

Developing an Evidence Based Approach To Follow Up in Breast Cancer

**By
David A. Montgomery
BSc (Med Sci) (Hons), MBChB (Hons), MRCS (Glasg)**

**A thesis submitted for the degree of
MSc (Med Sci)
to
The University of Glasgow**

**Research conducted in the
University Department of Surgery,
Glasgow Royal Infirmary
and the
Edinburgh Breast Unit
Western General Hospital
Edinburgh**

November 2007

© David A. Montgomery 2007

Summary of Thesis

After breast cancer, long term follow up is provided with a view to early detection of locoregional relapse, detection and amelioration of side effects of therapy and provision of psychosocial support to those requiring it. There is pressure within the United Kingdom from the National Institute for Clinical Excellence (NICE) to reduce the amount of follow up provided after breast cancer. There is reluctance on the part of clinicians to adopt the NICE guidelines as they are perceived to be based on limited evidence. At the same time, guidelines in other countries continue to advocate long term follow up, sometimes for life.

The evidence base for follow up is explored in the first section of this thesis. In chapter one, a systematic review of the literature is undertaken to establish what evidence exists from randomised controlled trials. The randomised controlled trial is the gold standard for comparing one intervention with another, and any evidence for the benefit of routine clinic visits from randomised trials would be of particular value. Trials which have compared differing frequencies of traditional follow up visits or have compared traditional clinic visits with a novel alternative are included. The impact of these different strategies on relapse detection, survival and quality of life is investigated.

In chapters 2 and 3, the pattern and timing of potentially treatable locoregional relapse and the contribution of regular clinical examination to detection of such relapse is explored. This information would help to establish the value of routine clinics in terms of detecting relapse, and whether there was a time after diagnosis and treatment that risk of relapse was low enough to allow safe discharge. In chapter 2, retrospective analyses of two local cohorts is

undertaken to explore this issue. A systematic review of the literature is presented in chapter 3 incorporating both the data presented in chapter 2 as well as any other evidence available from previously published retrospective analyses.

In section 2 of this thesis, the options for alternative follow up are explored. Very little work has been done to establish what women expect from follow up, either in terms of the amount of follow up they expect to receive or their perceptions of the purpose of follow up. In chapter 4, the results of a survey of a cohort of women undertaken prior to attending for their first review visit after completion of therapy is presented.

In the final chapter of this thesis, a prospective cohort study into the acceptability and feasibility of an alternative method of follow up is presented. In this study, an automated computer telephone system was used to deliver a well validated quality of life questionnaire to women in their own homes with the aim of remotely identifying women who were having significant problems with either psychosocial concerns or side effects of treatment and therefore identify those with ongoing problems who required to come back to clinic. In this way, those patients with no ongoing problems would be spared a potentially stressful clinic visit and the number of patients coming back to clinic would be reduced, allowing more time to attend to the needs of the few women brought back with concerns.

Summary of Results

The analysis in chapter 1 reveals that there are only 5 randomised trials of alternative follow methods, and only 2 trials differing frequencies of visits. None is of sufficient size to

establish whether routine clinic visits are necessary for relapse detection or overall survival. All suggest that clinic visits have limited impact on quality of life, and may even be less valuable than some alternative methods of follow up for diagnosing anxiety and depression in these women.

Potentially treatable relapse occurs at a constant rate to ten years and beyond. Very few relapses are diagnosed by routine clinical examination, the majority being diagnosed by mammography or symptoms. Mammography is of increasing importance. Patients with clinically detected relapse do no better than those relapses detected in other ways, and there is some evidence that, for some types of relapse, they may do worse.

Chapter 4 reveals that most women expect some follow up, but their expectations are that follow up will be more frequent but for a shorter duration than is currently provided in our unit. When informed of the inefficiency of routine follow up in terms of relapse detection, most women still choose to come back to routine follow up, but a large number state that they would be happy not to come back to clinic.

An alternative method of follow up is shown to be acceptable to a large proportion of women, and valuable in detecting psychological concerns among women after treatment for breast cancer.

The implications of all the findings presented in this thesis for the planning of follow up care in the future are discussed.

| Contents | Page |
|---|-------------|
| SUMMARY | 2 |
| INDEX TO TABLES | 9 |
| INDEX TO FIGURES | 11 |
| INDEX TO APPENDICES | 13 |
| ACKNOWLEDGEMENTS AND DEDICATION | 14 |
| DECLARATION | 16 |
| LIST OF PUBLICATIONS | 16 |
| GENERAL INTRODUCTION: | 18 |
| 1. Current Guidelines | 22 |
| 2. Distant Relapse | 25 |
| 3. Locoregional Relapse | |
| a. Survival | 31 |
| b. Timing | 34 |
| c. Detection of locoregional relapse – contribution of mammography | 37 |
| d. Detection of locoregional relapse – contribution of clinical examination | 44 |
| 4. Additional Aims of Follow Up | |
| a. Adverse effects of treatment – lymphoedema | 50 |
| b. Adverse effects of treatment – non-lymphoedematous arm dysfunction | 55 |
| c. Adverse effects of treatment – hormone therapy | 56 |
| d. Quality of Life | 61 |
| 5. Alternative Follow Up | 67 |
| 6. Summary | 71 |

