districts throughout the period studied, while rates in North Worcestershire **(M26)** dropped dramatically in **1987** compared with **1983.** Large changes in practice were also apparent in Coventry **(M17),** Dudley **(M18),** Wolverhampton *(M22)* and South Staffordshire **(M27).**

An increase in conservative surgery was seen across the region, shown in Figure 5.5, with the regional average rate rising from **20.7%** to **38.5%.** Inter-district variation in conservative surgery rates was significant, ranging from **7.7%** in Hereford *(M02)* to **26.4%** in South Staffordshire **(M27)** in **1983** (Chi square **54.9** df **14** p<0.0001) and from **24.2%** in Walsall **(M21)** to **50.6%** in South Staffordshire in **1987** (Chi square **59.9** df **14** p<0.0001). In general, districts with high rates in **1983 also** had high rates in **1987,** but the marked increases in the conservative surgery rate in **1987** in Wolverhampton **(M22)** and North Worcestershire **(M26)** may reflect significant changes in practice.

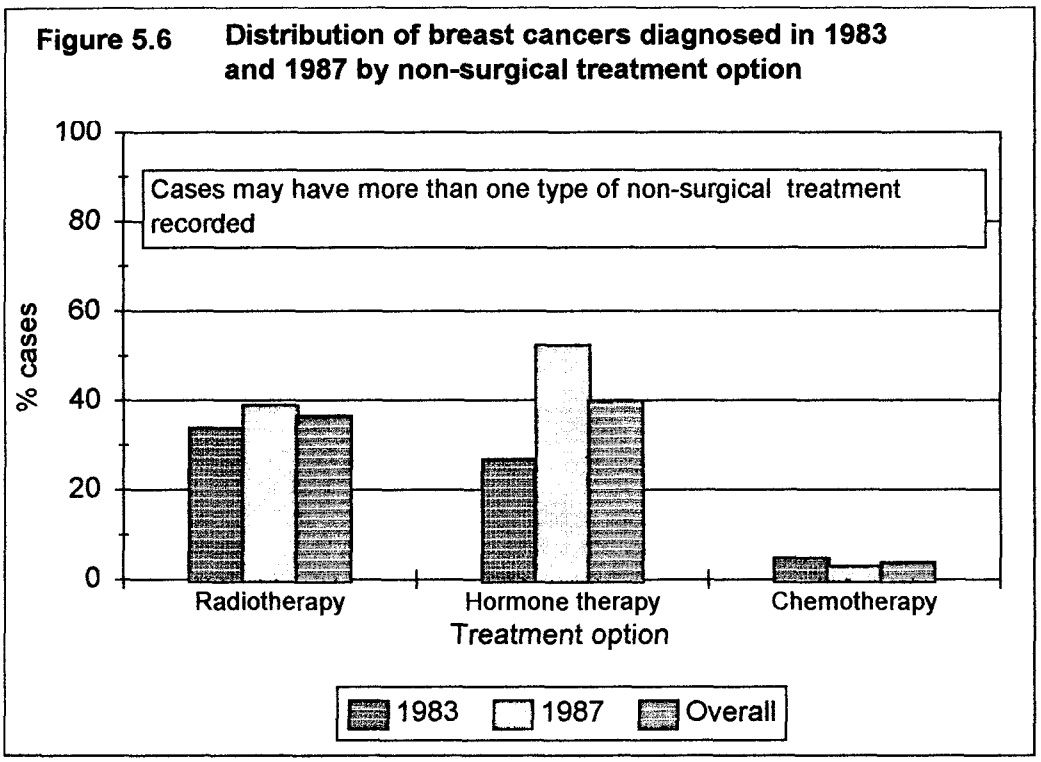
Figure 5.5 **Conservative surgery rates by DHA for breast cancers diagnosed in the West Midlands in 1983 and 1987**



5.2.2 Non-surgical oncology

Table A8.2 in Appendix Eight describes the non-surgical recorded for cases in the study group. Recorded non-surgical oncology is also shown in Figure 5.6. For the years combined, 1859 of the 5077 cases studied (36.6%) had recorded radiotherapy, 2038 (23.5%) hormone therapy, and 196 (3.9%) chemotherapy.

Figure 5.6 Distribution of breast cancers diagnosed in 1983 and 1987 by non-surgical treatment option



5.2.2.1 Radiotherapy

The completeness of non-surgical treatment data is related to the means by which data are notified to the WMCIU. Radiotherapy data are usually submitted directly by the five radiotherapy departments in the West Midlands, and the active communication links that exist between these departments and the WMCIU mean that completeness of radiotherapy data is believed to be reasonably high. Recorded radiotherapy rates rose from 33.9% in 1983 to 39.1% in 1987 (Chi square 14.5 df 1 $p<0.0001$). Use of adjuvant radiotherapy with conservative surgery was recommended by the King's Fund guidelines to reduce the risk of recurrence. Table A8.5 in Appendix Eight provides the treatment data for cases undergoing conservative surgery. Figure 5.7 shows that, of 1520 cases with recorded conservative surgery, 51.0% of 506 cases also received radiotherapy in 1983, rising to 59.1% of 1014 cases in 1987 (Chi square 17.4 df 1 $p<0.0001$). This rise

together with the associated increase in conservative surgery in 1987 compared with **1983 (20.7%** to 38.5%) suggests that the King's Fund recommendations did have an effect on practice in the West Midlands region. Of similar significance is the increase in hormone therapy in combination with radiotherapy for women undergoing conservative surgery. This change is again completely in line with the King's Fund recommendations.

Figure 5.7 **Recorded *non-surgical oncology* for conservative surgery cases diagnosed in the West Midlands in 1983 and 1987**

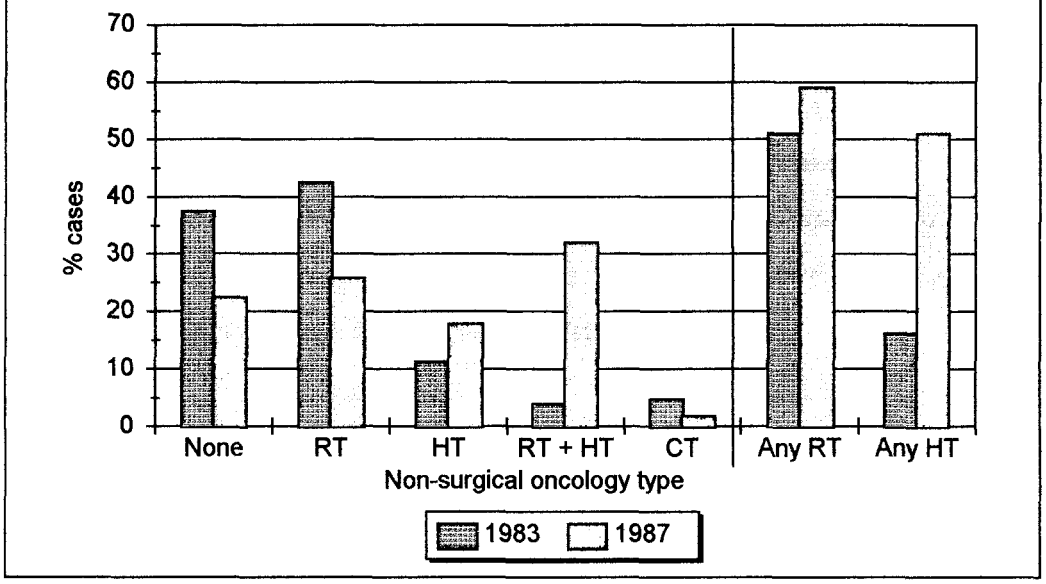
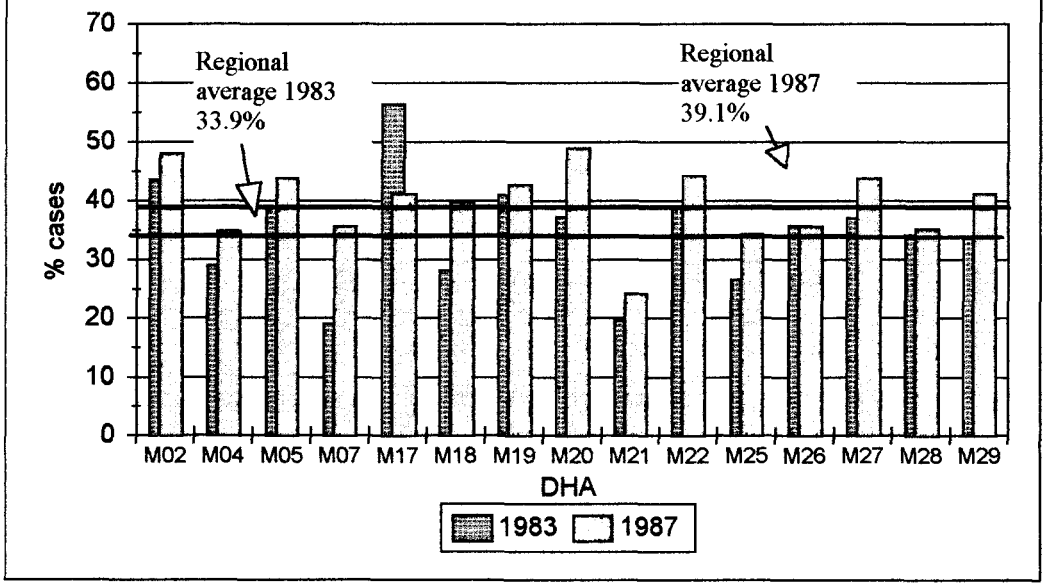


Figure 5.8 **Radiotherapy rates by DHA for breast cancers diagnosed in the West Midlands in 1983 and 1987**

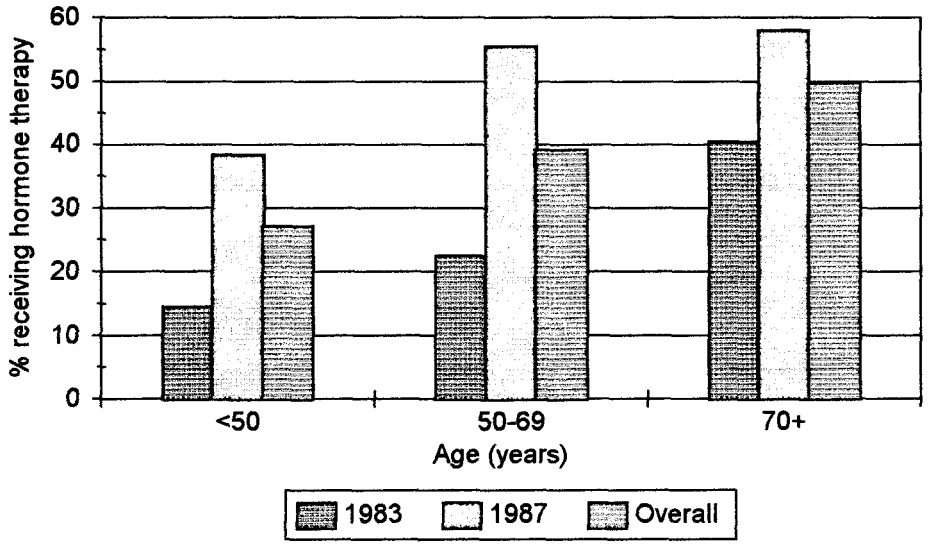


Recorded radiotherapy rates varied across the region, shown in Figure **5.8,** with the regional average rising from **33.9%** in **1983** to **39.1%** in **1987. In 1983** rates ranged **from 19.0%** in North Staffordshire **(M07)** to **56.5%** in Coventry **(M17)** (Chi square **92.7** df **14** p<0.0001). Rates in **1987** ranged from **24.2%** in Walsall **(M21)** to **48.1%** in Hereford **(M02)** (Chi square **33.3** df **14** p=0.003).

5.2.2.2 Hormone therapy

Hormone therapy is often prescribed within, and recorded by radiotherapy departments, and if this is the case hormone therapy is notified to the WMCIU at the same time as radiotherapy. Hormone therapy may also be prescribed in other clinical settings including out-patient appointments and primary care, from which it is known that notification of data to the WMCIU is incomplete. It is thus likely that recorded hormone therapy data are incomplete, and variation in recorded rates between DHAs, which are not presented, is likely to reflect differences in data completeness.

Figure 5.9 **Distribution of breast cancers diagnosed in 1983 and 1987 by age and recorded hormone therapy**



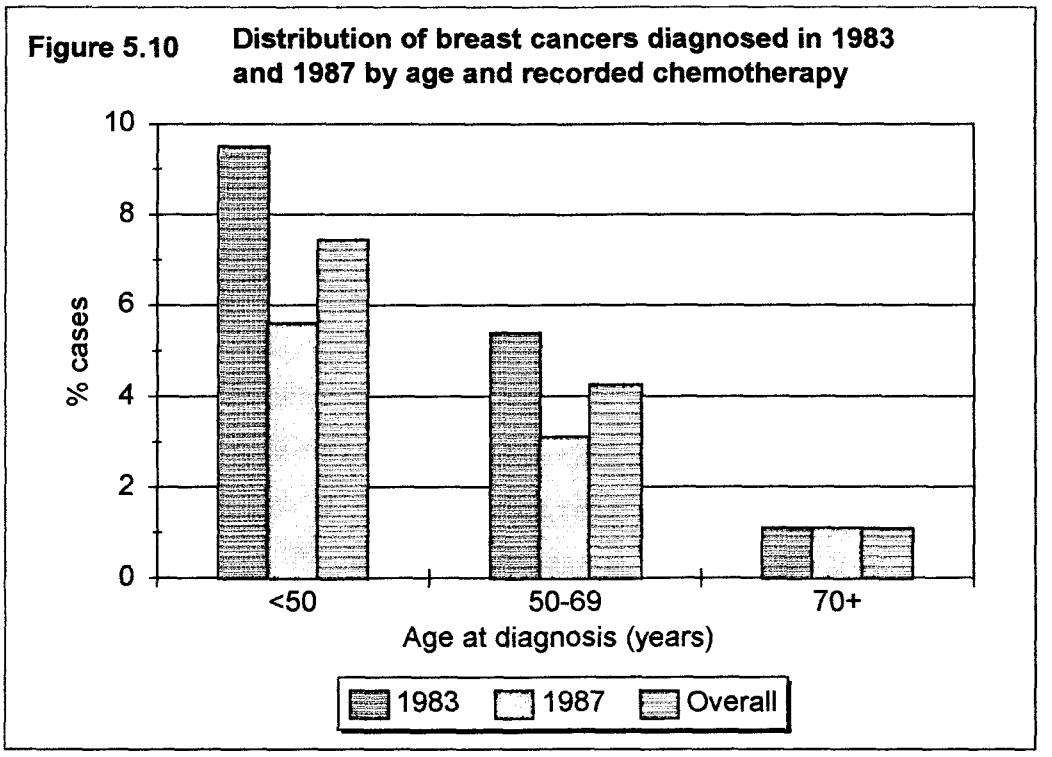
Despite this under-recording, it is probable that some of the changes in recorded hormone therapy over the period studied are due to genuine changes in clinical practice. Hormone therapy was recorded for **658 of 2444** cases in **1983 (26.9%),** rising to **1380 of 2633** in **1987 (52.4%)** (Chi square **329.2** df **1 p<0.0001),** suggesting that the King's

Fund guidelines were acted upon. Older women were more likely to have recorded hormone therapy, shown in Figure **5.9,** in line with the guidelines. In **1983** hormone therapy was recorded for **14.5%** of the youngest women, rising to **40.5%** in the oldest (Mantel Haenzel Chi square **145.9** df 1 p<0.0001). In **1987** rates rose from **38.3%** in the youngest to **58.0%** in the oldest group (Mantel Haenzel Chi square **59.4** df 1 p<0.0001).

5.2.2.3 Chemotherapy

Chemotherapy may be prescribed within radiotherapy departments, but may also be prescribed in other in-patient and out-patient settings, from which notification of data to the WMCIU is again known to be less complete. In this analysis, of the **196** cases with recorded chemotherapy, **106 (54.1%)** had this treatment notified to the WMCIU by **a** radiotherapy department. Despite the King's Fund guidelines recommending chemotherapy, recorded rates fell significantly in the West Midlands, being recorded for 118 women in **1983 (4.8%)** and **78** women in **1987 (3.0%)** (Chi square **11.4** df 1 p=0.0007). These observations must be viewed with great caution as the completeness of out-patient chemotherapy data is known to be poor.

Figure 5.10 **Distribution of breast cancers diagnosed in 1983 and 1987 by age and recorded chemotherapy**



Some differences in usage between groups were, however, apparent. Figure 5.10 shows that, in line with the guidelines, in both years younger women were significantly more likely to have recorded chemotherapy. In 1983 chemotherapy rates fell from 9.5% in the youngest to 1.1% in the oldest group (Mantel Haenzel Chi square 51.2 df 1 p<0.0001) and in 1987 from 5.6% in the youngest to 1.1% in the oldest women (Mantel Haenzel Chi square 26.2 df 1 p<0.0001).

5.2.3 Treatment variation according to stage

The King's Fund guidelines recommended that stage and age should be considered when allocating appropriate treatment. Recorded data suggest that age was taken into account, for example, older women were more likely to have hormone therapy recorded and less likely to undergo conservative surgery in both years. To examine whether treatment differed with stage, treatment recorded for early (stage '0, '1' and '2a'), and locally advanced (stage '2b' or '3') invasive breast cancers was examined. To reduce possible confounding by age the analysis was restricted to the 1139 early and locally advanced stage cancers dagnosed in 50-69 year olds.

Treatment combinations are presented in Figure 5.11. Each cell represents a combination of surgery, radiotherapy (RT), chemotherapy (CT) and hormone therapy (HT). The percentage of cases with each combination is shown, with darker shading indicating a larger contribution to the total. The likely incompleteness of hormone therapy and chemotherapy data should be considered when interpreting the grids.

In 1983 recorded treatment differed significantly with stage, with 291 of the 296 early cancers (98.3%) having surgery compared to 266 of the 315 advanced cancers (84.3%) (Chi square 34.7 df 1 p<0.0001). Conservative surgery was recorded for 65 early stage (22.0%), but only 24 (7.6%) advanced cancers (Chi square 24.1 df 1 p<0.0001). Hormone therapy was recorded for only 37 early cancers, but for 87 advanced cancers (12.6% vs. 27.6% Chi square 20.6 df 1 p<0.0001), and radiotherapy was also more often recorded for advanced disease (29.2% vs. 53.9% Chi square 37.9 df 1 p<0.0001). Chemotherapy was recorded for only 18 early cancers, and for 13 advanced cancers (6.2% vs 4.1% Chi square 0.8 df 1 p=0.36).

Figure 5.11 Treatment by stage for 1139 women aged 50-69 years

Early stage cancers 1983 (n=296)

	RT+		RT -		
Conservative surgery	12.8	0.7	0.0	6.1	HT-
	1.4	0.0	0.0	1.0	HT+
Mastectomy	8.8	0.7	4.1	52.4	HT-
	3.4	0.0	0.7	6.1	
No surgery	0.0	0.0	0.0	0.0	HT+
	1.4	0.0	0.0	0.3	HT-
	CT-	CT+		CT-	

Early stage cancers 1987 (n=292)

	RT+		RT -		
Conservative surgery	13.4	0.0	0.0	2.1	HT-
	19.5	0.7	0.3	7.2	HT+
Mastectomy	2.4	0.0	0.0	27.7	HT-
	6.2	0.0	0.0	17.8	
No surgery	0.7	0.0	0.0	0.3	HT+
	1.0	0.0	0.0	0.3	HT-
	CT-	CT+		CT-	

Locally advanced cancers 1983 (n=315)

	RT+		RT-		
Conservative surgery	3.8	0.3	0.0	1.6	HT-
	0.6	0.0	0.0	1.3	HT+
Mastectomy	34.0	0.0	2.2	25.7	HT-
	6.0	0.3	1.0	6.7	
No surgery	5.7	0.3	0.0	5.7	HT+
	2.9	0.0	1.0	1.0	HT-
	CT-	CT+		CT-	

Locally advanced cancers 1987 (n=236)

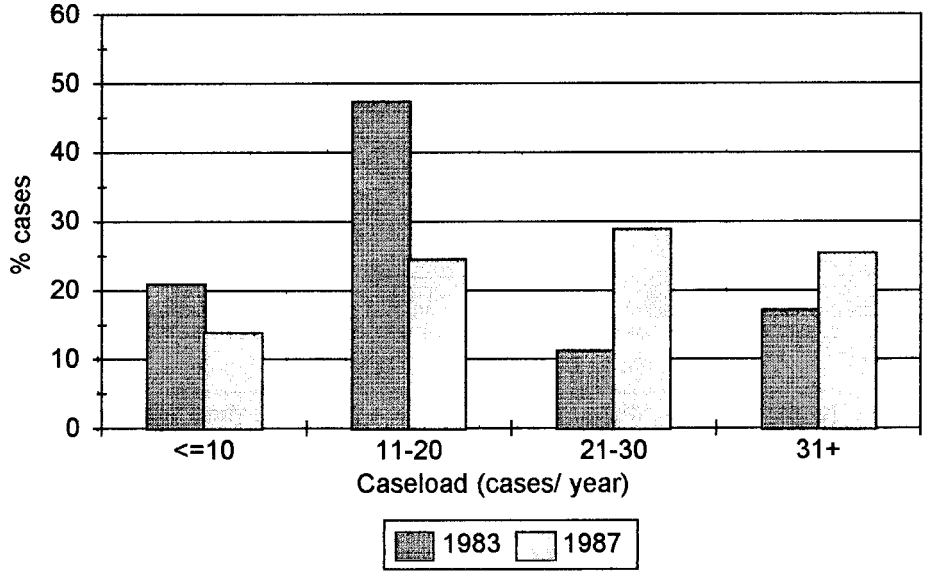
	RT+		RT-		
Conservative surgery	7.2	0.0	0.0	2.1	HT-
	10.2	0.0	0.0	4.2	HT+
Mastectomy	9.3	0.0	0.4	11.4	HT-
	16.1	0.8	0.4	16.1	
No surgery	6.8	1.3	1.7	6.4	HT+
	0.8	0.8	0.4	3.4	HT-
	CT-	CT+		CT-	

Recorded treatment also varied with stage in 1987, and the increased use of conservative surgery, hormone therapy and radiotherapy was apparent in both early and advanced cancers. 97.7% of the 292 early cancers had surgery compared to 74.4% of the 236 advanced cancers (Chi square 60.4 df 1 $p<0.0001$). Conservative surgery was recorded for 126 early (43.2%), but for 56 (23.7%) advanced cancers (Chi square 20.9 df 1 $p<0.0001$). Hormone therapy was recorded for 52.7% of early cancers, rising to 64.0% for advanced cancers (Chi square 20.6 df 1 $p<0.0001$). Radiotherapy was more commonly recorded for advanced disease, being recorded for 128 early, and 126 late stage cancers (43.9% vs. 53.3% Chi square 17.4 df 1 $p<0.0001$). Chemotherapy was rarely recorded, but later stage cancers were more likely to have recorded chemotherapy than early disease (1.0% vs. 5.9% Chi square 8.6 df 1 $p=0.003$). Changes in adjuvant treatment between 1983 and 1987 are also apparent. For example, of the 368 advanced cancers receiving mastectomy, 46.0% also received radiotherapy in 1983 compared to 51.9% in 1987 (Chi square 0.95 df 1 $p=0.33$), and 18.4% had recorded adjuvant hormone therapy in 1983 compared to 61.2% in 1987 (Chi square 67.2 df 1 $p<0.0001$).

5.2.4 Surgical caseload

Tables A8.7 to A8.10 describe treatment recorded for the study group according to surgical caseload. The annual caseload of surgeons treating cases diagnosed in 1983 and 1987 is shown in Figure 5.12. Caseload was available for 1800 of the 1859 cases diagnosed in 1983 with surgery recorded (96.6%), and for 1850 of the 1996 surgically treated cases in 1987 (92.7%). Caseload could not be ascertained for cases where the treating surgeon was unidentified. **Of** cases diagnosed in 1983 for which surgical caseload was available, 320 (17.8%) received treatment from a high caseload surgeon (treating >30 cases/year) compared to 507 (27.4%) of cases diagnosed and treated in 1987 (Chi square 48.2 df 1 p<0.0001). However, even in 1987, almost 75% of cases were treated by surgeons with an annual caseload of <=30 new cases per year.

Figure 5.12 **Distribution of breast cancers diagnosed in 1983 and 1987 by surgical caseload of the treating surgeon**

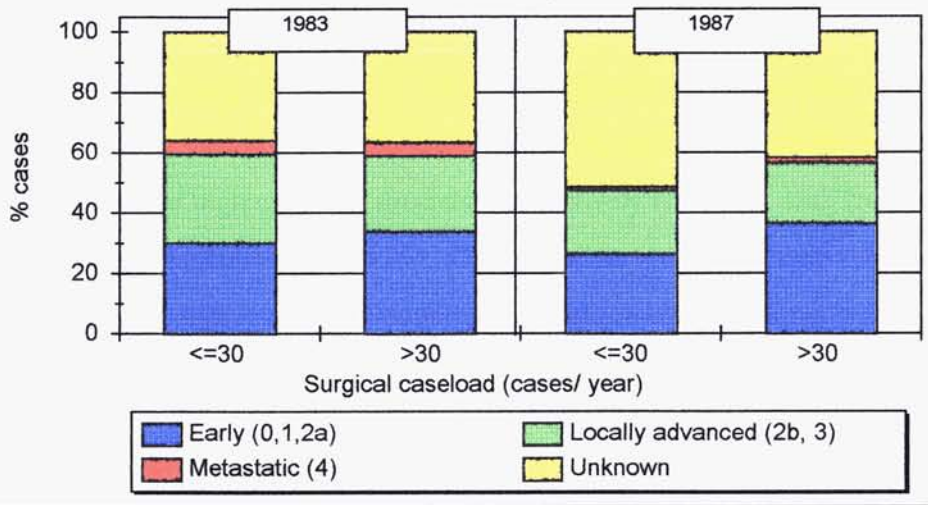


Stage distributions amongst the 1480 cases treated by a low caseload (<=30 cases/year) surgeon and the 320 treated by a high **(>30** case/year) caseload surgeon in 1983, and for the 1343 and 570 cases treated by low and high caseload surgeons respectively in 1987 are shown in Figure 5.13. In 1983 there was little difference in stage distribution according to caseload, with an average of 36.1% of cases (35.5% low caseload, 36.3% high caseload) remaining unstaged after using the staging algorithm. **Only** 4.4% of cases in each caseload group were metastatic at diagnosis, a low proportion which reflects the

use non-surgical oncology alone for very advanced cases, which would thus not receive any treatment from a surgeon.

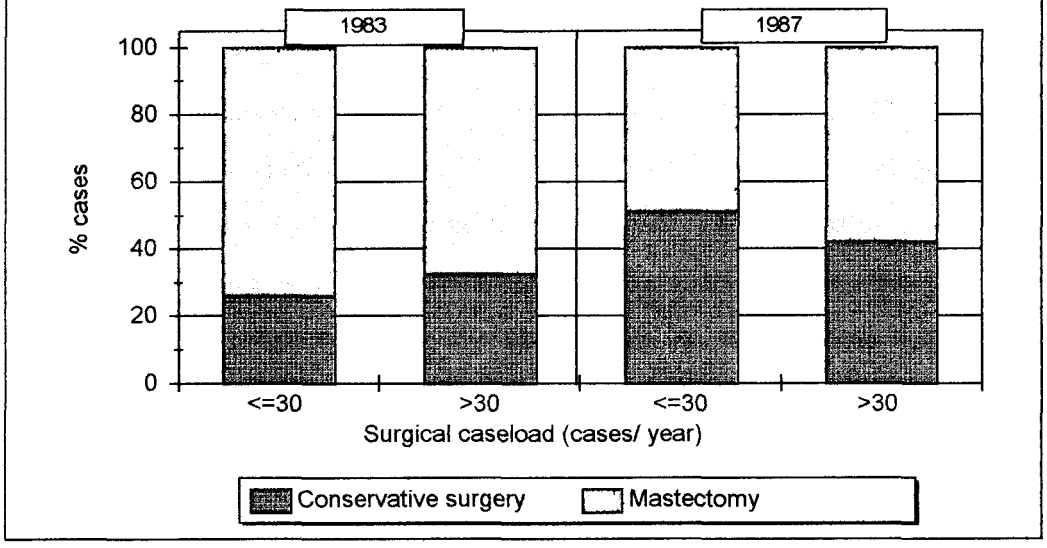
For cases diagnosed in 1987 differences in stage distribution according to caseload of the treating surgeon can be seen for early stage and unstaged cancers. 27.0% of low caseload, and 36.3% of high caseload cases were early stage. Proportions of locally advanced, and metastatic cancers were similar between groups. More cases treated by low caseload surgeons remained unstaged (51.0%) compared to 41.4% of cases treated by high caseload surgeons. The increased proportion of early stage cancers treated by high caseload surgeons may, however, reflect the higher completeness of staging data amongst these cases, reflected in the lower unstaged proportion.

Figure 5.13 **Distribution of breast cancers diagnosed in 1983 and 1987 by stage and surgical caseload**



Comparison of surgery recorded for the 1105 early stage cancers treated by low and high caseload surgeons (Figure 5.14, and Table A8.10 in Appendix Eight) revealed that in 1983 there was no significant variation in treatment according to caseload (Chi square 2.0, df 1 p=0.15), with mastectomy being the more popular option for both groups. In 1987 low caseload surgeons used conservative surgery more often, with 51.4% of cases treated by a low caseload surgeon receiving conservative surgery compared to 42.4% of those treated by a high caseload surgeon (Chi square 3.94 df 1 $p<0.046$). This suggests that, although conservative surgery increased in both groups, in 1987 conservation surgery was used more readily by surgeons treating fewer cases.

Figure 5.14 **Surgical treatment for early stage breast cancers diagnosed in 1983 and 1987 by surgical caseload**



Amongst cases with conservative surgery the rise in recorded radiotherapy from 1983 to 1987 was apparent for both low and high caseload surgeons. However, there was no difference in the frequency of recorded radiotherapy according to caseload in either year. Radiotherapy was recorded for 50.5% and 54.0% of cases diagnosed in 1983 receiving surgery from low and high caseload surgeons respectively (p=0.59), and, for 1987 cases, 60.1% and 59.3% respectively (p=0.89). The rise in recorded hormone therapy for this group from 1983 to 1987 was also apparent in both caseload groups, but once again, there was no difference in the frequency of recording according to caseload.

5.3 SURVIVAL OF BREAST CANCERS DIAGNOSED IN 1983 AND 1987

5.3.1 Findings from univariate analysis

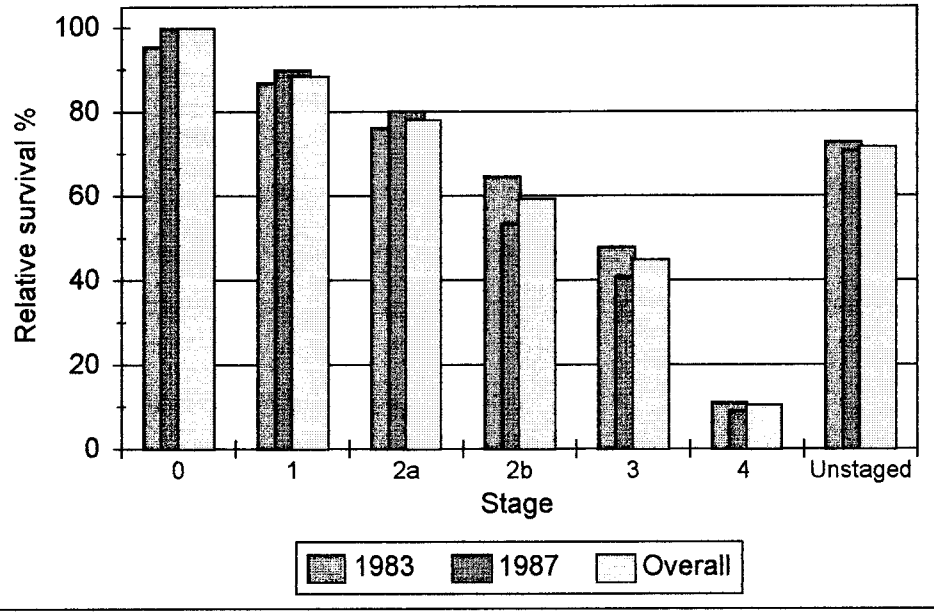
The five year relative survival rates in this section are presented in Tables A8.11 to A8.21 in Appendix Eight. The 33 cases registered from a death certificate only and 16 cases registered from a post mortem report only in 1983 (2%) were excluded, as were the 14 and 9 cases respectively diagnosed in 1987 (0.9%), leaving 2395 and 2610 cases diagnosed in 1983 and 1987 respectively eligible for analysis. The five year relative survival rate for all breast cancers diagnosed in 1983 was 62.7% [95% CI 60.4 - 65.0], rising significantly to 66.4% [64.2 - 68.6] for those diagnosed in 1987 (Chi square 4.4, df 1, p= 0.04). Five year survival rates for invasive breast cancers (excluding sarcomas)

rose from **62.3%** [60.0 - **64.7]** in **1983** to **66.0% [63.3 - 67.9]** in **1987,** although this rise just failed to achieve significance (Chi square **3.5** df 1 p=0.06).

5.3.1.1 Survival by stage

Survival by stage is shown in Figure **5.15.** Survival rates rose from **10.4%[95% CI 6.1 - 14.7]** for stage **'4'** cancers to **88.4% [84.4 - 92.4]** for stage '1'. For stage '0' cancers, which were ***in situ*** and had negative nodes, there were **2** breast cancer deaths amongst **1983** cases, suggesting incomplete removal of the tumour and progression to invasive cancer, or that the diagnosis **of *in situ*** cancer was incorrect. There were no deaths amongst ***in situ*** cases diagnosed in **1987** leading to relative survival in excess of **100%.**

Figure 5.15 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by stage**



Survival rates for stage '1' and stage '2a' cases were slightly higher in **1987** than in **1983,** and that of stage **'4'** cancers worsened slightly from **1983** to **1984,** although none of these differences was significant. For stage '2b' cases, five year relative survival rates dropped from **64.6%** in **1983** to **53.5%** in **1987** ($p<0.05$). Stage '2b' is made up of both larger node negative cases (>=50mm), and smaller node positive cases (>=20mm and <50mm). The proportion of node positive stage **'2b'** cases was similar in each year (8.5% **1983, 8.6% 1987).** The average size of tumours in this stage group were also similar between years both for node negative and node positive cases. It was thus unlikely that

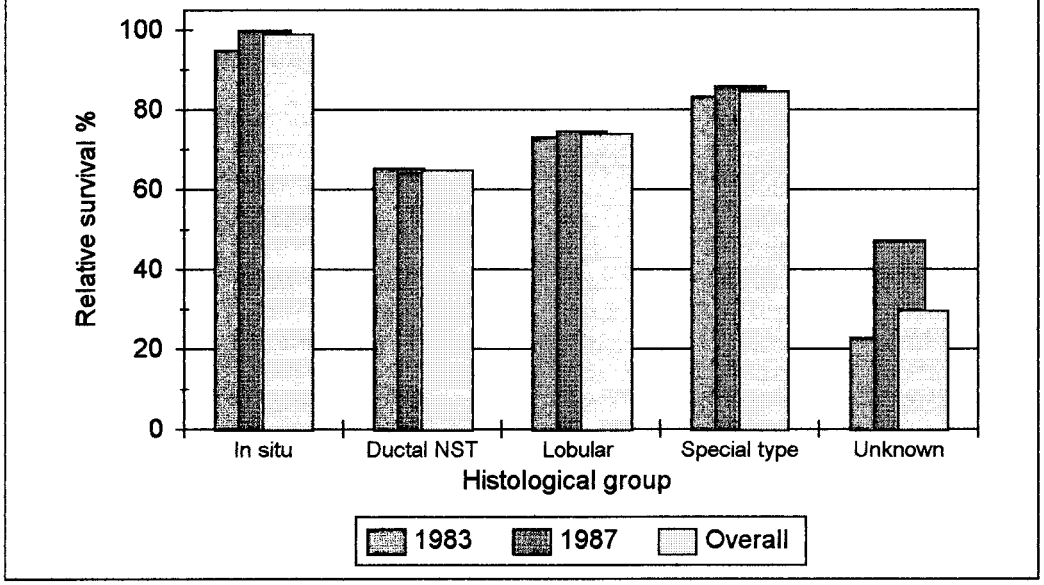
the observed survival difference between years for stage '2b cases reflected a genuine difference in the size or nodal status of cancers making up this stage group. However, poorer completeness of nodal recording in 1987 may have resulted in exclusion of some node negative cases from this stage group.

Stage '3' cases from 1983 fared better than those from 1987 (48.0% 1983, 40.6% 1987, $p<0.05$). In this study stage '3' included stage **'3a'** and '3b' cancers. There was little difference in the contribution made by stage '3b' in both years (90.9% 1983, 87.7% 1987). It is possible that the results for stage '2b' and stage '3' cancers reflect improved recording of staging information in 1987, perhaps due to an increase in the use of detailed investigations. Such an increase might lead to 'stage migration', as apparently localised cancers, which would previously be allocated an early stage are, as a result of more rigorous investigation, found to be more widespread and allocated a higher stage (Barr *et al.*, 1992). This migration would result in fewer cases being incorrectly allocated an early stage, so that survival would appear to improve for early stages (stages '1' and '2a' in this study), but worsen for higher stages (stages '2b, '3', and '4). It is also possible that these difference in survival by stage reflect changes in other factors which influence survival such as treatment regimen or patient-related factors.

5.3.1.2 Survival by histological group

The five year relative survival rates according to histology are shown in Figure 5.16, from which sarcomas were omitted as there were less than 10 cases in either year. In both years survival varied significantly between histological groups. For 1983 cases of unknown histology, survival was only 22.8%, but of 1987 cases in this group survival was better at 47.2%, possibly reflecting increased use of non-surgical therapy for women not undergoing surgery. *In situ* cancers, allocated stage '0 if nodes were recorded as negative, displayed very high survival rates. Survival for lobular carcinomas was high at 73.2%, and survival for special type carcinomas was as expected even higher at 84.5% overall. Invasive ductal carcinomas of no special type, which made up the majority of cancers, had an overall survival of 64.2%.

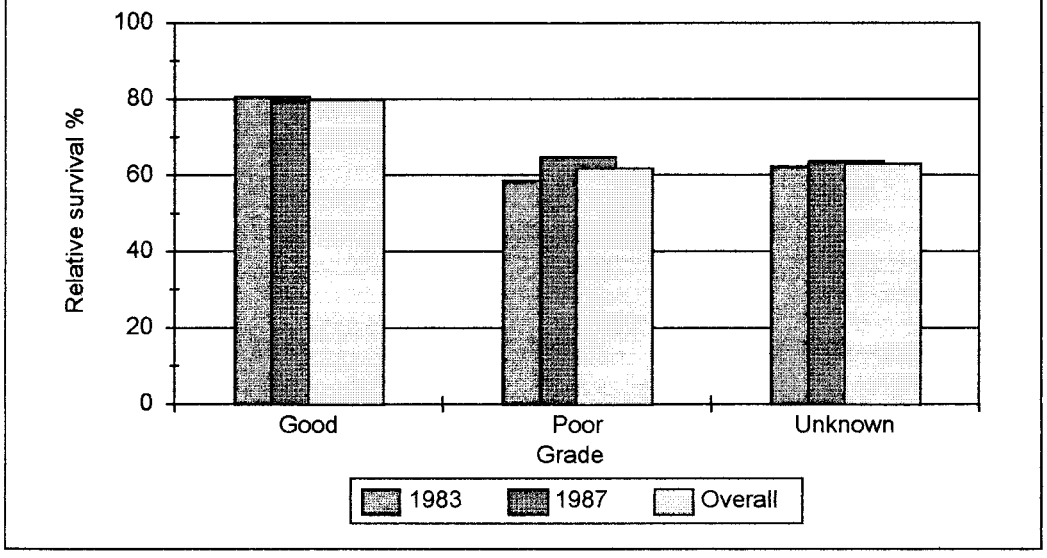
Figure 5.16 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by histological group**



5.31.3 Survival by grade

Five year survival rates by grade are shown in Figure **5.17.** Cases allocated a "good" grade had a five year survival rate of **79.9% [75.6 - 84.2]** overall, whereas "poor" grade cancers had a survival rate of only **61.7%** ***[58.5*** **- 64.9].** The majority **of** cases **(66.1%** overall) did not have grade recorded and displayed a survival rate of ***63.0%*** overall, similar to the poor grade cancers.