| | |
|---|-----------|
| AIMS OF THE THESIS | 74 |
| SECTION 1 | 76 |
| CHAPTER 1: THE IMPACT OF ALTERNATIVE FOLLOW UP METHODS ON RELAPSE DETECTION, SURVIVAL AND QUALITY OF LIFE – A SYSTEMATIC REVIEW OF THE LITERATURE | 77 |
| 1. Introduction | 77 |
| 2. Methods | 94 |
| 3. Results | 98 |
| 4. Discussion | 108 |
| 5. Summary | 114 |
| CHAPTER 2: PATTERN AND TIMING OF LOCOREGIONAL RELAPSE, METHOD OF DETECTION AND IMPACT ON SURVIVAL – RETROSPECTIVE ANALYSES OF TWO LOCAL COHORTS | 115 |
| 1. Introduction | 115 |
| 2. Edinburgh cohort | 117 |
| a. Patients and Methods | 117 |
| b. Results | 119 |
| 3. Glasgow cohort | 128 |
| c. Patients and Methods | 128 |
| d. Results | 130 |
| 4. Discussion | 135 |
| 5. Summary | 140 |

CHAPTER 3: PATTERN OF LOCOREGIONAL RELAPSE, METHOD OF DETECTION
AND IMPACT ON SURVIVAL – A SYSTEMATIC REVIEW OF THE

| | |
|-----------------|-----|
| LITERATURE | 141 |
| 1. Introduction | 141 |
| 2. Methods | 145 |
| 3. Results | 150 |
| 4. Discussion | 162 |
| 5. Summary | 169 |

SECTION 2 171

CHAPTER 4: PATIENTS' EXPECTATIONS FOR FOLLOW UP IN BREAST CANCER

| | |
|-----------------|-----|
| 1. Introduction | 172 |
| 2. Methods | 175 |
| 3. Results | 177 |
| 4. Discussion | 182 |
| 5. Summary | 188 |

CHAPTER 5: AUTOMATED TELEPHONE FOLLOW UP AFTER BREAST CANCER –

| | |
|---|-----|
| DESIGN AND INTRODUCTION OF A NEW SYSTEM | 190 |
| 6. Introduction | 190 |
| 7. Methods | 195 |
| 8. Results | 213 |
| 9. Discussion | 232 |
| 10. Summary | 241 |
| GENERAL DISCUSSION | 244 |
| OVERALL CONCLUSIONS | 259 |
| FURTHER WORK | 261 |
| REFERENCES | 262 |
| APPENDICES | 277 |

Index to Tables

| TABLE | PAGE |
|--|------|
| 1.1 Characteristics of studies included in the first systematic review - Alternative methods of follow up in breast cancer | 100 |
| 1.2 Methodological quality ratings of studies included in the first systematic review – Alternative methods of follow up in breast cancer | 102 |
| 2.1 Patient characteristics and original pathology of the entire Edinburgh study population | 118 |
| 2.2 Patient characteristics and original pathology of all patients from the Edinburgh study population in whom relapse was diagnosed | 121 |
| 2.3 Site of relapse and method of detection – Edinburgh cohort | 123 |
| 2.4 Clinico-pathological features of the recurrence by method of detection - Edinburgh cohort | 127 |
| 2.5 Age and original treatment – Glasgow cohort | 130 |
| 2.6 Annual hazard rate for isolated loco-regional relapse – Glasgow cohort | 132 |
| 3.1 Methodological quality ratings of studies included in the second systematic review – pattern of locoregional relapse, method of detection and impact on survival | 148 |
| 3.2 Characteristics of studies included in the second systematic review – pattern of locoregional relapse, method of detection and impact on survival | 153 |
| 3.3 Method of detection of all locoregional relapses for each study included in the second systematic review | 154 |

| | | |
|-----|---|-----|
| 3.4 | Method of detection of all ipsilateral breast relapses for each conservation surgery study included in the second systematic review | 157 |
| 3.5 | Method of detection of all axillary relapses for each study included in the second systematic review | 157 |
| 3.6 | Method of detection of all contralateral breast relapses for each study included in the second systematic review | 158 |
| 4.1 | Expected frequency of clinic visits among patients in the patient expectations of follow up study | 178 |
| 5.1 | Characteristics of patients enrolled into the feasibility/acceptability of automated telephone follow up study | 214 |
| 5.2 | Mean scores for baseline questionnaire delivered on paper for patients in the feasibility/acceptability study | 215 |
| 5.3 | Mean scores for follow up questionnaire delivered using the computer telephone system for patients in the feasibility/acceptability study | 217 |
| 5.4 | Mean difference between paper administered questionnaire score and computer telephone administered score for each of the questionnaire subscales | 227 |
| 5.5 | Pearson correlation coefficients for the degree of correlation between paper and telephone scores for each of the questionnaire subscales and for calculated overall scores | 228 |
| 5.6 | Change in scores between the baseline paper version and the telephone questionnaire completed at ten months in patients in the feasibility/acceptability study | 230 |

Index to Figures

| FIGURE | PAGE |
|---|------|
| 2.1 Overall survival for all 1312 patients in the Edinburgh cohort | 119 |
| 2.2 Cause specific survival for all 1312 patients in the Edinburgh cohort | 120 |
| 2.3 Annual incidence of relapse per hundred women at risk in the Edinburgh cohort | 122 |
| 2.4 Survival from original operation in all women from the Edinburgh cohort who suffered ipsilateral breast relapse | 125 |
| 2.5 Survival from relapse in all women from the Edinburgh cohort who suffered ipsilateral breast relapse | 126 |
| 2.6 Overall survival for all 198 patients in the Glasgow cohort | 131 |
| 2.7 Cause specific survival for all 198 patients in the Glasgow cohort | 131 |
| 2.8 Annual incidence of relapse per hundred women at risk in the Glasgow cohort | 132 |
| 2.9 Survival from relapse in all women from the Glasgow cohort who suffered relapse | 134 |
| 2.10 Survival from original operation in all women from the Glasgow cohort who suffered relapse | 135 |
| 3.1 Survival from original operation in patients suffering ipsilateral breast relapse in studies included in the second systematic review | 160 |
| 3.2 Survival from original operation in patients suffering axillary relapse in studies included in the second systematic review | 161 |

| | | |
|------|---|-----|
| 3.3 | Survival from original operation in patients suffering contralateral breast relapse in studies included in the second systematic review | 162 |
| 5.1 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Physical well being | 222 |
| 5.2 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Social well being | 222 |
| 5.3 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Emotional well being | 223 |
| 5.4 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Functional well being | 223 |
| 5.5 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Breast cancer subscale | 224 |
| 5.6 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Endocrine subscale | 224 |
| 5.7 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - Arm subscale | 225 |
| 5.8 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - FACT B overall | 225 |
| 5.9 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - FACT G overall | 226 |
| 5.10 | Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient - FACT ES overall | 226 |

Index to Appendices

| APPENDIX | PAGE |
|---|------|
| 1.1 Main search string used to identify studies – first systematic review of alternative methods of follow up in breast cancer | 278 |
| 1.2 Supplementary search string used in CancerLit | 279 |
| 1.3 Supplementary search string used in Web of Sciences | 280 |
| 1.4 Details of additional authors used in the first systematic review of alternative methods of follow up in breast cancer | 281 |
| 1.5 QUOROM statement for the first systematic review of alternative methods of follow up in breast cancer | 282 |
| 3.1 Main search string used to identify studies – second systematic review of pattern of relapse, method of detection and impact on survival | 284 |
| 3.2 Details of additional authors used in the second systematic review of pattern of relapse, method of detection and impact on survival | 285 |
| 4.1 Complete questionnaire and information sheet used in the study on patient expectations for follow up in breast cancer | 286 |
| 5.1 Seven initial questions used as part of the automated telephone follow up study | 289 |
| 5.2 Complete FACT-B questionnaire with arm and endocrine subscales which formed the main part of the automated telephone questionnaire | 290 |

Acknowledgements

I would like to thank the following people for their support in this project:

Professor T. G. Cooke (Professor of Surgery, Glasgow Royal Infirmary) for giving me this research project, acting as my supervisor and for providing me with support and encouragement throughout.

Mr J. Michael Dixon (Consultant surgeon and senior lecturer, Edinburgh Breast Unit) for giving me unlimited access to his patient data, which allowed me to undertake the analysis in chapter 2 which in turn contributed markedly to the review in chapter 3. Thanks also to **Gill Kerr** and **Wilma Jack** for their assistance in obtaining case notes and additional data in respect of the Edinburgh data.

Ms Katherine Krupa (Consultant Surgeon, Paisley Royal Alexandra Hospital) for acting as joint author on the systematic reviews presented here as chapters 1 and 3, and for sitting through hours of reference searches and reading countless abstracts. Thanks also for much general help and advice given throughout this project.

Professor Adri Voogd (Professor of Epidemiology, University of Maastricht, Holland) for giving me access to his data presented in the systematic review carried out in chapter 3, and for translating his databases from Dutch into English, and to all the other authors who took the time to respond to my requests for data even if none were available.

Mr Christopher R Wilson (Consultant Surgeon, Western Infirmary, Glasgow) for much help and advice during the planning of this project.

Flora MacInnes, Eileen Kesson and Paul Burton for their help and for much advice in constructing the new patient database upon which the telephone follow up system hangs.

Thomas Slaven and all at AxSys technology for being so helpful during the construction of the new telephone system and database.

Helen Morrow (Manager of Business Operations at FACIT) for permission to use the FACT questionnaires in this research, and for advice on scoring these questionnaires.

Finally, thanks to **Dr Donald C. McMillan** (University Department of Surgery, Glasgow Royal Infirmary) for always having an open door and an encouraging word, as well as general advice on statistical methods.

Dedication

I would like to dedicate this thesis to Katherine, whose support has been invaluable, and to Michael, whose imminent arrival gave me the fear needed to complete this thesis timeously.

Declaration

I declare that the work presented in this thesis has been solely carried out by me, except where indicated below:

Ms Katherine Krupa carried out independent literature searching and read abstracts in parallel with my own searches for both of the systematic reviews conducted in this thesis. This is in keeping with the advice that such searches should be conducted independently by two authors.

The design of the patient database and selection of the questionnaire for use in the telephone follow up section of this project were carried out by me with advice from Eileen Kesson, Paul Burton and Flora MacInnes. The database and the phone engine were built by AxSys technology.