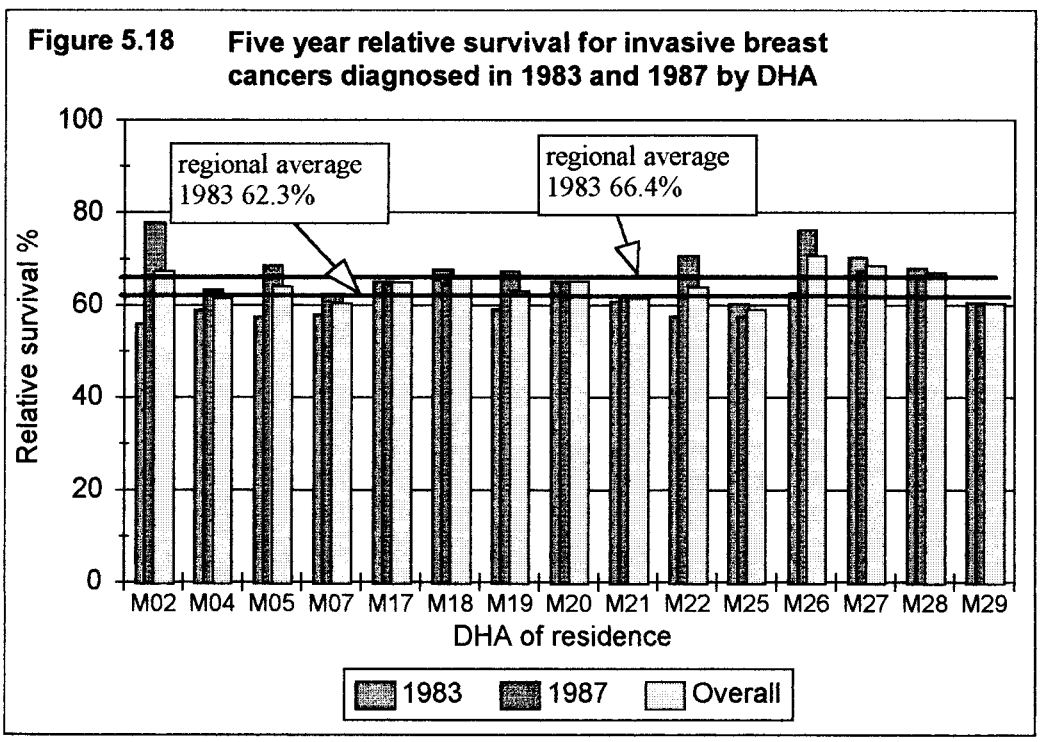
Figure 5.17 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by grade**



5.3.1.4 Survival by DHA

To examine whether the high survival rate for *in situ* cancers could have been responsible for inter-district variation in survival, rates were calculated for all breast cancers, and again for invasive carcinomas only. In the second analysis, *in situ* cancers and sarcomas were excluded, but cases without histology were included. The decision to include those without histology was based on the poor survival observed in this group which suggested that most of these cancers were invasive. There was a high degree of correlation in both years between the rank order by survival for each **DHA** for all cancers compared to that for invasive carcinomas **(1983** cases Spearman rank correlation coefficient **0.95** p<0.0001, **1987** cases **0.97** p<0.0001). Since inter-district variation remained after removal *in situ* cancers and sarcomas, these cases were excluded from all survival analyses thereafter. Survival rates for invasive breast cancers in each **DHA** are shown in Figure **5.18.**

Figure 5.18 **Five year relative survival for invasive breast cancers diagnosed in 1983 and 1987 by DHA**



For cases diagnosed in **1983** survival for invasive carcinomas ranged from 55.9% in Hereford **(M02)** to **70.5%** in South Staffordshire **(M27).** None of the **DHAs** were significantly different from the regional rate, partly due to the small numbers in each **DHA.** For cases diagnosed in **1987** rates ranged from **57.5%** in South Birmingham

(M25) to 77.9% in Hereford (M02), both of which again failed to differ significantly from the regional rate.

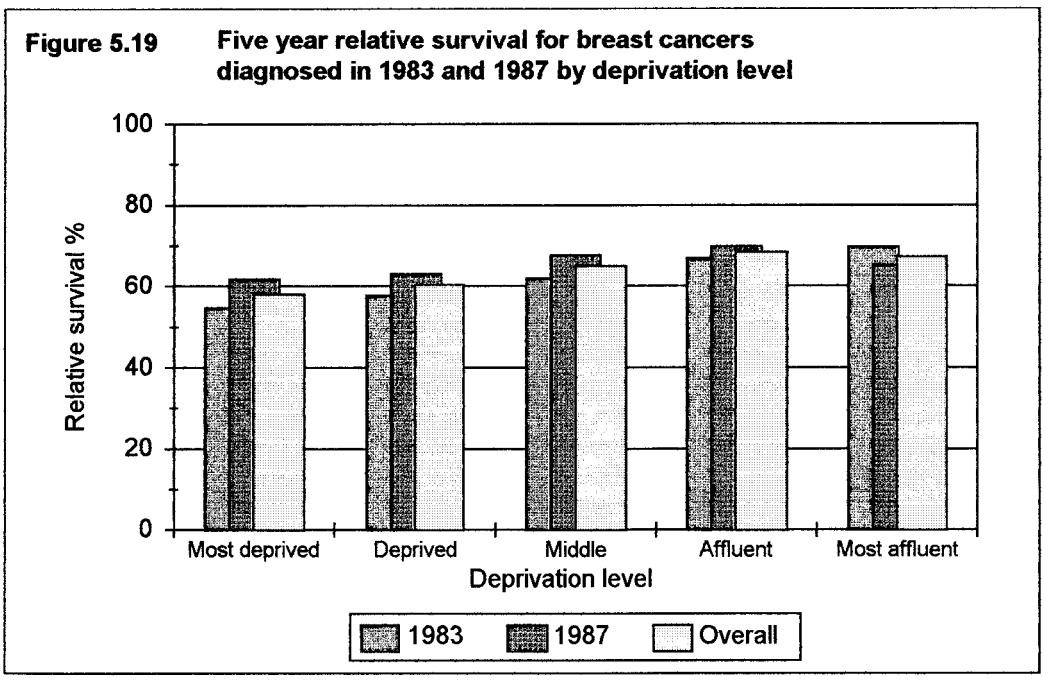
5.3.1.5 Survival by age

The highest survival rates were observed in women aged *<50* years (67.8% [65.0 - 70.7], with the two older groups displaying lower rates of 62.4% [60.2 - 64.6] for 50 - 69 year olds and 63.3% [59.7 - 66.8] for the oldest women. Variation in survival between age groups was significant (Pearson Chi square 13.8 df 2 p=0.001). Within age groups, however, the year groups studied did not differ significantly from each other.

5.3.1.6 Survival by deprivation level

Five year survival rates according to the Townsend group are shown in Figure 5.19. For 1983 cases the variation in survival with deprivation level was significant, with a clear trend in survival with the more deprived groups displaying poorer survival rates. The trend was less clear for 1987 cases, and although more deprived groups fared worse than the more affluent groups, the variation across deprivation levels just failed to achieve significance (Pearson Chi square 9.5 df **4** p=0.05).

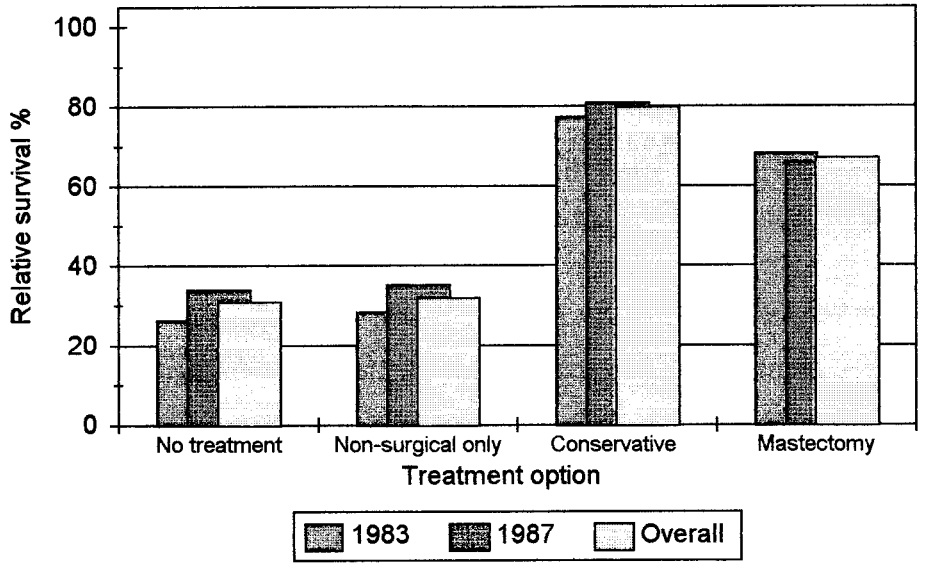
Figure 5.19 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by deprivation level**



5.3.1.7 Survival by treatment type

Survival rates according to treatment type are shown in Figure 5.20.

Figure 5.20 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by treatment option**



The highest survival rates were observed in cases undergoing conservative surgery, with survival rates of **77.2%** for **1983** and **80.8%** for 1987. Cases undergoing mastectomy displayed lower survival rates of **68.1%** for **1983** and 65.5% for 1987. Amongst cases with only non-surgical treatment or with no treatment data recorded survival was poor.

Figure 5.21 **Five year relative survival for convervative surgery cases by non-surgical treatment option**

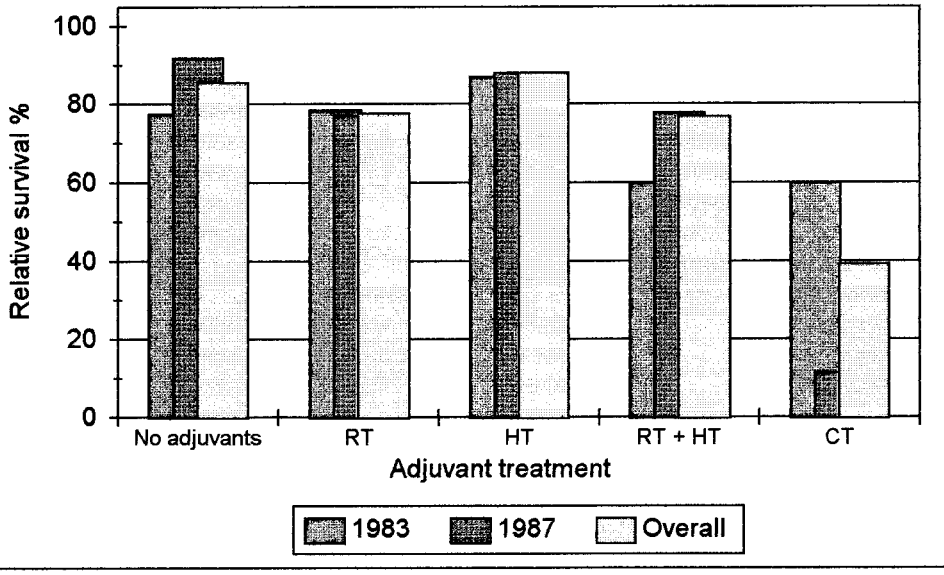
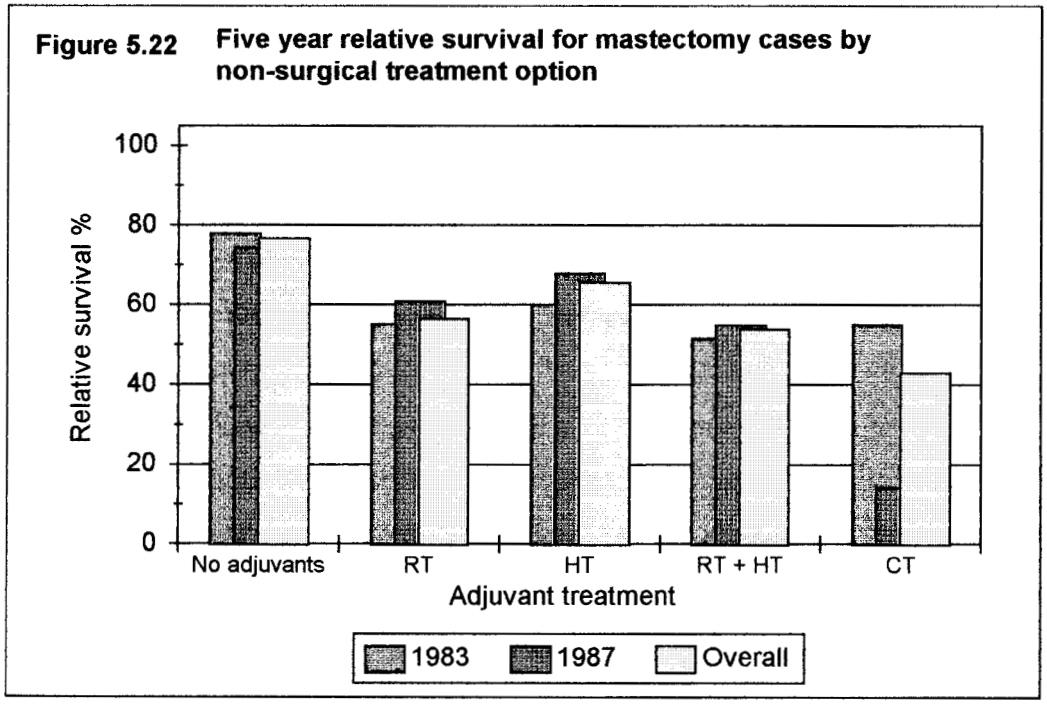


Figure **5.21** shows survival rates for cases with recorded conservative surgery according to the type of adjuvant therapy. For **1983,** cases with recorded hormone therapy alone fared the best, followed by those without adjuvant therapy and those with radiotherapy **only** recorded. For **1987** cases the highest survival was observed in those without recorded adjuvant therapy, with those with recorded hormone therapy faring almost as well. Those with radiotherapy recorded alone or with hormone therapy displayed slightly lower survival rates. In both years cases with chemotherapy recorded fared poorly.

Figure **5.22** shows survival for mastectomy cases according to non-surgical therapy. Variation in survival between adjuvant therapy options was significant in both years. The best survival was observed in those with no adjuvant therapy recorded. For **1983** cases survival rates in those with recorded adjuvant therapy ranged from **60.1%** for those with hormone therapy only recorded to 5**1.6%** for those with both radiotherapy and hormone therapy recorded. For **1987** cases where adjuvant therapy was recorded the best survival was seen in those with hormone recorded only, with those with recorded chemotherapy displaying poor survival. Survival for patients with recorded chemotherapy in **1987** was poorer than seen for this group for cases diagnosed in **1983.**

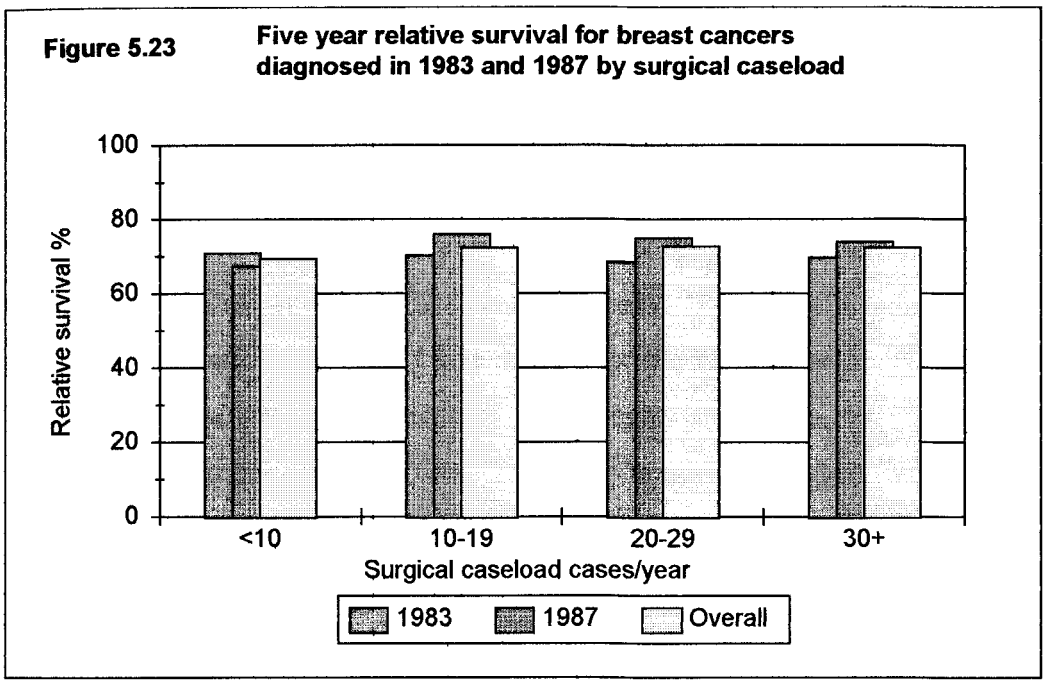
Figure 5.22 **Five year relative survival for mastectomy cases by non-surgical treatment option**



5.3.1.8 Survival by surgical caseload

Survival rates according to surgical caseload are shown in Figure 5.23, Survival variation according to caseload failed to achieved significance in either year. There was no clear trend in survival with increasing caseload, although of cases diagnosed in 1987, those treated by the lowest caseload surgeons displayed the poorest survival.

Figure 5.23 **Five year relative survival for breast cancers diagnosed in 1983 and 1987 by surgical caseload**



5.3.2 Findings from multivariate analysis

Multivariate analyses were carried out on cases for which all the factors examined in univariate analyses were recorded. This restricted the study population to 1041 cases, 20.8% of the original group of 5005 eligible for survival analysis. Of the 1041 cases there were almost equal numbers in terms of year of diagnosis, with 519 cases diagnosed in 1983 (49.9%), and 522 in 1987 (50.1%), with no significant difference in the likelihood of cases in either year being selected (Chi square 1.46 df 1 p=0.23). Selected cases were significantly younger than those unselected, with 271 of the 1041 selected (26.0%), and 1483 of the 3964 unselected cases (37.4%) being aged 70 years or more (Chi square 46.4 df 1 p<0.0001). Furthermore, the selected group did not represent each DHA equally, with the proportion of cases from each DHA included varying from 11.2% for Walsall to 34.5% for Warwickshire (Chi square 13.7 df 14 p=0.008). Selected cases were less deprived than unselected cases, with 189 (16.2%) of selected,

and 748 of the 3944 unselected cases for which deprivation level was available (19.0%) being in the most deprived group (Chi square 4.09 df 1 p=0.04). These observations reflect the differing data quality across the region described in Chapter Four, with data quality being poorest for the oldest and most deprived women. All DHAs were represented in the selected study group, however, and 26.0% of selected cases were aged 70 years or more.

The poorer survival identified in the univariate analysis for older and more deprived women suggested that selected cases could be biased towards better survivors. Amongst the 4910 invasive cancers, five year survival rates for the 1041 selected cases was 63.4% compared to 64.2% for the 3869 unselected cases (Pearson Chi square 0.36 df 1 p=0.55). In addition, a Cox model was created for the *5005* cases into which selection for the multivariate analysis was entered as a categorical variable, but failed to achieve significance (Wald 1.09 df 1 p=0.30). Thus despite the biases identified, selection for multivariate analysis was not associated with better survival, and thus performing multivariate analysis on the selected cases was considered to be valid.

Table 5.1 describes models in which stage, grade, age, histology, deprivation level, treatment, DHA and year were entered singly into Cox models.

Table 5.1 Statistics for models in which covariates were entered singly

Covariate	No. levels	Models in which covariates entered singly		Models in which covariate entered in addition to stage	
		Wald	p	Wald	p
Stage	5	253.27	<0.0001	-	-
Grade	2	18.33	<0.0001	6.04	0.01
Age	3	3.53	0.17	4.08	0.13
Histology	3	4.71	0.09	3.73	0.16
Deprivation level	5	12.98	0.01	9.39	0.05
Treatment group	6	102.26	<0.0001	33.62	<0.0001
DHA	15	18.75	0.18	29.95	0.01
Year	2	2.18	0.14	0.43	0.51

When entered alone, stage was the most significant factor, followed by treatment. Grade was also highly significant. Deprivation level was significant (p=0.01), but age, histology, DHA and year failed to achieved significance. When the process was repeated

using a baseline model including stage, treatment remained highly significant while the significance of grade and deprivation level was reduced. Histology, age and year remained non-significant. DHA became significant after adjustment for stage, although non-significant alone, indicating that, stage for stage, survival covaried significantly with DHA.

Table **5.2** shows statistics for the overall model. Stage, treatment group, DHA, grade **and** deprivation band were selected for inclusion in the model. Stage was the most significant factor, followed by treatment group. DHA was selected as the next factor for entry, followed by grade and finally deprivation level. The significance of the model based on the overall change in the log likelihood was high ($p<0.0001$).

Table 5.2 Model statistics for the overall model

Covariate	Step number	-2 Log Likelihood	Chi Square - change in -2LL from previous step	p
Baseline	0	6391.60	-	-
Stage	1	6169.40	222.18	<0.0001
Treatment group	2	6,137.69	31.74	<0.0001
DHA	3	6110.89	26.79	0.02
Grade	4	6,103.92	6.97	0.01
Deprivation band	5	6,093.92	10.01	0.05
Overall change in -2LL		297.68	405.79	<0.0001

Table 5.3 also describes the overall model. Relative risks for each risk factor are given in comparison to reference categories of stage '1' cancers, treated with surgery alone, of good grade, in the most affluent area. To aid interpretation of calculated relative risks, the reference category selected for each covariate was the category likely to have the best prognosis. The DHA selected as the reference category, Hereford **(M02),** was chosen simply because it is numerically the first DHA listed. Interpreting relative risks for individual DHAs in comparison to this reference DHA is of limited use, and for this reason relative risks for every individual DHA are not provided in Table 5.3.

Table 5.3 Relative risks for covariates in the overall model				
Covariates	**Level**	**Relative risk**	**Lower 95% CI**	**Upper 95% CI**
Stage	1	1	-	-
	2a	1.24	0.88	1.74
	2b	2.25	1.62	3.14
	3	3.39	2.46	4.67
	4	8.52	5.68	12.79
Treatment group	Surgery only	1	-	-
	Medical only/ no treatment	2.17	1.51	3.13
	Surgery + chemotherapy	2.56	1.70	3.87
	Surgery + hormones	1.16	0.86	1.57
	Surgery + radiotherapy	1.44	1.11	1.87
	Surgery + hormones + radiotherapy	1.47	1.08	1.99
Grade	Good	1	-	-
	Poor	1.37	1.1	1.72
Deprivation level	Most affluent	1	-	-
	Affluent	1.05	0.79	1.39
	Middle range	0.77	0.56	1.05
	Deprived	1.16	0.87	1.55
	Most deprived	1.21	0.86	1.69
DHA	DHA was significant (p=0.02). Relative risks were calculated relative to Hereford (M02). Only North Staffordshire (M07) had a relative risk significantly different 2.03 95% CI 1.08 - 3.84			

Compared to stage '1', stage '2a' cases were at a non-significant elevated risk, but the risk to other stages was increased significantly with advancing stage. Poor grade cancers were at significantly higher risk than good grade (RR 1.37 [1.1 - 1.721). Deprivation groups other than middle range were at non-significant elevated risk compared to the most affluent, whereas those in the middle range were at a non-significantly reduced risk. Results for treatment were, at first glance, surprising. **As** expected, those without surgery were at elevated risk compared to those with surgery alone **(RR** 2.17 [1.51 - 3.131). However, all the surgical groups with recorded adjuvant treatment were also at elevated risk, with only the surgery + hormone therapy group displaying a non-significant elevation (RR 1.16 [0.86 - 1.571). This is at variance with the weight of evidence in support of the benefits of adjuvant therapy for breast cancer. This result probably reflects incompleteness in the recording of chemotherapy and hormone therapy data, particularly for women who received these treatments in out-patient or primary care settings in the absence of radiotherapy. It is likely that such women had a better

prognosis than those requiring treatment at radiotherapy centres, and thus that chemotherapy and hormone therapy data recording was biased towards women with poorer health. Lack of chemotherapy and hormone therapy data is probably associated with a better prognosis, giving rise to the observed results. Relative **risks** for each **DHA** were calculated in comparison to M02 as an arbitrary reference category, and are not shown. Only M07 displayed a significantly elevated risk (RR 2.03 [1.08 - 3.841).

Separate stage-specific models were produced using forward stepwise selection to assess the influence of other covariates within stage. The stage-specific models are described in Table 5.4.

Table 5.4 Model statistics for overall models created for each stage

Stage for which model was fitted	Initial -2LL	Final -2LL	Chi Square for change in -2LL	Covariates selected in model
1	534.72	534.72	0	None
2a	1,073.04	1,055.39	19.138 p = 0.0018	Treatment group
2b	1,147.29	1,147.29	0	None
3	1,486.18	1,446.23	47.195 p <0.0001	Age, grade, treatment group
4	479.92	479.92	7.09 p = 0.03	Histological group

For stage '1' and '2b', none of the other covariates influenced survival. For stage '2a' cases treatment group was significant. For stage '3' cases, age, grade, and treatment group were significant, and for stage '4' cases, only histological group was significant. Table 5.5 shows relative risks from the overall stage-specific model for stage '2a' cases.

Table 5.5 Relative risks for overall model for 294 stage '2a' cases

Covariate	Level	Relative risk	Lower 95% CI	Upper 95% CI
Treatment group	Surgery only	1	-	-
	Non-surgical only or no treatment	3.71	1.52	9.03
	Surgery + chemotherapy	2.39	1.04	5.50
	Surgery + hormones	1.02	0.53	1.99
	Surgery + radiotherapy	2.14	1.27	3.63
	Surgery + hormones + radiotherapy	2.15	1.18	3.91

Those without surgery were at significantly elevated risk compared to those with surgery alone recorded. Surgical groups for which adjuvant treatment were recorded were at elevated risk, with all groups being at significantly higher risk except for the

surgery + hormone therapy group (RR 1.02 95% CI 0.53 - 1.99). Table 5.6 shows the relative risks from the stage-specific model for stage '3' cases.

Table 5.6 Relative risks for overall model for 238 stage '3' cases				
Covariate	**Level**	**Relative risk**	**Lower 95% CI**	**Upper 95% CI**
Age	< 50 years	1	-	-
	50 - 69 years	0.73	0.44	1.21
	70+ years	1.16	0.67	2.02
Treatment group	Surgery only	1	-	-
	Medical only/ no treatment	3.24	1.87	5.61
	Surgery + chemotherapy	4.53	2.00	10.23
	Surgery + hormones	1.15	0.70	1.87
	Surgery + radiotherapy	1.51	0.93	2.45
	Surgery + hormones + radiotherapy	1.49	0.84	2.66
Grade	Good	1	-	-
	Poor	2.04	1.33	3.12

Although age was significant, neither of the two older age groups displayed a relative risk significantly different from women aged < 50 years. Poor grade cancers were at elevated risk compared to good grade cancers (RR 2.04 95% CI 1.33 - 3.12). The results for treatment group were similar to those found in other models, with the surgery + hormone therapy group being the only category with a relative **risk** not significantly elevated compared to the reference category of surgery alone. Table 5.7 shows relative **risks** from the overall stage-specific model for stage *'4'* cases.

Table 5.7 Relative risks for overall model for 75 stage '4' cases				
Covariate	**Level**	**Relative risk**	**Lower 95% CI**	**Upper 95% CI**
Histological group	Ductal carcinoma NST	1	-	-
	Lobular carcinoma	0.59	0.18	1.89
	Special type carcinoma	4.13	1.21	14.15

The reduced risk for lobular carcinoma was non-significant. The significantly elevated risk found for special type carcinomas, commonly associated with a good prognosis, probably reflected the fact that there were only 3 special type carcinomas out of 75 cases. Where models are based on so few data their clinical relevance must be viewed with caution.

To further assess the influence of surgical type (conservative surgery or mastectomy) and surgical caseload on survival, the **933** cases with recorded surgery, and for which caseload and all other data items were available, were selected. Stage, surgical type, adjuvant therapy type, annual surgical caseload, age, grade, histological type, deprivation level, **DHA,** and year of diagnosis were entered singly as a categorical covariates, resulting in the models described in Table 5.8.

Table 5.8 Statistics for models created using surgical cases in which covariates were entered singly

Covariate	No. levels	Models in which covariates entered singly	
		Wald	p
Stage	5	199.1	<0.0001
Surgery type	2	3.6	0.06
Adjuvant therapy type	5	25.9	<0.0001
Caseload	4	0.52	0.92
Grade	2	16.2	0.0001
Age	3	1.5	0.46
Histology	3	4.0	0.14
Deprivation level	5	11.1	0.26
DHA	15	17.4	0.24
Year	2	1.5	0.22

Surgical caseload was non-significant (p=0.90), and surgical type also failed to achieve significance (p=0.06). Only stage, adjuvant therapy group, grade and deprivation level were significant and were selected for inclusion in an overall model using forward stepwise selection. The relative risks for this model were similar to those identified in the model for all cases (Table **5.2),** with risk increasing with worsening stage and grade cancers, and relative risks being of similar magnitude to those presented in Table **5.3.** Once again, the elevated relative risks for women with recorded adjuvant therapy suggest incompleteness in adjuvant therapy data recording. Although deprivation level was significant in the model, no individual deprivation sub-category achieved significance in comparison to the most affluent women.

5.4 SUMMARY

Examination of treatment recorded at the WMCIU for breast cancers diagnosed in 1983 and 1987 revealed changes in recorded treatment over this period which were in agreement with the King's Fund recommendations

- mastectomy rates fell from 55.4% to 37.3% with a simultaneous rise in recorded conservative surgery rates from 20.7% to 38.5%. In both years younger women were more likely to have conservative surgery than older women
- despite probable incompleteness of data, a rise in recorded hormone therapy from 26.9% to 52.4% was identified, with older women more likely to have recorded hormone therapy in both years
- recorded radiotherapy rates rose from 33.9% to 39.1%. This was most apparent amongst women receiving conservative surgery, in whom rates of adjuvant radiotherapy rose from 51.0% to 59.1%
- incompleteness of chemotherapy data meant that little could be concluded regarding use of adjuvant chemotherapy. Very low rates were identified in both years, being recorded most often for the youngest women
- in both years both age and stage were determining factors in allocating treatment
- the proportion of patients treated by high caseload surgeons (defined as those treating >30 cases/year) rose from 17.8% to 27.4%, but even in 1987 nearly 75% received surgery from surgeons with lower caseloads, with 15% treated by surgeons with a caseload of <10 cases/year
- although high caseload surgeons used conservative surgery less often for early stage disease in 1987 than low caseload surgeons, there was little difference in recorded adjuvant radiotherapy for conservative surgery cases according to surgical caseload, suggesting that use of adjuvant therapy did not differ markedly

In addition, in both years rates of recorded mastectomy, conservative surgery and radiotherapy varied significantly across DHA. Although incompleteness in staging data prevented examination of this variation more closely, is likely that this variation reflects, at least in part, variation in clinical practice.

Univariate survival analysis identified the following associations between the suspected prognostic factors entered into analyses and five year relative survival

- five year survival rates for all breast cancers rose significantly from **62.7%** to **66.4%,** with the rise for invasive cancers from **62.3%** to **66.0%** just failing to achieve significance
- stage was a highly significant factor, with five year relative survival rates rising from **10.4%** for stage '4' cases to **88.4%** for stage '**1**' cases
- special type and lobular carcinomas displayed more favourable survival than ductal carcinomas of no special type which made up the majority of cancers. As expected, *in situ* cancers displayed high survival rates, and tumours of good grade displayed better survival than poor grade tumours
- relative survival varied significantly across DHAs, with poorer survival rates identified for older compared to younger women, and for women living in deprived areas of the region compared to more affluent districts
- women without recorded surgery displayed poor survival compared to those with mastectomy or conservative surgery. However, amongst those receiving surgery, recorded adjuvant therapy of any type appeared to confer a survival disadvantage
- no trend in survival according to surgical caseload was identified

Multivariate analyses were performed, restricting the study population to **1041** of the **5005** cases eligible for survival analysis (20.8%) by excluding those cases which did not have all the data items entered into univariate analysis were recorded.

- all DHAs, ages and deprivation levels were represented in the **1041** cases selected, but these cases were significantly younger and less deprived than unselected cases, with unequal representation for each DHA in the selected group
- this biasing effect reflected poorer data quality amongst older and more deprived women, and differing data quality across the region
- despite this bias towards younger and less deprived women in the selected group, selection for the multivariate analysis was not associated with better survival

The following factors were identified as being significant in determining survival in the multivariate analysis

- stage was highly significant after adjustment for all other factors, with the risk of death for stage '**4**' cases being over 8 times greater than that of stage '**1**' cases

- grade was significant in the overall model, although less significant than stage. Poor grade cancers were at a 37% higher risk than good grade cancers
- although survival differences according to tumour histology and age were revealed in univariate analyses, these factors were not significant in the overall model

  deprivation level was the least significant factor in the overall model, with the most deprived women having a 21% higher **risk** compared to the most affluent.
- DHA was not significant when entered alone but was retained in the overall model
- treatment type was highly significant, with those with no recorded surgery being at higher risk than of those with surgery alone

  the apparent survival disadvantage conferred by recorded adjuvant therapy confiimed the results of the univariate analyses. This finding may reflect incompleteness of hormone therapy and chemotherapy data for some women with a good prognosis who contaminated the "no adjuvant therapy" group which was used as the comparison group in the multivariate analysis
- in a model constructed for cases with recorded surgery, neither surgical type (mastectomy or conservative surgery) nor surgical caseload achieved significance, a result in accordance with the univariate analysis

CHAPTER SIX

RESULTS: AN EVALUATION OF THE NHSBSP

6.1 INTRODUCTION

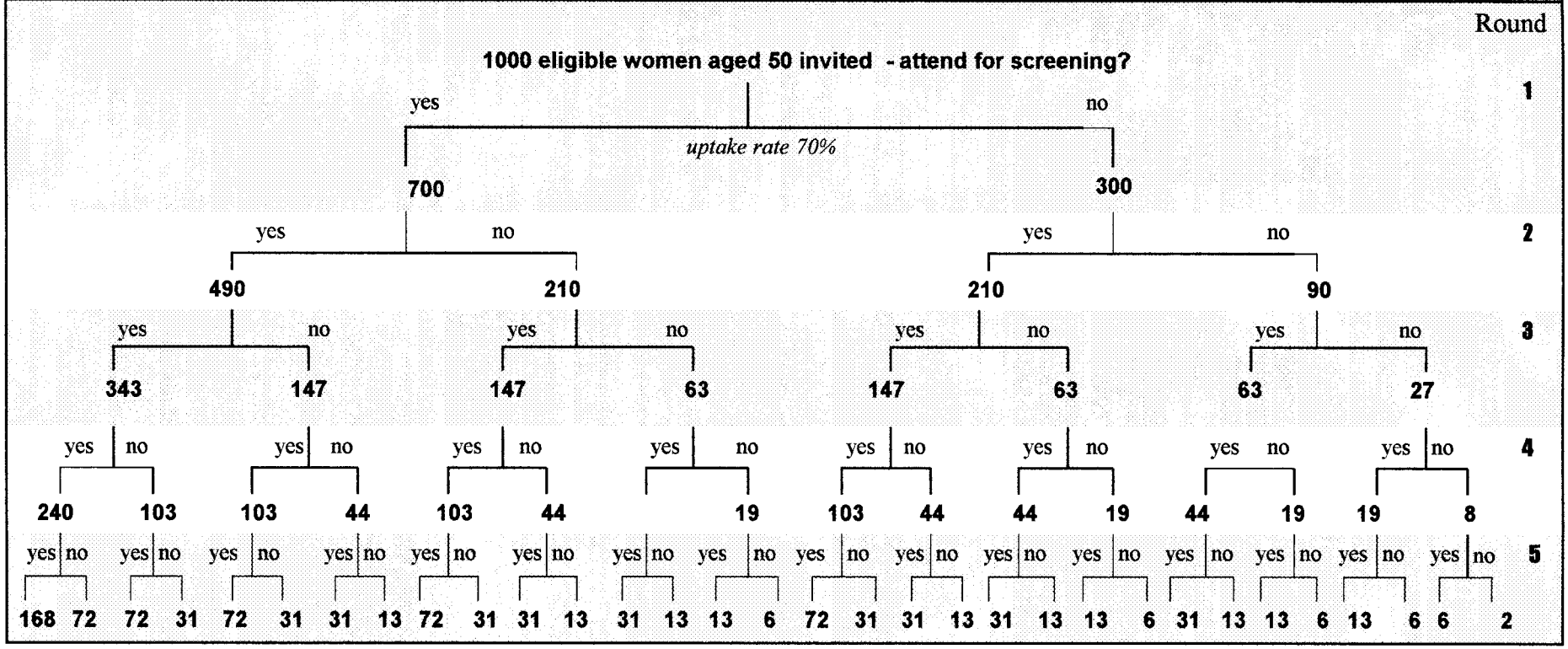
In this Chapter outcome measures used to evaluate the NHSBSP are described using data from the West Midlands, and suggestions made on how their calculation and use might be improved. To assess the impact of screening in the West Midlands a retrospective technique was used to allocate screening histories to breast cancers arising in women eligible for screening. Cases identified as possible interval cancers in this process were then used in an examination of the definition, classification and implications of interval breast cancers, including the effects of operational factors on interval rates and the predicted impact of a shorter, two year screening round.

6.2 A CRITICAL ASSESSMENT OF OUTCOME MEASURES USED TO EVALUATE THE NHSBSP

6.2.1 Uptake rate

The current minimum target uptake rate is 70% (NHSBSP, 1993b), a quality standard which has remained unchanged since the introduction of screening (NHSBSP, 1988). However, if satisfactory cumulative uptake over several rounds is to be achieved the 70% minimum target may be inadequate. If 70% uptake is applied to a hypothetical group of 1000 women entering the programme at age 50 years, and continued through the five screening rounds for which this group normally would be eligible, the predicted pattern of attendance that would result is presented in detail in Figure 6.1 and summarised in Figure 6.2. After five rounds, based on 70% attendance, only 16.8% of eligible women would have attended all five screens, and only 36.0% a total of four screens. Of those attending a total of four times, 60% would have included a lapsed attendance, so that two or more screens would be at least six years apart. The length of time between screens is critical to the success of screening and the occurrence of such lapsed attendance reduces the potential mortality reduction achieved. Of those attending three times, 60% would have included a lapsed attendance, and 16.7% of these women would not have attended any two consecutive screens.

Figure 6.1 Pattern of predicted cumulative attendance to screening based on 70% uptake after five rounds



Round 1 - women invited aged 50 years
Round 2 - women invited aged 53 years
Round 3 - women invited aged 56 years
Round 4 - women invited aged 59 years
Round 5 - women invited aged 62 years

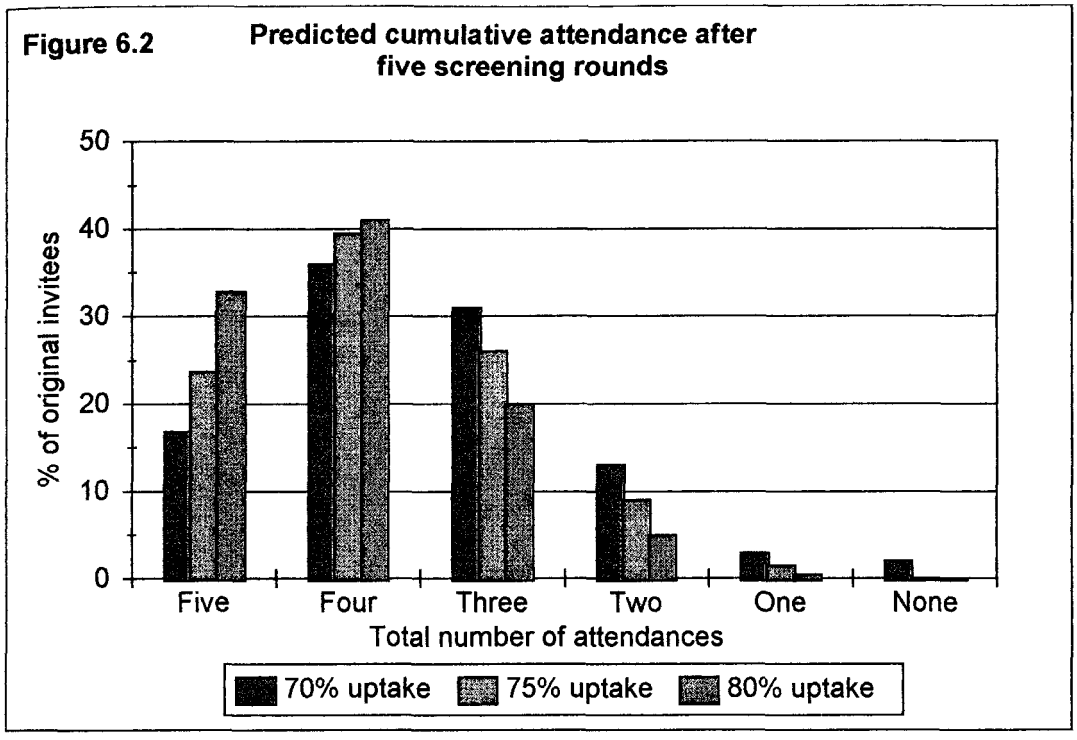
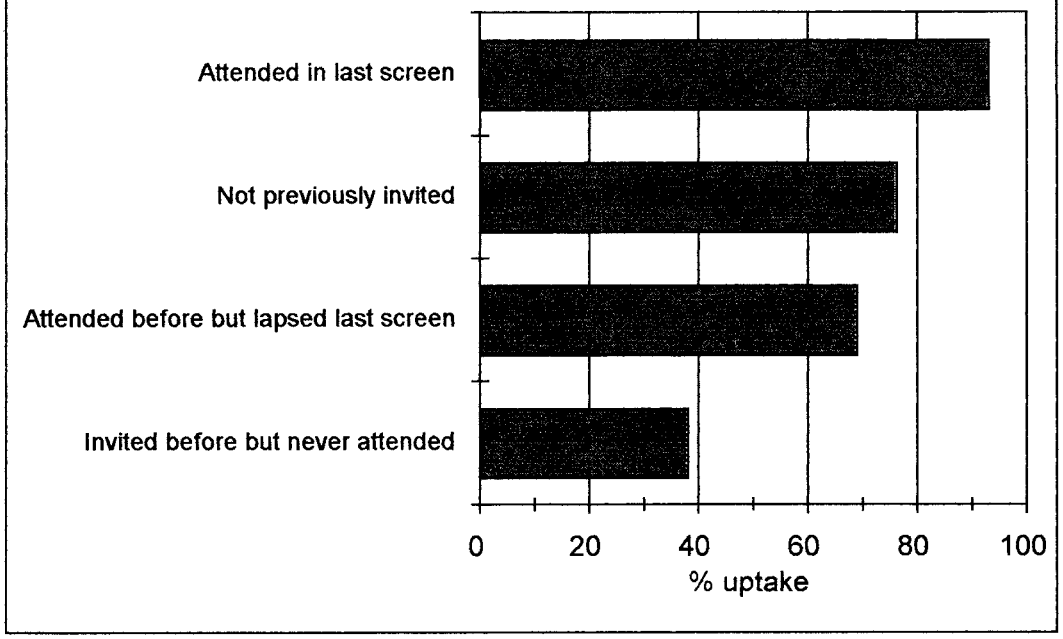


Figure **6.2** also demonstrates the impact of applying a **75%** and 80% uptake rate to cumulative attendance rates. If a target uptake rate of **75%** is applied, which was set as an achievable quality standard in **1993** (NHSBSP, **1993b),** but has since been withdrawn (Radiological Big **'18',** personal communication), over five rounds, nearly **25%** of women would attend all five screens, and assuming 80% uptake, over **30%** would attend in all rounds. To achieve an overall **70%** uptake after **5** rounds, with attendance at all 5 rounds, an uptake rate of **93.2%** at each round would be required. This simplistic model assumes that the probability of attending remains static regardless of previous screening history. The real situation, however, is more complex, as demonstrated by the achieved uptake rates according to screening history for a single West Midlands screening unit shown in Figure **6.3.** For women who attended in the previous round an uptake rate of **92.9%** was achieved and uptake for first time invitees (76.0%) was also high. Uptake for previous attenders who defaulted in the previous round was lower at **68.7%,** but for women who, despite being invited, had never attended previously, uptake was low at **37.9%.** The current standard for uptake rates set by the **NHSBSP** does not take previous screening history into account, being set as a single standard. The simplistic model based on **70%** uptake shows that, in order to achieve the high cumulative attendance needed to produce the desired impact on mortality, a much stricter target which more closely reflects the real

situation needs to be set for re-attendance rates. In addition, the current KC62 report only allows attendance at the previous invitation to be ascertained. Monitoring cumulative attendance over three or more rounds will thus inevitably involve complex evaluation techniques unless the KC62 report can be modified to allow screening history to be described in more detail.

Figure 6.3 **Achieved attendance rates according to previous attendance history 1994/95**



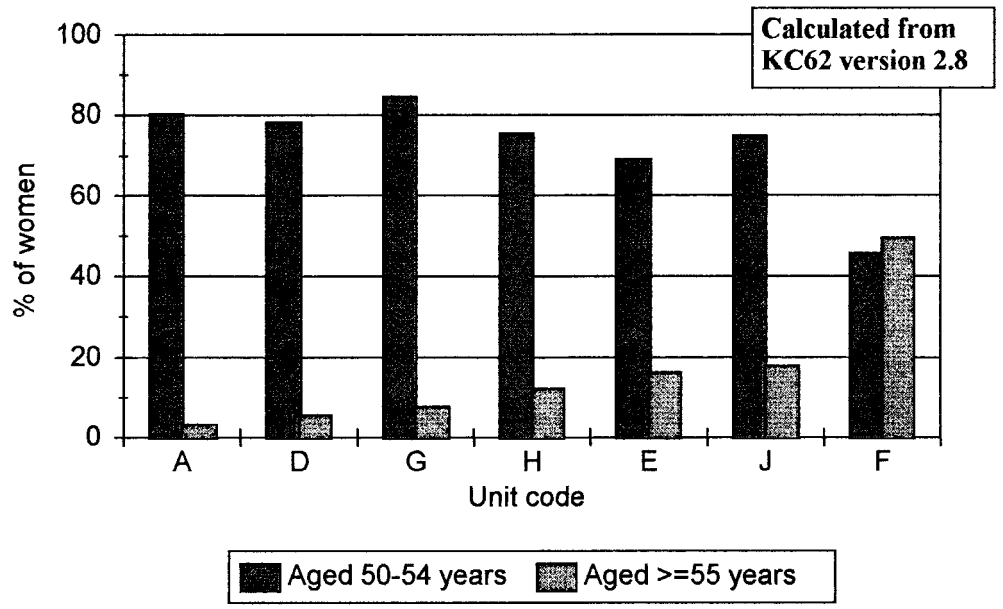
6.2.2 Cancer detection rates and the Standardised Detection Ratio

Essential to the effectiveness of screening is the detection of sufficient invasive cancers. **Up** until 1996 non-invasive (DCIS) and micro-invasive breast cancers were included in the calculation of cancer detection rates for 50-64 year olds **(NHSBSP,** 1993b). The inclusion of non-invasive and micro-invasive cancers in the total artificially inflated the rates achieved, since it is the only detection of potentially fatal invasive cancers that can lead to the mortality reduction desired.

Detection rates are currently calculated separately for prevalent and incident screens with quality standards set to reflect the higher rates expected at prevalent screens. However, invasive cancer detection rates (ICDRs) should be higher for older age groups as the underlying incidence of cancer increases with age. 50-64 year old or all ages ICDRs must be interpreted in relation to the age structure of the population from which they are

calculated to avoid misinterpretation of the achieved rate. The 1996 standards (Radiological Big '1**8',** personal communication) assume that first time invitees are 50-52.9 years old, and that the incident population is aged 53-64.9 years now that most units are past their prevalent round. Data on 50-52.9 year olds are not available from a KC62 report but data for 50-54 year olds shown in Figure 6.4 demonstrate that although such assumptions may be reasonable for some units, e.g. unit **A,** for others such as unit F the prevalent screening population is clearly not predominantly aged 50-54 years old. Similarly, the all ages incident ICDR may not only represent the 53-64.9 year age group.

Figure 6.4 **Age distribution of prevalent women screened at West Midlands units in 1994/95**

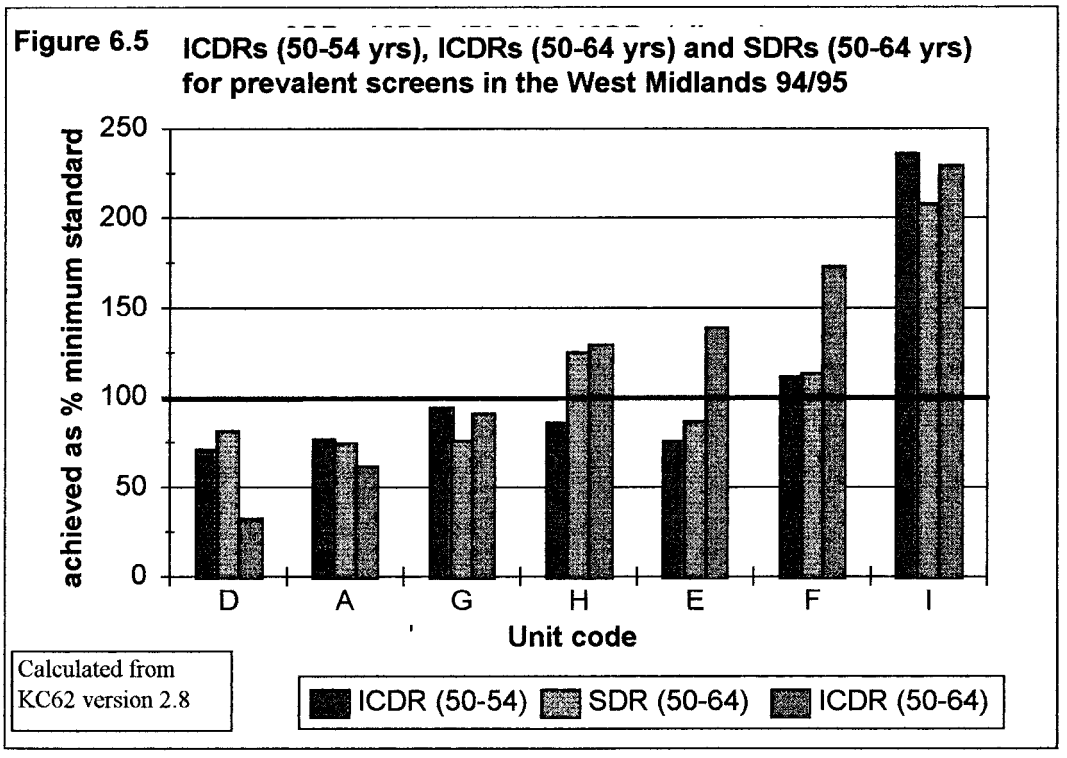


The currently recommended method of calculating ICDRs is to include cancers detected in the 50-54 year age band for prevalent screens, but to use the 50-64 year value for incident screens. There are several problems associated with this approach. The format of a KC62 report version 2.8 is shown in Appendix Two. The data required to produce a 50-54 year ICDR are unavailable from the KC62 version 2.8, and must be estimated. This is because the number of invasive cancers **is** not stratified by age for each Table in the KC62. The currently recommended method of estimating the number of 50-54 years invasive cancers is to assume that the overall proportion of invasive cancers in 50-54 years olds provided in Table ***G*** is representative of the invasive proportion in this age group in Table **A** (prevalent screens). The validity of this adjustment is questionable when the total

number of cancers detected by a Unit may be small, and the KC62 report needs to be altered to give rates for 50-52.9 year olds directly so that such adjustment becomes unnecessary. In addition, cancers detected at prevalent screens in women not aged 50-54 years are currently excluded from the calculation of the ICDR, and cannot therefore contribute to this measurement of screening performance at all.

An alternative measure to the ICDR has been developed using cancer detection rates achieved by the Swedish Two Counties Study to estimate expected rates according to *5* year age band (Blanks ***et al.,*** 1996a). Standardised Detection Ratios (SDRs) can then be calculated to adjust for the influence of differing age distributions in the screened population. **An** SDR of 1.0 indicates parity with the Swedish Two Counties Study and thus would suggest that the mortality reduction achieved in the Swedish Two Counties Study should be achievable in the **NHSBSP.** The advantages of using an SDR over the ICDR are illustrated by the data in Figure 6.5.

Figure 6.5 **ICDRs (50-54 yrs), ICDRs (50-64 yrs) and SDRs (50-64 yrs) for prevalent screens in the West Midlands 94/95**



In Figure 6.5 the achieved outcome measure is presented as a percentage of the minimum standard, which for the SDR is 0.75. At Unit D, the 50-64 years cancer detection rate calculated from the KC62 report is less than 40% of the minimum quality standard, but the SDR, which adjusts for the large number of younger women in the screening

population of unit D, is much closer to the minimum target, although still not achieving it. At unit E, on the other hand, the equivalent all ages ICDR gives the impression of a very effective screening unit, but the SDR, which takes into account the older population, falls short of the minimum. The 50-54 year ICDRs are, in each case, close to the equivalent SDRs in terms of the minimum standard, and it is questionable whether any additional information is gained by calculating both the 50-54 ICDR and the SDR particularly when the 50-54 year ICDR relies on an estimated correction factor of questionable reliability.

The effect of background geographical variation in incidence on SDRs has been examined nationally, and the correction factor needed to adjust for background variation estimated at a maximum of 15% (Blanks ***et al.***, 1996b). Although the influence of variation in background incidence on invasive cancer detection rates and SDRs is recognised by the NHSBSP as a problem, use of individual correction factors by Units is currently considered too complicated to be of value (Blanks, personal communication). However, with the variation in background incidence rates across the screening catchment areas of the West Midlands (illustrated in Chapter Four) ranging from 173.7/100,000 to 196.1/100,000 in 50-64 year olds, a maximum variation of 13%, the importance of such correction factors should not be under-estimated.

SDRs for incident screens are not currently calculated by the NHSBSP due to the impact of under-detection in previous rounds. For example, a Unit producing few false negative results in the first round might achieve a high prevalent SDR compared to a Unit with a higher false negative rate. At the next incident screen, however, the situation might be reversed, as the Unit with a low false negative rate will detect only genuine incident cancers, while the second Unit might achieve a higher SDR by also detecting cancers that were missed at the previous screen. The need for incident SDRs is increasing since the majority of the women currently being screened are incident, but until a method is devised that allows for this "overspill" effect, incident SDRs may be misleading, and the development of the incident SDR is ongoing (Blanks, personal communication).

6.2.3 The small cancer detection rate and non-invasive cancer (DCIS) rate

The prognostic characteristics of screen detected cancers can be used as surrogate end points to indicate whether screening is likely to be successful. Results of the Swedish Two

Counties Study showed a correlation between increasing size and increased nodal involvement, with size being the most important prognostic factor (Tabar *et al.*, 1992). Other factors such as grade are useful, but inter-observer variation in allocation of grade (Duffy *et al.*, 1991; Sloane *et al.*, 1994) may make comparison of absolute rates of tumours by grade misleading.

Definition of the small cancer detection rate has undergone several revisions since this measure was first introduced **(NHSBSP,** 1988). DCIS and micro-invasive cancers were included in the invasive cancer total until 1993 **(NHSBSP,** 1993b), but the 1996 QA Standards (Radiological Big '18', personal communication) exclude micro-invasives from totals for invasive cancers and include them with *in situ* cancers. The 1989 and 1993 standards defined "small" as <=10mm, but the 1996 standards define small as <10mm and 4 5 m m in line with European mainland practice. In addition the 1996 standards for small cancer rates are divided into prevalent and incident, and minimum and expected values.

The number of changes in this outcome measure have led to problems. Unless pre-1995 data are reproduced using KC62 version 2.8 no trends can be identified as previous version of the KC62 report used <=10mm to define small cancers rather than <10mm. Furthermore, calculated measures provided by the KC62 version 2.8 for small cancer detection rates are now incorrect as they include micro-invasives in the small cancer detection rates. The "new" measures will thus appear poorer than the previous KC62 version 2.8 measures if a unit has detected a significant proportion of micro-invasive cancers. Figure 6.6 demonstrates the effect of excluding micro-invasive cancers from the small cancer detection rate. At Units B, C, D, H, I and J, no micro-invasives were diagnosed and the change in definition makes no difference. At Unit A the detection rate of cancers <15mm when calculated with micro-invasives included exceeds the minimum standard, but fails to achieve it when micro-invasives are removed.

In addition to the absolute rate of small cancers detected, the proportion of small cancers is important (Day *et al.*, 1988). The proportion of small invasive cancers (<10mm and <15mm) detected in 1994/95 are shown in Figure 6.7. One Unit (unit D) achieved the recommended 30% target for cancers <10mm, and four units, units A, C, E and F met the 50% target for cancers <15mm.

Figure 6.6 **Small cancer detection rates (<15mm) for prevalent screens in the West Midlands 94/95**

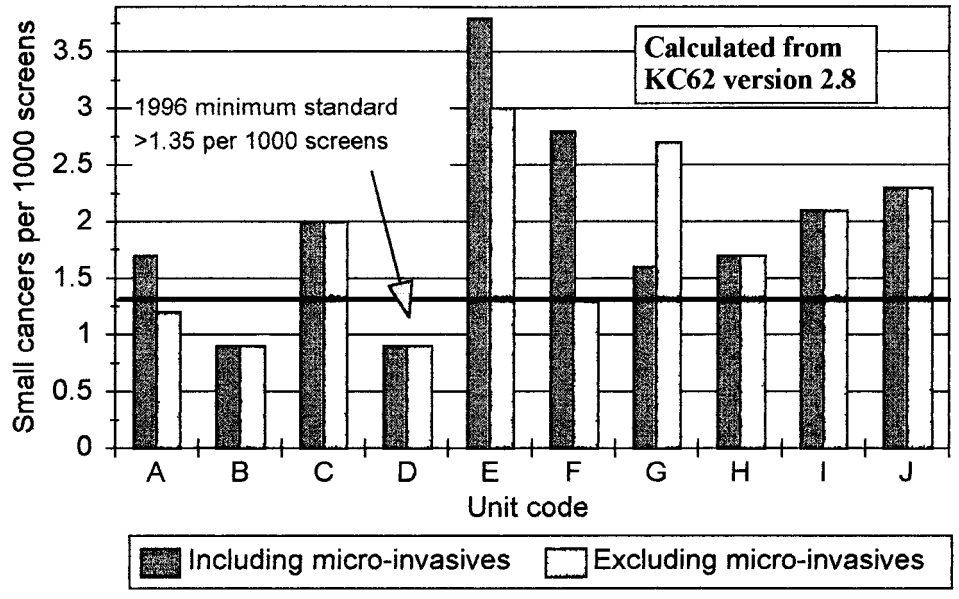
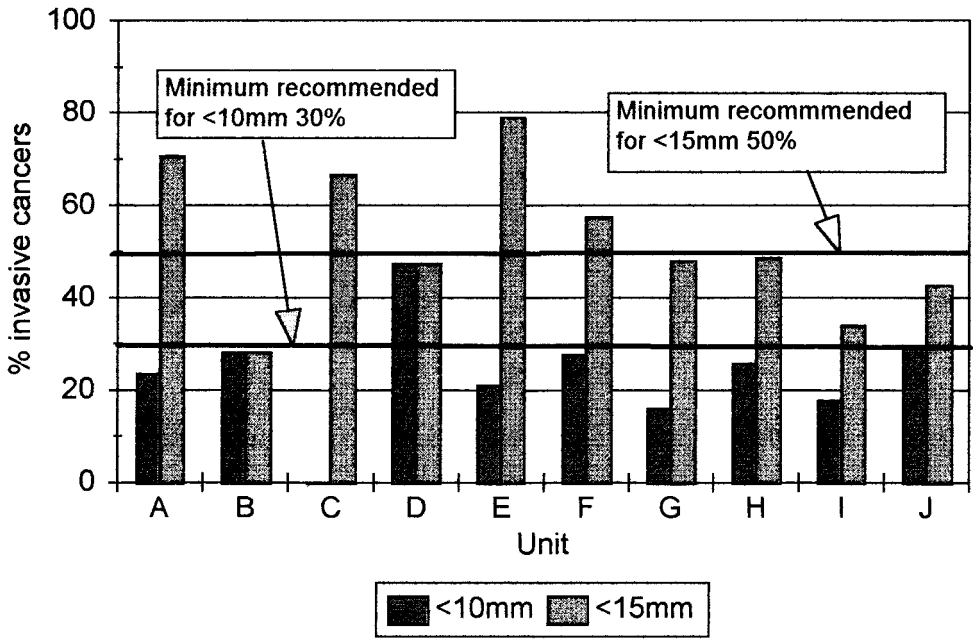


Figure 6.7 **Small invasive cancer proportions (<10mm and <15mm) for prevalent screens in the West Midlands 94/95**

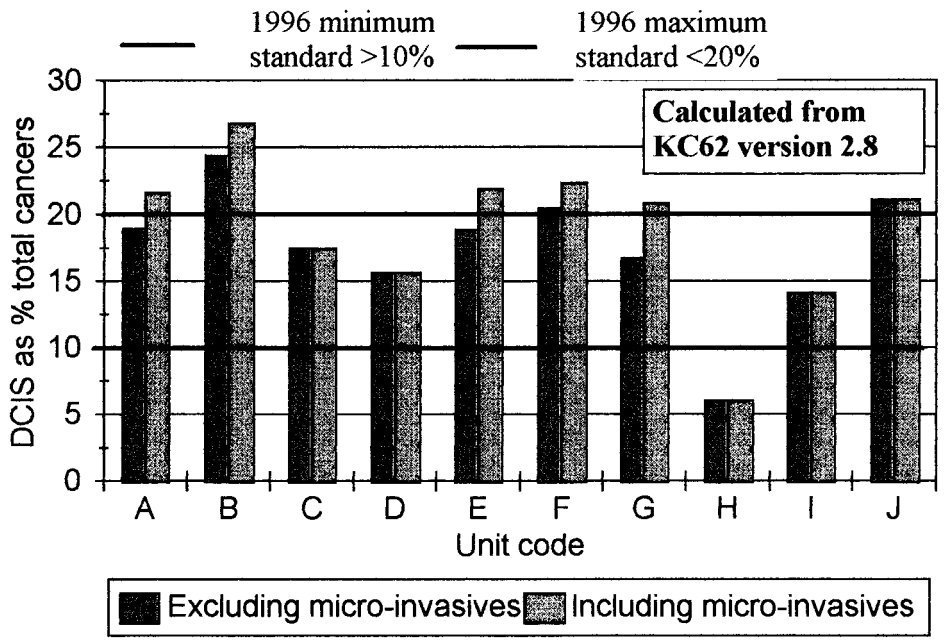


Although not currently routinely calculated as an outcome measure by the **NHSBSP,** the small cancer proportion is simple to calculate and provides additional information about the likely success of screening. To be successful, not only should large numbers of small

cancers be detected by screening, but large, more advanced cancers should be rarely found. Like the invasive cancer detection rate, the small cancer detection rate is dependent on the age structure of the screening population. A small cancer SDR has recently been developed and may replace the small cancer detection rate (Blanks *et al.,* 1995c).

The detection of non-invasive cancers is assessed by the DCIS rate, which is misleadingly named as it in fact includes all *in situ* breast cancers including the lobular variant (LCIS). The DCIS rate is not calculated by KC62 version 2.8. Recently the definition of the DCIS rate has been changed to include micro-invasive cancers. The 1996 QA standards stipulate a lower (>10%) and upper (<20%) limit for screen detected *in situ* disease. The upper limit is significant as a high DCIS rate may indicate over-diagnosis or misclassification of invasive disease. Figure 6.8 shows calculated DCIS rates for the West Midlands.

Figure 6.8 **DCIS rates +/- micro-invasve cancers for all screens in the West Midlands 94/95**



When micro-invasive cancers are included in DCIS rates units A, E, and G produce rates exceeding the maximum standard whereas the achieved rates excluding micro-invasives were within the recommended limits. Unit H detected less than the minimum rate of DCIS, and units B, F and J exceeded the maximum standard, regardless of the calculation method. A further difficulty stemming from the change in interpretation of micro-invasive

cancers is that cancer registries include micro-invasive cancers with invasive cancers, making comparison of data in relation to this outcome measure from the NHSBSP and cancer registries difficult. A possible solution to this difficulty that is being proposed for national use is the use of a separate behaviour code for micro-invasive tumours as specified by the SNOMED classification system (American College of Pathologists, 1980) by pathologists thus allowing their identification.

6.2.4 Outcome measures for surgical procedures

Measures using open biopsy data are calculated by QASS reports (described in Appendix Two) as well as from the KC62 version 2.8. An obstacle to the use of open biopsy data is that the KC62 version 2.8 does not distinguish between open surgical biopsies and other histological diagnostic procedures such as trucut or wide bore needle biopsy. Thus open biopsy data cannot be separated from pre-operative histology. Confusion has been created by the provision of conflicting standards for open biopsy rates and benign open biopsy rates in a variety of publications, as demonstrated in Table 6.1.

Table 6.1 Discrepancies between minimum standards set for open biopsy outcome measures in different NHSBSP publications

Measure	Surgical QA Guidelines	Surgical QA Guidelines (Updated)	NHSBSP Objectives 1993	Guide to Good Office Practice 1995	QA Standards 1996
	NHSBSP 1992	NHSBSP 1996a	NHSBSP 1993b	NHSBSP 1995b	Radiological 'Big 18'
Benign open biopsy rate	prevalent round <9.0 per 1000 incident round <6.0 per 1000	prevalent screens <1.5 per 1000 incident screens <1.0 per 1000	prevalent screens <5.0 per 1000 incident screens <3.5 per 1000	prevalent round <5.0 per 1000 incident round <2.0 per 1000	prevalent screens <2.7 per 1000 incident screens <2.0 per 1000
Open biopsy rate	<1.5% women screened	omitted	omitted	prevalent screen <3.5 per 1000 incident screen <1.5 per 1000	omitted

Not only has the misuse of the term "rounds" persisted in the 1995 "Good Office Practice" definitions, but the maximum target for benign open biopsies is greater than the maximum allowed total open biopsies rate. The updated QA Guidelines for Surgeons (NHSBSP, 1996a) has replaced the incorrect usage of "rounds" with the correct "screens", but omits to give a target for the open biopsy rate. The total open biopsy rate is not calculated by the KC62 version 2.8, cannot be derived from it, and is not included in the 1996 QA standards. Inconsistencies such as these highlight the need for clear and comprehensive

guidelines on the definition and use of outcome measures and targets, and especially those measures used by different professional groups within the NHSBSP.

Until 1996 the malignant:benign open biopsy ratio (M:B ratio) was used as an outcome measure, reflecting the need to minimise the use of benign biopsy, particularly for women without breast cancer. The M:B ratio is Calculated by the QASS report, is provided by the KC62 version 2.8 report and it is also possible to derive a M:B ratio for open biopsies from the QASS report. The updated version of the Surgical Guidelines (NHSBSP, 1996a) has withdrawn the use of this measure. The main reason for withdrawal is the difficulty of allowing for small numbers when assessing a calculated M:B ratio. An apparently poorer result may be produced for units which strives to minimise the total open biopsies performed, but produces an apparently better result for units undertaking a larger total of open biopsies, but for which the proportion of benign biopsies is lower. Furthermore, the M:B ratio is calculated by the QASS report using a different method that that used to produce the same measure by the KC62 report. It is dangerous to have two standard reports available, each of which calculates apparently the same outcome measure for the same time period, but which in fact use different definitions and produce different results. The positive predictive value (PPV) of open biopsy was also used in the past as an outcome measure but has also been withdrawn due to the effect of small numbers.

The avoidance of over-treatment is an objective of the NHSBSP. Surgical outcome measures relating to treatment can only be calculated from QASS reports. The 1992 QA Guidelines for Surgeons set a standard of >50% of cancers of <=15mm to be treated with conservative surgery, but the updated version published in 1996 omitted this outcome measure. Unfortunately the QASS was not designed to provide the data for this outcome measure, as treatment data for tumours <=15mm cannot be identified.

6.2.5 Early recall rate

The standard set for early recall rates attempts to ensure that use of this procedure is minimised. It is calculated as an overall rate, not split into prevalent and incident screens. The current target of 1% of those screened was achieved all units in the West Midlands in 1994/95. This standard was revised from an earlier standard of 0.25% which was

suggested in 1995 (Radiological 'Big 18', personal communication). In 1994/95 some West Midlands units produced early recall rates of more than 3 times this lower standard.

Placing women on early recall direct from screening should not occur, and instances of this procedure are monitored by the regional **QA** Committee. The early recall from screening rate is calculated by the KC62 version 2.8, but as all outcome measures are presented to one place of decimals, if a value of <0.04 is produced, the outcome measure is incorrectly presented as 0.0. In addition, the current definition of this measure exclude women appearing in Table D, who are invited to screening as a result of early recall. Although such women should not be placed on early recall again, this procedure has been recorded in the West Midlands. Excluding women appearing in Table D of the KC62 version **2.8,** who arrive at screening on early recall, from calculations masks this occurrence, which should be identified rapidly and investigated to minimise the likely distress caused by repeated invitations to early re-screen.

6.2.6 Rate of pre-operative diagnosis by cytology

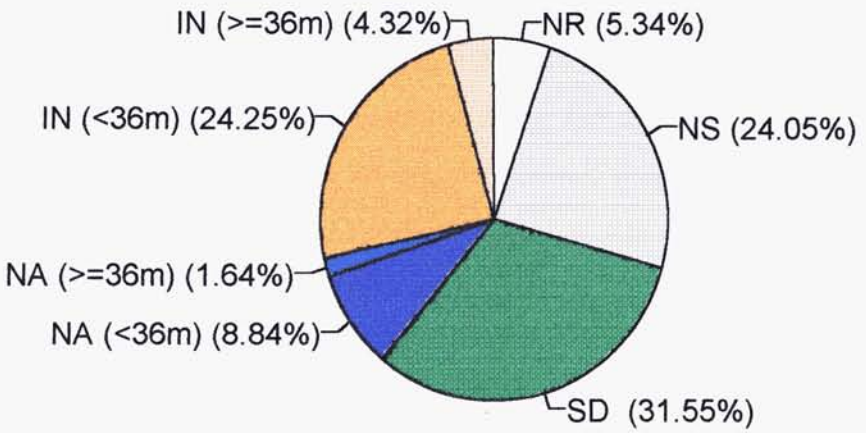
A target of >70% for pre-operative diagnosis by cytology or FNAC was included in the quality standards suggested by the Radiological 'Big 18' in 1996. In addition a minimum standard for the rate of pre-operative diagnosis of impalpable lesions by cytology and needle histology was set at **>30%.** However, the KC62 version 2.8 does not provide data describing whether lesions are palpable or impalpable, and does not provide the results of needle histology separately from those resulting from open biopsy. The only measure which can currently be calculated for the KC62 version 2.8 is the rate of pre-operative diagnosis by cytology for all cancers. The updated **QA** Guidelines for Surgeons (NHSBSP 1996a) set a standard for pre-operative cytological diagnosis of >60% for palpable and impalpable lesions. No unit achieved the lower standard of *>60%* in 1994/95, with some units demonstrating very low rates. It could be that such units use needle histology instead of fine needle aspiration cytology, making it difficult to draw conclusions on the basis of this outcome measure until a relevant standard is provided. However, it is unlikely that large differences can be accounted for in terms of needle histology **as** this procedure was known to be not extensively used in the West Midlands in 1994/95. To address this issue a national audit of pre-operative diagnosis of screen detected cancers has recently been undertaken by the Radiological 'Big 18'.

6.3 THE DERIVATION OF SCREENING HISTORIES

A total of 973 cancers were identified as "eligible" for detection by screening in the prevalent round at the Warwickshire, Solihull and Coventry Breast Screening Unit. The screening histories were allocated to each of these as shown in Table 6.2 and Figure 6.9.

Table 6.2 Screening history for cancers eligible for screen detection		
Screening history	**No. cancers**	**%**
Not registered at the unit	52	5.34
Not screened - excluded by GP	13	1.34
Not screened - invited after diagnosis	221	22.71
Screen detected at routine screen	289	29.70
Screen detected at GP/Self referral screen	10	1.03
Screen detected at early recall screen	8	0.82
Interval (<36 months from last screen)	236	24.25
Non-attender <36 months from last invitation	86	8.84
"Late invitation" interval (36 - 59.9 months from last screen)	42	4.32
Non-attender 36 - 59.9 months from last invitation	16	1.64
Total cancers	973	100

Figure 6.9 **Screening histories for breast cancers eligible for screen detection in the prevalent round**



NR Not registered **IN Interval**
NS Not screened **SD Screen detected**
NA Non attender

52 cancers arose in women who were not registered at the unit or at neighbouring units, suggesting inaccuracy or incompleteness of the FHSA population register. 13 cancers arose in women who were excluded from screening by their GP. Over 22% were diagnosed before the first screening invitation. Arrangements for screening in the prevalent round were unusual in that women aged 50 to ***64*** years were all invited for the first time, whereas in subsequent rounds those invited for the first time are largely aged 50-52.9 years. The increasing risk of breast cancer with age might therefore have been expected to result in higher incidence amongst older women awaiting screening compared to younger women. However rates of cancers diagnosed before screening were 30.9/10,000 in 50-54 year olds, 30.5/10,000 in 55-59 year olds, and 25.9/10,000 in 60-64 year olds. Thus the unusually high proportion of older women invited in the first round did not explain the relatively high number of cancers detected before the women could be screened. Incidence of "before screening" cancers rose over the first round from 17.9/10,000 in 1990, to 25.8/10,000 in 1991 and 33.8/10,000 in 1992. This might reflect increasing awareness of breast cancer over this period due to the NHSBSP and media interest causing women to present symptomatically before being called for screening.

A total of 307 cancers (31.55% of the total eligible) were screen detected, of which **94.1%** were found at routine screens, 8 at early recall, and 10 cancers were found at screens resulting from a self or GP referral in women eligible for routine screening. There were a total of 278 interval cancers including 236 arising at <36 months from the last screen, and a further 42 "late invitation" intervals arising at >36 months from the last screen in women too old to be re-invited, or who, although eligible for re-invitation, did not receive anther appointment prior to diagnosis.

Of the 102 cancers diagnosed in non-attenders, 86 arose within 36 months of the last invitation, and a further 16 cancers were diagnosed at 36-59.9 months from the last invitation. Of the 86 non-attenders diagnosed <36 months, 13 (15%) were diagnosed within **3** months of invitation. Although the NHSBSP predicts the occurrence of "programme provoked" cancers (NHSBSP, 1993a), no definition of these cancers is currently available. Given the short interval between the invitation and diagnosis it is possible that, in at least some of this 15% of cancers, the invitation precipitated presentation into the symptomatic breast service.

6.4 AN EXAMINATION OF THE DEFINITION AND CHARACTERISTICS OF INTERVAL CANCERS

6.4.1 The effect of exclusion criteria on interval cancer rates

Figure 6.10 shows the means by which, of 278 possible interval cancers, 213 were identified as "core" intervals. 86% of "core" intervals were known to the unit; the additional 30 being identified after cross-reference with records at the WMCIU. Table 6.3 gives further details of the 65 cancers excluded from the core set and the 11 '*core" intervals detected in "high risk" women.

Figure 6.10 Decision tree used in the definition of "core" interval breast cancers

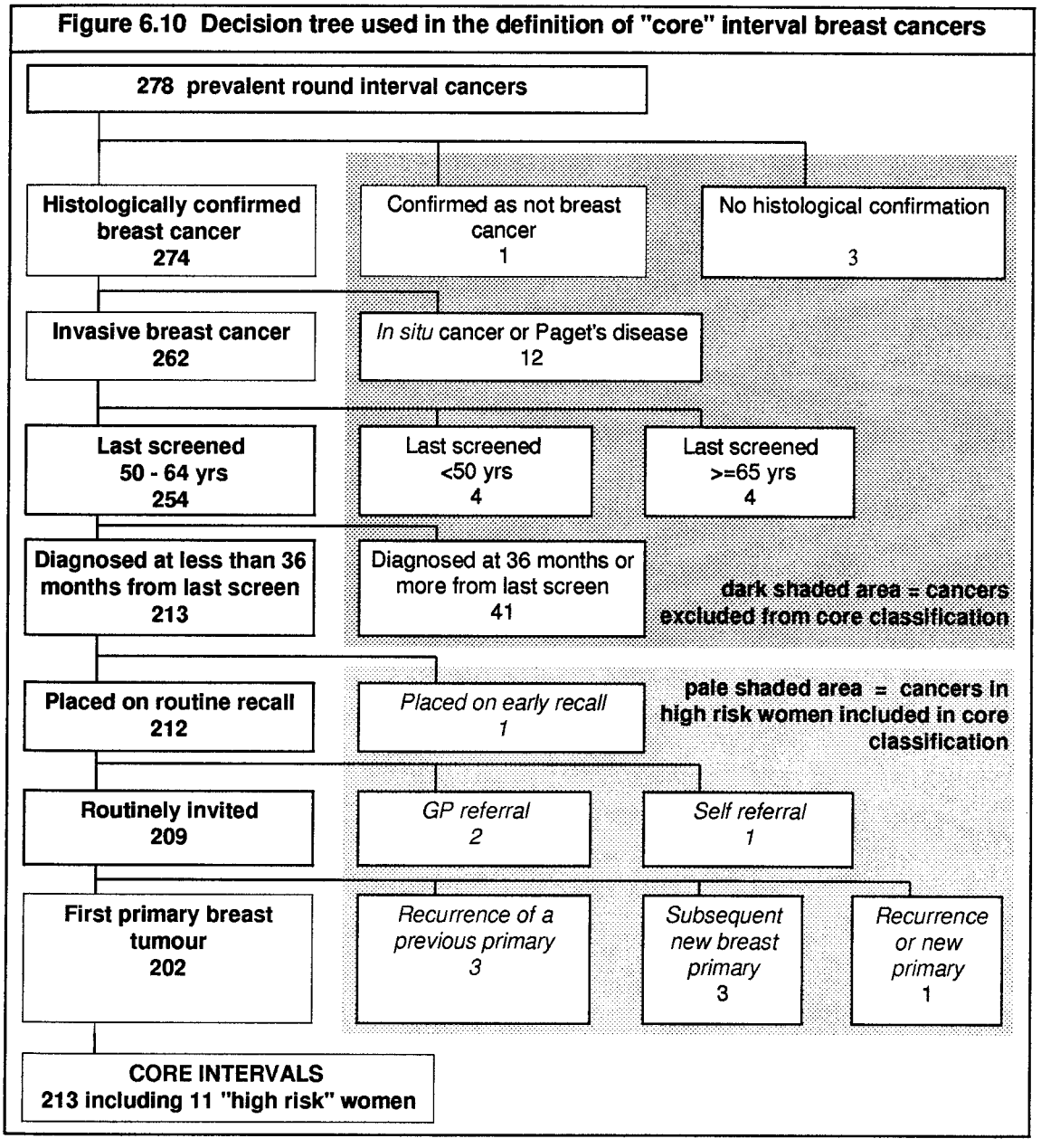


Table 6.3 Cancers excluded from core interval classification or in "high risk" women			
Exclusion criterion	**Total excluded**	**Other exclusion criteria or risk factors included in total**	**Included in analysis**
Not confirmed as breast cancer	4	1 confirmed lung cancer metastasis 3 without histological confirmation	No
Non-invasive cancer	12	10 *in situ* breast carcinoma. 2 Paget's disease (one at >36 months)	No
Underage at screening (<50 yrs)	4	1 subsequent new breast tumour, 1 recurrence, 1 GP referral 1>36 months	No
Overage at screening (>= 65 yrs)	4	3 self referrals	No
Interval screen to diagnosis >= 36 months	41	1 recurrence 1 GP referral, 1 possible recurrence	No
Risk factor type	**No. cancers**	**Comments**	
Early recall	1	-	Yes
Referral type	3	2 GP referrals, 1 self referral	Yes
Not first breast lump	7	3 subsequent new breast primaries, 3 recurrences, 1 possible recurrence	Yes

Four cases were due to lack of histological confirmation, or because the lump was not a breast cancer. Twelve non-invasive cancers were excluded, including 1 case that arose at 36 months or more. Of 41 cancers excluded on the grounds of an interval of 36 months or more, 10 had been identified as intervals by the Unit and none had been re-invited prior to diagnosis. A further 8 cases were excluded on the grounds of age at screening.

The effect of exclusion criteria on calculated interval cancer rates is shown in Figure 6.11. If no exclusions are applied, the rate of interval cancers arising from the first round was 47.1/10,000 women. Exclusion of cancers without histological confirmation, those found not to be breast cancers, non-invasive cancers and cancers in those not aged 50-64 years each produced small reductions in the interval rate. The largest drop, from 43.1/10,000 to 36.1/10,000 was produced by restricting interval cancers to those diagnosed at <36 months from the last screen. The rate produced using the most restrictive criteria (34.1/10,000) was 27.6% lower than the highest rate from these data. To ensure consistency it is vital that everyone calculates interval cancer rates using the same set of criteria. It is also important that the exclusion criteria used do not artificially reduce the apparent rate of interval cancers as is currently the case where only histologically verified cancers diagnosed within 36 months are included in most studies (Woodman *et al.,* 1995; Day *et al.,* 1995; Asbury *et al.,* 1996).

Figure 6.11 **Effect of additive exclusion criteria on prevalent round interval cancer rates**

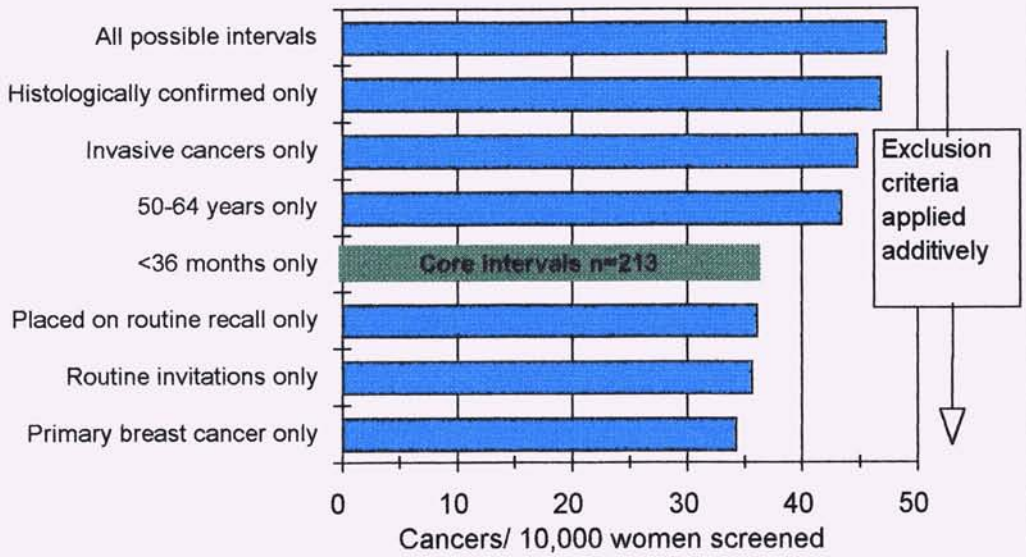


Figure 6.12 **Impact of changing definition of interval cancers on achieving QA standards**

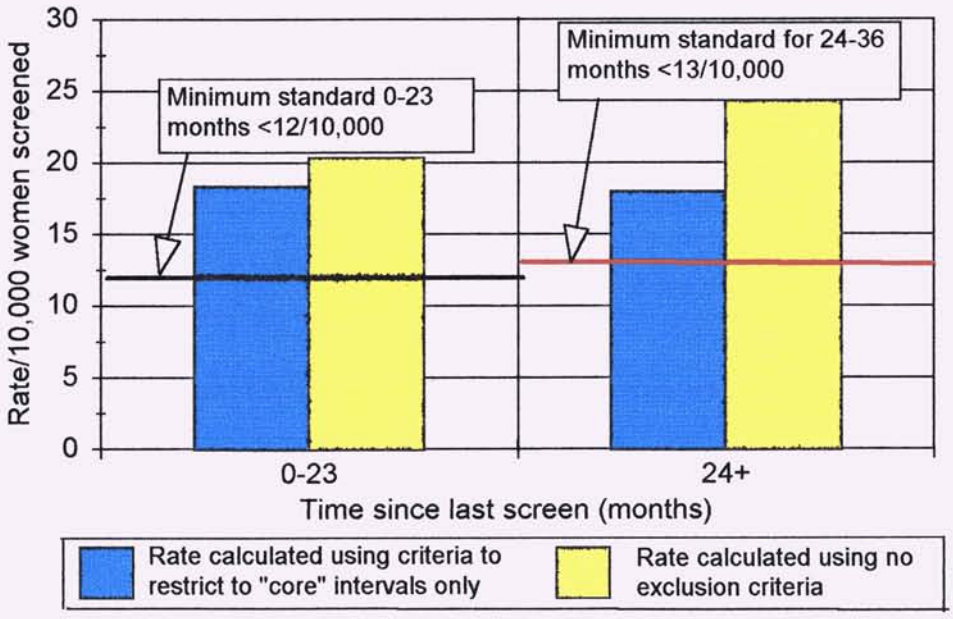


Figure 6.12 demonstrates the impact of changing the definition of interval cancers in relation to the current quality standards. These standards are <12/10,000 women screened occurring up to 23 months from the last screen, and <13/10,000 in the third year after

screening (Radiological Big '18', personal communication). This second standard assumes that screening units are universally operating a 36 month screening round. It is clear from Figure 6.12 that not only are core interval rates in excess of the standards, but if "non-core" intervals are included, the observed interval cancer rate occurring after 24 months **is** over twice the current standard in the unit examined in this study.