The work carried out in this thesis has so far led to the following publications and presentations:

Publications

- Montgomery DA, Krupa K, Cooke TG.
Alternative Methods of Follow Up in Breast Cancer: A Systematic Review of the Literature. **British Journal of Cancer** 2007; 96(11): 1625-32
- Montgomery DA, Krupa K, Jack WJL, Kerr GR, Kunkler IH, Thomas J, Dixon JM.
Changing Pattern of the Detection of Loco-regional Relapse in Breast Cancer: The Edinburgh Experience. **British Journal of Cancer** 2007; 96(12): 1802-7
- Montgomery DA, Krupa K, Cooke TG.
Follow up in breast cancer: Does routine clinical examination improve outcome? A systematic review of the literature. **British Journal of Cancer** 2007; in press

Editorials

- David Montgomery.
Long-term follow-up of early-stage breast cancer patients treated by breast-conserving surgery should be based on mammography
Senology Oncology and Women's Health newsletter. www.senology.it; 3rd July 2007
- J. Mike Dixon, David A Montgomery.
Evidence based breast cancer follow up.
British Medical journal. In press

News Articles

- Women treated for early breast cancer should be followed for at least ten years.
British Medical Journal 2007;334:1240

Oral Presentations

- David A. Montgomery, Katherine Krupa, Wilma Jack, Gill Kerr, Ian Kunkler, Jeremy Thomas, Mike Dixon.
Changing Pattern of Recurrence Detection in Breast Cancer: The Edinburgh Experience.
Presented to the 10th annual **Nottingham International Breast Cancer Conference.** September 18th-20th 2007
- David A. Montgomery, Katherine Krupa, Wilma Jack, Gill Kerr, Ian Kunkler, Jeremy Thomas, Mike Dixon.
Changing Pattern of Recurrence Detection in Breast Cancer: The Edinburgh Experience.
Presented to the annual meeting of the **Association of Surgeons of Great Britain and Ireland.** Manchester, 18th April 2007.
- David A. Montgomery
Evidence for follow up in breast cancer.
West of Scotland Managed Clinical Network in Breast Cancer Meeting, Ayr Hospital, 19th January 2007

Poster Presentations

- David A. Montgomery, Katherine Krupa, Timothy G. Cook
Alternative methods of follow up in breast cancer: A systematic review of the literature.
Presented to the 10th annual Nottingham International Breast Cancer Conference. September 18th-20th 2007

General Introduction

Introduction

Breast cancer is an important and increasingly prevalent clinical condition. In 1972, the American Cancer Society reported that this disease affected some one in fourteen women in the Western world[Seidman H, 1972]. In the three and a half decades since the publication of that report, there has been an increase in the proportion of women diagnosed with breast cancer such that now one in every nine of all women in the UK will be affected at some stage in their life. In real terms, this equates to some 3646 new diagnoses of breast cancer in Scotland in 2002[Scottish cancer registry, 2005], an increase of almost 50% from the 2480 cases diagnosed in Scotland in 1980.

While the incidence of this disease has increased dramatically over the past few decades, in the same period there has been a marked reduction in overall mortality from breast cancer[Scottish cancer registry, 2005]. The 10 year survival of all breast cancer sufferers aged between 15 and 74 at diagnosis was 41.5% in the late seventies and early eighties. By contrast, some 60.1% of all women diagnosed in the first half of the 1990s were alive and well ten years after diagnosis[Scottish cancer registry, 2005] and this figure continues to improve. The overall result of increasing incidence coupled with continually improving survival has been an increasing prevalence of breast cancer, and the impact of this has been very apparent in clinical practice.

The treatment of breast cancer requires significant resources. Every patient with a breast cancer requires assessment at a clinic, investigation, diagnosis and planning of treatment. Most require an operation and many also need additional adjuvant therapy of some type. The

successful diagnosis and treatment of this condition is not the end of the story. Breast cancer is a chronic condition which can recur at any time. Surveillance has therefore traditionally been offered with a view to detecting recurrent disease as early as possible, thus enabling the institution of further treatment.

Historically, follow up care has placed great demands on the available resources. Traditional follow up was time consuming for both patients and clinicians, with multiple visits being carried out for prolonged periods after completion of treatment. Under the American Society of Clinical Oncology guidelines from 1999, patients would undergo follow up for ten years in total, and would initially attend every three months for routine check-up[American Society of Clinical Oncology, 1999]. The numbers of additional tests and investigations provided varied between units but has in the past included frequent blood tests, ultrasound scans, chest x-rays and bone scans designed to detect recurrent disease early.

The burden on health care providers of supplying follow up care to patients treated for breast cancer, both in financial terms and in terms of clinical time, has been considerable. Yet the majority of patients with breast cancer do not relapse. Only 11% of patients treated in our practice will relapse in the first 5 years (personal communication). There are potential savings to be made in both time and resources if less intensive follow up methods can be shown to be equally safe and effective compared with the more traditional approach of frequent visits and multiple tests.

Recent guidelines from the National Institute for Clinical Excellence (NICE) in the UK have recommended greatly reduced frequency and duration of follow up [National Institute for Clinical Excellence, 2002]. These recommendations appear to have been driven by a desire to reduce the burden on clinics rather than being based on any emerging new evidence that reduced follow up is safe. As a consequence, the NICE guidelines have been unpopular with clinicians and have led to a raft of alternative guidelines being produced across the UK by various local managed clinical networks. This has done little to clarify what is required for follow up in breast cancer.