6.4.2 Variation in core interval cancer rates with time after screening and age

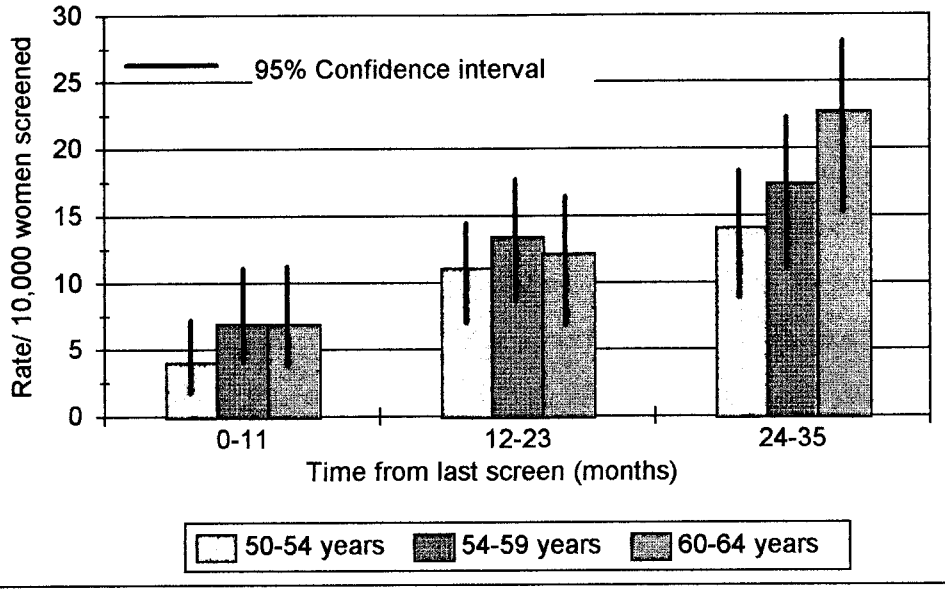
The overall core interval cancer rate was 36.1 per 10,000 women screened. Taking only these core interval cancers into consideration, Table 6.4 shows that, as in other studies, interval rates increased with the t h e from the last screen (Tabar *et al.,* 1987; Peeters *et al.,* 1989; Woodman *et al.,* 1995).

Table 6.4 Core interval cancer rates arising from the first round by year of screening and time from last screen to diagnosis

Year of last screening	Time since last screen to diagnosis (months) (number and rate per 10,000 women screened)						Total 0-35 months	
	0-11		12-23		24-35			
	No	Rate (95% CI)	No	Rate (95% CI)	No	Rate (95% CI)	No	Rate (95% CI)
1st April 1989 - 31st March 1990	10	6.2 (2.4 - 10.1)	25	15.5 (9.5 - 21.6)	24	14.9 (9.0 - 20.9)	59	36.7 (27.3 - 46.0)
1st April 1990 - 31st March 1991	14	6.5 (3.1 - 9.9)	30	13.9 (8.9 - 18.8)	45	20.8 (14.7 - 26.8)	89	41.1 (32.6 - 49.6)
1st April 1991 - 31st March 1992	11	5.2 (2.1 - 8.2)	17	8.0 (4.2 - 11.8)	37	17.4 (11.8 -23.0)	65	30.6 (23.1 - 38.0)
Total 1989 -1992	35	5.9 (4.0 - 7.9)	72	12.2 (9.4 - 15.0)	106	18.0 (14.5 - 21.4)	213	36.1 (31.3 - 40.9)

Incidence in 50-64 year olds in the resident population rose by 1% per year from 1980-87, a smaller rise than the 2% per year seen nationally (Quinn and Allen, 1995). The underlying incidence rate in 1992 in the catchment area of the unit studied was estimated at 21.5/10,000. **As** a proportion of the underlying incidence, core interval cancer rates were 27.6% at 0-11 months, 56.7% at 12-23 months, and 82.5% at 24-35 months from the last screen, similar to the North West (Woodman *et al.,* 1995) and East Anglia (Day *et al.,* 1995). Rates of core interval cancers by age at screening are shown in Figure 6.13.

Figure 6.13 **Incidence of "core" interval cancers by time since, and age at, last screening**



Interval cancer rates varied little with age in the first two periods, but in the 23-35 month period, rose from 14.1/10,000 in 50-54 year olds and 17.4/10,000 in 55-59 year olds to 22.7/10,000 in 60-64 year olds, although the differences were not statistically significant.

6.4.3 Radiological classification of core interval cancers

Table 6.5 describes the radiological classification of "core" interval cancers.

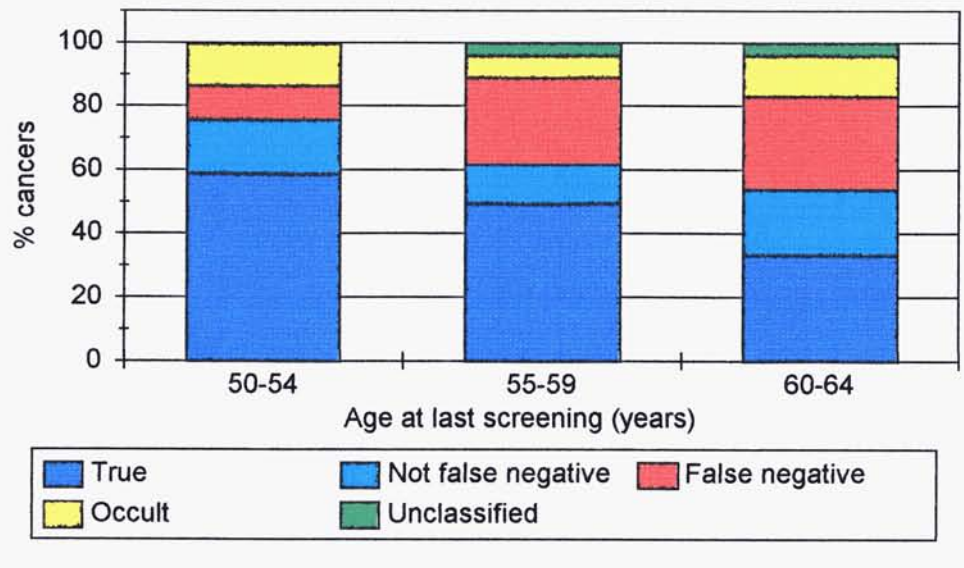
Table 6.5 Radiological classification of core intervals by time from last screen

Radiological classification	Time since last screen to diagnosis (months) (number and rate per 10,000 women screened)						Total 0-35 months	
	0-11		12-23		24-35			
	No	Rate (95% CI)	No	Rate (95% CI)	No	Rate (95% CI)	No	Rate (95% CI)
False negative	12	2.0 (0.9 - 3.2)	13	2.2 (1.0 - 3.4)	23	3.9 (2.3 - 5.5)	48	8.1 (5.8 - 10.4)
True interval	7	1.2 (0.3 - 2.1)	35	5.9 (4.0 - 7.9)	52	8.8 (6.4 - 11.2)	94	15.9 (12.7 - 19.1)
Occult	7	1.2 (0.3 - 2.1)	6	1.0 (0.2 - 1.8)	8	1.4 (0.4 - 2.3)	21	3.6 (2.0 - 5.1)
"Not false negative"	3	0.5 (0.0 - 1.1)	15	2.5 (1.3 - 3.8)	15	2.5 (1.3 - 3.8)	33	5.6 (3.7 - 7.5)
Unclassifiable	6	1.0 (0.2 - 1.8)	3	0.5 (0.0 - 1.1)	8	1.4 (0.4 - 2.3)	17	2.9 (1.5 - 4.3)
Total	35	5.9 (4.0 - 7.9)	72	12.2 (9.4 - 15.0)	106	18.0 (14.5 - 21.4)	213	36.1 (31.3 - 40.9)

Only 22.5% of core interval cancers were false negative, 44.1% were true interval cancers and 9.9% were radiologically occult. A further 15.5% were classed as "not false negative" on the basis of the screening mammogram alone, and only 8.0% remained unclassifiable due to lack of diagnostic mammograms. These proportions are similar to those identified elsewhere (Simpson *et al.*, 1995; Peeters *et al.*, 1989). As expected, rates of false negative and occult interval cancers roses only slightly over the three year period, but rates of true interval cancers rose steadily. When only false negative and true core interval cancers were considered, the proportion of true interval cancers rose from 36.8% in the first 12 months and 72.9% at 12-23 months to 69.3% in the 24-35 month period, proportions which are similar to those reported from the North West (Asbury *et al.*, 1996).

The radiological classifications of core interval cancers by age group is shown in Figure 6.14. The highest false negative proportion was seen in the oldest women, making up 29.7% of cancers in 60-64 year olds compared to 10.3% in 50-54 year olds. The highest proportion of true interval cancers was seen in the youngest women, in whom 55.2% of cancers were true intervals.

Figure 6.14 **Radiological classification of "core" interval cancers by age at screening**



The total rate of true interval cancers arising in the first year after screening, made up of the true interval cancers identified, plus a proportion of the "not false negative" cancers, was estimated at 1.45/10,000 using the method described in Chapter Two section 2.4.2.

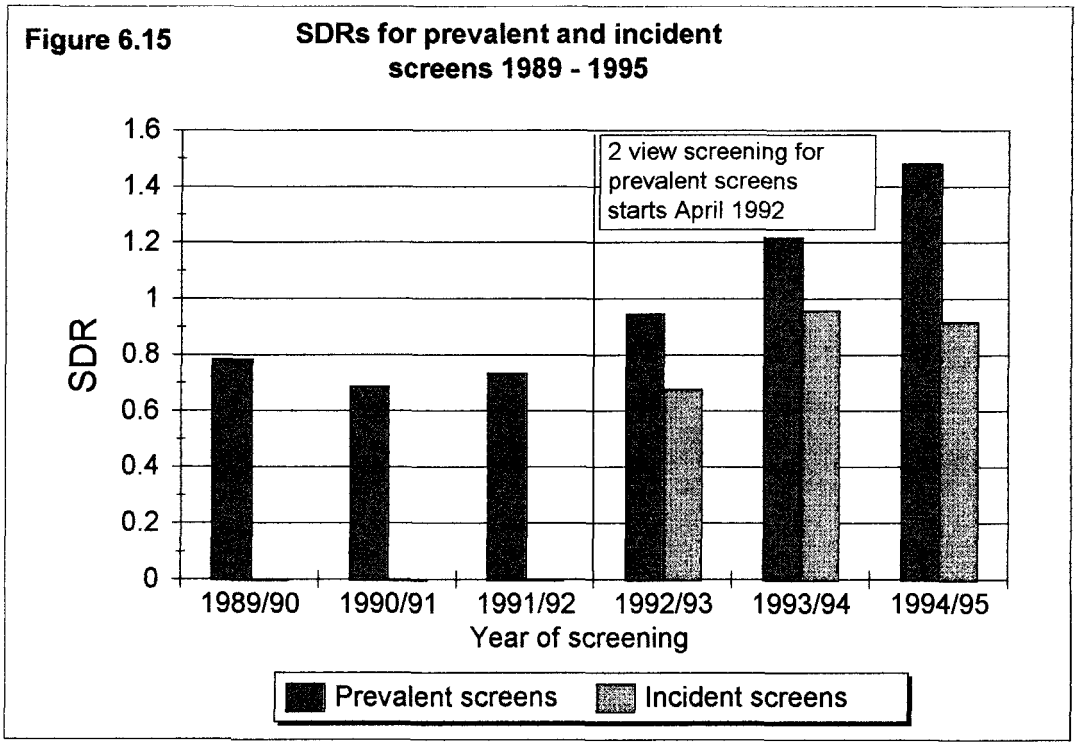
Assuming that this total rate of true interval cancers remained constant with time, an estimated 79 cancers might have detectable at two years from the last screen. Thus the proportion of core interval cancers that could have been detected at two years was estimated at 37.3%. Operating a two year screening round might thus bring about a **37.3%** reduction in the interval cancer rate based on the data used in this investigation.

6.4.4 Other factors influencing core interval cancer rates

Any improvement in screening sensitivity relative to underlying incidence will reduce the interval cancer rate. This includes true intervals, since the classification "true" does not mean that the cancer was necessarily undetectable at the last screen, but that it was not detected by the method used. Techniques which improve sensitivity, such as two view screening which is already used for prevalent screens and which has been shown to increase the detection of small cancers (Blanks ***et al.,*** 1996c), may allow such cancers to be detected by screening and thus not arise as interval cancers.

Figure **6.15** shows **SDRs** for first round prevalent screens, and for prevalent and incident second round screens (1 April 1992 - **31** March 1995) at the Warwickshire, Solihull and Coventry Breast Screening Unit.

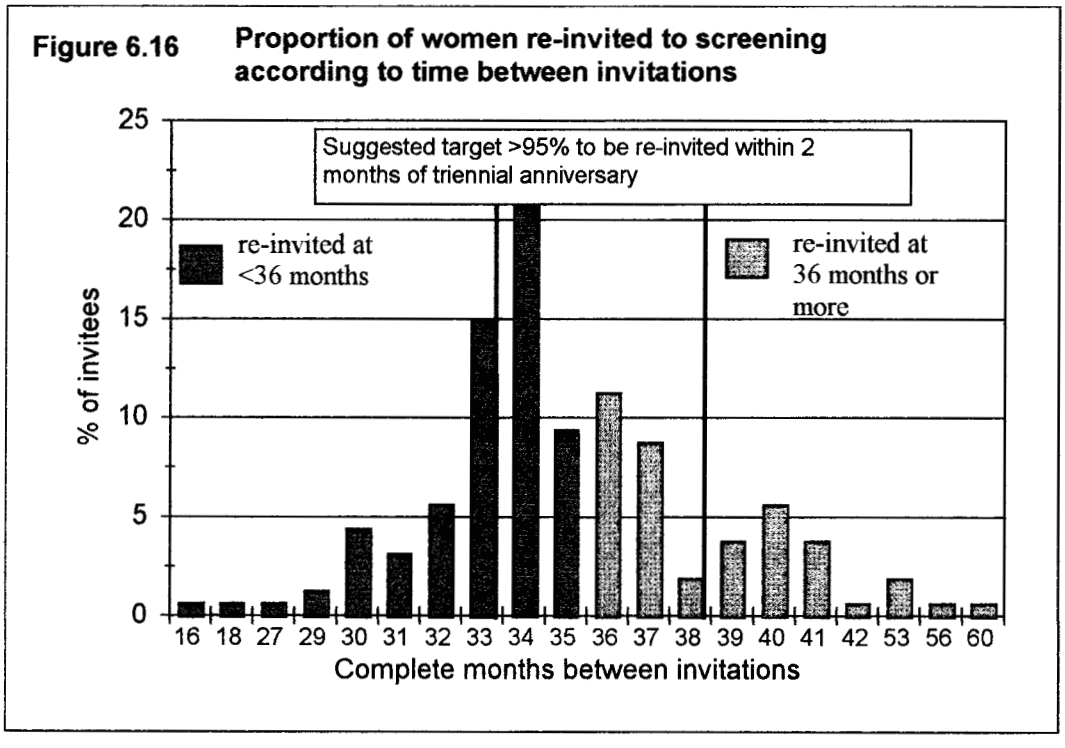
Figure 6.15 **SDRs for prevalent and incident screens 1989 - 1995**



Prevalent **SDRs** rose from an average of **0.74** in the first round when single-view mammography was used, to **1.48** in the third year following the introduction of two-view mammography in April **1992.** Although a proportion of this increase is undoubtedly due to the use of two-view mammography, from **1992** the parallel but smaller rise in incident **SDRs** is consistent with an underlying, general improvement in screening sensitivity as the programme matured. Fewer interval cancers would thus be expected following screens in the second round from this unit without any changes to the round length.

Figure **6.16** describes the time between screening invitations for **155** women with incident second round screen-detected cancers. **A** quality standard suggested by the **NHSBSP** is that **>95%** of women should be re-invited within **2** months of the tri-ennial anniversary (Radiological Big '**18',** personal communication). The average screening round for the unit was **35.8** months, but only **83 (51.9%)** were re-invited within the **2** month target zone. Of the 155 women, **72.5%** were re-invited within ***36*** months of the last screen, with **27.5%** re-invited at **36** months or more, and **17.3%** re-invited at **>39** months.

Figure 6.16 **Proportion of women re-invited to screening according to time between invitations**



Re-screening at regular intervals is critical to success of any screening programme aimed at detecting diseases of insidious onset (Wilson and Jungner, **1968).** The Forrest Committee decided that a three year screening round was appropriate, but stressed that

more research was required in this area (HMSO, 1986). Running a population based screening programme so that all women are invited and re-invited at three-year intervals is, however, extremely difficult. Even if the average round length for a unit is 36 months, the round length for individual women may vary substantially, so that some women experience much longer or shorter screening rounds. Furthermore, data with which to evaluate round lengths for individual women are not easily available.

6.5 SUMMARY

The NHSBSP **QA** Service has developed a number of outcome measures to act as performance indicators and early indicators of success for the NHSBSP. Many are calculated by the "KC62" report software used to collate screening activity data for the Department of Health. This investigation identified a number of problems associated with these outcome measures which reduce their usefulness.

some measures calculated by the KC62 reports are inconsistent with the methods recommended by the NHSBSP, for example small cancer detection rates are currently calculated excluding micro-invasive cancers, but these cancers are included in the small cancer detection rates calculated by KC62 reports

- some measures that could be Calculated from the raw data provided by a KC62 report are not calculated, for example the DCIS rate

  some calculated measures cannot be reproduced from raw data on KC62 reports, for example the benign open biopsy rate

  some measures used by the NHSBSP are not calculated by, and cannot be derived from KC62 reports, for example the open biopsy rate
- some measures which are no longer recommended by the NHSBSP are provided by KC62 reports, for example the malignant:benign open biopsy ratio

  changes in the recommended definitions used to calculate outcome measures have been a source of confusion, in particular in relation to the invasive cancer detection rate. Current recommendations are to calculate the prevalent invasive cancer detection rate for women aged 50-54 years only even though a significant proportion of women at attending some units are not in this age group. In addition, it is assumed that most of the women in the 50-54 year age group are 50-52.9 years old, although the validity of this assumption cannot be assessed from a KC62 report,

and in order to perform the calculation, the number of invasive cancers detected in women aged 50-54 must be estimated since it is not supplied by the KC62 report

- the SDR attempts to adjust invasive cancer detection rates for the age distribution of the screened population. However, the SDR is currently only calculated for prevalent women, who now make up a small proportion (around 20%) of those screened at many units
- as well as changes to the definitions of outcome measures, some of the quality standards set for these measures have undergone repeated revisions as the **NHSBSP** has progressed, leading to more confusion, for example in the case of cancer detection rates. In addition, for some measures such as the benign open biopsy rate, inconsistency exists in the quality standards set by different professional groups within the NHSBSP
- some quality standards have remained unchanged although needing revision, for example the uptake rate which has remained unchanged since 1988 at 70%. When 70% uptake is applied to a hypothetical screening population over five rounds, only 16.8% of women attend all five screens, and only **36%** a total of four screens. To improve the utility of this outcome measure and help to ensure high cumulative uptake, a higher quality standard for re-uptake rates is required and it must be made possible to identify which screening round women were last invited and/or screened

To assess the impact of the NHSBSP on a selected West Midlands population a retrospective technique was devised to identify a cohort of women with cancers eligible for screen detection in the prevalent round and to allocate screening histories to the 973 cancers in this group.

- only 31.6% of cancers were screen detected. 94.1% of these were found at routine screens, 3.2% as a result of self or GP referral and 2.6% at an early recall screen
- 22.9% were diagnosed before the screening invitation. Incidence rates in this group were similar according to age at invitation, but rates rose over the screening round
- 28.6% of cancers arose as interval cancers after a negative screen, of which 84.9% arose at <36 months from the last screen, and 15.1% were diagnosed at >=36 and ***<60*** months from the last screen, in women either ineligible for re-invitation due to age, or in women who, although eligible, failed to receive a re-invitation prior to diagnosis

- 10.5% arose in non-attenders of which 15% were diagnosed within 3 months of the invitation date and which could be "programme provoked" cancers
- 5.3% arose in women never invited for screening, indicating that the accuracy and completeness of the population register needs to be improved

An examination of the definition and interpretation of the interval cancers identified that the definition of interval cancers needs classification. In this investigation

- using the most inclusive criteria, the rate of interval cancers arising from the prevalent screening round was 47.1/10,000 women screened
- of the 278 interval cancers identified, 10.8% were unknown to the screening unit and were only identified after cross-matching data from the WMCIU and the screening unit
- by excluding cancers without histological confirmation, non-invasive cancers, cancers arising in women not aged 50-64 years at screening, and cancers arising at >=36 months from the last screen, the rate fell by 23% to 36.1/10,000. The largest contribution to this drop came from exclusion of cancers arising at >=36 months from the last screen
- interval cancers fulfilling these criteria were termed "core" intervals
- rates of core interval cancer rose from 5.9/10,000 at 0-11 months, to 12.2/10,000 at 12-23 months and 18.0/10,000 at 23-35 months from the last screen
- expressed as proportions of the expected underlying incidence, estimated as 21.5/10,000, the incidence of core intervals rose from 27.6% at 0-11 months, to 56.7% at 12-23 months and 82.5% at 23-35 months from the last screen
- in the third year after screening the highest rate of core intervals was identified in women aged 60-64 years at screening
- radiological classification of core intervals identified 44.1% true intervals, 22.5% false negative and 9.9% occult. Lack of diagnostic mammograms meant that 15.5% could only be classified as "not false negative" and a further 8.0% were unclassifiable
- from this investigation it was estimated that moving to a two year round would lead to a 37.3% drop in interval cancers

Although a shorter screening round would reduce the rate of interval cancers, improvement in screening sensitivity would also lead to reduced interval rates including

rates of true interval cancers. SDRs have risen for the unit studied from 1989 to 1995, suggesting improvements in sensitivity including that brought about by the introduction of two-view screening in 1992. In addition, although the NHSBSP assumes that a three year round is currently in operation, this is often not the case. In this investigation, despite an average round length for the unit of 35.8 months, only 51.9% of women were re-invited to screening within two months of the triennial anniversary, with 27.5% re-invited at 36 months or more from the last screen. High interval cancer rates may not indicate that a three year round is insufficient, but that the three year round is not being maintained, and although a two year round would reduce interval cancers rates, there are a number of alternative solutions to high interval cancer rates might be more effective and less costly than shortening the screening round and which should be explored before implementing a two year screening round.

CHAPTER SEVEN

DISCUSSION AND CONCLUSIONS

7.1 IMPROVING THE AVAILABILITY AND VALIDITY OF ROUTINELY COLLECTED PROGNOSTIC DATA

7.1.1 The availability and validity of prognostic data recorded at the WMCIU

The WMCIU is ideally placed to undertake population-based studies of breast cancer in the West Midlands since it is unique in recording detailed data for all breast cancers diagnosed in the region. For evaluations to be valid, the availability of data must be high and data recorded at the WMCIU must be complete and accurate. This study demonstrated that it was broadly possible to undertake examine the main issues relating to breast cancer using routinely collected data. A number of problems associated with routinely recorded data were identified, and a range of solutions were identified which not only improved the usefulness of data available for this study, but will result in improvements in the quality of data recorded at the WMCIU in the future.

The completeness and accuracy of recording tumour characteristics at the WMCIU was found to be generally high. A high proportion of histologically verified cases, and a low proportion of death certificate registrations are indicators of high data validity (Parkin et ***al.*, 1994).** It was thus encouraging to find that, of the 5077 breast cancers examined, only **9.6%** had no histological confirmation, and only **1.4%** of cases were registered from a death certificate or post mortem report only. This compares with the South East where, even in **1991, 17%** of breast cancers had no histological verification and **13%** were diagnosed from death certificates only (Thames Cancer Registry, **1994).** With the improvements in data coding and storage introduced at the WMCIU as a result of this study, the data recorded at the WMCIU were considered to be of sufficiently high quality for analytical purposes despite the limitations imposed by missing data.

An audit of staging data recorded at the WMCIU revealed that less than 5% of breast cancers diagnosed in 1983 had stage recorded on the WMCIU's old computer database. This, however, was only a reflection of the similarly small proportion of cases with stage

recorded in the case notes. Although the completeness of the recording of stage at the WMCIU was poorer than for other types of prognostic data, the accuracy with which staging data were recorded was very high at 96.3%. The main reason for absence of recorded stage at the WMCIU was the failure of Trusts to send staging data to the WMCIU. Lack of recorded stage at cancer registries is a well recognised problem. Only 0.89% of breast cancers diagnosed in 1982-86 in the South West region were staged, and 0.28% in the Wessex region (Bristol Cancer Epidemiology Unit, 1995). In the South East only 24% of 334 breast cancers diagnosed in 1990 had stage recorded (Chouillet *et ai.,* 1994). In addition, a study at the Thames Cancer Registry (TCR) comparing data recorded for bladder cancers at the TCR with data from hospital notes concluded that, although data recorded by the TCR were generally reliable, stage was largely unreliable due to missing data in the hospital notes and misclassification of recorded stage at the TCR (Gulliford *et al.,* 1993). The problem of missing and unreliable staging data is thus not restricted to the WMCIU or to breast cancer.

The methods used to classify and store prognostic data can have a large impact on the utility of the data recorded. **A** significant problem identified in the recording of tumour characteristic data on the old WMCIU database was misuse of the codes for local node status. This meant that cases for which no information about the local nodes was available in the notes were recorded as "nodes negative", resulting in cases with unknown nodal status being indistinguishable from those for which the nodes were recorded in the notes as negative. The poorer survival for cases with unknown nodal status compared to those for which nodes were known to be negative suggests that at least some of these cases had positive nodes, and demonstrates that the grouping of cases in this way limited the usefulness of prognostic data recorded on the old WMCIU database. The development of the new WMCIU database GRACE during the course of this project provided the opportunity for improvements in coding rules and to the database structure identified as a result of this study to be implemented.

Of the 2444 breast cancers diagnosed in 1983 in this study, 34.9% had no size and 28.6% had no nodal status recorded in the case notes. These proportions are similar to those reported in Scotland, where of 3786 breast cancers diagnosed in 1980-88, 28.6%

had no size, and 33.1% no nodal data recorded (Gillis and Hole, 1996). In the present study, 69.5% of cases had no grade recorded. This lack of grading has also been reported in Yorkshire, where, of 12,861 breast cancers diagnosed between 1979 and 1988, grade was not recorded for 47% (Sainsbury et al., 1995b). The lack of these prognostic data in the present investigation reflects lack of data in the notes since, in 745 of 1387 cases (53.7%) without a pathological size, and 476 of 1322 cases (36.0%) without pathological node status recorded, the pathology report sent to the WMCIU did not record these prognostic data. The lack of data in the West Midlands was therefore seen to stem partly from a lack of pathological investigation and/or data recording by Trusts as well as failure to supply data to the WMCIU. While improvements to coding rules and database design have enhanced the utility of data recorded by the WMCIU, the potential of data recorded is restricted by the proportion of missing data.

Various methods of increasing the staged proportion have been used by cancer registries. Manual combination of recorded tumour characteristic data to retrospectively allocate stage has been undertaken in the South East, where the staged proportion amongst 338 breast cancers diagnosed in 1990 rose from 24% to 47.4% (Choiullet ***et al.,*** 1994). In the Wessex region, peripatetic data collection officers have been employed since 1987 to visit hospitals and collect data with which to allocate stage to breast cancers using a defined protocol (Wessex Cancer Intelligence Unit, 1992). This process has led to a rise in the staged proportion from 0.28% for 1982-86 to 62.5% in 1987-91 (Bristol Cancer Epidemiology Unit, 1995). The high completeness and accuracy of tumour characteristic data identified in the present study allowed the development of a staging algorithm combining tumour characteristics to allocate stage to breast cancers at the WMCIU. This resulted in 63.9% of the 2444 cases diagnosed in 1983 being staged, a 13-fold increase in the staged proportion. The algorithm provided a means of routinely allocating a stage for breast cancers recorded at the WMCIU and was able to produce a significant rise in the staged proportion. The algorithm needs to be tested in other populations, and its utility was restricted by the completeness of data sent to the WMCIU, in particular with regard to local node status and tumour size, necessitating the incorporation of a number of assumptions which may not always be justified. However, the availability of the staging algorithm at the WMCIU, in conjunction with improved

coding and data storage on GRACE, allows large series of breast cancers to be staged quickly and consistently. The same methodology is to be introduced for other cancers, based on the pilot carried out is this study of breast cancer.

The ideal solution to lack of staging data at cancer registries would be for stage to be ascertained and recorded for all breast cancers by a clinician at diagnosis, and for these data to be transferred to the WMCIU. However, until this is the case, strategies to allocate stage, such as the staging algorithm, have an important role to play. There are two barriers to recording prognostic data at the WMCIU - failure to ascertain and/or record prognostic data at diagnosis, and failure to supply data to the WMCIU. In order to meet the requirements of the Minimum Data Set for Cancer Registration (MDSCR), not only must data supply to cancer registries be efficient, but Trusts must collect and store complete staging data in an accessible form. In response to the Calman-Hine report (HMSO, 1995), a Regional Cancer Implementation Team was set up in 1995 to visit those Trusts in the West Midlands applying for Cancer Unit and Cancer Centre status. The Cancer Implementation Team considered the availability and quality of information collection systems as a key quality criterion. The importance of data collection has been further recognised in the Clinical Outcome Group (COG) guidelines for breast cancer (NHS Executive, 1996a) produced in response to the Calman-Hine report. These guidelines recommended that purchasers of breast cancer services should monitor patient outcomes, stipulating that staging data were desirable, and pointing out both the importance of cancer registries and the need to improve information collection systems.

The use of peripatetic data collectors in Wessex to improve the availability of data, including stage, for breast and colorectal cancers is unfeasible in the West Midlands due to the large geographical size and population density of the region. Furthermore, in Wessex, even with access to the hospital notes, 37.5% of breast cancers remained unstaged because staging data were not recorded by clinicians (Bristol Cancer Epidemiology Unit, 1995). **A** different approach is being applied in the West Midlands as part of the Calman-Hine implementation process. A region-wide oncology information system is currently being set up to allow direct supply of clinical, pathological and treatment data from Trusts to the WMCIU. **In** addition, several

projects are currently being undertaken aimed at improving the availability of pathological data. The WMCIU is planning to ask pathology laboratories to send special reports detailing additional nodal procedures to the WMCIU in an attempt to improve the completeness of nodal data recording. An audit of pathology reports sent to the WMCIU is being carried out to assess the completeness and quality of a range of prognostic data including size, nodal status and grade in relation to guidelines provided by the NHSBSP (NHSBSP, 1996a) and BASO (BASO, 1995). With the introduction of a region-wide oncology information collection system the timeliness, accuracy and completeness of data should be improved particularly in clinical data fields. However, region-wide audit will be required to monitor the quality of data provided and to ensure the comparability coding systems and the consistency of their utilisation. This new data collection system, in combination with the improvements in data coding and storage introduced with the new database GRACE, and the availability of the staging algorithm, should mean that the reliability of data recorded at the WMCIU will continue to improve and allow more detailed evaluation of breast cancer in the West Midlands in the future.

7.1.2 The availability of prognostic data recorded by the NHSBSP

Analysis of prognostic data collated on HQA reports supplied by West Midlands screening units revealed similar problems to those seen with pathological data at the WMCIU. In 1994/95 three of the nine screening units supplying data had no size recorded for *>5%* of cancers, failing the quality standard of 5% set by the Radiological 'Big 18' (Radiological 'Big 18', personal communication), with rates varying across the region from 0% to 6.3%. This lack of data is not isolated to the West Midlands, with the national proportion of screen-detected cancers with size missing reported as 12% for 1990-93 (Moss *et al.*, 1995). In the present study five units had no grade specified for **>5%** of cases, with rates varying from 0% to 14.7%. On average 42.1% of cancers had no local node data recorded, with rates varying from 10.3% to 75.0% across the region. Whether lack of data reflected reluctance to sample nodes, or failure to supply pathology data to screening units was not investigated in the present study but will be followed up. However, in the West Midlands programme, completeness of prognostic data increased from 1992 to 1995, indicating that improvements in the ascertainment of the data or their reporting to screening units occurred over this period.

The availability of prognostic data according to screening history revealed marked differences. Histological confmation, size and grade were available more often for screen detected cancers compared to those detected symptomatically as interval cancers, cancers in non-attenders or cancers diagnosed prior to the start of screening. The availability of local node status, however, was particularly poor amongst prevalent screen detected cancers, although incident screen detected cancers had more complete local node data. Lack of nodal data probably reflects failure of pathology laboratories to supply follow-up pathology reports to screening units as well as to the WMCIU. This lack of prognostic data is a barrier to the evaluation of the impact of screening. Any examination of survival for breast cancers according to screening history must not only incorporate complex adjustments to allow for lead time bias, but is also dependent on the availability of complete, high quality prognostic data. For example, in a study of survival for breast cancer patients in the UK Trial of Early Detection of Breast Cancer, **high** levels of data incompleteness were identified which limited the ability of the study to assess the impact of the trial. Of the 2995 cancers studied, although the authors did not report exact numbers of cancers with missing data, from the graphs provided it can be estimated that approximately 5% had no size recorded, and nodal status was absent for 30%, with data incompleteness being more marked for cancers presenting symptomaticaily (Moss *et al.,* 1994).

Screening aims to reduce mortality in the population invited, an aim reflected in the Health of the Nation (HoN) target for breast cancer of a 25% reduction in mortality between 1990 and the year 2000 in women aged 50 - 69 years (HMSO, 1992). In the West Midlands, crude mortality rates in women aged 55-69 years, the group used by the NHSBSP to monitor mortality, fell by 4.9% between 1983 and 1987, and by 8.9% between 1988 and 1994. It unlikely that screening, which was not fully implemented in the region until 1991, is responsible for much of this reduction, which is more likely to reflect an increase in the use of adjuvant therapy. The situation in the West Midlands mirrors a national trend in falling mortality prior to the introduction of screening in 1988 (Quinn and Allen, 1995; ONS, 1996). Although welcome, this background mortality drop complicates the evaluation of screening, since not only is there an inevitable delay before any mortality reduction due to screening can be expected, adjustments will be

necessary to account for the underlying mortality reduction when assessing the impact of screening. In addition to monitoring mortality, surrogate endpoints such as a shift towards early stage breast cancers in the invited population can be used as early indicators of success for the NHSBSP (Bull *et al.,* 1991), and evidence of such a stage shift has been identified in Wessex (Wight *et al.,* 1993). Detecting a stage shift requires the combined efforts of the NHSBSP and cancer registries in identifying breast cancers in the invited population, and is hampered if prognostic data are incomplete. Data completeness indicators are now included in the KC62 reports submitted annually by each unit to the Department of Health, and it is the responsibility of the regional Quality Assurance Reference Centre (QARC) to liaise with units in order to minimise data incompleteness. The NHSBSP QA service in the West Midlands also includes assessments of data completeness in annual **QA** visits to each unit, and recommendations to improve data completeness and quality are made as a result. Further improvements to data quality have resulted from siting the QARC within the WMCIU allowing ready combination of cancer registration and breast screening data.

7.2 REDUCING THE IMPACT OF BREAST CANCER IN THE WEST MIDLANDS

7.2.1 Treatment patterns for breast cancers in the West Midlands

Before examination of treatment recorded at the WMCIU was undertaken, the reliability of recorded treatment data was assessed. While treatment data describing surgery and radiotherapy were considered to be reasonably complete and accurate, those describing hormone therapy and chemotherapy were less robust. Until the completeness of adjuvant treatment data supplied to the WMCIU is improved, which is an aim of the region-wide oncology information collection system currently being developed, recorded rates of hormone therapy and chemotherapy at the WMCIU need to be viewed with caution. The problem with incompleteness of adjuvant chemotherapy and hormone therapy data at the WMCIU is largely a reflection of the failure to collect routine outpatient adjuvant therapy data within Trusts. Unless out-patient chemotherapy and hormone therapy data are included in day case minimum data sets (as is the case in a small minority of West Midlands hospitals) the only record of the patient's treatment will be a hand-written

entry into the patient's detailed medical records. **As** part of the Calman-Hine implementation programme in the West Midlands, all Trusts applying for Cancer Centre or Cancer Unit status are being advised to install an adjuvant therapy prescribing database on which all chemotherapy episodes or hormone treatments (in, out and day case) will be recorded. Direct downloading of data to the WMCIU will then ensure high completeness for all Trust-based chemotherapy and hormone procedures in the future. Tracking hormone therapy prescribed by GPs will, however, still remain a problem.

The changes in recorded treatment identified in this study were in agreement with the King's Fund guidelines on the management of breast cancer (King's Fund, 1986) suggesting that these guidelines may have had a significant impact on clinical practice in the West Midlands. **A** decrease in recorded mastectomy rates over this period from **55.4%** to 37.3% was mirrored by a simultaneous increase in conservative surgery rates from 20.7% to 38.5%. Reports from elsewhere in the UK indicate that this fall in the mastectomy rate was not isolated to the West Midlands (Morris ***et*** *al.,* 1992), with rates in Yorkshire falling from 70% to 44% in the period 1978-92 (Sainsbury *et al.,* 1995a). In the West Midlands older women were less likely to undergo conservative surgery in both years, a result also reported by the Yorkshire region (Yorkshire Cancer Organisation, 1995). Despite the incompleteness of adjuvant chemotherapy and hormone therapy data in the present study, a rise in recorded hormone therapy was seen, particularly in older women, and recorded radiotherapy rates also increased. In line with the King's Fund guidelines, women receiving conservative surgery were more likely to have recorded adjuvant radiotherapy in 1987 than 1983, with 46.3% of cases having recorded radiotherapy in 1983 compared to 57.9% in 1987.

In the West Midlands chemotherapy was recorded most often for women aged less than 50 years, a situation also reported in Yorkshire, suggesting that chemotherapy may have been directed more at pre-menopausal women, in line with the King's Fund guidelines. However, because of the known problems with notification of out-patient chemotherapy data, little could be concluded regarding the use or changes in the use of adjuvant chemotherapy, although the very low rates of recorded chemotherapy in the West

Midlandsmirror those reported from the South East (Chouillet *et al.,* 1994) and Yorkshire (Yorkshire Cancer Organisation, 1995).

In addition to the clear changes in treatment between 1983 and 1987, in both years there were differences in recorded treatment for early stage and advanced cancers, and for older and younger women, suggesting that stage and age were determining factors in allocating treatment. In the West Midlands, the release of the King's Fund guidelines co-incided with a shift in clinical practice towards the recommendations. The extent to which these changes can be directly attributed to the guidelines is unknown, but nonetheless, the ability to detect significant changes after a relatively short time is encouraging. However, despite clear improvements in treatment in the West Midlands over this period, clinical practice still differed from the King's Fund guidelines in some respects, for example, even in 1987 over 42% of women undergoing conservative surgery had no adjuvant radiotherapy recorded. This type of divergence from the King's Fund guidelines has been reported elsewhere in the UK (McCarthy *et al.,* 1991; Choiullet *et al.,* 1994). An additional cause for concern is the significant variation in the use of conservative surgery, mastectomy and radiotherapy across the West Midlands both in 1983 and 1987, suggesting inconsistency in clinical preference. Variations in treatment have also been reported by other UK regions (Choiullet *et al.,* 1994; Sainsbury *et al.,* 1995a) and a recent survey of Trusts revealed differences in access to a variety of treatments (Cancer Relief Macmillan Fund, 1996), indicating that the problem of inconsistency in clinical practice is a national one which still needs to be addressed.

7.2.2 Surgical caseload, treatment and survival

Cancer care in the UK is being reorganised with breast cancer care concentrated into multi-disciplinary specialist teams (HMSO, 1995; NHS Executive, 1996a). To monitor improvements following these changes valid markers of multi-disciplinary specialist care are needed. Surgical caseload is currently being used as a marker for specialist care in the COG recommendation that **all** breast referrals should be to specialist breast teams with a throughput of at least 100 cases/year (NHS Executive, 1996a).

Evidence that treatment from specialist teams leads to improved survival was provided by a recent Scottish study, with five year survival rates for breast cancer patients treated

by specialist surgeons being 9% higher than for those treated by non-specialists (Gillis and Hole, 1996). Gillis and Hole identified specialists based not on surgical caseload, but by the fulfilment of a number of criteria, including involvement in dedicated breast clinics, associations with pathologists and oncologists, and maintenance of patient records. Although treatment data were not investigated, Gillis and Hole suggest that the improved outcomes may reflect use of adjuvant therapies.

Associations between patient throughput and patient outcomes have been reported for a number of cancers including breast cancer (Stiller *et al.,* 1994). **A** report from Yorkshire estimated that, of the 12,861 breast cancer patients diagnosed between 1979 and 1988, those treated by surgeons with a caseload of more than 30 cases a year had a 15% reduction in the risk of death compared to those treated by surgeons with lower caseloads (Sainsbury *et al.,* 1995b). In the present study no association between survival and surgical caseload was identified in both a univariate analysis and in a multivariate model which adjusted for other factors including stage. Sainsbury and colleagues suggest that the survival advantage that they identified resulted, at least in part, from increased use of adjuvant therapy in this group (Sainsbury *et al.,* 1995b). In the present study high caseload surgeons were less likely to use conservative surgery for early disease in 1987, suggesting that clinical practice was altered in line with the King's Fund guidelines more rapidly by surgeons treating fewer cases. However, no difference in recorded radiotherapy or hormone therapy rates was apparent amongst conservative surgery cases according to caseload. Thus the present study provides no evidence of differences in adjuvant treatment according to caseload, and indeed suggests that such high caseload surgeons were slower in changing clinical practice in line with the King's Fund guidelines than their lower caseload colleagues.

The King's Fund guidelines also recommended that breast cancer care should be delivered by multi-disciplinary specialist teams (King's Fund, 1986). In the present study it was not possible to identify specialist surgeons or multi-disciplinary working, but an increase in surgical caseload was identified, with the proportion of cases treated by "high" caseload surgeons (treating >30 cases/year) rising from 17.8% in 1983 to 27.4% in 1987. Despite this rise, nearly three quarters of women diagnosed in 1987 were

treated by low caseload surgeons. With the introduction of the NHSBSP, and the site specialisation demanded by the Calman-Hine report and BASO it might be expected that this trend in increasing surgical caseload has continued since 1987. However, a recent survey of breast surgeons in the UK found that, although the average annual caseload was high at 66, one third of surgeons saw 25 or fewer new cases per year. Furthermore, half of the 425 responders estimated that they spent less than 20% of their time on breast surgery, with 90% declaring another specialist interest in addition to breast surgery (Harries ***et*** *al.,* 1997). Thus the shift towards site specialisation demanded by the Calman-Hine report must be encouraged if all breast cancer patients in the UK are to receive treatment from a surgeon with a dedicated specialist interest in breast cancer.

The improved care supplied by specialist teams is unlikely to reflect caseload alone, more probably being the result of a multi-disciplinary approach, with increased clinical expertise, specialist interest, and effective communication between professional groups involved in patient care (Dixon, 1996). Although an association between high caseload and specialisation was identified in one study (Sainsbury ***et*** *al.,* 1995a), it seems that caseload alone may be invalid when used as the sole marker of high quality, multi-disciplinary care (Baum, 1996; Sikora, 1996). Monitoring improvements in breast cancer care is vital if the recommendations made by the Calman-Hine report are to have their intended impact. If caseload is to be used as a proxy for specialist care, the limitations of such a simplistic approach must be recognised. Surgical caseload would be better used in conjunction with other indicators of quality such as those used in a recent national survey (Harries ***et*** *al.,* 1997). Routine monitoring of improvements in care must be undertaken on a regular basis and will require detailed and up to date knowledge of working practices within Trusts. This process will require close co-operation between those undertaking the monitoring function, such as cancer registries, and Trusts undertaking the care of breast cancer patients.

7.2.3 Improving survival from breast cancer in the West Midlands

To implement effective strategies to reduce the impact of breast cancer in the West Midlands, and to achieve maximum impact with scarce NHS resources, patterns of disease in the region must be recognised so that actions taken to reduce the impact of

the disease can be targeted effectively. Incidence rates of invasive breast cancer varied significantly across the West Midlands prior to the introduction of screening in 1988. Similar variation has been reported elsewhere in the UK, for example in Yorkshire (Yorkshire Cancer Organisation, 1995). Differing incidence rates suggest variation in the prevalence of risk factors for the disease, but as no preventative measures are currently available, efforts to reduce the impact of the disease and eliminate inequalities take the form of early diagnosis and improvements in treatment. In the UK, breast cancer screening is offered to women aged 50-64 years by the NHSBSP, and, for those outside this age range, through access to a variety of diagnostic facilities.

Re-screening variation in mortality and survival across the West Midlands suggests variation in prognostic characteristics and/or treatment across the region which need to be taken into account when implementing strategies to enable the HoN target for breast cancer to be met (HMSO, 1992). In the present study the small but significant rise in the five year survival rate from 62.7% in 1983 to 66.3% in 1987 suggests a trend in improving survival in the region. However, year of diagnosis was not retained in a Cox model when entered alone or in combination with other prognostic factors. This is likely to reflect differences between the two survival analysis techniques. Only five years of follow up were included in the univariate analysis, whereas up to a maximum of twelve years of follow up were included in the multivariate analysis. Thus although five year relative survival rates improved significantly between 1983 and 1987, longer term cause specific survival did not improve. This is likely to reflect the tendency of apparently localised breast cancer to metastasise even after a prolonged disease-free period. Survival analysis incorporating more than five years of follow up may thus be more useful than five year survival rates in assessing the outcome for breast cancer patients.

Absence of prognostic data reduced the proportion of cases for which survival analysis could be performed. Selection for multivariate analysis, for which all the data items included in univariate analyses were required, restricted the study group to only 20.8% of the 5005 cases eligible for univariate survival analysis. Absence of grade and stage were mainly responsible for this restriction, which potentially reduced the degree to which the findings from *Cox* models produced could be generalised. Although selected

cases were significantly younger and less deprived than unselected cases, there was no difference in five year relative survival rate for the unselected and selected groups. In addition, a *Cox* model into which selection for the multivariate analysis was entered as the only variable provided no evidence that selection was associated with better survival. Therefore, despite this restriction of the sample size imposed by absence of data, the ***Cox*** models produced for the 1041 cases were considered to be reasonably generalisable to the total study group. However, that this restriction was necessary clearly demonstrates the need for improved routine recording of prognostic data.

In this study the significance of a number of established prognostic factors was assessed for a West Midlands population. **A** clear trend in five year relative survival rates was seen according to stage, with survival rates rising from 10.4% for stage '4' cases to 88.4% for stage '1' cases. In the univariate analysis stage '**3**' cases from 1983 fared better than those from 1987 (48.0% 1983, 40.6% 1987, $p<0.05$). It is possible that this result reflected improved recording of prognostic data in 1987, perhaps due to an increase in the use of detailed investigations. Such an increase might lead to 'stage migration', as apparently localised cancers, which would previously be allocated an early stage are, following more rigorous investigation, found to be more widespread and allocated a higher stage **(Barr *et al.,*** 1992). This migration would result in fewer cases being incorrectly allocated an early stage, so that survival would appear to improve for early stages (stages '1' and '2a' in this study), but worsen for higher stages (stages '2b, '3', and '4'). It is also possible that these difference in survival by stage reflect changes in other prognostic factors such as treatment type or patient-related factors. Stage was the strongest predictor of survival in the overall *Cox* model, suggesting that early diagnosis is the key to improved outcome for breast cancer patients in the West Midlands.

The well established significance of tumour grade (Elston ***et al.,*** 1991) was also confirmed in this study group in both univariate and multivariate analysis. In the present study screen detected cancers were both smaller and of lower grade than symptomatic cancers. There is some support for the theory that tumours become metastatic as they get larger, a process termed phenotypic progression (Ponten ***et al.,*** 1990) and thus that by arresting tumour development when tumours are small, progression to aggressive

disease may be avoided. Thus, in addition to detecting cancers at an earlier stage, it is possible that early diagnosis may also improve survival by reducing the proportion of high grade tumours. Whether phenotypic progression occurs in breast cancer is the subject of continuing debate. A recent Finnish study assessed the aggressiveness of 739 breast cancers according to screening history (Hakama *et al.,* 1995). The identification of similar levels of aggressiveness amongst the pre-screening control group compared to incident screen-detected cancers was interpreted as evidence that the aggressiveness of tumours did not worsen during the pre-clinical detectable phase, and was put forward by the authors as evidence against phenotypic progression. However, concerns over the interpretation of the study findings have been raised (Day *et al.,* 1995; Paci, 1995; Alexander *et al.,* 1995), focusing on the relatively low sensitivity of the Finnish screening programme. Further research into the possible occurrence of phenotypic progression will be required if this issue is to be resolved.

The previously reported significance of histological type (Ellis *et al.,* 1992) was confirmed in the present study, with invasive carcinomas of "special type" and lobular carcinomas displaying high five year survival rates. However, only 6.6% of the invasive carcinomas examined were of "special type", with a further 8% being lobular carcinomas, and thus the favourable survival of these histological types has little impact on the overall population survival. These proportions are similar to those reported elsewhere (Merseyside and Cheshire Cancer Registry, 1994).

Survival varied across the West Midlands, an observation similar to variations reported elsewhere in the UK (Yorkshire Cancer Organisation, 1995), with poorer survival identified in women aged >50 years compared to younger women. In the 5077 cases analysed, 46.7% arose in women aged 65 years or more, with an association between increasing age and poorer stage. Screening alone cannot reduce breast cancer mortality throughout the population, as currently screening is restricted to 50-64 year olds. Women aged 65 years or more are not currently invited for screening, although such women are able to request screening if they wish. Screening women aged 65 years or more has been shown to be effective in reducing mortality in Sweden (Chen, 1995). Following calls from the charity Age Concern, the NHSBSP is considering extending the

age range for routine screening up to 69 years. Demonstration projects assessing the efficacy of screening women aged 65-69 were initiated in 1996 at two pilot sites to run for three years, with first results expected by the year 2000. In the meantime, women aged 65 or more are being encouraged to request screening (NHSBSP and Age Concern, 1996). Recommendations have been issued by the NHSBSP that written information sent by the NHSBSP to women should include information on accessing the service for those about to become ineligible for routine invitation. The rising risk of breast cancer with age, and the decreasing density of breast tissue which makes mammography more effective, means that breast cancer is more easily detected by screening in this age group than in younger women. Widening the age range to include 65-69 year olds would provide a means of significantly reducing breast cancer mortality in a greater proportion of the population than is currently eligible for routine screening.

In the study group, the most deprived women displayed the poorest survival, a trend established by a number of other studies (Vernon *et al.,* 1985; Karjalainen ***et*** *al.,* 1990; Kogevinas ***et*** *al.,* 1991; Shrijvers ***et*** *al.,* 1995). However, within the population there was not only a significant association between increasing age and more advanced stage, but also between more advanced stage and increasing deprivation. Poorer prognosis amongst deprived women has been identified elsewhere (Richardson, 1992), although not in all studies (Carnon *et al.,* 1994). Although the poorer survival amongst deprived groups is likely to reflect poorer prognostic characteristics in these women, in the present study, deprivation was retained in the multivariate Cox model after adjustment for stage, grade, and treatment type. This finding suggests that the influence of deprivation cannot be entirely attributed to variation in other prognostic characteristics entered into the model. In the present study it was not possible to adjust for Co-morbidity, and it is thus possible that underlying ill health could be responsible to some extent for the observed trend. Actions aimed at reducing breast cancer mortality must take account of associations between deprivation and survival. Lower screening uptake amongst deprived women (Hoare *et al.,* 1993) suggests that screening is likely to have less impact amongst deprived women. Efforts to increase screening uptake are encompassed in health promotion activities undertaken by the NHSBSP (NHSBSP, 1995f) and should be particularly targeted at women living in deprived areas. Promoting

early diagnosis for those outside the screening age range is also vital. Projects targeting areas of low uptake in the West Midlands are underway in collaboration with the Geographical Information System (GIS) project team at the WMCIU.

In the present study cases without recorded surgery fared worse than those surgically treated. Although women with recorded conservative surgery fared slightly better than mastectomy patients in the univariate analysis, there was no difference between these groups after adjustment for other factors in the multivariate model, an observation in line with other studies (EBCTCG, 1994, Fisher *et al.*, 1995; Riley *et al.*, 1995). However, the apparent lack of survival advantage conferred by all types of adjuvant therapy in the present study was unexpected given the weight of evidence supporting the benefits of these therapies (EBCTCG, 1992). This outcome is likely to reflect incompleteness of adjuvant chemotherapy and hormone therapy data, as well as possible selection of patients with a generally poor outlook for adjuvant therapy. This survival disadvantage was present even for surgical cases with recorded radiotherapy, who displayed a **44%** elevation in risk compared to those with surgery alone after adjustment for other prognostic factors. This result was particularly unexpected given that the completeness of radiotherapy data is considered to be high at the WMCIU due to the strong links with the five radiotherapy centres in the West Midlands. It is possible that this finding came about because the group for which no adjuvant therapy was recorded was contaminated with cases which, despite receiving adjuvant hormone and/or chemotherapy, did not have this reported to the WMCIU, possibly reflecting that they did not require any treatment at a radiotherapy department. If this was the case the favourable survival for such cases would elevate the overall survival of the "no adjuvant therapy" group, which was used as the reference group against which the survival for those with recorded adjuvant therapy was assessed. Thus identification of apparent survival disadvantages for cases with recorded adjuvant therapy in comparison to the "no adjuvant therapy" group may reflect this contamination rather than a genuine survival disadvantage. No association between survival and surgical caseload was identified in either univariate or multivariate analysis in contrast to results from elsewhere (Yorkshire Cancer Organisation, 1995; Sainsbury *et al.*, 1995a; Gillis and Hole, 1996).

The completeness of treatment data should, in addition to prognostic data, be improved with the introduction of a region-wide oncology information collection system. Monitoring treatment and prognostic data will play a crucial role not only in identifying improvements in care, but in evaluating whether such changes have an impact on patient outcome. With the current activities aimed at improving the quality of both prognostic and treatment data supplied to the WMCIU, the WMCIU is ideally placed to undertake this monitoring and evaluating function and thus to contribute to the improvement of outcomes for breast cancers patients in the West Midlands.

7.3 MAXIMISING THE IMPACT OF THE NHSBSP IN THE WEST MIDLANDS

7.3.1 Improving the usefulness of outcome measures

Monitoring the performance of the NHSBSP is a vital function of the **QA** service and, from the outset of screening in the UK, the need to calculate performance indicators which could be compared with quality standards was recognised. In addition, as previously discussed, measures which act as early indicators of success in terms of likely mortality reduction are essential so that the impact of the programme can be assessed. **A** number of outcome measures are currently calculated by the "KC62" software used by the programme to compile reports of screening activity for the Department of Health, and, in addition, further outcome measures not included in the KC62 report are recommended by the NHSBSP (Radiology 'Big 18', personal communication).

For outcome measures to be useful and valid the data with which outcome measures are calculated must be complete and accurate, and the method of calculation must be valid and consistently applied. Targets set in relation to each outcome measure should reflect both the need to readily identify unacceptably poor performance (the minimum standard), and to allow prediction of whether the programme is likely to bring about the mortality reduction desired (the expected standard). This study identified a number of problems associated with the use of currently available outcome measures.

Some difficulties arise due to the lag time between the introduction of changes in the definition of outcome measures, and incorporation of changes into the KC62 software.

For example, whereas micro-invasive cancers were previously included in small cancer detection rates (NHSBSP, 1993b), current definitions exclude them (Radiological 'Big 18', personal communication). However, the small invasive cancer detection rates calculated by KC62 reports extant in 1996/97 still include micro-invasive cancers. This inconsistency causes confusion and can give the appearance that acceptable small cancer detection rates are being achieved when this is not the case. If KC62 reports are to continue to calculate outcome measures then all the measures which can be calculated from the raw data presented in these reports, which are also currently recommended by the NHSBSP, should be provided. Currently the DCIS rate is not calculated by a KC62 report, although it can be calculated from the raw data provided. It is also currently impossible to calculate an overall open benign biopsy rate because needle biopsies are included with open biopsies in the raw data. Furthermore the overall open biopsy rate is quoted in national performance reports (Moss *et al.,* 1995; NHSBSP, 1995d), but is not provided by, and cannot be calculated from KC62 reports. Different publications produced by the NHSBSP provide different quality standards for the same outcome measures, for example those for benign open biopsies (NHSBSP, 1992; 1993b; 1995b; 1996). Further inconsistency in the use of outcome measures by different professional groups within the NHSBSP is demonstrated by the fact that the malignant:benign open biopsy ratio is still calculated by the KC62 report despite the recommendation by the Radiological 'Big 18' that this measure be withdrawn as a performance indicator due to biasing effect of small numbers. However, the malignant:benign biopsy ratio was retained in the 1996 update of the Surgical **QA** Guidelines (NHSBSP, 1996b).

The outcome measures calculated by KC62 reports are intended to provide a ready source of **QA** data for the NHSBSP. However, until the problems described are rectified, outcome measures calculated by KC62 reports must be treated with caution. **A** review of the KC62 software is ongoing, and it is hoped that changes to the software which eliminate these difficulties will be made **as** soon as possible. In addition, when new KC62 software is released, previous years of data should be reproduced using the new version of the software. This would allow identification of trends in outcome measures to evaluate progress. Of particular importance for smaller screening units is the

aggregation of data across years to counteract the effect of small numbers, which is not possible unless data for several years are produced using consistent methods.

Of critical importance to the success of screening is the achievement of adequate compliance rates. The quality standard for this measure has remained unchanged at a minimum of 70% since screening was introduced (NHSBSP, 1988). However, when this 70% rate is applied to a hypothetical screening population over five screening rounds, the resultant cumulative attendance rates reveal that this quality standard needs to be raised and split into prevalent and incident standards. The simple model used predicted that, if the 70% standard was met in all rounds, only 16.8% of women would attend in all five rounds, and 36% in four rounds. In particular, a stricter standard for the reattendance rate is required if this outcome measure is to be useful in improving screening performance. Given that reattendance rates across the West Midlands in 1994/95 ranged from 85% to 96%, a minimum standard of 85% for the reattendance rate, with an expected standard of 90% would not be unreasonable. Measuring cumulative uptake in the NHSBSP over four or more rounds will be difficult. **A** recent study from the Njimegen screening programme calculated cumulative attendance over 17 years of two-yearly screening, and identified that, although compliance in the first round was high at 88%, cumulative uptake after 8 and 9 rounds fell to 39% and 24% respectively (Scaf-Klomp ***et al.,*** 1995). Lower uptake rates have been achieved in the NHSBSP, with overall uptake rates for all screens in 1993/94 reported as 72% (NHSBSP, 1995e) rising slightly to 76.7% in 1994/95 (NHSBSP, 1996b) and it is likely that achieved cumulative uptake in the UK will be lower than that in Njimegen. However, the current format of the KC62 report only provides information about attendance in the current round and at the previous two invitations. Calculations of cumulative uptake over more than three rounds will require complex analysis of data describing screening histories recorded by screening units, and is likely to be lengthy process. These issues needs to be addressed by the NHSBSP so that a routine method of calculating cumulative uptake over several rounds can be made available.

The calculation and interpretation of measures assessing whether sufficient cancers are being detected has undergone many changes. In the past, **all** breast cancers detected by

screening,-both invasive and *in situ,* were included in the cancer detection rate (NHSBSP, 1988). As it is invasive cancers which are potentially fatal, *in situ* cancers are now excluded from calculations which are, in addition, split for prevalent and incident screens (Radiological 'Big 18', personal communication). The effect of the age of the screened population on the achieved detection rates has resulted in further changes to the recommended calculation method. The NHSBSP currently recommends the calculation of invasive cancer detection rates for prevalent women restricted to those aged 50-54 years only (Radiological 'Big 18', personal communication). While the assumption that most women attending for the first time are aged 50-52.9 years is reasonable for well established units, is not always valid. In addition, it is currently necessary to estimate the number of invasive cancers in this age group as it is not supplied by the KC62 report. This restriction also excludes cancers detected in prevalent women aged 55-64 years, which although a minority, are nevertheless part of the programme. In contrast, the NHSBSP currently recommends that incident invasive cancer detection rates are calculated for 50-64 year olds, whereas if this were consistent with the calculation of prevalent invasive cancer detection rates, women aged 55-64 years only would be included.

The Standardised Detection Ratio (SDR) has recently been introduced as an outcome measure by the NHSBSP in an effort to adjust observed cancer detection rates according to the age distribution of the invited population. The SDR utilises expected cancer incidence rates obtained from the Swedish Two Counties Trial on the assumption that, given that the UK programme achieves comparable cancer detection, mortality reduction similar to the Swedish trial should be realised (Blanks *et al.,* 1996). More recently this method has been extended to the detection of small cancers, with the development of SDRs for small (<15mm) and larger (>=15mm) cancers (Blanks *et al.,* 1995~). Maintaining a high quality screening programme is essential if breast cancer mortality is to be reduced, and the detection of sufficient cancers is both a crucial performance indicator, and a marker of likely mortality reduction for the programme. Hence, differing baseline incidence rates between districts have implications for the evaluation of the NHSBSP, as they could influence both the apparent cancer detection rates achieved by different screening units and rates of interval cancers.

Those units situated in low baseline incidence areas might be likely to achieve lower detection rates than those in high baseline incidence areas independently of the quality of screening. In a recent study background correction factors relative to the national average were calculated for a sample of NHSBSP catchment areas to allow for variation in baseline incidence when interpreting observed cancer detection rates or SDRs (Blanks *et al.,* 1996b), with the largest correction factor bringing about a 15% change. The average correction factor for West Midlands programmes was estimated at only 3%, ranging from -5% to 6% across the region. These correction factors were calculated using geographical information obtained for women aged 45 years or more in 1968-85 (Swerdlow *et al.,* 1993), and, as Blanks and his colleagues point out, calculation of correction factors is more accurate if based on detailed geographical information relating to women in the screening age group. In the present study, correction factors for 50-64 years olds relative to the West Midlands average incidence rate in 1983-87 varied between catchment areas from -7.2% to 5.1%. Further work examining baseline incidence is being undertaken in collaboration with the GIS project team at the WMCIU.

Currently the NHSBSP recommends that SDRs are only calculated for prevalent screens. Incident SDRs are complicated by the impact of the detection of cancers at the incident screen which were missed at the prevalent screen (termed "overspill" cancers). Thus a screening unit with poor sensitivity might achieve a low SDR in the prevalent round, but, in the subsequent round, achieve a higher SDR for incident women as cancers missed at the last screen are detected. Such a unit would also be expected to display a high interval cancer rate due to the symptomatic presentation of false negative cancers. A study of data from the Dutch screening programme analysed the previous screening mammograms of 44 women with incident screen detected cancer, and concluded that in 9% the cancer was apparent on the previous mammogram, and that in a further 48% "minimal signs" which could have indicated cancer were present (van Dijck *et al.,* 1993). Although based on a small sample, these findings suggest that this overspill effect may be considerable. Development of the incident SDR is ongoing and it is likely that a method for calculating this measure will be available for routine use in the near future, although the impact of any overspill effect must be taken into account when assessing achieved SDRs (Blanks, personal communication). As around 80% of women

screened at many units currently attend incident screens the prevalent SDR provides insight into screening performance for a minority of women screened. An incident SDR must become a routine QA measure as soon as possible. A simple method of assessing the overspill effect would be to present incident SDRs along with the prevalent SDR relating to the screening year three years earlier. For example, overspill would be a likely explanation for a unit achieving a low prevalent SDR in 1990 and a high incident SDR for the same women returning for their re-screen in 1993, which also displayed a high interval cancer rate arising from the prevalent screens in 1990. Such calculations require the identification of the screening round in which women were invited to attend. Now that many units are in their third screening round, separating women invited in the second and third rounds is difficult, requiring complex manipulation of data from the screening office computer. A means of readily identifying the screening round, according to both the screening unit and the woman, in which women are invited will become more important as the programme matures. This need must be addressed at a national level.

The rising incidence in the screening age group prior to the introduction of screening in the West Midlands has been seen nationally (Quinn and Allen, 1995), and also has implications for the NHSBSP. When setting cancer detection targets it is necessary to estimate the expected incidence of cancer in the absence of screening. This estimation was undertaken in this study and elsewhere (Woodman *et al.,* 1995, Blanks *et al.,* 1996a), with the incidence of interval cancers arising in each 12 month period following a negative screen presented as the proportion of the expected underlying incidence. In a recent study the rate of expected increase was greatest for the oldest women (Prior *et al.,* 1996), suggesting that expected rates should be calculated separately for each age band rather than for the age group 50-64 years. In the present study the expected incidence rate used in calculations was for the 50-64 year age group, an approach also used elsewhere (Woodman *et al.,* 1995). It is important that the increasing complexity surrounding the assessment of cancer detection rates, SDRs and interval cancers should not reduce the usefulness and accessibility of QA data calculated for the NHSBSP.

Since being introduced in 1987, the NHSBSP has continuously developed the methods used to evaluate progress. The NHSBSP recognised at the outset that further research

was needed into some aspects of the programme, and has incorporated changes to evaluation methods based both on the results of research, and in the light of experience gained by screening units since 1987. For example, the quality standard for the early recall rate, set at a maximum of 0.25%, was subsequently raised to 1% as nationally achieved early recall rates revealed the earlier target to be too low (Radiological Big '18', personal communication). In addition, the format and content of KC62 reports has been developed since the start of screening, although the version current in 1997 still requires considerable amendment. While the readiness to incorporate new ideas into the methods used to evaluate the programme is a strength of the NHSBSP, it also has the potential to cause confusion for those working in the programme. To avoid this:

- clear and up to date definitions of recommended outcome measures and the relevant quality standards should be circulated to all those involved in the evaluation of the NHSBSP
- changes from previous definitions must be identified, and advice supplied as to the date from which new definitions apply
- when new quality standards are set these must be circulated rapidly to all those involved in evaluation of the NHSBSP, with clear information on how these differ from the previously used standards, and from what date the new standards apply
- where measures are being used to assess the programme, the sources of the data and the definitions used to calculate them must be clearly stated
- where assumptions are made in the calculation of outcome measures these must be validated, and if found to be invalid remedial action must be taken
- finally, guidance on the interpretation of outcome measures should be provided for those who may want to be informed of the progress of their local screening service but who are unfamiliar with the definitions used, such as purchasers

7.3.2 The impact of screening on breast cancer in the West Midlands

Evaluation of a population based screening programme is more difficult that that of a randomised controlled trial of screening due to the lack of a defined control population. However, where population based screening has been introduced in a staggered fashion it is sometimes possible to construct a control group from women not invited in the early years of the programme. For example, a recent study of the Finnish programme

identifieda-control group of women not selected for invitation in the introductory phase of the programme (Hakama ***et*** *al.,* 1997). Comparison of death rates amongst women invited and uninvited for screening revealed a significant reduction in risk for those invited. However, as the Finnish programme has matured women in this retrospectively constructed control group have been invited for screening and thus identifying a control population for assessment of more recent screening performance will not be possible. Other means of assessing the impact of population based screening,are thus required.

In the present study a retrospective technique was designed to identify a cohort of women with breast cancer for whom the cancer could potentially have been detected by screening in the prevalent screening round. Although the catchment area of only one screening unit was considered, this unit is the largest in the West Midlands, accounting for nearly 20% of women aged **50-64** years in the region. Women in the cohort were followed up until diagnosis and the screening history classified. Of the 973 breast cancers studied, only 31.6% were screen detected, and 22.9% arose in women known to the screening unit but diagnosed before the screening invitation. Despite the increasing risk of breast cancer with age, incidence rates in this group did not rise with age at invitation. It is possible that, for some of these women, the arrival of a screening invitation precipitated presentation into the symptomatic service, and thus that some of this group of cancers were "programme provoked', although the extent of such an effect could not be assessed in this study. This investigation focused on the prevalent round, and the proportion of women diagnosed while eligible for screening but prior to invitation is expected to have fallen now that the programme has completed the prevalent round across the West Midlands. Whether this has occurred can only be ascertained by performing similar analyses for later years.

In the sample investigated 278 cancers (28.6%) arose after a negative screen. The majority of these (84.9%) were interval cancers arising at <36 months from the last screen. Also included in the 28.6% were breast cancers arising at 36-59 months in women no longer eligible for routine invitation due to age, and in women re-invited to screening at more than 36 months, at which time the cancer was already diagnosed. **A** further 10.5% arose in non-attenders, of which 15% were diagnosed within three

months of the invitation. It is possible that, for these women, although the screening invitation was not taken up, it provided the impetus for seeking medical attention. Thus these 15% could, in addition to some of those diagnosed prior to the invitation date, include some of the programme provoked cancers, although once again it was not possible to explore this issue further in this study. In addition, 5.3% of breast cancers arose in women not known to the screening programme, indicating inaccuracy in the population register used to identify eligible women.

A study of the South East Thames screening programme examined data from the prevalent round (Garvican ***et*** *al.,* 1996). No established protocol is available for undertaking this type of study and methodological differences exist between the scheme used to identify the eligible population and allocate screening histories in South East Thames and **in** the present study. In the present study a cohort of women eligible for having their breast cancer detected in the prevalent round was identified retrospectively and followed up. In South East Thames a cross-sectional analysis of cancers arising between 1988 and 1993 was used, although the study focused on the screening histories of cancers presenting in 1991 and 1992. The likely impact of this difference in methodology is demonstrated by the example of a women invited for screening during the prevalent round in 1991, and diagnosed with an interval cancer in 1993 after the prevalent round had finished. In the present study this cancer would be included in the study cohort. However, had this study been cross-sectional, and included only cancers diagnosed within the prevalent round, this cancer would have been excluded. **A** cross sectional methodology thus excludes cancers arising after the period of screening has ended and may provide an over-optimistic view of the impact of screening.

Of 4202 breast cancers diagnosed in 1991-92 in South East Thames, 47.9% were screen detected, 10% were interval cancers, 6.6% arose in non-attenders, 17.4% were diagnosed prior to invitation, 7.4% were unknown to the screening service, and, at the time of publication, the screening histories of 10.7% had yet to be checked (Garvican ***et*** *al.,* 1996). Both the South Thames East programme, and the Warwickshire, Solihull and Coventry Screening Service (WSCSS) achieved most **QA** standards for the prevalent screening round. Invasive cancer detection rates in both studies exceed the minimum

standard, being 5.1/1000 in the South East and the slightly lower rate of 4.1/1000 at the WSCSS, although differences in the age structure of the eligible populations could account for this difference. The SDR for the WSCSS for the prevalent round was 0.73, just under the target of 0.75, although SDRs have risen since then, being, for example, 1.48 in 1994/95. As the data presented by South East Thames were not stratified by age, and as no prevalent round SDR was provided, it was not possible to assess the influence of age differences on the difference in achieved cancer detection rates for South East Thames and the WSCSS. It is possible that lower sensitivity in the WSCSS could account for the lower screen detected proportion seen. However, the difference in study methodology is likely to account to some extent for the higher proportion of screen detected cancers reported by South East Thames (47.9%) compared to 36.1% identified in this study, and the lower proportions of interval cancers in South East Thames.

Despite the methodological differences, both studies reveal that, although programmes may meet the quality standards set by the NHSBSP, the proportion of screen detected cancers is low, indicating low programme sensitivity (Garvican ***et al.,*** 1996). This low sensitivity has serious implications for the likely impact of the programme on the population. Authors from the South East estimated that, as only 40% of breast cancers arose in women in the age range eligible for screening, and just under 50% of breast cancers in eligible women were screen detected, only 20% of women could have an early diagnosis due to screening (Garvican ***et al.,*** 1996). In the present study only 36.1% of cancers in eligible women were screen detected in the present study, and 31% of cancers diagnosed pre-screening arose in 50-64 year olds. Thus, from these figures, only 10% of breast cancer patients potentially benefit from early diagnosis by screening. Even if the 25% HoN mortality reduction were to be achieved in this 10%, only a 2.5% population-wide mortality drop would be achieved. If the NHSBSP can only achieve such a small mortality reduction the cost effectiveness of the programme must be reviewed. Further work examining the proportions of cancers according to screening history in the West Midlands will be carried out.

Also of concern is that, like the West Midlands, South East Thames identified a significant proportion of cancers in women not known to the screening service, making

up 7.4% of cases in South East Thames and 5.3% of cases in the present study. Women are omitted from the Prior Notification Lists (PNLs) utilised by the NHSBSP if they are not registered with a GP. In addition, if the details on the PNL are incorrect, for example, if addresses or ages are incorrect, invitations will not be received by women. Although it is difficult to increase the proportion of women registered with GPs, it may be possible to increase the proportion of such women who self refer for screening by means of health promotion initiatives. In addition, improving the accuracy of population registers is vital (Bowling, 1989; Bickler, 1993). The establishment of computerised **links** between health authorities and the screening service in some parts of the country, for example in Coventry, has improved the quality of the population register in these areas considerably (Wheaton, personal communication).

To assess the impact of screening it is not only necessary to establish the proportions of breast cancers according to screening history in the eligible population, but also to examine the prognosis of these cancers. In this study pre-screening rates of non-invasive cancer (DCIS) were low, making up only 2% of the pre-screening control group. Amongst screen detected cancers, however, 6.9% of prevalent and 14.1% of incident breast cancers were DCIS. The proportion of screen detected DCIS has been reported nationally as 18% for 1990-93 (Moss ***et*** *al.,* 1995), with *20.5%* of screen detected cancers in South East Thames in 1988-93 being **DCIS** or micro-invasive cancer (Garvican ***et*** *al.,* 1996). Although these rates are within or near to the current quality standard range of 10%-20% (Radiological 'Big 18', personal communication), the high DCIS proportion amongst screen detected cancers may be a cause for concern. The proportion of DCIS that progress to invasive disease has been estimated at 2550% (Page, 1995). Thus, a significant proportion of screen detected DCIS may not progress to invasive cancer if left untreated, raising doubts as to the benefits of screening for these women. **A** study of the outcome for women with screen detected DCIS is being designed in the West Midlands NHSBSP.

In the present study, as elsewhere, screen detected cancers had a generally more favourable prognosis in terms of size, nodal status and grade than cancers diagnosed pre-screening (Duffy ***et*** *al.,* 1991; Tabar ***et*** *al.,* 1992; Crisp ***et*** *al.,* 1993; Wight ***et*** *al.,*

1993; Moss *et al.,* 1994). The prognostic characteristics of cancers in non-attenders were particularly poor. Although the prognosis of interval cancers were more favourable than those of pre-screening cancers they were poorer than for the screen detected group. This suggests that cancers detected after the introduction of screening, with the exception of cancers in non-attenders, should display better survival than cancers detected pre-screening. However, detection by screening of high proportions of special type cancers may indicate length bias, in which screening preferentially detects those types of cancer which have lowest aggressive potential. If this is the case aggressive tumour types are more likely to be diagnosed symptomatically than by screening, and are also more likely to present at a later stage, as both this study and others have identified poorer prognostic characteristics amongst cancers diagnosed symptomatically compared to screen detected cancers (Duffy *et al.,* 1991; Tabar ***et*** *al.,* 1992; Crisp *et al.,* 1993; Wight ***et*** *al.,* 1993, Moss ***et*** *al.,* 1994). The occurrence of length bias would reduce the ability of the NHSBSP to bring about a fall mortality in eligible women. However, the identification of high proportions of special type tumours in the NHSBSP may reflect an increase in specialist knowledge amongst pathologists increasingly involved in the NHSBSP compared to those working only in the symptomatic service.

In this study, although screen detected cancers were of favourable prognosis, cancers arising outside the programme were of poorer prognosis. The favourable prognosis of interval cancers in comparison to pre-screening controls suggests an increase in breast awareness in women who have been screened. However, interval cancers had a poorer prognosis than screen detected cancers. Unless screen detected cancers make up the bulk of breast cancers in the eligible population, screening is likely to have less impact on survival and mortality in the population than expected. In addition, there is some evidence that overdiagnosis and length bias may be occurring in the programme, both of which reduce the potential benefit resulting from the programme. The NHSBSP has put much effort into developing internal quality assurance and improvements to the management of the service are ongoing (NHSBSP, 1995d). However, although attention is being focused on interval breast cancers, the occurrence of other breast cancers outside the NHSBSP in the eligible population has received less scrutiny. Unless the NHSBSP can increase the proportion of breast cancers detected by screening it is

unlikely to achieve it's aims. Analyses similar to those undertaken here will be critical in assessing whether, since the prevalent round, the situation has become more favourable. Before further investigations are performed, however, it is important that a consistent investigative technique is identified so that studies conducted in different parts of the UK will be comparable.

7.3.3 The definition and interpretation of interval breast cancers

Although the significance of interval breast cancers has always been recognised by the NHSBSP (Day *et al.*, 1989) there still remains ambiguity in the definition of interval breast cancers. In this investigation the impact of changing the definition was explored using the 278 interval cancers identified. When restricted to histologically confirmed, invasive breast cancers, occurring in women aged 50-64 years at screening, and arising at <36 months from the last negative screen, the total fell to 213 (termed "core" intervals). These restrictions led to a drop in the overall interval cancer rate arising from the prevalent round from a maximum of 47.1/10,000 women screened to 36.1/10,000.

The largest change in the calculated interval cancer rate was brought about by the exclusion of cancers arising at >36 months from the last screen. This has serious implications since, although the NHSBSP assumes that a three year round is currently operating for all women, this is not always the case. In the sample analysed 27.5% of women were re-invited at 36 months or more from the last screen, with 17.3% waiting over 39 months for a re-invitation. Where a woman, although eligible for re-invitation, was not offered another appointment prior to the diagnosis of a cancer symptomatically at >36 months from her last screen, it is arguable whether this should be excluded from interval cancer rates. Occurrence of such cancers represents a failure of the NHSBSP and to exclude such cases from the figures may mask a significant problem.

From this investigation it is clear that changing the definition of interval cancers can have a significant impact on calculated rates. In addition, of the 278 interval cancers, 30 were unknown the screening unit and were only identified by cross-checking information between the screening unit and the WMCIU, highlighting the need for effective information exchange. To assist future assessments the definition of an interval cancer must be made clear and applied in a standardised fashion throughout the NHSBSP. To

facilitate identification of interval cancers, and all breast cancers arising in women eligible for screening, the exchange of information between the cancer registries and the NHSBSP is vital. A modem link based in the WMCIU allowing access to the NHSBSP computer system by staff at the QAFC is currently being introduced and will allow ready comparison of NHSBSP and cancer registration data.

This investigation identified high interval cancer rates which rose as the time from the last screen increased, giving an overall rate of 36.1/10,000 women screened for core interval cancers arising from the prevalent round. This high rate supports findings other studies (Woodman ***et*** *al.,* 1995; Day ***et*** *al.* 1995a). In the present study interval cancer rates, expressed as proportions of the estimated underlying incidence, rose from 27.6% at 0-11 months from screening, to 56.7% at 12-23 months and 82.5% at 24-35 months. These rates are similar to the findings reported for East Anglia (Day ***et*** *al.,* 1995a), and for the North West (Woodman ***et*** *al.,* 1995). The West Midlands, East Anglia and the North West all display proportionate interval rates roughly double those of 17%, 30% and 56% respectively reported by the Swedish Two Counties trial (Day ***et*** *al.,* 1995a). Authors in East Anglia predicted a 18% mortality fall based on comparison of the proportionate interval and compliance rates achieved and those reported by the Swedish Two Counties study (Day ***et*** *al.,* 1995a). Given the similar proportionate incidence and compliance rates in the West Midlands a comparable mortality drop might also be expected in this region, less than the 25% HoN target (HMSO, 1992).

Although there is no doubt that rates of interval cancers in the UK are too high, how this situation should be addressed is the subject of much debate. Radiological classification of interval cancers which provides information about whether or not the cancer could have been detected at the previous screen is vital in the assessment of interval cancers, and should be performed using a standardised set of criteria. Radiological classification relies on comparison of the screening film with a mammogram taken at diagnosis. In the absence of diagnostic mammography, interval cancers remain unclassifiable, unless an estimated radiological classification system such as that used in this study is employed. For example, in a recent study of interval cancers in the North West, only the 53% of cancers with a diagnostic mammogram were included in the study (Asbury ***et*** *al.,* 1996).

In the present investigation, cancers without diagnostic mammograms were classified by examination of the screening mammogram alone using a scheme devised by Simpson and his colleagues (Simpson *et al.,* 1995). Using this method a radiological classification was available for 92% of the 213 core interval cancers.

Of the core intervals in the study, 44.1% were true intervals, 22.5% were false negative (present on the previous mammogram) and 9.9% were occult. Lack of diagnostic mammograms meant that 15.5% could only be classified as "not false negative", and were made up of a mixture of true and occult interval cancers. Only 8% of core interval cancers were radiologically unclassifiable. These proportions are similar to those reported by Simpson and colleagues for a series of 167 interval cancers of 46% true intervals, 26% false negative, 11% occult and 16% unclassifiable (Simpson *et al.,* 1995).

Interval cancers arising in the third year represented 49% of interval cancers in this study and in a series reported by the North West (Woodman *et al.,* 1995). Authors in the North West reported that 80% of interval cancers arising in the third year were true intervals (Asbury *et al.,* 1996) compared to an estimated 68% in the present study. The high proportion of true interval cancers identified in the third year has prompted calls to shorten the screening round to two years (Woodman *et al.,* 1995; Asbury *et al.,* 1996). However, the present study estimated that a possible 37.3% of interval cancers might be avoided by detecting them at screening using a shorter two year round.

There are two means of reducing interval cancer rates. Improving screening sensitivity, leading to the detection of more cancers relative to the underlying incidence rate, should reduce the incidence of false negative interval cancers. The second strategy, that of implementing a shorter screening round, is based on the assumption that the current 3 year round is too long in relation to the sojourn time of many breast cancers. Where the sojourn time is less than the round length, it is possible for breast cancers to arise and become symptomatic in the time between one negative screen and before the next screen is due. Screening at a shorter interval should detect some of these cancers and thus reduce the rate of interval cancers. In the present study there was little difference in interval cancer rates according to age at diagnosis for cancers arising in the first 24 months from screening, but in the 24 - 35 month period interval cancer rates were

highest in-the 60-64 year age group. This is in contrast to a report from the North West which identified a lower risk of interval cancers in older women compared to those aged 50-54 years (Threlfall *et al.,* 1997). This difference might be explained by the use of separate expected incidence rates for each age group in the North West compared to the single figure used in the present study. Threlfall and colleagues suggest that the higher risk of interval cancers in younger women might be addressed by introduction of a shorter screening round this age group. However, any shortening of the screening round would be difficult and costly, and if different round lengths were used for different age groups, the organisational complexity would be considerable. Until larger series of interval cancers are available from which proportionate incidence rates for each age group can be calculated, it seems unlikely that such drastic changes to the organisation of the screening programme are justified on the basis of the current research findings.