In this thesis, an attempt has been made to bring some clarity to the issue of follow up after breast cancer. Patients are brought back to clinic to screen for potentially treatable relapse, to monitor for side effects of treatment and for the detection and amelioration of psychosocial problems in the wake of diagnosis and treatment. The author has sought to establish, by a combination of systematic review of the literature and analysis of relapse in local cohorts, the contribution of routine clinic visits to the achievement of these aims, and also to explore what alternatives to traditional follow up have been suggested and subjected to evaluation in a randomised controlled trial. In the second section of the thesis, the author has explored patients' views on follow up and their expectations when they attend clinic visits. Finally, a prospective cohort study has been conducted to evaluate a novel method of follow up which may provide a more effective method of achieving the aims of follow up while reducing the burden on the clinic system.

In the remainder of this introduction, various key guidelines for follow up which currently exist are described, along with some discussion of the limited evidence which underpins these guidelines. The problems and complications associated with diagnosis and treatment of breast cancer are also described in order to highlight the problems which can arise in patients and which should be detected by an effective follow up regime.

Current Guidelines

Numerous organisations provide guidelines for the follow up of women after breast cancer. These organisations recognise that follow up should have several of aims, which are clearly defined by the National Institute for Clinical Excellence in England and Wales (NICE)[National Institute for Clinical Excellence, 2002]. A key aim remains the early detection of potentially treatable local relapse. NICE currently do not recommend follow up for the detection of metastatic disease [National Institute for Clinical Excellence, 2002]. In addition, follow up should be directed towards detecting and ameliorating adverse effects of therapy, in particular lymphoedema, and providing psychological support to women at need[National Institute for Clinical Excellence, 2002]. While these objectives are common to all of the major international guidelines, there is tremendous variation in what follow up is recommended to achieve these aims.

In the United Kingdom there has been a recent trend among producers of guidelines towards reducing the amount of follow up recommended. The potential cost savings to be made by more limited follow up has prompted NICE to recommend that patients should be discharged after only two to three years of routine visits, although patients should continue to get

mammography after this time and have open access to breast care nurses in the event of a problem[National Institute for Clinical Excellence, 2002]. The Association of Breast Surgery at the British Association of Surgical Oncology (BASO) recommend a follow up duration almost twice that recommended by NICE, but with discharge from routine follow up still only at five years after treatment for breast cancer[The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005].

In contrast, the recent guidelines published in North America do not advocate reducing follow up. The American Society of Clinical Oncology (ASCO) in their 2006 guidelines continues to recommend 3 to 6 monthly visits for the first three years, 6 to 12 monthly visits for two years then annual visits after five years with no advice on discharging patients[Khatcheressian JL et al., 2006]. The Canadian Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer is somewhat more pragmatic regarding the frequency of visits, suggesting that these should be tailored to individual patients needs, but their recommendations explicitly state that follow up should be provided indefinitely[Grunfeld E et al., 2005].

The advice regarding mammographic frequency is equally varied. Mammography is recognised as important in the Canadian guidelines, with mammograms recommended annually[Grunfeld E et al., 2005]. They do concede that there is little high level evidence to support this contention. Similar recommendations are made by ASCO[Khatcheressian JL et al., 2006]. In contrast, NICE suggest that the yield of mammography is low, and that local networks should decide their own policy[National Institute for Clinical Excellence, 2002]

and BASO also state that an ideal frequency of mammography has not been established, and that one to two yearly mammography should be undertaken for up to ten years after diagnosis[The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005].

There is clearly disagreement between the guidelines with regards to just how much follow up is required. Yet common to all the guidelines is a particular emphasis on providing the most intensive follow up during the first three to five years after diagnosis and treatment, with either reduced frequency of visits or discharge to the general practitioner after this time. This results from the perception that recurrent disease is most common in the first three to five years after treatment[The ATAC Trialists' Group: 2005;Saphner T et al., 1996;Hussain ST et al., 1995]. Also common to the guidelines is that mammography remains the only recommended investigation, but there appears to be little evidence for how frequently mammography should be carried out. The result is that, while mammography is recommended annually in North America[Grunfeld E et al., 2005;Khatcheressian JL et al., 2006], in the UK this investigation can be carried out less frequently than this within the framework of both the NICE and BASO guidelines[National Institute for Clinical Excellence, 2002;The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005].

While it is desirable to reduce the impact of follow up provision on hospital resources, this should not be at the cost of poorer outcomes for women with breast cancer. If routine clinic visits and frequent mammography are necessary for the detection of relapse, then the more

minimal British guidelines are unsafe. Of course, if follow up can safely be reduced, then the guidelines from North America are unnecessarily overcautious, and will lead to excessive impact on both time and financial resources.

In establishing which of these is the case, and in planning any alternative follow up after breast cancer, it is necessary to establish the contribution of routine clinic visits to achieving the various aims of follow up. The timing and pattern of potentially treatable relapse should be established, and the contribution of routine clinical examination to detection of such relapse must be defined. Equally important is whether clinical examination confers survival advantage. The effectiveness of routine outpatient follow up on the detection of adverse effects of treatment and psychological problems related to the diagnosis and treatment of breast cancer should also be established. There follows a discussion of how fully these questions have been answered in the literature to date.