Crucial to the current debate over whether a shorter round is required is the presence of some confusion over the meaning of the radiological classification "true" interval cancer. True interval cancers are those which were not detectable on the previous screening film which was taken using the screening method current at the time. However, this does not necessarily mean that the tumour was absent at the time of the previous screen. With improved sensitivity, for example, using a cranio-caudal X-ray view in addition to the oblique view, some cancers which would be missed by an oblique view alone may be visible. Had the original screening been an oblique view only, on which the cancer was not shown, this cancer would be classified as a true interval. Thus adding an additional view may lower the rate of true and false negative interval cancers. The improved sensitivity of two view screening is borne out by a study of prevalent screening activity for 1994/95, which identified that screening units using two view screening for prevalent women achieved a small cancer SDR (<15mm) 42% greater than those units using one view screening (Blanks *et al,* 199%). The difference in SDRs for non-invasive and micro-invasive cancers was only 3%, and for larger invasive cancers (>=15mm) was only 7%. Thus, use of two view screening improves sensitivity such that, in particular, more small cancers are detected. **As** it is the detection and treatment of these cancers which has the potential to bring about a mortality reduction, introduction of two view screening for incident screens in addition to prevalent screens seems desirable. The

findings reported by Blanks and colleagues confirm earlier work comparing the effect of one and two view screening, which identified higher cancer detection rates for two view screening, and in addition, identified a reduction in the referral rate, so that by using two view screening, specificity as well as sensitivity is improved (Moss *et al.,* 1995).

Interpreting interval cancers is a retrospective process and changes that will cause a fall in interval cancer rates may already be in place, although this reduction is not yet detectable. There is evidence that sensitivity in the West Midlands programme has risen, with SDRs increasing since 1992. Improvements in sensitivity are expected with the introduction of two view screening for all prevalent screens since August 1995 (Wald *et al.,* 1995). Thus the NHSBSP may already be detecting more cancers and potentially reducing future interval cancer rates. The extension of two-view screening to incident screens would be likely to improve sensitivity throughout the programme and should be considered, especially as incident women currently make up the majority of women screened. In addition, double reading of films has been shown to increase sensitivity without a large increase in cost (Brown *et al.,* 1996; Seradour ***et*** *al.,* 1996). Double reading counteracts inter-observer variation in the interpretation of mammograms and, in a recent study, resulted in 12% more cancers being detected than would have been diagnosed by a single reader (Parham *et al.,* 1996). Lack of radiologist time makes routine double reading difficult, and although a number of screening services have trained members of radiography staff to act as first-line readers, many screening services in the UK do not undertake any double reading.

Although the NHSBSP is currently based on a **3** year screening round, many women experience a longer round length. Thus, interval cancer rates may not indicate that a three year round is insufficient, but that the three year round is not being maintained. The NHSBSP is currently under pressure to extend the age for routine invitation to include women aged 65-69 years. The NHSBSP represents a large commitment of scarce NHS resources. Although a two year round would undoubtedly reduce interval cancers rates, there are a number of alternative solutions to high interval cancer rates, such as improvements in sensitivity which are already in place, which may be more effective and less costly than shortening the screening round.

7.4 CONCLUSIONS

High incidence and mortality from breast cancer make it a major health problem in the West Midlands and in the UK as a whole. As no method of preventing the disease is yet available, actions to reduce its impact focus on reducing mortality and improving survival. Stage at diagnosis, the type of treatment received, and a number of biological factors including tumour grade have been identified as important factors in determining survival from breast cancer.

Any actions taken to reduce mortality and improve survival must be monitored and evaluated to ensure effectiveness. The true impact of such actions cannot be ascertained from randomised clinical trials or small studies, but require population based studies. Currently cancer registries are the only source of population based data for such studies in the UK, and, furthermore, cancer is the only disease for which such population based registers exist. To be useful, population based data must be of high quality. In this study, few cases had stage recorded on the WMCIU database but this reflected the small number for which stage had been sent to the WMCIU. However, where data were supplied to the WMCIU, the accuracy and completeness with which tumour characteristics were recorded at the WMCIU were very high. This allowed the development of a staging algorithm combining tumour characteristic data to allocate TNM stage to breast cancers. This staging algorithm requires testing against other populations, and incorporates a number of assumptions which may not always be justified. Nonetheless the staging algorithm is a useful tool allowing a TNM stage to be allocated at the WMCIU for breast cancers thus increasing the utility of the data for analytical purposes. Similar staging algorithms are soon to be developed for other cancers. Furthermore, improvements to the coding rules and database structure made as a result of this study will improve data quality for all cancers.

This study demonstrated that it was broadly possible to undertake evaluation of the main issues relating to breast cancer using routinely collected data. The limitations of cancer registry data identified in this study are mostly imposed by inadequacies in the supply of data from Trusts, and the extent to which such important prognostic data are stored in a readily accessible form within Trusts. This is apparent not only for prognostic data but

for adjuvant chemotherapy and hormone therapy which are recognised as being key to improved patient outcome. The Calman-Hine Report identified the need to monitor the effects of restructuring cancer services on outcome. High quality data detailing all stages of patient care, from diagnosis to initial treatment and follow up, are required for this process which necessitates multi-disciplinary data collection. These data should be collected by Trusts so that they can monitor the management of their own patients. Transmission to the cancer registry then becomes a routine component of the Trust's data collection processes. When this process is in place, the type of quality assurance developed in this study will become an integral part of the WMCIU's data collection, evaluation and audit role of the cancer care available in the West Midlands.

The evaluation of the impact of the NHSBSP is hampered by the lack of a control population, and survival analysis comparing screen-detected cancers with symptomatic cases will always be complicated by lead time bias. However, comparison of screening data with cancer registry data allows indicators of the likely-effect of screening to be derived. In the sample of pre-screening cancers examined stage was confirmed as the most significant prognostic factor, with treatment type, tumour grade, and level of deprivation also identified as important. The NHSBSP aims to reduce mortality by detecting breast cancers early. In the sample examined, only 31.6% of cancers in eligible women aged 50-64 years were screen detected. In combination with the observation that, in the pre-screening population, only 31% of cancer patients were aged 50-64 years, this means that only 10% of breast cancer patients directly benefit from early diagnosis by screening. Even if the 25% HoN target mortality reduction was achieved for screen-detected cancers, this would translate into a 2.5% population-wide reduction. Thus, although the prognostic characteristics of screen detected cancers were clearly better than those of breast cancers diagnosed outside the screening programme, the impact of this on breast cancer mortality in the entire population is likely to be smaller than expected. The favourable prognosis of interval breast cancers compared to cancers diagnosed pre-screening could reflect an increased breast awareness in women who have been screened, and thus could, in this respect, represent a beneficial effect of the programme. However, with increased sensitivity, at least some of these cancers might have been detected at the previous screen when at an earlier stage. The negative

psychological impact of the diagnosis of an interval cancer should not be ignored although little research has been carried out in this area. The NHSBSP represents a massive input of scarce NHS resources, and efforts to increase the potential population benefit must be made. Further assessment of the proportions of cancers detected by screening using the method developed in this study will now be carried out to evaluate whether improvements since the prevalent screening round have been made.

The NHSBSP represents the first model in the NHS of a structured multi-disciplinary service in which quality assurance is a core activity. The existence of the NHSBSP has undoubtedly had a beneficial knock-on effect on the symptomatic breast service and in the development of services provided for other cancers. However, the NHSBSP QA system stills needs refining. The current KC62 statistical report does not provide all the data needed for the calculation of the range of outcome measures used by the NHSBSP. The definition and layout of the KC62 report is currently under review, and the problems identified in this study have been fed back to the NHSBSP to facilitate this review process. This study identified a number of problems associated with the definition of outcome measures and their associated quality standards. For example, the minimum standard for the uptake rate is set at 70% for all screens. Cumulative uptake over several rounds must be high if the desired mortality reduction is to be achieved, and should be considered when setting the standard. This study demonstrated that, after five rounds, 70% of women would attend all five screens only if an uptake rate of 93.2% were to be achieved in each round, and thus the 70% standard currently applied to each year and round of screening is inadequate and must be raised. It is important that an iterative process is established between those calculating and using outcome measures, and those setting quality standards. Auditing a service against inappropriate standards or using invalid outcome measures can be counterproductive and may provide misleading information. As a result of this study, problems identified with outcome measures and quality standards have been fed back to the NHSBSP via the West Midlands QA Committee members.

Much attention is currently being focused on interval breast cancers as a QA measure for the NHSBSP. This study identified a need for clarification of definition of interval

cancers. The current definition excludes cancers arising at 36 months or more from the last screen. These cancers themselves signify an important problem, reflecting variation in the round length operated by the NHSBSP, which is assumed in the current definition of an interval cancer to be a uniform 3 years. Excluding these cancers in this study reduced the observed interval cancer rate from 43.1/100,000 to 36.1/100,000. Thus, for meaningful interpretation of interval cancer data, factors that influence interval cancer rates must be considered in the collation of national data.

Finally, improving treatment is a crucial element in improving survival for breast cancer patients. Recent guidelines from **BASO** and the COG stress the importance of multi-disciplinary care within specialist **teams,** with use of appropriate surgery and adjuvant therapy. Such guidelines were initiated in the King's Fund guidelines published in 1986. In the current study clear improvements in treatment in line with the guidelines were evident. However, variation in treatment across the region indicated possible inconsistency in clinical practice. This variation will now be examined in more recent years using standard population based cancer registration data. The lack of survival advantage for women treated by high caseload surgeons, or of a difference in the use adjuvant therapy according to caseload, is contrary to the recommendation made by the COG. Further work will be now be carried out in this area to assess whether caseload is a valid indicator of high quality care and improved patient outcome.

APPENDIX ONE

THE TNM/PTNM TUMOUR CLASSIFICATION SYSTEM FOR BREAST CANCER

Adapted from the TNM Atlas, Third edition, 2nd revision 1992 UICC

The classification applies only to breast carcinomas. There should be histological confirmation of the disease. The anatomical subsite of origin is not considered in the classification. In the case of multiple simultaneous tumours in one breast, the tumour with the highest T category should be used for classification. Simultaneous bilateral breast cancers should be classified simultaneously. The separate T, N and M components are combined using the rules in Table **A1.** 1 to produce TNM stages.

T - Primary Tumour

TX Primary tumour cannot be assessed

T0 No evidence of primary tumour

Tis Carcinoma in situ; intraductal carcinoma, or lobular carcinoma in situ, or Paget's disease of the nipple with no underlying invasive tumour

Note: Paget's disease associated with an underlying invasive tumour is classified according to the size of the invasive tumour

T1 Tumour **2** cm or less in greatest dimension

- T1A 0.5 cm or less in greatest dimension
- T1B More than 0.5 cm but not more than 1 cm in greatest dimension
- T1C More than 1 cm but not more than **2** cm in greatest dimension

T2 Tumour more than **2** cm but not more than 5 cm in greatest dimension

T3 Tumour more than 5 cm in greatest dimension

T4 Tumour of any size with direct extension to chest wall or skin

Note: The chest wall includes ribs, intercostal muscles and serratus anterior muscle but not pectoral muscle

- T4A Extension to the chest wall
- T4B Oedema (including peau d'orange) ***or*** ulceration of the skin ***or*** satellite skin nodules confined to the same breast
- T4C Both T4A and T4B
- T4D Inflammatory carcinoma

Inflammatory carcinoma is characterised by diffuse, brawny induration of the skin, usually with no underlying palpable mass. Pathological classification requires examination of the primary carcinoma with no gross tumour at the resection margins. A case can be classified pT if there **is** microscopic tumour in the margin. The pT categories correspond to T categories. When classifling pT the size is that of the invasive component. If there is a large ***in situ*** component **(e.g.** 4cm) and a small invasive component (e.g. 0.5 cm) the tumour is coded pT1A.

The regional (local) lymph nodes

1. Axillary (ipsilateral) nodes, interpectoral (Rotter's) nodes and lymph nodes along the axillary vein and its tributaries which may be divided into the following levels:

   i. Level I (low axilla): lymph nodes lateral to the lateral border of the pectoralis minor muscle

   ii Level II (mid axilla): lymph nodes between the medial and lateral borders of the pectoralis minor muscle and the interpectoral (Rotter's) nodes

   iii. Level III (apical axilla): lymph nodes medial to the medial margin of the pectoralis minor muscle including those designated to subclavicular, infiaclavicular or apical

   Note: intramammary nodes are coded as axillary nodes.

2. Internal mammary (ipsilateral): lymph nodes in the intercostal spaces along the edge of the sternum in the endothoracic facia.

Any other lymph node metastasis is coded as a distant metastasis (M1), including supraclavicular, cervical, or contralateral internal mammary lymph nodes.

N - Clinical Regional Lymph Nodes

NX Regional lymph nodes cannot be assessed (e g. previously removed)

N0 No regional lymph node metastases

N1 Metastasis to movable ipsilateral axillary node(s)

N2 Metastasis to ipsilateral axillary node(s) fixed to one another or to other structures

N3 Metastasis to ipsilateral internal mammary lymph nodes

pN - Pathological Regional Lymph Nodes

The pathological classification requires resection and examination of at least the low axillary lymph nodes. Such a resection should include six or more nodes.

pNX Regional lymph nodes cannot be assessed (not removed or previously removed)

pN0 No regional lymph node metastasis

pN1 Metastasis to movable ipsilateral axillary node(s)

- pN1A Only micro-metastasis (none larger than 0.2 cm)
- pN1BI Metastasis to one to three lymph nodes, any more than 0.2 cm and all less than 2.0 cm in greatest dimension
- pN1BII Metastasis to four or more lymph nodes, any more than 0.2 cm and all less than 2.0 cm in greatest dimension
- pN1BIII Extension of tumour beyond the capsule of a lymph node metastasis less than 2.0 cm in greatest dimension .
- pN1BIV Metastasis to lymph node 2.0 cm or more in greatest dimension

pN2 Metastasis to ipsilateral nodes fixed to one another or other structures

pN3 Metastasis to ipsilateral internal mammary lymph nodes

M - Distant Metastases

MX No information available about metastases

M0 No distant metastases present including metastasis to the supraclavicular, cervical or contralateral internal mammary nodes

M1 Distant metastasis to any site (including brain, bone, liver, lungs and skin), or to supraclavicular, cervical or contralateral internal mammary nodes

Table A1.1 Combination of stage components to produce TNM stage

Stage	T	N	M
0	Tis	N0	M0
1	T1	N0	M0
2a	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
2b	T2	N1	M0
	T3	N0	M0
3a	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1	M0
	T3	N2	M0
3b	T4	any N	M0
	any T	N3	M0
4	any T	any N	M1

APPENDIX TWO

STANDARD DATA REPORTS AVAILABLE FROM THE NATIONAL BREAST SCREENING SYSTEM

Figures A2.1 to A2.5 describe standard reports available from the National Breast Screening System from which data examined in this investigation were obtained.

A2.1 Format of a KC62 Version 2.8 Report

Figures A2.1 to A2.3 show the format of KC62 Version 2.8 report using 1994 - 1995 data from the West Midlands (Unit identity code has been removed). The KC62 version 2.8 is made up of 10 tables. The first 8 Tables (A to F2) have an identical format, and describe screening activity for women with different to the screening histories, divided into prevalent (first time) screens, incident screens (re-screens), women invited to screening as a result of early recall, and women screened following self or GP referral.

Table A	Prevalent screens, no previous invitation
Table B	Prevalent screens, defaulted previous invitation
Table C1	Incident screens, last screened within 5 years
Table C2	Incident screens, last screened at 5 years or more
Table D	Early recall screens
Table E	Prevalent self or GP referrals
Table F1	Incident self or GP referrals last screened within 5 years
Table F2	Incident self or GP referral last screened at 5 years or more

The format of Table A is shown in Figure A2.1 as an example of Tables A - F2. Table G shown in Figure A2.2 reports the characteristics of cancers identified by screening. Table H shown in Figure A2.3 supplies a number of outcome measures. In addition, part of table H also supplies data completeness indicators.

Figure A2.1 KC62 report - outcome of screens for first time screens (Table A)

KC62 2.8 1st invitation for routine screening — Screening Office Code: M X X

Women invited 1st April 1994 to 31st March 1995

Table A

Age at 1st Appt	Line no.	No. women invited	Lost to follow-up tech. inad. mamm-ogram	No. screened (techn. ad.)	Outcome of initial screen: Not known	Routine recall	Early recall	Referred for assess-ment or direct to histology	Final outcome of assessment: Failed to attend for assess-ment or histology	Outcome of assess-ment not known	Routine recall	Early recall	Cancer
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<=44	1	0	0	0	0	0	0	0	0	0	0	0	0
45-49	2	543	0	479	0	441	0	38	0	0	34	1	3
50-54	3	6416	3	5291	0	4967	0	324	0	0	269	16	38
55-59	4	727	1	306	0	286	0	20	0	0	15	3	1
60-64	5	587	0	216	0	208	0	8	0	0	5	1	2
>=65	6	16	0	4	0	4	0	0	0	0	0	0	0
All ages	7	8289	4	6296	0	5906	0	390	0	0	323	21	44

Age at 1st Appt	Line no.	Cancer diag. without Cytol. or Histol.	Outcome of Assessment up to and including Cytology and Diagnostic Histology: Up to and including Cytology: Referred for Cyto-logy	Not referred for diagnostic histology: No result recorded /Inad. result	Routine recall	Early recall	Cancer	Referred for diag. Histol.	Up to and including Diagnostic Histology: Total Referred for Diag. Histol.	No result recorded /Inad. result	Benign/ Normal: Routine recall	Early recall	Cancer
		(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
<=44	8	0	0	0	0	0	0	0	0	0	0	0	0
45-49	9	0	9	0	5	0	0	4	4	0	1	0	3
50-54	10	0	87	0	32	9	20	26	36	0	18	0	18
55-59	11	0	6	0	2	1	1	2	3	0	2	1	0
60-64	12	0	2	0	0	0	0	0	0	0	0	0	0
>=65	13	0	0	0	0	0		0	0	0	0	0	0
All ages	14	0	104	0	39	10	21	32	43	0	21	1	21

Figure A2.2 KC62 report - characteristics of cancers diagnosed as a result of screening (Table G)

KC62 2.8 Women with cancer diagnosed as a result of screening Screening Office M X X

Table G

Cancers diagnosed at screens included in Tables A - F2

Age at 1st Appt	Line no.	Total number of women with cancer (1)	Invasive status not known (2)	Non-invasive or possibly micro-invasive (3)	Definitely micro-invasive (4)	Invasive size (mm) <10 (5)	>=10 & <15 (6)	>=15 & <20 (7)	>=20 & <50 (8)	>=50 (9)	Size not known (10)	Total invasive (11)
<=44	1	0	0	0	0	0	0	0	0	0	0	0
45-49	2	3	0	0	0	1	0	0	2	0	0	3
50-54	3	51	0	7	0	6	8	15	12	2	0	43
55-59	4	35	0	5	0	8	11	6	5	0	0	30
60-64	5	36	0	6	0	6	10	7	7	0	0	30
50-64	6	122	0	18	0	20	29	28	24	2	0	103
>=65	7	3	0	0	0	0	2	0	1	0	0	3
All ages	8	128	0	18	0	21	31	28	27	2	0	109

Cancers diagnosed at screens included in Tables A - F2

Type of screen	Line no.	Total no. of women with cancer (1)	Invasive status not known (2)	Non-invasive or possibly micro-invasive (3)	Defin -itely. micro-invasive (4)	Invasive size (mm) <10 (5)	>=10 & <15 (6)	>=15 & <20 (7)	>=20 & <50 (8)	>=50 (9)	Size not known (10)	Total invasive (11)
A 1st invitation	9	44	0	5	0	7	6	13	11	2	0	39
B 1st screen previous NA	10	11	0	2	0	3	0	4	2	0	0	9
C1 Routine rescreen (<5yr)	11	60	0	9	0	10	19	10	11	0	0	50
C2 Routine rescreen (>=5yr)	12	3	0	1	0	1	1	0	0	0	0	2
D Early recall	13	2	0	1	0	0	1	0	0	0	0	1
E Self/GP referral (1st time)	14	5	0	0	0	0	1	3	3	0	0	5
F1 Self/GP referral (<5yr)	15	3	0	0	0	0	3	0	0	0	0	3
F2 Self/GP referral (>=5yr)	16	0	0	0	0	0	0	0	0	0	0	0

Figure A2.3 KC62 report - outcome measures (Table H)

KC62 2.8 Report for the period 1 - Apr - 94 to 31 - Mar - 95
Table H Outcome measures

	Uptake rate % of invited	Referral rate % of screens	Ben. biopsy rate % of screens	Benign diag. histol rate % of screens	Inv. cancer detect'n rate per 10,000 screens	Detect'n rate of inv. cancers <10mm per 10,000 screens	Detect'n rate of invasive cancers <15mm per 10,000 screens	Referral rate for cytology % of screens	Referral rate for diag. hist. % of screens	Malignant: benign ratio (x : 1)		Early recall rate after initial screen % of screens	Early recall rate after assess-ment % of screens
										Open biopsy	All diag hist.		
A	76.0	6.2	0.3	0.3	61.9	11.1	20.6	1.7	0.7	1.0	1.0	0.0	0.33
B	37.9	6.3	0.5	0.5	120.2	40.1	40.1	2.8	1.3	1.5	1.5	0.0	0.53
C1	92.9	3.1	0.1	0.1	36.3	7.3	21.0	0.6	0.2	2.0	2.0	0.0	0.09
C2	68.7	6.8	0.5	0.5	97.1	48.5	97.1	1.9	1.0	1.0	1.0	0.0	0.49
D	100.0	81.5	3.7	3.7	370.4	0.0	370.4	7.4	11.1	2.0	2.0	0.0	0.0
E	-	7.1	0.9	0.9	28.5	0.0	28.5	2.8	1.1	0.3	0.3	0.0	0.0
F1	-	4.3	0.3	0.3	76.7	0.0	76.7	1.3	0.5	1.0	1.0	0.0	0.0
F2	-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0

Data Completeness Indicators

	Assessment result not known (% of referred)	Cytology result not known (% of referred)	Histology result not known (% of referred)	Invasive status of cancer and/or size not known (% of cancers)	Invasive status of cancer not known (% of cancers)	Size not known (% of invasive cancers)
A	0.0	0.0	0.0	0.0	0.0	0.0
B	0.0	0.0	0.0	0.0	0.0	0.0
C1	0.0	0.0	0.0	0.0	0.0	0.0
C2	0.0	0.0	0.0	0.0	0.0	0.0
D	0.0	0.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0
F1	0.0	0.0	0.0	0.0	0.0	0.0
F2	0.0	0.0	0.0	0.0	0.0	0.0

A2.2 Format of a Surgical QA (QASS) report

Figure A2.4 Surgical QA (QASS) report

QASS Data for period	1 - Apr - 94	to	31 - Mar - 95	Screening Office Code	
Prevalent Screen				*shaded section repeated for incident screens*	
No. of women screened	7274		No. included with open episodes	0	
No. of women referred for assessment	454		% of women screened	6.2	
No. of women having surgical biopsy	61		% of women screened	0.8	
No. of women with cancer	56		No./10,000 screened	77	
No. of women with cancer <=15mm	24		No./10,000 screened	33	
No. of women with DCIS	7		% of cancers detected	12.5	
No. of women with cancers of special type*1	19		% of invasive cancers	38.8	
No. of women with bilateral cancer	2		Malignant to benign ratio	1.8:1	
No. of women with cancer and no cytology/histology	0		FNAC malignant, no histology		0
No. of women having FNA and no cytology	22		FNAC malignant confirmed by histology		26

BIOPSIES	FNAC	WBN	Surgical	Total
Benign *2	35	4	31	70
Malignant	26	0	30	56
Unknown	0	0	0	0
Total	61	4	61	126

Localisation Biopsy	No. of with diagnostic biopsy	68	No. of benign lesions >20g	6
	Average weight benign lesions	18.9	Range	4 to 201

INVASIVE CANCERS		Size (mm)	<=10	11-20m	21-50	>50mm	Size n/a	Total
All invasives *4			28	56	17	2	1	104
Type	Special types		6	9	2	0	0	17
	Invasive ductal*3		20	44	15	0	0	79
	Invasive lobular		2	1	0	2	0	5
	Medullary		0	2	0	0	0	2
	Unknown		0	0	0	0	0	0
Axillary node status	Node positive		1	7	11	1	0	20
	Unknown		0	0	0	0	0	0
Grade (invasive ductal)	1		9	9	2	0	0	20
	2		10	24	9	0	0	43
	3		1	11	4	0	0	16
	Unknown		0	0	0	0	0	0
Treatment *4	Conservative		25	42	7	0	0	74
	mastectomy		3	14	10	2	1	30
	Medical/other		0	0	0	0	0	0
	Unknown		0	0	0	0	0	0

IN SITU CANCERS		Localised/not assessed	Diffuse/ multicentric	Total
Treatment	Conservative	11	0	11
	Mastectomy	6	1	7
	Medical/other	0	0	0
	Unknown	0	0	0

*1 cancers of special type, total includes IDC grade 1
*2 FNAC biopsies benign total includes C1, C2, C3 and C4
*3 Invasive ductal totals include IDC grade 1
*4 all invasives totals include cancers with no histology

A2.3 Format of a Histology QA (HQA) report

Figure A2.5 Histology QA (HQA) report

HQA — All Clients (including <49 but not non-attenders or intervals)

Screening Start Date — 1 - Apr - 94 — Totals Reports

Screening End Date — 31 - Mar - 95

Breakdown of histology results

	Benign/normal not radial scar	Radial scar/ complex sclerosing lesion	Atypical hyperplasia	Carcinoma in situ/ poss micro	Micro-invasive	Invasive carcinoma	Details not recorded
No. cases	47	8 (8)*	3	18	0	105	0

Breakdown of invasive cancers

Tumour type	Ductal NST	Tubular (pure)	Lobular (pure)	Mucinous (pure)	Medullary	Tubular mixed	Other mixed	Other primary	Other malignant
No. cases	80	15	5	3	1	0	0	1	1

Size ranges of invasive cancers

Size (mm)	1-10	10.01-15	15.01-20	20.01-50	Over 50	Size unknown
No. cases	29	35	22	17	2	0

Prognostic factors of invasive cancers

Grade	I	II	III	Not assessed	Not specified
No. cases	33	42	17	13	0

Node status	Node -ve	Node +ve	Not sampled/ cleared
No. cases	66	20	19

Vascular inv.	Not present	Present	Not specified
No. cases	79	10	16

* total no. radial scars including atypia or malignancy

APPENDIX THREE

RULES FOR RECORDING BREAST TUMOUR CHARACTERISTICS ON THE OLD WMCIU PDP11 DATABASE

The following rules allow translation of breast tumour characteristics recorded at the West Midlands Cancer Intelligence Unit (WMCIU) on the old PDP11 database. Coding rules were described in an instruction booklet produced by the WCMIU in 1985 *"Instructions on filling in the RCR12A and RCR12L"*. Further information about the use of codes was derived from discussions with abstractors in 1994. All categories were coded based on the information relating to the time of diagnosis.

A3.1 Tumour size

One field was provided for recording the maximum diameter of the tumour. Pathological size was selected in preference to other sizes. No rules were provided on the selection of size if both radiological and clinical size were available without pathological size. Sizes were recorded in millimetres. Although not specified in the instructions, abstractors reported that where pre-operative tumour shrinking treatment had been given, or the tumour had been excised in several pieces, or the resection margins were not clear of malignancy, the tumour size was not recorded and a default code of '999' was entered.

A3.2 Disease extent

The field that coded disease extent also coded multiple lesions and diagnosis at *post mortem*. Only one code could be entered in this field. The instructions stated that the "multiple lesions" or "found at post mortem" should be coded if either was appropriate, but gave no advice as to which of these to code if both were appropriate.

'1' = Limited to the site of origin
Tumours which had not invaded into the overlying skin, the chest wall, pectoral muscles, or the nipple, were not ulcerating or fungating, and did not have any information suggestive of invasion into surrounding tissues.

'2' = Invading neighbouring structures
Tumours that displayed any of the following characteristics

Ulceration of the skin

Oedema of the overlying skin including peau d'orange

Fungation or imminent fungation

- Erosion of the nipple

Involvement of the chest wall or the pectoral or "underlying" muscles.

Involvement of the overlying skin including involvement of the epidermis or sweat glands, but not lower layers of the skin including the dermis. Tumours that extended up to the skin were coded as 'limited to site of origin'. This information could be clinical or pathological in origin.

Associated with satellite skin nodules confined to same breast.

Inflammatory carcinoma. The term inflammatory was often used as a clinical description of appearance, or in relation to infiltrates associated with a tumour, and to use this code pathological confirmation of inflammatory malignancy was required

'3' = Extent unknown
No information was available in relation to tumour extent

'4' = Multiple lesions
More than one discrete mass was found in the same breast

'5' = Found at *post mortem*
Tumours diagnosed at *post mortem*

A3.3 The status and mobility of the local lymph nodes

The local lymph nodes were defined as the ipsilateral axillary nodes. The presence or absence of malignant involvement of the local nodes was recorded separately for clinical and "at operation" data. The clinical local nodes field recorded the mobility of the nodes.

A3.3.1 Clinical local nodes

Information used for coding the clinical local nodes field was derived from reports in the case notes describing clinical examination of the patient at diagnosis.

'1' = No nodes
Clinical examination was made of the axilla and the clinician stated that there was no involvement. Instructions stated that "if nodes were not mentioned code 'none' if the notes are adequate". The term "adequate" was not defined.

'2' = Nodes present and mobile
Clinical examination was made of the axilla and surrounding area and the clinician stated that the nodes were present (involved), and mobile.

'3' = Nodes present and fixed
Clinical examination was made of the axilla and surrounding area and the clinician stated that the nodes were present (involved) and fixed.

'4' = Nodes present and mobility unknown
Clinical examination was made of the axilla and surrounding area and the clinician stated that the nodes were present (involved). No information describing nodal mobility was available.

'5' = Insufficient information
Insufficient information was available describing the status of the local nodes.

A3.3.2 "at operation" local nodes data

Data recorded in the "at operation" local nodes field were derived from pathology reports and summaries written in the case notes reporting observations made during surgery.

'1' = Nodes negative
One or more local nodes was sampled and pathologically confirmed as negative.

'2' = Nodes present and confirmed
One or more local nodes was sampled and pathologically confirmed as positive.

'3' = Nodes present but not confirmed
One or more local nodes was observed to be suspicious, but no further investigations were reported.

'4' = Nodes not examined
No observations at operation or pathological investigations were reported.

A3.4 Distant lymph nodes

Cervical, supra-clavicular, contralateral axillary and contralateral internal mammary nodes are considered to be distant lymph nodes. The coding rules for the distant nodes were identical in field structure and coding rules to the local lymph nodes.

A3.5 Presence or absence of metastases

The presence or absence of metastases was recorded separately on the PDP11 for clinical and "at operation" data. Data describing the skin, bone, brain, liver, peritoneum, contralateral breast, other sites were recorded separately. Carcinomatosis was recorded only if no more specific data were available.

A3.5.1 Clinical metastases

'1' = No clinical evidence of metastases
If there was no mention of metastases, and no specific tests were performed looking for metastases, it was assumed that the clinician did not think there were any.

'2' = Positive clinical evidence of metastases
The patient was examined at diagnosis and the clinician was certain that a specific site was involved. Test results e.g. from chest X-rays, liver function tests, full blood counts or bone scans were used to code the clinical fields. Other information, such as the subsequent treatment of a site with radiation, was also used to support the use of a positive clinical code.

'3' = Insufficient information
Insufficient evidence was available to support definite clinical involvement at a metastatic site.

A3.5.2 "at operation" metastases

'1' = negative "at operation"
One or more suspected sites were sampled at operation, and pathologically confirmed as negative.

'2' = positive "at operation"
One or more suspected sites were sampled at operation, and pathologically confirmed as positive.

'3' = suspected but not confirmed
One or more suspected sites were mentioned at operation but no further investigations were reported.

'4' = not examined
No metastatic sites were mentioned at operation and no pathological investigations were reported.

APPENDIX FOUR

RULES FOR RECORDING TUMOUR CHARACTERISTICS ON THE AUDIT DATABASE AND GRACE

These rules were devised for the review of case notes and recording of breast cancer data for the audit of staging and prognostic data discussed in Chapter Three. Many of the changes to the previous coding system were retained for the new computer system GRACE. Where there are differences between the audit database and GRACE, both coding systems are given. All fields are coded from data relating to the time of diagnosis.

A4.1 Tumour size (for use on the audit database and GRACE)

There were 3 size fields on the audit database, which have been retained for GRACE; one for pathological, one for clinical and one for radiological size. The greatest dimension should be recorded in mm. Since clinicians differ in the units they use, care must be taken to convert units to mm. If an *in situ* and an invasive tumour occur together, only the size of the invasive tumour should be recorded. There are some circumstances in which it is inappropriate to record pathological size. These are listed in Table A4.1 If no sizes is given or the size is invalid, '999' should be entered.

Table A4.1 Circumstances in which pathological tumour size is invalid
When a patient has been given pre-operative radiotherapy to reduce the size of the tumour
When pre-operative therapy (e.g. Tamoxifen) has been used to reduce tumour size
When a tumour has not been completely resected i.e. the pathology report indicates positive resection margins (particularly common with intraductal breast tumours).
When the pathological dimensions are available for a biopsy and for a subsequently resected tumour - the two values cannot be added together. FNAs or needle biopsies remove such small samples that these do not invalidate the pathological size of the remaining tumour.

Table A4.2 lists the descriptions and equivalent sizes for use on GRACE.

Table A4.2 Clinical size descriptions of size and equivalent sizes for use on GRACE	
Description of tumour size	**Equivalent size (mm)**
Pea	9.99
Peanut (shelled)	14.99
Bean, plum stone	19.99
Walnut, conker	29.99
Golf ball	39.99
Plum	49.99

A4.2 Disease extent (for use on the audit database and GRACE)

Data should be obtained from pathology reports, unless the tumour was so obviously invasive that it was apparent clinically. Some terms are purely clinical and not useful for classification - terms such as *tethered, puckered, dimpled* etc. should be ignored.

0 = *in situ*

This category is appropriate if the tumour has been pathologically described as *in situ*. It should be used in cases of Paget's disease if it occurs without an underlying tumour. If it occurs with an underlying tumour, the categorisation depends on the characteristics of the tumour. Paget's disease with and without an underlying invasive tumour can be distinguished through the ICDM code:

condition	ICDM	extent
Paget's disease alone	85403	0
Paget's + *in situ* intraduct carcinoma	85403	0
Paget's + underlying invasive carcinoma	85413	1,2,3,4 or ?
Paget's + separate invasive carcinoma	code for invasive carcinoma + pre-malignant condition	

A case with Paget's disease should not be designated as invading neighbouring structures (i.e. extent '2') in the absence of other evidence.

1 = limited to site of origin

This classification should be used for tumours where it is stated that there is no invasion of surrounding tissues, e.g. descriptions such as "circumscribed" or "well defined". This category is appropriate where tumours have invaded fat, blood or lymphatic vessels, nervous tissue, or pectoral muscle as these are part of the breast tissue. It should not be used where tumours have invaded the overlying skin, the chest wall or the nipple, or are ulcerating or fungating.

2 = invading neighbouring structures

This classification should be used for tumours which have one or more of the following properties indicated in the cases notes.

Ulceration of the skin.

Oedema of overlying skin including peau d'orange.

Fungation or imminent fungation

Erosion of the nipple (not Paget's)

Pathological chest wall involvement. Pathological investigation of the chest wall only occurs when a radical mastectomy is performed, and is

now uncommon. If information about chest wall involvement comes from clinical descriptions such as *"involvement of deep tissues"*, then category '4' should be used. If X-rays confirm the clinical description, then category '2' should be used. Involvement of the pectoral muscles is *not* a criterion for extent '2' and should be recorded as extent '1'.

Invasion of the skin. This information may be clinical or pathological. A distinction should be made between *"extending into.."*, and *"extending up to.."* the skin. *"Extending into.."* the skin means that the tumour has invaded the skin and category '2' should be used. *"Extending up to.."* the skin may not indicate invasion. In this case if no other evidence of invasion is present, use category '4'.

Occurrence of skin nodules confined to same breast. Skin nodules should be recorded in the multiple lesions field.

Inflammatory carcinoma. Only if the term inflammatory is supported by a pathological confirmation should this category be used. If inflammatory carcinoma is mentioned, but the histology report is negative, and the tumour was completely resected, category '3' should be used. In the absence of a histology report, disease extent category '4' should be used.

Total destruction of breast tissue with skin involvement. If there is no mention of skin involvement, category '4' should be used.

3 = probably limited to site of origin
There is no specific statement that the tumour is limited to the site of origin, but there is no evidence suggesting that it has invaded surrounding tissues.

4 = possibly invading neighbouring structures
There is a clinical description suggesting that the tumour may have invaded surrounding tissues, but the pathological or other evidence needed to use extent '2' is unavailable. This category is appropriate in cases with clinical descriptions such as *"extending to chest muscle"* or *"involves the deep tissues"*, but where no pathological investigation of the chest wall, or a chest X-ray was performed. Cases with a clinical diagnosis of inflammatory carcinoma without pathology should be classified as extent '4'. This category is appropriate if the description "extending up to the skin" is used. For the audit only, cases described as "*tethered to the skin*" of *"fixed"* were also included in this category.

5 = micro-invasive - for GRACE only
Subsequent to the audit of breast cancer data described in Chapter Three, some amendments to the rules for GRACE were made in relation to the recording of data available from the NHS Breast Screening Programme. One data item of

interest is the presence of micro-invasion. This classification is appropriate if the pathology report specifically states that the tumour is micro-invasive.

'?' = don't know, i.e. known as not known
The pathological data available are insufficient for any of the previous five categories. This category should only be used if a pathology report or post mortem report is available. In other cases the field should be left blank.

A4.3 Multiple lesions and 'found at post mortem' (for use on GRACE)

There are two additional fields on the audit database and GRACE in which information previously classified in the extent field is recorded. In both cases the extent of the primary tumour is recorded in the 'DISEASE_EXT' field.

The MULT_LESIONS field records the presence of multiple lesions including satellite skin nodules

The WHY_NOTREAT field records the reason why a tumour was not treated, which includes 'found at post mortem'.

A4.4 Local lymph nodes

The local nodes include the ipsilateral axillary nodes, the interpectoral (Rotter's) nodes, and the ipsilateral internal mammary lymph nodes. A single field was created on the audit database to code both clinical and pathological information together in a hierarchical fashion. This system was not continued in the final set of rules for use on GRACE, as separate clinical and pathological node data are required for some purposes.

A4.4.1 Local nodes categories (for use on the audit database)

Four fields were created to record data describing the local nodes.

a. Local node status (for use on the audit database)

Data may be clinical or pathological. All data describing the local nodes were recorded in one field. Pathological data, where available, replaced clinical data in this field.

'1' = local nodes positive and pathologically confirmed
A pathology report was available and specifically stated that one or more of the sampled local nodes examined was positively confirmed as being involved.

'2' = local nodes negative and pathologically confirmed
A pathology report was available and stated that all the nodes examined were not involved. This classification includes nodes described as being "reactive".

'3' = local nodes clinically positive
At diagnosis the local nodes were found to be enlarged or palpable. This category was used for cases where there was clinical suspicion, but no mention of the local nodes was made on a pathology report.

'4' = local nodes clinically negative
At diagnosis the nodes were not enlarged or palpable (e.g. 'no nodes'). This category was used if no clinical data were given, or they were unclear.

'X' = insufficient information
This classification was used where the patient was investigated either clinically or pathologically, but insufficient data describing the local nodes were available.

b. Local node mobility (for use on the audit database)
Where the local nodes were positive either clinically or pathologically, data describing fixation or mobility of the nodes to surrounding tissues were recorded in a separate field.

'F' = fixed
The pathological or clinical description indicate that involved nodes were fixed.

'M' = mobile
The pathological or clinical description indicate that involved nodes were mobile

If no data were available the field was left blank.

A4.4.2 Local nodes (for use on GRACE)
Separate fields were provided to record clinical and pathological data .

a. Clinical local nodes (for use on GRACE)

1 = local nodes clinically negative
At diagnosis the local nodes were not enlarged or palpable. This classification must not be used if no clinical data are given or the data are ambiguous.

2 = local nodes clinically positive and mobile
At diagnosis the local nodes are enlarged or palpable and mobile or the local nodes were involved and mobile.

3 = local nodes clinically positive and fixed
At diagnosis the local nodes were enlarged or palpable and fixed or the local nodes were involved and fixed.

4 = local nodes clinically positive, mobility unknown
At diagnosis the local nodes were enlarged or palpable or involved but no data concerning the mobility of the local nodes were available.

b. Pathological local nodes (for use on GRACE)

1 = local nodes stated to be pathologically negative
Pathology report states that all the sampled local nodes were not involved. This classification includes nodes described as being "reactive".

2 = local nodes stated to be pathologically positive
A pathology report states that one or more of the sampled local nodes was pathologically involved.

3 = local node sampling mentioned but not reported
Clinical notes state that local nodes were sampled but node status was not mentioned on a pathology report.

A4.5 Distant lymph nodes

Cervical, supra-clavicular, contralateral axillary and contralateral internal mammary nodes are distant lymph nodes. During the audit a single field was created to code both clinical and pathological data in an identical fashion to that used for local node data. This system was not continued on GRACE, as it is sometimes necessary to separate clinical and pathological data. Two fields were created for GRACE which were identical in structure and coding rules to those described in relation to the local nodes.

A4.6 Presence or absence of metastases

A single field was created on the audit database to code both clinical and pathological data together in a hierarchical fashion. This system was not retained on GRACE, as for some tasks it is necessary to separate clinical and pathological metastases data.