Distant Relapse

Follow up of patients after treatment for breast cancer has usually focussed on the detection of recurrent disease. This has historically included surveillance for metastatic disease in addition to locoregional relapse. As a result, breast cancer follow up has in the past been a very labour intensive process. In addition to clinical examination and mammography which now form the basis of most follow up regimes, numerous other tests and investigations were undertaken by clinicians during routine follow up visits. These included blood tests such as full blood count, liver function tests and tumour markers, as well as radiological investigations such as chest x-ray, liver ultrasound and bone scans. While clinical

examination and mammography were undertaken to survey for locoregional relapse and new contralateral tumours, these other investigations were designed to detect distant relapse before it became symptomatic. It was hoped that this would allow for targeted therapy with a view to prolonging life.

One questionnaire based study in 1991 revealed how commonplace some of these other investigations were. 197 oncologists practicing in the United States were surveyed. Most of those responding reported the use of routine blood tests at least every six months. Chest x-ray was carried out at least annually by the majority of respondents also. Liver ultrasound and bone scanning undertaken less commonly, but still featured in the follow up protocols of some clinicians in up to 10% and 30% respectively of high risk patients[Loomer L et al., 1991].

By the beginning of the 1990s, there was growing evidence that such invasive and intensive monitoring was at the very least poorly cost effective and many clinicians were beginning to consider it unnecessary. It had already been reported that signs and symptoms were the first indicators of relapse in 92% of patients[Scanlon EF et al., 1980]. In order to detect just eight asymptomatic relapses, \$89446 was spent on 391 chest x-rays, 299 bone scans, 44 liver ultrasounds and multiple blood profiles on their study cohort of just 99 patients[Scanlon EF et al., 1980]. This analysis was conducted in 1980; the cost today of such intensive monitoring would be significantly higher.

Other authors have conducted similar analyses, with comparable results. With intensive follow up regimes comprising frequent visits, a variety of blood tests and numerous radiological investigations, clinical signs or symptoms were nevertheless the first sign of recurrent disease in between 78% and 95% of those developing relapse in the 1970s and 1980s[Valagussa P et al., 1981;Ormiston MC et al., 1985;Ojeda MB et al., 1987;Mansi JL et al., 1988;Muss HB et al., 1988;Winchester DP et al., 1979]. While these figures are a percentage of all relapses, both metastatic disease and locoregional, several of the studies report that all the loco-regional relapses diagnosed presented with signs or symptoms[Muss HB et al., 1988;Valagussa P et al., 1981;Mansi JL et al., 1988;Winchester DP et al., 1979]. This would intuitively be the case, as local relapse should be clinically apparent after mastectomy, and certainly no routine investigations would detect relapse in the chest wall. It was only upon introduction of breast conserving surgery that asymptomatic, non-palpable relapse began to be detected using mammography and the studies above were conducted prior to the wide scale adoption of breast conserving techniques. It is likely that all of those relapses detected by routine investigations in the above studies were distant disease.

Such follow up was labour intensive, expensive and a significant burden on patients' and clinicians' time and effort. The yield in terms of detecting asymptomatic and potentially treatable relapse was poor. Such follow up continued as it was unclear whether the detection of distant relapse using these methods had any influence on patients' outcome. Until this time, the published work in this area had been mainly retrospective or simple observational. It was impossible to recommend reducing the intensity of follow up while some studies were reporting detection of asymptomatic relapse by routine investigations in up to 22% of

patients and the possibility of a survival advantage in these patients had not yet been discounted, although such a survival advantage did not seem likely[Ormiston MC et al., 1985].

It was clear that a prospective evaluation of the impact of such intensive surveillance after breast cancer on patients' overall survival was required. Two such trials were published in 1994, back to back in the same issue of the Journal of the American Medical Association, and provided the first randomised, prospective evidence against the accepted practice of intensive investigation during follow up.

The GIVIO investigators presented a series of 1320 women with stage I, II and III unilateral primary breast cancer[The GIVIO investigators, 1994]. A study group was seen 3 monthly at clinic and underwent regular bone scan, liver ultrasound, chest x-ray and biochemical profile. A control group received 3 monthly clinical examinations also, but only clinically indicated tests. Both groups had annual mammography. Follow up was for a mean 71 months.

Metastatic disease was detected in 31% of asymptomatic patients in the intensively observed group, and in only 21% of the group screened by clinical examination and mammogram alone. In those who had metastatic disease detected, there was no significant difference in the mean time to detect this disease (53.39 months in the intensive group versus 54.07 months in the clinical examination and mammogram alone group).

Perhaps the most significant finding was that there was no difference between the two groups in terms of overall 5 year survival. There was a difference in the proportion of people in each group diagnosed with recurrent disease. This difference was not commented on by the authors but, if significant, may have reflected missed occult metastatic disease in the less intensively studied group. A longer follow up period would have allowed these occult relapses to declare themselves and may have had impact on the overall survival over, for example, ten years.

Despite this, the authors concluded that there was no justification for actively investigating for asymptomatic distant relapse. There was no evidence that finding such disease altered outcome, but the extra investigations had significant implications both on hospital resources and on patients' time and well being.

The second of these trials was by Del Turco and colleagues[Del Turco MR et al., 1994]. Here, 1243 consecutive patients, also with invasive unilateral cancers but no distant metastasis, were studied over a minimum follow up period of five years. The results not only confirmed the findings of the GIVIO investigators, but went further.

Like the GIVIO trial, there was no difference in time to detect local relapse between the two groups. There was similarly no difference in overall survival. There were significantly more isolated intrathoracic and bone metastases detected in the intensively studied group, and therefore an overall reduced time to diagnosis of metastatic disease in those who had been subjected to intensive investigation.