A4.6.1 Metastases (for use on the audit database)

All metastatic sites were recorded in one field. Pathological data confirming the presence of metastases, where available, replaced clinical data in this field.

'1' = the presence of metastases was confirmed pathologically
Metastatic sites were examined, and a pathology report indicated that the site(s) were involved.

'2' = the absence of metastases was confirmed pathologically
All suspected metastatic sites were examined and found to be negative.

'3'= metastases were present at the time of clinical diagnosis
Data are present which support a clinical diagnosis of metastases. These include symptoms at presentation, chest X-rays, liver function tests, liver scans, bone scans, other imaging techniques or biochemical tests. Treatment of a site also indicates clinical suspicion.

'4'= no evidence of metastases at time of clinical diagnosis
If all the suspected metastatic sites were examined and found to be negative this category is appropriate. If there is no mention of metastases, and there have been

no specific tests performed looking for metastases, although it can be assumed that the clinician did not think the tumour had spread, and a '4' cannot be entered in this field.

'?' = clinical suspicion of metastases but no supporting evidence
If there was clinical suspicion of metastases, e.g. symptoms such as bone pain or an enlarged liver, but no further investigations were carried, use this category. This is also appropriate for cases where symptoms commonly associated with the presence of metastases were suspected clinically, but no investigations were carried out, and death occurred within a few months of diagnosis.

A4.6.2 Metastases (for use on GRACE)

Clinical and pathological data describing the presence or absence of metastases are recorded in separate fields on GRACE.

a. Clinical metastases (for use on GRACE)

'1' = ovary (not used for breast cancer cases)

'2' = present
The patient was examined and metastases identified. Test results e.g. chest X rays, liver function tests, full blood counts or bone scans may support the diagnosis. If clinical evidence is weak, there may be other supporting data such as details of treatment of a site with radiation. This category can be used for metastases found at operation which were not sampled pathologically.

'3' = insufficient information
A site was suspected of being metastatically involved, but no unequivocal evidence is available. This category is appropriate where common symptoms such as back ache were present, but no investigations were performed, and the patient died from breast cancer within a few months of diagnosis.

b. Pathological metastases (for use on GRACE)

'1' = pathologically negative (all sites)
All suspected metastases were pathologically examined and found negative.

'2' = pathologically positive (any site)
One or more suspected metastatic sites was found to be pathologically positive.

APPENDIX FIVE

QUALITY ASSURANCE MEASURES USED IN CHAPTER THREE

The following formulae define Quality Assurance (QA) measures presented in Chapter Three. Table A5.1 provides the QA measures calculated in Chapter Three for separate T, N, and M staging components, and for full TNM stages.

Definitions

Completeness	The proportion of cancers with data present in the notes that had data on the PDP11 regardless of the accuracy of transfer
Accuracy	The proportion of cancers with data in the notes that had that data transferred correctly onto the PDP11 database
Mismatched data	The proportion of cancers with data in the notes and data on the PDP11 database for which the specific values recorded did not match (calculated for tumour size only)
Added data	The proportion of cancers with data on the PDP11 database that did not have data in the notes

In the formulae that follow, both the numerator and the denominator refer to number of breast cancers which fulfil the specified criteria.

A5.1 QA measures for staging data

$$\text{completeness} = \frac{[\text{stage in the notes and any stage recorded on the PDP11}]}{[\text{stage in the notes}]} * 100$$

$$\text{accuracy} = \frac{[\text{stage in the notes same as stage on the PDP11}]}{[\text{stage on the PDP11}]} * 100$$

$$\text{added data} = \frac{[\text{stage on the PDP11 and no stage in the notes}]}{[\text{stage on the PDP11}]} * 100$$

A5.2 QA measures for tumour size

completeness = [size in the notes and any size recorded on the PDP11] *100
[size in the notes]

accuracy = [size in the notes same as size on the PDP11] *100
[size in the PDP11]

mismatched data = [size in the notes not same as size on the PDP11] *100
[size in the notes and any size recorded on the PDP11]

A5.3 QA measures for local and distant nodal status

completeness = [nodal data in the notes and any nodal data on the PDP11] *100
[nodal data in the notes]

accuracy = [nodal data in the notes same as nodal data on the PDP11] *100
[nodal data on the PDP11]

added data = [nodal data on the PDP11 but no data in the notes] *100
[nodal data on the PDP11]

A5.4 QA measures for disease extent

completeness = [extent in the notes and extent recorded on the PDP11] *100
[extent in the notes]

accuracy = [extent in the notes same as extent on the PDP11] *100
[extent on the PDP11]

added data = [extent on the PDP11 but no extent data in the notes] *100
[extent on the PDP11]

5.5 QA measures for metastases

Completeness and % added data were calculated for positive metastases only.

completeness = [metastasis data in the notes and on the PDP11] *100

— [metastasis data in the notes]

added data = [metastasis data on the PDP11 but not in the notes] *100

[metastasis data on the PDP11]

Table A5.1 Quality assurance measures for staging data held at the WMCIU							
Staging data	**% cases with data on PDP11**		**% cases with data in case notes**		**Completeness of transfer**	**Accuracy of transfer**	**% added data on PDP11**
	No.	**%**	**No.**	**%**	**%**	**%**	**%**
T component							
T0	0	0	0	0	-	-	-
T1	33	22.8	32	21.8	46.9	93.3	57.6
T2	58	40.0	61	41.5	65.6	100.0	31.0
T3	34	23.4	37	25.2	56.8	100.0	38.2
T4	20	13.8	17	11.6	64.7	100.0	40.0
Overall T	**145**	**100.0**	**147**	**100.0**	**59.2**	**98.9**	**40.0**
N component							
N0	83	60.6	68	51.5	58.8	100.0	49.4
N1	43	31.4	54	40.9	57.4	93.5	32.6
N2	5	3.6	4	3.0	50.0	100.0	60.0
N3	6	4.4	6	4.5	66.7	100.0	33.3
Overall N	**137**	**100.0**	**132**	**100.0**	**58.3**	**97.4**	**43.8**
M component							
M0	119	93.7	64	91.4	50.0	100.0	72.2
M1	8	6.3	6	8.6	83.3	100.0	37.5
Overall M	**127**	**100.0**	**70**	**100.0**	**52.9**	**100.0**	**70.0**
TNM stage							
0	26	22.0	0	0.0	-	-	100.0
1	15	12.7	25	25.0	52.0	92.3	20.0
2 (2a, 2b)	21	17.8	22	22.0	36.4	87.5	61.9
3 (3a,3b)	29	24.6	34	34.0	61.8	100.0	27.6
4	27	22.9	19	19.0	63.2	100.0	55.6
Overall TNM	**118**	**100.0**	**100**	**100.0**	**54.0**	**96.3**	**54.2**

APPENDIX SIX

PROGRAMME LISTINGS FOR THE FOUR STAGING ALGORITHMS DESCRIBED IN CHAPTER THREE

The programme listings below describe the four staging Algorithms used in Chapter Three. Programmes were run using the package Epi Info -Public Domain Software for Epidemiology and Disease Surveillance, Version 5.0 (Dean *et al.*, 1990).

Staging Algorithm 1

```
* Cannot stage (calcstg=8) if DC only, Post Mortem only, known noncarcinomas (ICDM>8999)
* Both probably/possibly extent data taken to mean a definite finding, no subdivisions of T4
* Pathological size and node data preferred, but clinical data used otherwise
* Unless specified as fixed, local nodes are presumed mobile
* If mets not positive, and extent not 2 or 4, must have size and local node status to be staged
* The default value for M is 0 i.e. no mets
  define calct __
  define calcn __
  define calcm __
  define calcstg __
  calct="x"
  if sizeclin>50 and sizeclin<>999 then calct=3
  if sizeclin>20 and sizeclin<=50 then calct=2
  if sizeclin<=20 then calct=1
  if sizerad>50 and sizerad<>999 then calct=3
  if sizerad>20 and sizerad<=50 then calct=2
  if sizerad<=20 then calct=1
  if sizepath>50 and sizepath<>999 then calct=3
  if sizepath>20 and sizepath<=50 then calct=2
  if sizepath<=20 then calct=1
  if diseasee=2 or diseasee=4 then calct=4
  if diseasee=0 then calct="i"
  calcn="x"
  if localnod=4 then calcn=0
  if localnod=2 then calcn=0
  if localnod=3 then calcn=1
  if localnod=1 then calcn=1
  if localmob="f" then calcn=2
  if localnod=8 or localnod="x" then calcn="x"
  calcm=0
  if mets=1 or mets=3 or distantn=1 or distantn=3 then calcm=1
  if mets="x" then calcm=1
  calcstg="X"
  if calct="i" then calcstg="0"
  if calct=1 then calcstg="1"
  if calcn=1 then calcstg="2a"
  if calcn=2 then calcstg="3a"
  if calct=2 then calcstg="2a"
  if calcn=1 and calct=2 then calcstg="2b"
  if calcn=2 calcstg="3a"
  if calct=3 calcstg="2b"
  if calct=3 and diseasee<>"x" and calcn=1 then calcstg="3a"
  if calct=3 and diseasee<>"x" and calcn=2 then calcstg="3a"
  if calct="x" or calcn ="x" or diseaseext="'x" then calcstg="X"
  if calct=4 then calcstg="3b"
  if calcm=1 then calcstg="4"
  if ICDM>8999 or diagbasi="D" or diagbasi="P" then calcstg="8"
```

Staging Algorithm 2

```
* Cannot stage (calcstg2=8) if DC only, Post Mortem only, known noncarcinimas ICDM>8999
* Only definite extent data taken to mean a definite finding, subdivisions of T4 not used.
* Pathological size and node data are preferred, but clinical data  used otherwise.
* Unless specified as fixed, local nodes are presumed mobile.
* If mets not positive,extent not 2 or 4, must  have size and local node status to be staged.
* The default value for M is 0 i.e. no mets
   define calct2 __
   define calcn2 __
   define calcm2 __
   define calcstg2 __
   calct2="x"
   if sizeclin>50 and sizeclin<>999 then calct2=3
   if sizeclin>20 and sizeclin<=50 then calct2=2
   if sizeclin<=20 then calct2=1
   if sizerad>50 and sizerad<>999 then calct2=3
   if sizerad>20 and sizerad<=50 then calct2=2
   if sizerad<=20 then calct2=1
   if sizepath>50 and sizepath<>999 then calct2=3
   if sizepath>20 and sizepath<=50 then calct2=2
   if sizepath<=20 then calct2=1
   if diseasee=2 then calct2=4
   if diseasee=0 then calct2="i"
   calcn2="x"
   if localnod=4 then calcn2=0
   if localnod=2 then calcn2=0
   if localnod=3 then calcn2=1
   if localnod=1 then calcn2=1
   if localmob="f" then calcn2=2
   if localnod=8 or localnod="x" then calcn2="x"
   calcm2=0
   if mets=1 or mets=3 or distantn=1 or distantn=3 then calcm2=1
   if mets="x" then calcm2=1
   calcstg2="X"
   if calct2="i" then calcstg2="0"
   if calct2=1 then calcstg2="1"
   if calcn2=1 then calcstg2="2a"
   if calcn2=2  then calcstg2="3a"
   if calct2=2  then calcstg2="2a"
   if calcn2=1 and calct2=2 then calcstg2="2b"
   if calcn2=2 then calcstg2="3a"
   if calct2=3 then calcstg2="2b"
   if calct2=3 and calcn2=1 then calcstg2="3a"
   if calct2=3 and calcn2=2 then calcstg2="3a"
   if calct2="x" or calcn2="x" or diseaseext="x" then calcstg2="X"
   if calct2=4 then calcstg2="3b"
   if calcm2=1 then calcstg2="4"
   if ICDM>8999 then calcstg2="8"
   if diagbasi="D" or diagbasi="P" then calcstg2="8"
```

Staging Algorithm 3

```
*Cannot stage (calcstg3=8) if DC or Post Mortem only, known noncarcinomas (ICDM>8999)
* Only definite extent data is taken to mean a definite finding, subdivisions of T4 not used.
* Pathological size and node data are preferred, but clinical data used otherwise.
* Unless specified as "mobile", local nodes are presumed fixed.
* If mets not positive, extent not 2 or 4, must have size and local node status to be staged.
* The default value for M is 0 i.e. no mets
   define calct3 __
   define calcn3 __
   define calcm3 __
   define calcstg3 __
   calct3="x"
   if sizeclin>50 and sizeclin<>999 then calct3=3
   if sizeclin>20 and sizeclin<=50 then calct3=2
   if sizeclin<=20 then calct3=1
   if sizerad>50 and sizerad<>999 then calct3=3
   if sizerad>20 and sizerad<=50 then calct3=2
   if sizerad<=20 then calct3=1
   if sizepath>50 and sizepath<>999 then calct3=3
   if sizepath>20 and sizepath<=50 then calct3=2
   if sizepath<=20 then calct3=1
   if diseasee=2 then calct3=4
   if diseasee=0 then calct3="i"
   calcn3="x"
   if localnod=4 then calcn3=0
   if localnod=2 then calcn3=0
   if localnod=3 then calcn3=2
   if localnod=1 then calcn3=2
   if localmob="m" then calcn3=1
   if localnod=8 or localnod="x" then calcn3="x"
   calcm3=0
   if mets=1 or mets=3 or distantn=1 or distantn=3 then calcm3=1
   if mets="x" then calcm3=1
   calcstg3="X"
   if calct3="i" then calcstg3="0"
   if calct3=1 then calcstg3="1"
   if calcn3=1 then calcstg3="2a"
   if calcn3=2 then calcstg3="3a"
   if calct3=2 then calcstg3="2a"
   if calcn3=1 and calct3=2 then calcstg3="2b"
   if calcn3=2 then calcstg3="3a"
   if calct3=3 then calcstg3="2b"
   if calct3=3 and calcn3=1 then calcstg3="3a"
   if calct3=3 and calcn3=2 then calcstg3="3a"
   if calct3="x" or calcn3="x" or diseaseext="x" then calcstg3="X"
   if calct3=4 then calcstg3="3b"
   if calcm3=1 then calcstg3="4"
   if ICDM>8999 then calcstg3="8"
   if diagbasi="D" then calcstg3="8"
   if diagbasi="P" then calcstg3="8"
```

Staging Algorithm 4

```
* Cannot stage (calcstg4=8) if  DC only, PM only, known noncarcinimas (ICDM>8999)
* Only definite extent data is taken to mean a definite  finding, subdivisions of T4 not used.
* Pathological size and node data are preferred, but clinical data  used otherwise.
* Unless specified as "fixed", local nodes are presumed mobile.
* If mets not positive, extent not 2 or 4, must have size and local node status to be staged.
* The default value for M is 0 i.e. no mets  if mets uncertain assume negative
   define calct4 __
   define calcn4 __
   define calcm4 __
   define calcstg4 __
   calct4="x"
   if sizeclin>50 and sizeclin<>999 then calct4=3
   if sizeclin>20 and sizeclin<=50 then calct4=2
   if sizeclin<=20 then calct4=1
   if sizerad>50 and sizerad<>999 then calct4=3
   if sizerad>20 and sizerad<=50 then calct4=2
   if sizerad<=20 then calct4=1
   if sizepath>50 and sizepath<>999 then calct4=3
   if sizepath>20 and sizepath<=50 then calct4=2
   if sizepath<=20 then calct4=1
   if diseasee=2 then calct4=4
   if diseasee=0 then calct4="i"
   calcn4="x"
   if localnod=4 then calcn4=0
   if localnod=2 then calcn4=0
   if localnod=3 then calcn4=1
   if localnod=1 then calcn4=1
   if localmob="f" then calcn4=2
   if localnod=8 or localnod="x" then calcn4="x"
   calcm4=0
   if mets=1 or mets=3 or distantn=1 or distantn=3 then calcm4=1
   calcstg4="X"
   if calct4="i" then calcstg4="0"
   if calct4=1 then calcstg4="1"
   if calcn4=1 then calcstg4="2a"
   if calcn4=2 then calcstg4="3a"
   if calct4=2 then calcstg4="2a"
   if calcn4=1 and calct4=2 then calcstg4="2b"
   if calcn4=2 then calcstg4="3a"
   if calct4=3 then calcstg4="2b"
   if calct4=3 and calcn4=1 then calcstg4="3a"
   if calct4=3 and calcn4=2 then calcstg4="3a"
   if calct4="x" or calcn4="x" or diseaseext="x" then calcstg4="X"
   if calct4=4 then calcstg4="3b"
   if calcm4=1 then calcstg4="4"
   if ICDM>8999 then calcstg4="8"
   if diagbasi="D" then calcstg4="8"
   if diagbasi="P" then calcstg4="8"
```

APPENDIX SEVEN

PATIENT AND TUMOUR CHARACTERISTICS OF BREAST CANCERS DIAGNOSED IN 1983 AND 1987

Tables A7.1 to A7.13 present patient and tumour characteristic data for breast cancers diagnosed in 1983 and 1987 described in Chapter Four.

Table A7.1 Distribution of cases by age at diagnosis

Age (years)	1983		1987		Both years	
	No.	%	No.	%	No.	%
<50	526	21.52	603	22.9	1129	22.24
50-69	1088	44.52	1106	42.01	2194	43.21
70+	830	33.96	924	35.09	1754	34.55
Total	2444	100	2633	100	5077	100

Table A7.2 Distribution of cases by DHA

DHA	1983		1987		Both years	
	No.	%	No.	%	No.	%
M02	78	3.19	79	3	157	3.09
M04	100	4.09	146	5.55	246	4.85
M05	161	6.59	219	8.32	380	7.48
M07	221	9.04	227	8.62	448	8.82
M17	170	6.96	165	6.27	335	6.6
M18	135	5.52	186	7.06	321	6.32
M19	144	5.89	131	4.98	275	5.42
M20	94	3.85	94	3.57	188	3.7
M21	137	5.61	128	4.86	265	5.22
M22	119	4.87	113	4.29	232	4.57
M25	237	9.7	210	7.98	447	8.8
M26	109	4.46	157	5.96	266	5.24
M27	208	8.51	269	10.22	477	9.4
M28	231	9.45	276	10.48	507	9.99
M29	300	12.27	233	8.85	533	10.5
Total	2444	100	2633	100	5077	100

Table A7.3 Distribution of cases by age at diagnosis (in years) and DHA																
DHA	1983								1987							
	<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
M02	13	16.67	34	43.59	31	39.74	78	100	18	22.78	35	44.3	26	32.91	79	100
M04	21	21	39	39	40	40	100	100	36	24.66	59	40.41	51	34.93	146	100
M05	43	26.71	67	41.61	51	31.68	161	100	62	28.31	75	34.25	82	37.44	219	100
M07	43	19.46	96	43.44	82	37.1	221	100	46	20.26	108	47.58	73	32.16	227	100
M17	35	20.59	85	50	50	29.41	170	100	38	23.03	67	40.61	60	36.36	165	100
M18	32	23.7	64	47.41	39	28.89	135	100	35	18.82	85	45.7	66	35.48	186	100
M19	28	19.44	71	49.31	45	31.25	144	100	26	19.85	59	45.04	46	35.11	131	100
M20	26	27.66	42	44.68	26	27.66	94	100	23	24.47	37	39.36	34	36.17	94	100
M21	39	28.47	54	39.42	44	32.12	137	100	30	23.44	60	46.88	38	29.69	128	100
M22	26	21.85	54	45.38	39	32.77	119	100	28	24.78	51	45.13	34	30.09	113	100
M25	41	17.3	93	39.24	103	43.46	237	100	31	14.76	87	41.43	92	43.81	210	100
M26	26	23.85	61	55.96	22	20.18	109	100	42	26.75	64	40.76	51	32.48	157	100
M27	54	25.96	94	45.19	60	28.85	208	100	77	28.62	109	40.52	83	30.86	269	100
M28	44	19.05	105	45.45	82	35.5	231	100	64	23.19	119	43.12	93	33.7	276	100
M29	55	18.33	129	43	116	38.67	300	100	47	20.17	91	39.06	95	40.77	233	100
Total	526	21.52	1088	44.52	830	33.96	2444	100	603	22.9	1106	42.01	924	35.09	2633	100

Table A7.4 Distribution of cases by stage						
Stage	1983		1987		Both years	
	No.	%	No.	%	No.	%
0	12	0.49	15	0.57	27	0.53
1	230	9.41	264	10.03	494	9.73
2a	349	14.28	329	12.5	678	13.35
2b	235	9.62	221	8.39	456	8.98
3	461	18.86	327	12.42	788	15.52
4	275	11.25	123	4.67	398	7.84
Unstaged	882	36.09	1354	51.42	2236	44.04
Total	2444	100	2633	100	5077	100

Table A7.5 Distribution of cases by age and stage																
Age	0		1		2a		2b		3		4		Unstaged		Total	
	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%
1983																
<50	5	41.67	68	29.57	92	26.36	65	27.66	51	11.06	31	11.27	214	24.26	526	21.52
50-69	6	50	110	47.83	186	53.3	114	48.51	201	43.6	127	46.18	344	39	1088	44.52
70+	1	8.33	52	22.61	71	20.34	56	23.83	209	45.34	117	42.55	324	36.73	830	33.96
Total	12	100	230	100	349	100	235	100	461	100	275	100	882	100	2444	100
1987																
<50	8	53.33	63	23.86	73	22.19	58	26.24	35	10.7	7	5.69	359	26.51	603	22.9
50-69	5	33.33	129	48.86	163	49.54	103	46.61	133	40.67	61	49.59	512	37.81	1106	42.01
70+	2	13.33	72	27.27	93	28.27	60	27.15	159	48.62	55	44.72	483	35.67	924	35.09
Total	15	100	264	100	329	100	221	100	327	100	123	100	1354	100	2633	100
Both years																
<50	13	48.15	131	26.52	165	24.34	123	26.97	86	10.91	38	9.55	573	25.63	1129	22.24
50-69	11	40.74	239	48.38	349	51.47	217	47.59	334	42.39	188	47.24	856	38.28	2194	43.21
70+	3	11.11	124	25.1	164	24.19	116	25.44	368	46.7	172	43.22	807	36.09	1754	34.55
Total	27	100	494	100	678	100	456	100	788	100	398	100	2236	100	5077	100

Table A7.6 Distribution of cases by grade						
Grade*	1983		1987		Both years	
	No.	%	No.	%	No.	%
Good	177	7.24	382	14.51	559	11.01
Poor	568	23.24	596	22.64	1164	22.93
Unknown	1699	69.52	1655	62.86	3354	66.06
Total	2444	100	2633	100	5077	100

* Definition of grades

Good grade Bloom and Richardson grade I and II

Cancers described on a pathology report as "low grade", "well differentiated" or "moderately well differentiated";

Poor grade Bloom and Richardson grade III

Cancers described on a pathology report as "undifferentiated", "anaplastic", "poor grade" or "high grade"

Table A7.7 Distribution of cases grade and age																								
Grade	1983								1987								Both years							
	<50		50-69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Good	33	6.27	82	7.54	62	7.47	177	7.24	84	13.93	186	16.82	112	12.12	382	14.51	117	10.36	268	12.22	174	9.92	559	11.01
Poor	136	25.86	302	27.76	130	15.66	568	23.24	160	26.53	272	24.59	164	17.75	596	22.64	296	26.22	574	26.16	294	16.76	1164	22.93
Un-known	357	67.87	704	64.71	638	76.87	1699	69.52	359	59.54	648	58.59	648	70.13	1655	62.86	716	63.42	1352	61.62	1286	73.32	3354	66.06
Total	526	21.52	1088	44.52	830	33.96	2444	100	603	24.67	1106	45.25	924	37.81	2633	107.73	1129	100	2194	100	1754	100	5077	100

Table A7.8 Distribution of cases by deprivation level						
Deprivation level (range of Townsend scores)	1983		1987		Both years	
	No.	%	No.	%	No.	%
Most affluent (< -3)	491	20.09	426	16.18	917	18.06
Affluent (-3.0 to <-1)	486	19.89	493	18.72	979	19.28
Middle range (-1 to <1)	486	19.89	551	20.93	1037	20.43
Deprived (1 to <3.5)	504	20.62	528	20.05	1032	20.33
Most deprived (3.5+)	461	18.86	559	21.23	1020	20.09
Unknown	16	0.65	76	2.89	92	1.81
Total	2444	100	2633	100	5077	100

Table A7.9 Distribution of cases by deprivation level and age								
Deprivation level (range of Townsend scores)	<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%
Most affluent (< -3)	307	30.1	434	42.55	279	27.35	1020	100
Affluent (-3.0 to <-1)	241	23.35	445	43.12	346	33.53	1032	100
Middle range (-1 to <1)	218	21.02	437	42.14	382	36.84	1037	100
Deprived (1 to <3.5)	180	18.39	437	44.64	362	36.98	979	100
Most deprived (3.5+)	168	18.32	414	45.15	335	36.53	917	100
Unknown	15	16.3	27	29.35	50	54.35	92	100
Total	1129	22.24	2194	43.21	1754	34.55	5077	100

Table A7.10 Distribution of cases by deprivation level and stage																
Deprivation level (range of Townsend scores)	0		1		2a		2b		3		4		Unknown		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Most affluent (< -3)	11	1.08	112	10.98	134	13.14	104	10.2	126	12.35	62	6.08	471	46.18	1020	100
Affluent (-3.0 to <-1)	4	0.39	111	10.76	159	15.41	98	9.5	153	14.83	79	7.66	428	41.47	1032	100
Middle range (-1 to <1)	6	0.58	102	9.84	136	13.11	96	9.26	164	15.81	72	6.94	461	44.46	1037	100
Deprived (1 to <3.5)	2	0.2	75	7.66	115	11.75	87	8.89	183	18.69	87	8.89	430	43.92	979	100
Most deprived (3.5+)	4	0.44	90	9.81	131	14.29	66	7.2	156	17.01	92	10.03	378	41.22	917	100
Unknown	0	0.00	3	3.26	4	4.35	4	4.35	7	7.61	6	6.52	68	73.91	92	100
Total	27	0.53	493	9.71	679	13.37	455	8.96	789	15.54	398	7.84	2236	44.04	5077	100

Table A7.11 Distribution of cases by histological group						
Histological group	1983		1987		Both years	
	No.	%	No.	%	No.	%
Unknown	362	14.81	126	4.79	488	9.61
Ductal carcinoma not otherwise specified (NST)	1742	71.28	2099	79.72	3841	75.65
Lobular carcinoma	154	6.3	204	7.75	358	7.05
Sarcoma	6	0.25	3	0.11	9	0.18
Special type carcinoma	150	6.14	145	5.51	295	5.81
In situ carcinoma	30	1.23	56	2.13	86	1.69
Total	2444	100	2633	100	5077	100

Table A7.12 Proportion of cases with unknown histology by age									
Age (years)	1983			1987			Both years		
	Total cases	No.	%	Total cases	No.	%	Total cases	No.	%
<50	526	16	3.04	603	9	1.49	1129	25	2.21
50-69	1088	97	8.92	1106	37	3.35	2194	134	6.11
70+	830	249	30	924	80	8.66	1754	329	18.76
Total	2444	362	14.81	2633	126	4.79	5077	488	9.61

Table A7.13 Proportion of cases with unknown histology by deprivation level									
Deprivation level (range of Townsend scores)	1983			1987			Both years		
	Total cases	No.	%	Total cases	No.	%	Total cases	No.	%
Most affluent (< -3)	491	42	8.55	426	23	5.4	917	65	7.09
Affluent (-3.0 to <-1)	486	55	11.32	493	24	4.87	979	79	8.07
Middle range (-1 to <1)	486	82	16.87	551	13	2.36	1037	95	9.16
Deprived (1 to <3.5)	504	91	18.06	528	36	6.82	1032	127	12.31
Most deprived (3.5+)	461	89	19.31	559	27	4.83	1020	116	11.37
Unknown	16	3	18.75	76	3	3.95	92	6	6.52
Total	2444	362	14.81	2633	126	4.79	5077	488	9.61

APPENDIX EIGHT

TREATMENT OF AND SURVIVAL FOR BREAST CANCERS DIAGNOSED IN 1983 AND 1987

Tables A8.1 to A8.10 present treatment data and Tables A8.11 to A8.21 present five year relative survival rates described in Chapter Five.

A8.1 Treatment recorded for breast cancers diagnosed in 1983 and 1987

The treatment presented in Tables A8.1 to A8.10 is restricted to that received within 6 months of diagnosis.

KEY

Surgical treatment

None	No treatment data recorded, either surgical or non-surgical
NS	No surgery recorded, but non-surgical oncology recorded
CO	Conservative surgery recorded +/- adjuvant therapy
MA	Mastectomy recorded +/- adjuvant therapy

If a mastectomy is recorded following conservative surgery, with both operations taking place within 6 months of diagnosis, only the mastectomy has been reported in this investigation.

Non-surgical oncology

RT	Radiotherapy
HT	Hormone therapy
CT	Chemotherapy

Combined surgery and non-surgical oncology

RT	Radiotherapy only
HT	Hormone therapy only
RT + HT	Radiotherapy and hormone therapy
CT	Chemotherapy with or without radiotherapy and/or hormone therapy

Table A8.1 Overall treatment by age and year																								
Option	1983								1987								Both years							
	<50		50-69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
None	8	1.52	46	4.23	118	14.22	172	7.04	15	2.49	61	5.52	133	14.39	209	7.94	23	2.04	107	4.88	251	14.31	381	7.5
NS	27	5.13	136	12.5	250	30.12	413	16.9	23	3.81	135	12.21	270	29.22	428	16.26	50	4.43	271	12.35	520	29.65	841	16.56
CO	144	27.38	202	18.57	160	19.28	506	20.7	314	52.07	436	39.42	264	28.57	1014	38.51	458	40.57	638	29.08	424	24.17	1520	29.94
MA	347	65.97	704	64.71	302	36.39	1353	55.36	251	41.63	474	42.86	257	27.81	982	37.3	598	52.97	1178	53.69	559	31.87	2335	45.99
Total	526	100	1088	100	830	100	2444	100	603	100	1106	100	924	100	2633	100	1129	100	2194	100	1754	100	5077	100

Table A8.2 Overall non-surgical oncology by age and year																								
	1983								1987								Both years							
	<50		50 - 69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
RT	246	46.77	436	40.07	147	17.71	829	33.92	356	59.04	513	46.38	161	17.42	1030	39.12	602	53.32	949	43.25	308	17.56	1859	36.62
HT	76	14.45	246	22.61	336	40.48	658	26.92	231	38.31	613	55.42	536	58.01	1380	52.41	307	27.19	859	39.15	872	49.71	2038	40.14
CT	50	9.51	59	5.42	9	1.08	118	4.83	34	5.64	34	3.07	10	1.08	78	2.96	84	7.44	93	4.24	19	1.08	196	3.86

* some women received more than one type of non-surgical treatment

Table A8.3 Treatment by DHA and year																					
DHA	1983							1987							Both years						
	Total cases	MA		CO		RT		Total cases	MA		CO		RT		Total cases	MA		CO		RT	
		No.	%	No.	%	No.	%		No.	%	No.	%	No.	%		No.	%	No.	%	No.	%
Region	2444	1353	55.36	506	20.7	829	33.92	2633	982	37.3	1014	38.51	1030	39.12	5077	2335	45.99	1520	29.94	1859	36.62
M02	78	52	66.67	6	7.69	34	43.59	79	42	53.16	24	30.38	38	48.1	157	94	59.87	30	19.11	72	45.86
M04	100	55	55	25	25	29	29	146	52	35.62	60	41.1	51	34.93	246	107	43.5	85	34.55	80	32.52
M05	161	92	57.14	26	16.15	63	39.13	219	104	47.49	72	32.88	96	43.84	380	196	51.58	98	25.79	159	41.84
M07	221	105	47.51	32	14.48	42	19	227	84	37	61	26.87	81	35.68	448	189	42.19	93	20.76	123	27.46
M17	170	96	56.47	40	23.53	96	56.47	165	55	33.33	75	45.45	68	41.21	335	151	45.07	115	34.33	164	48.96
M18	135	92	68.15	18	13.33	38	28.15	186	82	44.09	61	32.8	74	39.78	321	174	54.21	79	24.61	112	34.89
M19	144	74	51.39	34	23.61	59	40.97	131	50	38.17	51	38.93	56	42.75	275	124	45.09	85	30.91	115	41.82
M20	94	52	55.32	20	21.28	35	37.23	94	36	38.3	40	42.55	46	48.94	188	88	46.81	60	31.91	81	43.09
M21	137	92	67.15	16	11.68	27	19.71	128	77	60.16	31	24.22	31	24.22	265	169	63.77	47	17.74	58	21.89
M22	119	82	68.91	16	13.45	46	38.66	113	51	45.13	45	39.82	50	44.25	232	133	57.33	61	26.29	96	41.38
M25	237	105	44.3	72	30.38	63	26.58	210	54	25.71	90	42.86	72	34.29	447	159	35.57	162	36.24	135	30.2
M26	109	70	64.22	19	17.43	39	35.78	157	54	34.39	73	46.5	56	35.67	266	124	46.62	92	34.59	95	35.71
M27	208	103	49.52	55	26.44	77	37.02	269	68	25.28	136	50.56	118	43.87	477	171	35.85	191	40.04	195	40.88
M28	231	134	58.01	57	24.68	79	34.2	276	104	37.68	101	36.59	97	35.14	507	238	46.94	158	31.16	176	34.71
M29	300	149	49.67	70	23.33	102	34	233	69	29.61	94	40.34	96	41.2	533	218	40.9	164	30.77	198	37.15

Table A8.4 Non-surgical oncology for 841 non-surgical cases by age and year																								
Option	1983								1987								Both years							
	<50		50-69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
RT*	9	33.33	19	13.97	10	4.00	38	9.20	6	26.09	15	11.11	7	2.59	28	6.54	15	30.00	34	12.55	17	3.27	66	7.85
HT**	7	25.93	56	41.18	183	73.20	246	59.56	3	13.04	55	40.74	212	28.52	270	63.08	10	20.00	111	40.96	395	75.96	516	61.36
RT+ HT	6	22.22	48	35.29	52	20.80	106	25.67	10	43.48	48	35.56	43	15.93	101	23.60	16	32.00	96	35.42	95	18.27	207	24.61
CT***	5	18.52	13	9.56	5	2.00	23	5.57	4	17.39	17	12.59	8	2.96	29	6.78	9	18.00	30	11.07	13	2.5	52	6.18
Total	27	100	136	100	250	100	413	100	23	100	135	100	270	100	428	100	50	100	271	100	520	100	841	100
Overall provision of radiotherapy and hormone therapy																								
RT	18	66.67	74	54.41	65	26.00	157	38.01	20	86.96	71	52.59	54	20.00	145	33.88	38	76	145	53.51	119	22.88	302	35.91
HT	18	66.67	111	81.62	237	94.80	366	88.62	15	65.22	116	85.93	261	96.67	392	91.59	33	66	227	83.76	498	95.77	758	90.13

* RT only

** HT only

*** CT includes patients with other adjuvant therapies in addition to CT e.g. CT + HT, CT + RT

Table A8.5 Non-surgical oncology for 1520 conservative surgery patients by age and year																								
Option	1983								1987								Both years							
	<50		50-69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
RT*	90	62.50	97	48.02	28	17.50	215	42.49	124	39.49	121	27.75	17	6.44	262	25.84	214	46.72	218	34.17	45	10.61	477	31.38
HT**	5	3.47	18	8.91	34	21.25	57	11.26	14	4.46	69	15.83	98	37.12	181	17.85	19	4.15	87	13.64	132	31.13	238	15.66
RT+HT	6	4.17	10	4.95	4	2.50	20	3.95	102	32.48	178	40.83	45	17.05	325	32.05	108	23.58	188	29.47	49	11.56	345	22.7
CT***	15	10.42	9	4.46	0	0.00	24	4.74	10	3.18	8	1.83	0	0.00	18	1.78	25	5.46	17	2.66	0	0	42	2.76
None	28	19.44	68	33.66	94	58.75	190	37.55	64	20.38	60	13.76	104	39.39	228	22.49	92	20.09	128	20.06	198	46.7	418	27.5
Total	144	100	202	100	160	100	506	100	314	100	436	100	264	100	1014	100	458	100	638	100	424	100	1520	100
Overall provision of radiotherapy and hormone therapy																								
RT	110	76.39	116	57.43	32	20.00	258	50.99	232	73.89	305	69.95	62	23.48	599	59.07	342	74.67	421	65.99	94	22.17	857	56.38
HT	15	10.42	29	14.36	38	23.75	82	16.21	119	37.90	255	58.49	143	54.17	517	50.99	134	29.26	284	44.51	181	42.69	599	39.41

* RT only
** HT only
*** CT includes patients with other adjuvant therapies in addition to CT e.g. CT + HT, CT + RT

Table A8.6 Non-surgical oncology for 2335 mastectomy patients by age and year																								
Option	1983								1987								Both years							
	<50		50-69		70+		Total		<50		50-69		70+		Total		<50		50-69		70+		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
RT*	78	22.48	190	26.99	38	12.58	306	22.62	48	19.12	45	9.49	9	3.50	102	10.39	126	21.07	235	19.95	47	8.41	408	17.47
HT**	14	4.03	56	7.95	49	16.23	119	8.80	42	16.73	150	31.65	94	36.58	286	29.12	56	9.36	206	17.49	143	25.58	405	17.34
RT + HT	25	7.20	45	6.39	11	3.64	81	5.99	46	18.33	87	18.35	36	14.01	169	17.21	71	11.87	132	11.21	47	8.41	250	10.71
CT***	30	8.65	37	5.26	4	1.32	71	5.25	20	7.97	9	1.90	2	0.78	31	3.16	50	8.36	46	3.9	6	1.07	102	4.37
None	200	57.64	376	53.41	200	66.23	776	57.35	95	37.85	183	38.61	116	45.14	394	40.12	295	49.33	559	47.45	316	56.53	1170	50.11
Total	347	100	704	100	302	100	1353	100	251	100	474	100	257	100	982	100	598	100	1178	100	559	100	2335	100
Overall provision of radiotherapy and hormone therapy																								
RT	118	34.01	246	34.94	50	16.56	414	30.60	104	41.43	137	28.90	45	17.51	286	29.12	222	37.12	383	32.51	95	16.99	700	29.98
HT	43	12.39	106	15.06	61	20.20	210	15.52	97	38.65	242	51.05	132	51.36	471	47.96	140	23.41	348	29.54	193	34.53	681	29.16

* RT only

** HT only

*** CT includes patients with other adjuvant therapies in addition to CT e.g. CT + HT, CT + RT

Table A8.7 Surgical caseload of the treating surgeon by year						
Surgical caseload (cases/year)	1983		1987		Both years	
	No.	%	No.	%	No.	%
<=10	390	15.96	276	10.48	666	13.12
11-20	880	36.01	490	18.61	1370	26.98
21-30	210	8.59	577	21.91	787	15.5
31+	320	13.09	507	19.26	827	16.29
Unknown	59	2.41	146	5.55	205	4.04
Not surgically treated	585	23.94	637	24.19	1222	24.07
Total	2444	100	2633	100	5077	100

Table A8.8 Surgical caseload of the treating surgeon by age for surgical cases only																			
Surgical caseload (cases/yr)	1983							1987							Both years						
	Total cases	<50		50-69		70+		Total cases	<50		50-69		70+		Total cases	<50		50-69		70+	
		No.	%	No.	%	No.	%		No.	%	No.	%	No.	%		No.	%	No.	%	No.	%
<=10	390	89	22.82	183	46.92	118	30.26	276	55	19.93	136	49.28	85	30.8	666	144	21.62	319	47.9	203	30.48
11-20	880	242	27.5	445	50.57	193	21.93	490	146	29.8	229	46.73	115	23.47	1370	388	28.32	674	49.2	308	22.48
21-30	210	50	23.81	97	46.19	63	30	577	168	29.12	252	43.67	157	27.21	787	218	27.7	349	44.35	220	27.95
31+	320	92	28.75	154	48.13	74	23.13	507	151	29.78	234	46.15	122	24.06	827	243	29.38	388	46.92	196	23.7
Unknown	59	18	30.51	27	45.76	14	23.73	146	45	30.82	59	40.41	42	28.77	205	63	30.73	86	41.95	56	27.32
Total	1859	491	26.41	906	48.74	462	24.85	1996	565	28.31	910	45.59	521	26.1	3855	1056	27.39	1816	47.11	983	25.5

Table A8.9 Stage distribution according to surgical caseload for surgical cases only												
Stage*	1983				1987				Both years			
	<=30 cases/yr		31+ cases/yr		<=30 cases/yr		31+ cases/yr		<=30 cases/yr		31+ cases/yr	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Early	450	30.41	109	34.06	362	26.95	184	36.29	812	28.76	293	35.43
Locally advanced	440	29.73	81	25.31	275	20.48	105	20.71	715	25.33	186	22.49
Metastatic	65	4.39	14	4.38	21	1.56	8	1.58	86	3.05	22	2.66
Unknown	525	35.47	116	36.25	685	51.01	210	41.42	1210	42.86	326	39.42
Total	1480	100	320	100	1343	100	507	100	2823	100	827	100

* early = stage '0', '1', '2a'

locally advanced = stage '2b', '3'

metastatic = stage '4'

Table A8.10 Treatment of early stage surgical cases by surgical caseload and year												
Treatment type	1983				1987				Both years			
	<=30 cases/yr		31+ cases/yr		<=30 cases/yr		31+ cases/yr		<=30 cases/yr		31+ cases/yr	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
CO	118	26.22	36	33.03	186	51.38	78	42.39	304	37.44	114	38.91
MA	332	73.78	73	66.97	176	48.62	106	57.61	508	62.56	179	61.09
Total	450	100	109	100	362	100	184	100	812	100	293	100

A8.2 Five year relative survival for breast cancers diagnosed in 1983 and 1987

Table A8.11 Five year relative survival rates by year			
Sample	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
All breast cancers	64.62 (63.01 - 66.23)	62.72 (60.41 - 65.03)	66.38 (64.15 - 68.61)
Invasive breast cancers	64.01 (62.38 - 65.64)	62.32 (60.00 - 64.64)	65.99 (63.34 - 67.89)