The authors concluded that early diagnosis of asymptomatic metastatic disease does nothing to alter prognosis. This was reflected in identical survival rates between both groups despite earlier diagnosis of metastatic disease those intensively followed. The authors conclude that early diagnosis of asymptomatic and incurable metastatic disease would be a source of undue distress for patients and would lead to the difficult dilemma of whether to treat such incurable but as yet asymptomatic disease with often very toxic therapies in patients with limited life expectancy and no conclusive evidence that such therapies prolonged life. These authors concluded that actively investigating for metastatic disease would cause undue distress with no survival benefit, and that such investigations should stop.

One criticism of both of these trials is that the length of follow up is short. Some commentators did suggest that 5 year survival was an inadequate measure of long term outcome after breast cancer. This criticism was answered when Del Turco et al published ten year follow up data showing identical survival rates between those followed up intensively compared with those followed up with only clinical examination and mammography over this longer period[Palli D et al., 1999].

The reports outlined above had all shown that more intensive follow up protocols required more clinical time and were much more costly yet were responsible for diagnosing relatively few asymptomatic relapses. The only potential advantage, improved survival for these few patients, was discounted by the findings of the GIVIO group and Del Turco et al. Both groups of authors recommended that such intensive follow up be abandoned, and that follow

up regimes be targeted toward detecting curable locoregional relapse or new contralateral disease only. This approach would provide marked savings both financially and in terms of clinical resources while reducing the potential for undue distress among patients. The major guidelines for follow up in the UK and North America now no longer recommend routine investigations designed to detect metastatic disease.

Locoregional Relapse

Survival after locoregional relapse

Local relapse after mastectomy was historically associated with poor outcome. In one early study, only 22% of women who developed a single skin lesion were alive and well at ten years. In the presence of multiple lesions, 5 year survival was only 10% with no patients remaining alive at 10 years[Fentiman IS et al., 1985]. The introduction and widespread use of breast conserving surgery had marked a significant change in the way early breast cancer was dealt with. But while the improved cosmetic results were welcomed, there was initial concern among clinicians that there may be greater risk of local relapse in the conserved breast than in a post mastectomy chest wall. If this was borne out in clinical practice, and such relapse was attended by the same dismal prognosis as after mastectomy, then this new technique would be accompanied by reduced disease free and overall survival.

Early investigators did report poorer prognosis in patients who had suffered local relapse than in those in whom local control was achieved at the first operation. Whelan and colleagues reported that ipsilateral breast tumour relapse after wide local excision was associated with a relative risk of mortality of 2.18 compared with those who had not suffered

local relapse[Whelan T et al., 1994]. They further demonstrated that relapse within the first year particularly was associated with poor outcome. Others reported even poorer outcome after local relapse, with one group suggesting that local relapse resulted in a relative risk of distant metastases of 4.9 (95% confidence interval 3.15-7.62) and death of 4.29 (95 % confidence interval 2.93-6.28) after 52 months median follow up compared with those who do not recur[Elkhuizen PHM et al., 1998] and a second group reporting a relative risk of death of 8.8 (95% confidence interval 4.6-16.8) upon developing a local relapse within five years compared with those who do not recur[Kemperman H et al., 1995]. So it would appear that local relapse was associated with increased mortality. There was some debate regarding the mechanism; was local relapse a risk factor for later distant spread and therefore mortality or was it actually a marker of disseminated disease?

Some clarity was brought to this issue in 1999 when Fortin and colleagues reported a relative risk for mortality of 3.6 at six years for those who relapse after local resection of breast cancer compared with those who have not suffered a relapse, a figure somewhere in the middle of those other published relative risks[Fortin A et al., 1999]. Among patients who suffered metastases, a different time to those metastases was observed in those with local failure as compared to those with local control. In patients with local control, the risk of metastatic disease peaked at two years after surgery. This had been observed by other investigators previously[Veronesi U et al., 1995]. In those who suffer a local relapse, time to metastases was between 5 and 6 years after the initial procedure[Fortin A et al., 1999]. This tends to suggest not that local relapse is a marker of disseminated disease, but that it is a second source of potential distant disease. This would, in turn, suggest that adequate early

treatment of local relapse could be associated with a similar rate of cure as the initial cancer. In keeping with this, several investigators demonstrated that cure was most likely in those with small relapses and no lymph node involvement[Clark DH et al., 1985;Veronesi U et al., 1995;Voogd AC et al., 2005].

There is some suggestion that local relapse after breast conserving therapy is not a homogenous entity, and may be biologically distinct from local relapse after mastectomy, at least in some cases. It has been reported that survival after a local relapse in patients treated by breast conserving therapy was greater in those who recurred after 2 years as compared with those who recurred within two years[Veronesi U et al., 1995]. van der Sangen and colleagues made a similar observation, but in their study mortality was higher in those who relapsed within five years of the original diagnosis[van der Sangen MJC et al., 2006]. This has led to suggestions that late relapse is biologically distinct from early relapse and may reflect a second primary rather than true recurrence of the primary cancer. As early as 1989 it had been suggested that, while 79% of relapses were within the vicinity of the tumour bed, the proportion that arose elsewhere in the breast increased with time from operation such that, by ten years, all ipsilateral breast relapses probably had to be considered new tumours[Kurtz JM et al., 1989]. These authors also confirmed a favourable prognosis with late relapse compared with early. Fewer relapses after 60 months were inoperable (1.4% v. 17%) and relapse after 60 months had a 5-year survival from relapse of 84% versus 61% in those who recurred within 60 months. In fact, patients who recurred in the ipsilateral breast more than five years from their original treatment had identical 15 year survival to those who had never suffered a local relapse[Kurtz JM et al., 1990].