Table A8.12 Five year relative survival rates for all breast cancers by stage			
Stage	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
0	100.75 (93.34 - 108.16)	95.55 (79.25 - 111.85)	104.76 -*
1	88.43 (84.43 - 92.43)	86.80 (80.89 - 92.71)	89.88 (84.44 - 95.32)
2a	78.08 (74.21 - 81.95)	76.10 (70.76 - 81.44)	80.25 (74.64 - 85.86)
2b	59.24 (54.12 - 64.36)	64.57 (57.58 - 71.56)	53.48 (46.07 - 60.89)
3	44.95 (40.79 - 49.11)	48.02 (42.54 - 53.50)	40.56 (34.21 - 46.91)
4	10.41 (6.08 - 14.74)	11.02 (6.90 - 15.14)	9.03 (3.36 - 14.70)
Unstaged	71.59 (69.17 - 74.01)	72.88 (69.04 - 76.72)	70.77 (67.65 - 73.89)

* no breast cancer deaths occurred within 5 years of diagnosis in this group therefore the standard error of the relative survival rate was zero

Table A8.13 Five year relative survival rates for all cancers by histological group			
Histological group	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
Not known	29.62 (24.22 - 35.02)	22.80 (17.02 - 28.58)	47.18 (35.34 - 59.02)
Ductal carcinoma of no special type	64.72 (62.91 - 66.53)	65.35 (62.70 - 68.00)	64.18 (61.68 - 66.68)
Lobular carcinoma	73.98 (68.38 - 79.58)	73.18 (64.81 - 81.55)	74.60 (67.07 - 82.13)
Special type carcinoma	84.49 (78.58 - 90.40)	83.18 (74.81 - 91.55)	85.84 (77.51 - 94.17)
In situ carcinoma	98.21 (92.03 - 104.39)	94.83 (81.66 - 108.00)	100.50 (94.24 - 106.76)

Table A8.14 Five year relative survival rates for all breast cancers by grade			
Grade	Overall	1983	1987
	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)
Good	79.86 (75.55 - 84.17)	80.50 (72.79 - 88.21)	79.53 (74.33 - 84.73)
Poor	61.68 (58.51 - 64.85)	58.51 (53.94 - 63.08)	64.80 (60.27 - 69.33)
Unknown	63.01 (60.98 - 65.04)	62.29 (59.45 - 65.13)	63.75 (60.84 - 66.66)

Table A8.15 Five year relative survival rates for invasive breast carcinomas by age			
Age group (years)	Both years	1983	1987
	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)
<50	67.83 (65.00 - 70.66)	66.15 (61.99 - 70.32)	69.34 (65.49 - 73.19)
50-69	62.39 (60.15 - 64.63)	61.26 (58.08 - 64.45)	63.50 (60.35 - 66.65)
70+	63.25 (59.72 - 66.77)	60.63 (55.54 - 65.72)	65.57 (60.68 - 70.46)
Pearson Chi Square 2 degrees of freedom	13.79 p=0.001	7.81 p=0.02	6.857 p=0.03

Table A8.16 Five year relative survival rates for invasive breast carcinomas by deprivation level			
Townsend group (range of scores)	Overall	1983	1987
	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)	Relative survival (%) (95% CI)
Most affluent (< -3)	67.25 (63.76 - 70.74)	69.76 (64.77 - 74.75)	65.07 (60.22 - 69.92)
Affluent (-3 to <-1)	68.38 (64.84 - 71.92)	66.89 (61.85 - 71.93)	69.83 (64.86 - 74.80)
Middle range (-1 to <1)	64.89 (61.24 - 68.54)	61.89 (56.59 - 67.19)	67.59 (62.58 - 72.60)
Deprived (1 to <3.5)	60.38 (56.61 - 64.15)	57.61 (52.24 - 62.98)	63.07 (57.77 - 68.37)
Most deprived (3.5+)	58.05 (54.18 - 61.92)	54.69 (49.35 - 60.03)	61.79 (56.19 - 67.39)
Pearson Chi Square 4 degrees of freedom	24.91 P<0.0001	20.77 p=0.0004	8.235 p=0.008

Table A8.17 Five year relative survival rates by DHA for all cancers and all invasive carcinomas						
DHA	Overall	1983	1987	Overall	1983	1987
	All cancers			All invasive carcinomas*		
Region	64.62 (63.01 - 66.23)	62.72 (60.41 - 65.03)	66.38 (64.15 - 68.61)	64.01 (62.38 - 65.64)	65.59 (63.33 - 67.85)	65.99 (63.34 - 67.89)
M02	67.84 (58.73 - 76.95)	55.93 (42.40 - 69.40)	**78.70 (67.06 - 90.34)**	**67.26 (58.06 - 76.46)**	**55.93 (42.40 - 69.46)**	**77.88 (65.98 - 89.78)**
M04	62.15 (54.88 - 69.42)	59.06 (47.43 - 70.69)	63.80 (54.57 - 73.03)	61.58 (54.25 - 68.91)	59.06 (47.73 - 70.69)	63.28 (53.85 - 72.71)
M05	65.10 (59.20 - 71.00)	59.24 (50.16 - 68.32)	69.32 (61.61 - 77.03)	63.98 (57.95 - 69.97)	57.55 (48.28 - 66.82)	68.58 (60.73 - 76.43)
M07	60.57 (55.05 - 66.09)	57.93 (49.98 - 65.88)	63.07 (55.43 - 70.71)	60.34 (57.79 - 65.89)	58.13 (50.17 - 66.09)	62.45 (54.74 - 70.16)
M17	65.68 (59.47 - 71.89)	65.70 (57.16 - 74.24)	65.66 (56.62 - 74.70)	64.90 (58.61 - 71.19)	65.19 (56.58 - 73.80)	64.59 (55.39 - 73.79)
M18	66.44 (60.17 - 72.71)	67.47 (58.03 - 76.91)	65.65 (57.17 - 74.13)	66.08 (59.69 - 72.47)	67.74 (58.22 - 77.26)	64.80 (56.20 - 73.40)
M19	62.98 (56.11 - 69.85)	58.99 (49.52 - 68.46)	67.42 (57.49 - 77.35)	62.98 (56.03 - 69.93)	59.12 (49.57 - 68.68)	67.33 (57.26 - 77.40)
M20	65.59 (57.25 - 73.93)	66.01 (54.61 - 77.41)	65.15 (52.95 - 77.35)	65.13 (56.72 - 73.54)	65.11 (53.53 - 76.69)	65.15 (52.95 - 77.35)
M21	62.50 (55.63 - 69.37)	61.55 (51.93 - 71.17)	63.50 (53.69 - 73.31)	61.46 (54.47 - 68.45)	60.87 (51.14 - 70.60)	62.09 (52.05 - 72.13)
M22	64.78 (57.44 - 72.12)	57.64 (47.24 - 68.04)	72.22 (62.06 - 82.38)	63.82 (56.36 - 71.28)	57.64 (47.21 - 68.07)	70.59 (60.05 - 81.13)
M25	60.14 (54.46 - 65.82)	61.82 (54.09 - 69.55)	**58.22 (49.85 - 66.59)**	**58.98 (53.21 - 64.75)**	**60.29 (52.45 - 68.13)**	**57.49 (48.99 - 65.99)**
M26	**70.97 (64.22 - 77.72)**	63.47 (53.09 - 73.85)	**76.42 (67.57 - 85.27)**	**70.56 (63.74 - 77.38)**	**62.67 (52.15 - 73.19)**	**76.23 (67.33 - 85.13)**
M27	69.05 (64.01 - 74.09)	**70.65 (63.16 - 78.14)**	67.96 (61.17 - 74.75)	68.65 (63.57 - 73.73)	**70.46** (62.95 - 77.97)	67.2 60.33 - 74.07)
M28	68.13 (63.07 - 73.19)	68.66 (61.21 - 76.11)	67.68 (60.78 - 74.58)	67.07 (61.88 - 72.26)	68.11 (60.52 - 75.70)	66.17 (59.05 - 73.29)
M29	60.45 (55.32 - 65.58)	60.26 (53.44 - 67.08)	60.46 (52.73 - 68.19)	60.42 (55.29 - 65.55)	60.54 (53.66 - 67.42)	60.27 (52.46 - 68.08)

*** including cancers with unknown histology**

Shaded boxes indicate that Chi Square DHA vs region gave a result with $p < 0.1$

Table A8.18 Five year relative survival for invasive breast carcinomas by surgical treatment option			
Surgical Treatment	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
None	30.92 (24.26 - 37.58)	26.32 (16.36 - 36.28)	33.95 (25.07 - 42.83)
No surgery	31.84 (27.93 - 35.75)	28.38 (23.09 - 33.67)	35.31 (29.56 - 41.06)
Conservative surgery	79.63 (77.01 - 82.25)	77.23 (72.48 - 81.98)	80.79 (77.65 - 83.93)
Mastectomy	67.03 (64.81 - 69.25)	68.13 (65.25 - 71.01)	65.46 (61.94 - 68.98)
Pearson Chi square 3 degrees of freedom	511.86 P<0.0001	285.78 p<0.0001	242.28 p<0.0001

Table A8.19 Five year relative survival rates for invasive breast carcinomas receiving conservative surgery by adjuvant therapy option			
Adjuvant treatment	**Both years**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
No adjuvant therapy	85.46 (79.90 - 91.02)	77.51 (68.68 - 86.34)	91.90 (84.99 - 98.81)
RT	77.64 (73.41 - 81.87)	78.53 (72.23 - 84.83)	76.93 (71.21 - 82.65)
HT	87.88 (80.56 - 95.20)	86.95 (70.99 - 102.91)	88.17 (79.96 - 96.38)
RT/ HT	76.88 (71.74 - 82.03)	59.92 (57.55 - 62.29)	77.92 (72.68 - 83.16)
CT	39.16 (24.07 - 54.25)	59.94 (39.67 - 80.21)	11.42 -*
Pearson Chi square 4 degrees of freedom	67.33 P<0.0001	7.835 p=0.01	125.35 p<0.0001

* 95% CI (-3.51 - 26.35) includes zero as the relative survival rate was based on only 18 cases of which 14 died within 2 years of diagnosis

Table A8.20 Five year relative survival for invasive breast carcinomas receiving mastectomy by adjuvant option			
Adjuvant option	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
No adjuvant therapy	76.73 (73.68 - 79.78)	77.93 (74.27 - 81.59)	74.26 (19.52 - 129.00)
RT	56.54 (60.35 - 70.73)	55.08 (49.06 - 61.10)	60.91 (50.68 - 71.14)
HT	65.62 (59.99 - 71.26)	60.13 (49.77 - 70.49)	67.95 (61.26 - 74.64)
RT + HT	53.83 (47.11 - 60.55)	51.64 (39.97 - 63.31)	54.89 (46.66 - 63.12)
CT	42.96 (32.90 - 53.02)	54.97 (42.81 - 67.13)	14.22 (1.28 - 27.16)
Pearson Chi square 4 degrees of freedom	97.23 p<0.0001	60.976 p<0.0001	69.506 p<0.0001

Table A8.21 Five year relative survival rates for invasive breast carcinomas by surgical caseload			
Surgical caseload of treating surgeon (new cases/year)	**Overall**	**1983**	**1987**
	Relative survival (%) (95% CI)	**Relative survival (%) (95% CI)**	**Relative survival (%) (95% CI)**
All surgically treated cases	71.92 (70.21 - 73.64)	70.06 (67.55 - 72.57)	73.81 (71.26 - 76.36)
<10	69.42 (65.17 - 73.67)	70.89 (76.42 - 76.36)	67.28 (60.56 - 74.00)
11 - 20	72.25 (69.43 - 75.07)	70.22 (66.68 - 73.76)	75.96 (71.30 - 80.72)
21 - 30	72.72 (68.91 - 76.53)	68.40 (60.85 - 75.95)	74.94 (70.53 - 79.35)
31+	72.27 (68.82 - 75.92)	69.70 (63.77 - 75.63)	73.88 (69.25 - 78.51)
Pearson Chi Square 3 degrees of freedom	1.683 p=0.064	0.471 p=0.93	5.290 p=0.15

REFERENCES

Alexander F E, Anderson T J
Population breast cancer screening
The Lancet 1995 vol. 345 p. 854

American College of Pathologists
Systematised Nomenclature of Medicine (SNOMED)
American College of Pathologists, Illinois, 1980

Andersson I, Aspregren K, Janzon L *et al*.
Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial
BMJ 1988 vol. 297 pp. 943-8

Asbury D, Boggis C R M, Sheals D, Threlfall A G, Woodman C B J
NHS breast screening programme: is the high incidence of interval cancers inevitable?
BMJ 1996 vol. 313 pp. 1369-70

Austoker J, Humphreys J
Practical Guides for General Practice 6 - Breast Cancer Screening
Oxford University Press, Oxford, 1988

Austoker J, Ong G
Written information needs of women who are recalled for further investigation of breast screening: results of a multicentric study
Journal of Medical Screening 1994 vol. 1 pp. 238-244

Barr L C, Baum M
Time to abandon TNM staging of breast cancer?
The Lancet, 1992 vol. 339 pp. 915-17

Baum M
Specialist surgeons and survival from breast cancer - large differences in survival are not explained
BMJ 1996 vol. 312 pp. 1155

Basnett I, Gill M, Tobias J A S
Variations in breast cancer management between a teaching and non-teaching district
European Journal of Cancer 1992 vol. 28A pp. 1945-50

BAS0 (Breast Surgeons Group of the British Association of Surgical Oncology)
Guidelines for Surgeons in the management of symptomatic breast disease in the UK
European Journal of Surgery 1995 vol. 21 Supplement A

BBG (Working Party for of the British Breast Group)
Provision of breast services in the UK: the advantages of Specialist Breast Units
British Breast Group 1994

Benson E A
Reservations about conservative surgery for early breast cancer
BMJ 1996 vol.312 pp. 845-846

Berrino F, Sant M, Verdecchia A, Capocaccia R (Editors) on behalf of the WHO International Agency for Research on Cancer European Commission
Survival of cancer patients in Europe - the EUROCARE study
IARC Scientific Publications No. 132 , Lyon, 1994

Bickler G, Sutton S.
Inaccuracy of FHSA registers: help from electoral registers
BMJ 1993 vol. 306 p.1167

Blanks R G, Day N E, Moss S M
Monitoring the performance of breast screening programmes: Use of indirect standardisation in evaluating the invasive cancer detection rate
Journal of Medical Screening 1996a vol. 3 pp. 79-81

Blanks R G, Day N E, Moss S M
Monitoring the performance of breast screening programmes: allowing for geographical variation in breast cancer incidence
Journal of Medical Screening 1996b vol. 3 pp. 82-84

Blanks R G, Moss S M, Wallis M G
A comparison of two view and on view mammography in the detection of small invasive cancers: results from the National Health Service breast screening programme
Journal of Medical Screening 1996c vol. 3 pp. 200-203

Bloom H J G, Richardson W W
Histological grading and prognosis in breast cancer
Br J Cancer 1957 vol. 11 pp. 359-377

Bodian C A
Benign breast diseases, carcinoma *in situ* and breast cancer risk
Epidemiologic Reviews vol. 15 number 1 pp. 177-187

Bofetta P, Merletti F, Winkelmann, Magnani C, Cappa A P M, Terracini B
Survival of breast cancer patients from Piedmont, Italy
Cancer Causes and Control 1993 vol. 4 pp. 209-15

Bowling **A,** Jacobson B
Screening: The inadequacy of population registers
BMJ 1989 vol. 298 pp. 545-546

Boyle P, Parkin D M
Statistical Methods for Cancer Registries (Chapter 11) in
Cancer Registration Principles and Methods
IARC Scientific Publication No. 95, Lyon, 1991

Bristol Cancer Epidemiology Unit (incorporating the South Western Regional Cancer Registry)
Cancer of the female breast in the South and West
Bristol Cancer Epidemiology Unit, University of Bristol, 1995

Brown J M, Benson E **A,** Jones M
Confirmation of a long-term prognostic index in breast cancer
The Breast 1993 vol. 2 pp. 144-147

Brown J, Bryan S, Warren R
Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms
The Breast 1996 vol. 312 pp. 809-12

Buell P
Changing incidence of breast cancer in Japanese-American Women
Journal of the National Cancer Institute 1973 vol. 51 pp. 1479-83

Bull **A,** Mountney L, Sanderson H
Use of a cancer register for monitoring and evaluating a breast cancer screening programme
Community Medicine 1989, vol. 3 number 3 pp. 220-224

Bull A, Mountney L, Sanderson H
Stage distribution of breast cancer: a basis for evaluation of breast screening programmes
British Journal of Radiology, 1991 vol. 64 pp. 516-519

Bundred N J, Morgan D **A,** Dixon J M
Management of regional nodes in breast cancer
BMJ 1994 vol. 309 pp. 1222-1225

Cancer Relief Macmillan Fund
Raising **Standards** in Breast Cancer Care - the Macmillan Directory of Breast Cancer Services in the UK (Second Edition)
Cancer Relief Macmillan Fund, 1996

Cancer Research Campaign (based on data supplied by Thames Cancer Registry)
Factsheet No. 6.2 , CRC, 1988

Carnon **A** G, Ssemwogerere **A,** Lamont DW et al
Relation between socio-economic deprivation and pathological prognostic factors in women with breast cancer
BMJ 1994 vol. 309 pp. 1054-1057

Carter C L, Allen C, Henson D E
Relation of tumor size, lymph node status and survival in 24,740 breast cancer cases
Cancer 1989 vol. 63 pp. 181-187

Chen H-H, Tabar L, Fagerberg G., Duffy S W
Effect of breast cancer screening after age 65
Journal of Medical Screening 1995 vol. 2 pp. 10-14

Chouillet A M, Bell C M J, Hiscox J G
Management of breast cancer in Southeast England
BMJ 1994 vol. 308 pp. 168-71

Coleman M P, Esteve J, Damiecki P, Arslan A, Renard H
Trends in cancer incidence and mortality
IARC Publication No. 121, Lyon, 1993

Collette H J, Rombach J J, Day N E
Evaluation of screening for breast cancer in a non-randomized study (the DOM project) by means of a case-control study
The Lancet 1984 vol. i pp. 1224-6

Cox D R
Regression models and life tables (with discussion)
Journal of the Royal Statistical Society 1972 vol. 34(B) pp. 187-220

Crisp W J, Higgs M J, Cowan W K, Cunliffe W J, Liston J, Lunt L G, Peakman D J, and Young J R
Screening for breast cancer detects tumours at an earlier biological stage
British Journal of Surgery 1993 vol. 80 pp. 863-865

Day N E, Williams D R R, Khaw K T
Breast cancer screening programmes: the development of a monitoring and evaluation system
British Journal of Cancer 1989 vol. 59 pp. 954-958

Day N E, McCann J, Camilleri-Ferrante, Britton P, Hurst P, Cush S, Duffy S on behalf of the Quality Assurance management Group of the East Anglian Breast Screening Programme
Monitoring interval cancers in breast screening programmes: the East Anglian experience
Journal of Medical Screening 1995a vol. 2 pp. 180-185

Day N E, Tabar L, Duffy S W, Chen H H
Population breast cancer screening
The Lancet 1995b vol. 345 p. 853

Day N E, Davies T W
Cancer registration: integrate or disintegrate?
BMJ 1996 vol. 313 p. 896

Dean J, Dean A, Burton A, Dicken R
Epi Info - Public Domain Software for Epidemiology and Disease Surveillance Version 5.0
Centers for Disease Control, Epidemiology Programs Office, Atlanta, Georgia, April 1990

Department of Health
Public Health Common Data Set 1992 incorporating indicators from the Health of the Nation: England - vol. II
Department of Health, 1993

Department of Health
Health of the Nation Technical Supplement
In press

Dixon J M
Surgery and radiotherapy for early breast cancer
BMJ 1995 vol. 311 pp. 1515-1516

Dixon J M
Specialist surgeons and survival from breast cancer - a multi-disciplinary approach is needed
BMJ 1996 vol. 312 pp. 1155

Donegan W L
Prognostic factors
Cancer Supplement 1992 vol. 70 no. 6 pp. 1755-1763

Duffy S W, Tabar L, Fagerberg G, Gad A, Grontoft O, South M C, Day N E
Breast screening, prognostic factors and survival - results from the Swedish two county study
British Journal of Cancer 1991 vol. 64 pp. 1133-1138

EBCTCG Early breast cancer triallists collaborative group
Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy
New England Journal of Medicine 1992 vol. 339 pp. 1-15, 71-85

EBCTCG Early breast cancer triallists collaborative group
Effects of radiotherapy and surgery in early breast cancer
New England Journal of Medicine 1995 vol. 333 pp. 1444-55

EDBCTG Early detection of breast cancer triallists group
Breast cancer mortality after ten years in the UK trial of early detection of breast cancer
The Breast 1993 2, 13-20

Elledge R M, Fuqua S A W, Clark G M, Pujol P, Allred D C
The role and prognostic significance of p53 gene alterations in breast cancer
Breast Cancer Research and Treatment 1993 vol. 27 pp. 95-102

Ellman R, Angeli N, Christians A et al
Psychiatric morbidity associated with screening for breast cancer
British Journal of Cancer 1989 vol. 60 pp. 781-4

Ellis I O, Galea M, Broughton N, Locker A, Blamey R W, Elston C W
Pathological factors in breast cancer 11. Histological type. Relationship with survival in a large study with long-term follow-up
Histopathology 1992 vol. 20 pp. 479-489

Elston C W, Ellis I O
Pathological factors in breast cancer I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up
Histopathology 1991 vol. 19 pp. 403-410

Farrow D C, Hunt W C, Samet J M
Geographical variation in the treatment of localised breast cancer
New England Journal of Medicine 1992 vol. 326 pp. 1097-101

Fricker J
Studies needed on prognosis of ductal carcinoma in situ
The Lancet 1996 vol. 347 p. 890

Fisher B, Constantino J, Redmond C, Fisher E, Margolese R, Dimitrov N, Wolnark N, Wickeram D L, Deutsch M, Ore L, Mamounas E, Poller W and Kavanagh M
Lumpectomy compared with lumpectomy and radiation therapy for the treatment of intraductal breast cancer
New England Journal of Medicine 1993 vol. 328 pp. 1581-6

Fisher B, Anderson S, Redmond C K, Wolmark N, Wickerham D L, Cronin W M
Reanalysis and results after 12 years of follow-up in a randomised clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer
New England Journal of Medicine 1995 vol. 333 pp. 1456-61

Foucar E
Carcinoma-in-situ of the breast: have pathologists run amok?
The Lancet 1996 vol. 347 pp. 707-7-8

Gallagher H S
Pathologic types of breast cancer: their prognoses
Cancer Supplement 1984 pp. 623-629

Garvican L, Littlejohns P
An evaluation of the prevalent round of the breast screening programme in South East Thames, 1988 - 1993: achievement of quality standards and population impact
Journal of Medical Screening 1996 vol. 3 number 3 pp. 123-128

Gelber R D, Cole B F, Goldhirsh A et al
Adjuvant chemotherapy plus tamoxifen compared with tamoxifen alone for postmenopausal breast cancer: meta-analysis of quality-adjusted survival
The Lancet 1996 vol. 347 pp. 1066-71

Gilchrist K W, Kalish L, Gould V E et al, (an ECOG Study)
Inter-observer reproducibility of histopathological features in stage II breast cancer
Breast Cancer Research and Treatment 1985 vol. 5 pp. 3-10

Gillis C R, Hole D J
Survival outcome of care by specialist surgeons in breast cancer: a study of 3786 patients in the West of Scotland
BMJ 1996 vol. 312 pp. 145-8

Greenstein G, Lamster I
Understanding diagnostic testing for periodontal diseases
Journal of Periodontology 1995 vol. 66 number 8 pp. 659-666

Gulliford M C, Bell J, Bourne H M, Petruckevitch
The reliability of Cancer Registry records
British Journal of Cancer 1993 vol. 67 pp. 819-821

Hakulinen T, Abeywickrama K H
A computer package for relative survival analysis
Computer Programs in Biomedicine 1985 vol. 19 pp. 197-209

Hakama M, Holli K, Isola J et al
Aggressiveness of screen-detected breast cancers
The Lancet 1995 vol. 345 pp. 221-24

Hakama M, Pukkula E, Heikkila M, Kallio M
Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study
BMJ 1997 vol. 314 pp. 864-7

Harries S A, Scrivener R, Lawrence R N, Fieldman N R, Kissin M W, Bates T
Workload of surgeons involved in the treatment of breast cancer
The Breast 1997 vol. 6 pp. 21-25

Haybittle J L, Blamey R W, Elston C W, Johnston J, Doyle P, Campbell F C, Nicholson R I and Griffiths K
A prognostic index in primary breast cancer
British Journal of Cancer 1982 vol. 45 pp. 361-6

Hermanek P, Gospodarowicz M K, Henson D E, Hutter R V P, Sobin L H [Editors]
Prognostic factors in Cancer
UICC, Geneva, Springer-Verlag, 1995

HMSO
Breast cancer screening - report to the Health Ministers of England, Wales., Scotland and Northern Ireland - the "Forrest" Report
HMSO, 1986

HMSO
The Health of The Nation - A Strategy for Health In England
HMSO, 1992

HMSO
A policy framework for commissioning cancer services - A report by the expert advisory group on cancer to the chief medical officers of England and Wales
(The "Calman-Hine report")
HMSO, 1995

Hoare T A, Johnson C M, Gorton R, Alberg C
Reasons for non-attendance for beast screening by Asian women
Health Education Journal 1993 vol. 51 number 4 **pp.** 157-161

Hurley S F, Kaldor J M
The benefits and risks of mammographic screening for breast cancer
Epidemiologic Reviews 1992 vol. 14 pp. 101-130

International Agency for Research on Cancer (IARC)
Cancer in Five Continents, Vol VI, IARC Scientific Publications No. 120
International Agency for Research on Cancer and International Association of Cancer Registries, Lyon, 1992

Information and Statistics Division, NHS in Scotland
Trends in Cancer Survival in Scotland 1968-1990
Information and Statistics Division, NHS in Scotland, Edinburgh, 1993

Karjalainen S, Pukkala M **A**
Social class as a prognostic factor in breast cancer survival
Cancer 1990 vol. 66 pp. 819-26

King's Fund Consensus Development Conference Statement
Treatment of primary breast cancer
BMJ 1986 vol. 293 pp. 946-947

Klemi P J, Joensuu H, Toikkanen S, Rasanen O, Tyrkko J, Parvinen
Aggressiveness of breast cancers found with and without screening
BMJ 1992 vol. 304 pp. 467-469

Knox E G
Evaluation of a proposed breast cancer screening programme
BMJ 1988 vol. 297 pp. 650-654

Kogevinas M, Marmot M G, Fox A J, Goldblatt P O
Socio-economic differences in cancer survival
Journal of Epidemiology and Community Health 1991 vol. 45 pp. 216-219

Lee-Feldstein A, Anton-Culver H, Feldstein P J
Treatment differences and other prognostic factors related to breast cancer survival
Journal of the American Medical Association 1994 vol. 271 number 15 pp. 1163-1168

Leon D A
Longitudinal study - social distribution of cancer
HMSO Series **LS** number 3, 1988

Lidbrink E, Elfving J, Frisell J, Jonsson E
Neglected aspects of false positive findings of mammography in breast cancer screening: an analysis of false positive cases from the Stockholm trial
BMJ 1996 vol. 312 pp. 273-6

Maguire P
Psychological aspects
BMJ 1994 vol. 309 pp. 1649-1652

Maher J
Management of adverse effects following breast radiotherapy
Royal College of Radiologists, London, 1995

Manchester Centre for Cancer Epidemiology
Cancer in the North West
Centre for Cancer Epidemiology, Manchester, 1992

McCarthy M, Bore J
Treatment of breast cancer in two teaching hospitals: a comparison with consensus guidelines
European Journal of Cancer 1991 vol. 27 number 5 pp. 579-582

McPherson K, Steel C M, Dixon J M
Breast cancer - epidemiology, risk factors, and genetics
BMJ 1994 vol. 309 pp. 1003-1153

Merseyside & Cheshire Cancer Registry
Breast Cancer Bulletin - Stimulating The Debate
Merseyside & Cheshire Cancer Registry, 1994

Miller W R, Ellis I O, Sainsbury J R C, Dixon J M
Prognostic Factors
BMJ 1994 vol. 309 pp. 1573-1576

Morris J, Farmer A, Royle G
Recent changes in the surgical management of T1/2 breast cancer in England
European Journal of Cancer 1992 vol. 28A number 10 pp. 1709-1712

Moss S M, Ellman R, Coleman D, Chamberlain J for the United Kingdom Trial of Early Detection of Breast Cancer Group
Survival of patients with breast cancer diagnosed n the United Kingdom trial of early detection of breast cancer
Journal of Medical Screening 1994 vol. 1 pp. 193-198

Moss S M, Michell M, Patnick J, Johns L, Blanks R, Chamberlain J
Results from the NHS breast screening programme 1990-1993
J Medical Screening 1995 vol. 2 pp. 186-190

Muir Gray J. A.
Planning for quality in the NHS breast screening programme
Quality In Health Care 1992 vol. 1 pp. 119-123

NHSBSP
Quality Assurance Guidelines for Mammography (The Pritchard Report)
NHSBSP, 1988

NHSBSP
Protocol for the trial of the management of screen detected ductal carcinoma *in situ* (DCIS) of the breast
NHSBSP, 1989

NHSBSP
Breast Cancer Screening 1991: Evidence and Experience since the Forrest Report 1991
NHSBSP, 1991

NHSBSP
Quality Assurance Guidelines for Surgeons in Breast Cancer Screening
NHSBSP 1992

NHSBSP
Guidelines on the Collection and Use of Breast Cancer Data, January 1993
NHSBSP Publication No. 26,1993a

NHSBSP
Objectives for the Breast Screening Programme, January 1993
NHSBSP, 1993b

NHSBSP
Pathology Reporting in Breast Cancer Screening, Second Edition, April 1995
NHSBSP Publication No.3, 1995a

NHSBSP
Guide to Good Office Practice 5 - Recording Biopsy Type
NHSBSP 1995b

NHSBSP
Quality Assurance in the NHS Breast Screening Programme, October 1995
NHSBSP,1995c

NHSBSP
Systematic Management of Quality for Breast Screening Units
NHSBSP 1995d

NHSBSP
NHS Breast Screening Programme Review - 1995
NHSBSP 1995e

NHSBSP
Messages about breast screening - guidelines for health promotion specialists, primary care teams and screening unit staff
NHSBSP publication No. 31, 1995f

NHSBSP
Quality Assurance Guidelines for Surgeons in Breast Cancer Screening
NHSBSP 1996a

NHSBSP
NHS Breast Screening Programme Review - 1996
NHSBSP 1996b

NHSBSP and Age Concern
65 Or Over? - You Are Still Entitled To Breast Screening (leaflet)
NHSBSP and Age Concern, 1996

NHS Executive
Cancer Guidance Subgroup of the Clinical Outcomes Group
Guidance for purchasers - Improving outcomes in breast cancer : the manual
keds, NHS Executive, 1996a

NHS Executive
Cancer Guidance Subgroup of the Clinical Outcomes Group
Guidance for purchasers - Improving outcomes in breast cancer : the research evidence
keds, NHS Executive, 1996b

Nystrom L, Rutqvist L E, Wall S, Lindgren A, Lindqvist M, Ryden S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabar L, Larsson L-G
Breast cancer screening with mammography: overview of the Swedish randomised trials
The Lancet 1993 vol. 341 pp. 973-78

Ong G, Austoker J, Brouwer A
Evaluation of the written information sent to women who are called back for further investigation of breast screening in the UK
Health Education Journal 1996 vol. *55* pp 413-429

ONS
Monitor - Population and Health
ONS, Series **MB** 1 Issue 96/2, London, November 1996

OPCS
Mortality Statistics - cause
Review of the Registrar General on deaths by cause, sex and age, in England and Wales, 1989 Series DH2 number 16
HMSO, London, 1991

OPCS
Cancer Statistics - Registrations
Registrations of cancer diagnosed in 1988, England and Wales
Series MI31 number 21
HMSO, London, 1994

Paci E
Population breast cancer screening
The Lancet 1995 vol. 345 pp. 853-4

Page D L, Steel C **M,** Dixon J M
Carcinoma *in situ* and patients at high risk of breast cancer
BMJ 1995 vol. 310 pp. 39-42

Palli D, del Turco ***MR,*** Buiatti E et al
A case-control study of the efficacy of a non-randomized breast cancer screening programme in Florence (Italy)
International Journal of Cancer 1986 vol. 38 pp. 501-4

Parham D M, Creagh-Barry **M,** Hill R A, Nicholas D S
Observer variation in the grading of screen detected mammographic abnormalities 1. Assessment of reproducibility
The Breast 1996 vol. 5 pp. 422-424

Parkin D M, Muir C S, Whelan S L, Gao Y-T, Feraly J, Powell J
Cancer In Five Continents (vol. 6)
IARC Publication number 120, Lyon, 1992

Parkin D M, Chen V W, Ferlay J, Galceran J
Comparability and Quality Control in Cancer Registration
IARC Technical Report No. 19, Lyon, 1994

Pater J L, Mores D, Loeb M
Survival after recurrence of breast cancer
Journal of the Canadian Medical Association 1981 vol. 124 pp. 1591-1595

Paterson R, 1948
in Treatment of Malignant Disease by Radium and X-rays
Edward Arnold, London, 1948

Peeters P H M, Verbeek A L M, Hendriks J H C L, Holland R, Mravunac M, Vooijs G P
The Occurrence of interval cancers in the Nijmegen screening programme
British Journal of Cancer **1989** vol. **59** pp. **929-932**

Ponten J, Holmberg L, Trichopoulos D, Kallioniemi O-P, Kvale G, Wallgren A, Taylor-Papadimitriou J
Biology and natural history of breast cancer
International Journal of Cancer **1990** Supplement *5* pp. **5-21**

Prior P, Woodman C B J, Wilson S, Threlfall A G
Reliability of underlying incidence rates for estimating the effect and efficiency of screening for breast cancer
Journal of Medical Screening **1996** vol. **3** pp. **119-122**

Quinn M, Allen E on behalf of the United Kingdom Association of Cancer Registries.
Changes in incidence of and mortality from breast cancer in England and Wales since introduction of screening
BMJ **1995** vol. **311** pp. **1391-5**

Rajakariar R, Walker R A
Pathological and biological features of mammographically detected invasive breast carcinomas
British Journal of Cancer **1995** vol. **71 pp. 150-154**

Richards M A, Smith I E, Dixon J M
Role of systemic treatment for primary operable breast cancer
BMJ **1994** vol. **309** pp. **1363-1365**

Richardson J L, Langholz B, Bernstein L, Burciaga C, Danley K, Ross R K
Stage and delay in breast cancer diagnosis by race, socio-economic status, age and year
British Journal of Cancer **1992** vol. **65** pp. **922-926**

Robbins P, Pinder S, de Klerk N, Dawkins H, Sterrett G, Ellis I and Elston C
Histological grading of breast carcinomas: a study of inter-observer agreement
Human Pathology **1995** vol. **26** number 8 pp. **873-879**

Roberts MM, Alexander FE, Anderson T J
Edinburgh trial of screening for breast cancer: mortality at seven years
The Lancet **1990** vol. **335** pp. **241-6**

Rodger A, Leonard R C, Dixon J M
Locally advanced breast cancer
BMJ **1994** vol. **309** pp. **1431-1433**

Rubens R D
Management of early breast cancer
BMJ **1992** vol. **304 pp. 1361-1364**

Sacks N P, Baum M
Primary management of carcinoma of the breast
The Lancet 1993 vol. 342 pp. 1402-1408

Sainsbury R, Rider L, Smith A, MacAdam A on behalf of the Yorkshire Breast Cancer Group
Does it matter where you live? - treatment variation for breast cancer in Yorkshire
British Journal of Cancer 1995a vol. 71 pp. 1275-1278

Sainsbury R, Haward B, Rider L, Johnston C, Round C
Influence of clinician workload and patterns of treatment on survival from breast cancer
The Lancet 1995b vol. 345 pp. 1265-70

Seradour B, Wait S, Jacquemier J, Dubac M
Dual reading in a non-specialised breast cancer screening programme
The Breast 1996 vol. 5 pp. 398-403

Shapiro S, Venet W, Strax P et al
in Periodic Screening For Breast Cancer: The Health Insurance Plan Project And Its Sequalae 1963-86
Johns Hopkins University Press, London, 1988

Shrijvers C T, Mackenbach JP, Lutz J-M, Quinn M J, Coleman M P.
Deprivation and survival from breast cancer
British Journal of Cancer 1995 vol. 72 pp. 738-743

Sikora K
Specialist surgeons and survival from breast cancer - breast cancer is a medical, not a surgical disease
BMJ 1996 vol. 312 pp. 1155

Simpson W, Neilson F, Young J R and the Northern Region Breast Radiology Audit Group
The identification of false negatives in a population of interval cancers: a method for audit of screening mammography
The Breast 1995 vol. 4 pp. 183-188

Sloane *et al* - UK National Co-ordinating Group for Breast Screening Pathology
Consistency of histopathological reporting of breast lesions detected by screening: findings of the UK National External Quality Assessment (EQA) Scheme
European Journal of Cancer 1994 vol. 30A number 10 pp. 1414-1419

Smart C. R., Hendrick R E., Rutledge J H., Smith R A
Benefit of Mammographic screening in women ages 40 to 49 years
Cancer 1995 vol. 75 number 7 pp. 1619-1619-1629
(Erratum - Cancer 1995 vol. 76 number 9 pp. 1680-1681)

Spittle M F
Brachial plexus neuropathy after radiotherapy for breast cancer
BMJ 1995 vol. 311 pp. 1516-1517

Steeg P S, De La Rosa A, Flatow U, Macdonald N J, Benedict M, Leone A
Nm23 and breast cancer metastasis
Breast Cancer Research and Treatment 1993 vol. 25 pp. 175-187

Stiller C A
Centralised treatment, entry to trials and survival
British Journal of Cancer 1994 vol. 70 pp. 352-362

Swerdlow A, dos Santos Silva I
Atlas of cancer incidence in England and Wales 1968-85
Oxford University Press, 1993

Tabar L, Fagerberg G, Duffy S W, Day N E, Gad A, Grontoft O
Update of the Swedish Two County program of mammographic screening for breast cancer
Radiologic Clinics of North America 1992 vol. 30 pp. 187-210

Thames Cancer Registry
Cancer in South East England - Cancer incidence, prevalence and survival in residents of the District Health Authorities in South East England
Thames Cancer Registry, 1994

Threlfall A G, Woodman C B J, Prior P
Breast screening programme: should the interval between tests depend on age?
The Lancet 1997 vol. 349 p. 472

Todd J H, Dowle C, Williams M R, Elston C W, Ellis I O, Hinton C P, Blamey R W, Haybittle J L
Confirmation of a prognostic index in primary breast cancer
British Journal of Cancer 1987 vol. 56 pp. 489-492

Torenson D
Risk factors for breast cancer
BMJ 1994 vol. 309 pp. 1662

Townsend P, Phillimore P, Beattie A
Health and deprivation: inequality and the North
London, Routledge, 1989

UICC Union Internationale Contre le Cancer
TNM Atlas - Illustrated Guide to the TNM/pTNM Classification of Malignant Tumours
Third **Edition** 2nd Revision
Springer-Verlag, New York Berlin & Heidelberg, 1992

United Kingdom Association of Cancer Registries
Cancer Registry Handbook
UKACR, Cambridge, 1994

van Dijck J **A A** M, Verbeek **A** L M, Hendriks J H C L, Holland R
The current detectability of breast cancer in a mammographic screening programme
Cancer 1993 vol. 72 number 6 pp. 1933-1938

Vernon S W, Tilley B C, Neale **A** V, Steinfeldt L
Ethnicity, survival, and delay in seeking treatment for symptoms of breast cancer
Cancer 1985 vol. 55 pp. 1563-1571

Wald N J, Chamberlain J, Hackshaw **A**
Consensus Statement - European Society of Mastology Conference on Breast Cancer Screening
Clinical Oncology 1994 vol. 6 pp. 261-268

Wald N J, Murphy P, Major P, Parkes C, Townsend J, Frost C
UKCCCR multicentric randomised controlled trial of one and two view mammography in breast cancer screening
BMJ 1995 vol. 311 pp. 1189-93

Watson J D, Sainsbury J R C, Dixon J M
Breast reconstruction after surgery
BMJ 1995 vol. 310 pp. 117-121

Wessex Cancer Intelligence Unit
Guide to Breast Cancer Staging For Research **Assistants**
Wessex Cancer Intelligence Unit, 1992

West Midlands Regional Health Authority Regional Director of Public Health and the West Midlands Regional Cancer Registry
Cancer and Health, NHS Executive West Midlands, 1995

World Health Organisation
International Classification of Diseases, Injuries and Causes of Death, Ninth Revision.
WHO, Geneva, 1977

Wight S F L, Bygrave S M, Moss S M, Smith J **A** E
The Impact of the Breast Screening Programme on the Utilisation of Health Service Resources
Proceedings of the EORTC Joint Breast Cancer Meeting, Nottingham, 1993

Wilson J M G, Jungner G
Principles and practice of screening for disease
World Health Organisation, 1968, WHO Public Health Paper No. 34

Woodman C B J, Threlfall **A** G, Boggis C R M, Prior P
Is the three year breast screening interval too long? Occurrence of interval cancers in the **NHS** breast screening programme's north western region
BMJ 1995 vol. 310 pp. 224-226

Wright C J, Mueller C B
Screening mammography and public health policy: the need for perspective
The Lancet 1995 vol. 346 pp. 29-32

Yorkshire Cancer Organisation
Cancer in Yorkshire - Cancer Registry Special Report Series 3 - Breast Cancer
Yorkshire Cancer Organisation, 1995

Zeneca Pharma
··Arimidex" Monograph
Zeneca Pharma, Wilmslow, Cheshire, 1995