So while early local relapse may be associated with reduced survival, the same cannot be said for all local relapses and particularly for small node negative relapses. As such, the early detection of these relapses should be a central aim of follow up in breast cancer. In order to design follow up schedules to meet this aim, it is important to fully appreciate when such relapses occur, and what components of follow up are required for detection of such relapse.

Timing of relapse

As discussed above, the recommended duration of follow up varies between organisations, with anything from 2-3 years follow up[National Institute for Clinical Excellence, 2002] through to indefinite[Grunfeld E et al., 2005] visits being recommended. This stems from a lack of consistency over what types of relapse are being analysed when studies on the timing of relapse are undertaken.

The incidence of metastatic disease peaks at around two to three years after initial treatment[Fortin A et al., 1999;Veronesi U et al., 1995]. Whether potentially treatable relapse follows a similar pattern is less clear from the literature. One group reported that most local relapse in a small cohort of 56 patients occurred early, with 37.5% occurring in the first two years and only 19.6% occurring after 5 years[Nishimura R et al., 2006]. This pattern has been reported by others, with the majority of local recurrences within a large Australian cohort occurring within 5 years of the original treatment[Elder EE et al., 2006] and 93% of all relapses in a study by Hussain and colleagues occurring in the first five years after treatment[Hussain ST et al., 1995]. This is in keeping with local relapse being more

predominant in the early period after treatment, and is reflective of what had been reported in the literature generally. These groups only analysed ipsilateral, true local relapses. A more detailed look at the issue reveals that this is less than the whole story.

Freedman and colleagues analysed the pattern of all potentially treatable relapses after conservation surgery and radiotherapy[Freedman GM et al., 2005]. They reported that the incidence of true, ipsilateral local relapse was constant at around 2-2.5% every five years until 15 years. They reported additionally that the number of new breast cancers elsewhere in the ipsilateral breast increased over time, with the greatest increase in incidence of these new tumours being after ten years, going from 2% of patients affected in the first ten years to 6% of all patients affected at 15 years. The proportion of patients suffering ipsilateral breast relapse in that study was therefore 13% at 15 years, with almost half being new tumours away from the original surgery site. The incidence of new contralateral cancers in that study also increased with time and was an again 13% after 15 years, identical to the percentage of women suffering ipsilateral relapse. Taken together, these figures suggest that the rate of potentially treatable relapse, rather than peaking early and falling off, may in fact be rising with time. These findings are in keeping with those of Veronesi and colleagues, who have reported that the risk of contralateral cancer is similar to the risk of ipsilateral breast relapse and that, by 20 years, the majority of ipsilateral relapses were new cancers in a different quadrant of the breast from the original tumour[Veronesi U et al., 2001;Veronesi U et al., 2002].

These findings suggest that the risk of a new, treatable event remains for at least 15-20 years, and probably beyond. The risk of a new cancer away from the site of the original primary or in the contralateral breast increases over time. These facts have obvious implications. Firstly, providing more frequent follow up during the first few years after treatment, as recommended widely in the guidelines[American Society of Clinical Oncology, 1999;National Institute for Clinical Excellence, 2002], is not in keeping with the apparent pattern of relapse. Certainly, metastatic disease is more common initially, but, as discussed above, such relapse will present with symptoms initially and routine investigation for metastatic relapse is not recommended. If detection of potentially treatable relapse is to be the focus of follow up, then follow up cannot be limited in duration, but must be continued for at least ten years.

One argument presented for more frequent follow up in the first few years after relapse would be that the prognosis of patients who recur locally in the first few years is poorer than those who recur later. They may have more aggressive disease, and so very early detection may improve outcome. There are two problems with this argument. Firstly, that the current practice is to provide frequent follow up in these first few years and despite this the outcome is still poorer in those who recur early. Secondly, while it is agreed that early relapse has a poor prognosis, there is little agreement as to what constitutes early relapse. Several authors have suggested relapse before two years has a poor prognosis, others have stated that relapse before five years has a poorer relapse than after[Veronesi U et al., 1995;van der Sangen MJC et al., 2006;Kurtz JM et al., 1990;Kurtz JM et al., 1989]. This issue will be confused by differences between the authors involved as to how locoregional relapse is defined, in

particular whether new ipsilateral cancers away from the original operation site are included as locoregional relapse and whether isolated axillary relapse is included.

From the few studies which do exist and which have clearly detailed the incidence of all types of relapse, both ipsilateral breast and axillary as well as contralateral breast, it appears that there is no peak in incidence of treatable relapse. There is no evidence for a period where follow up should be focussed on, and similarly no evidence to suggest a time when patients could be safely discharged. Any component of follow up which is shown to be of benefit in terms of detecting potentially treatable relapse should therefore be continued indefinitely.

The main components of follow up are mammography and routine clinical examination. Their contribution to detection of potentially treatable relapse is discussed below.

Detection of Locoregional Relapse – Contribution of Mammography

The use of routine mammography during follow up in patients with breast cancer has logical appeal. Locoregional relapse after mastectomy is a problem which can be confirmed with examination and biopsy. Certainly, there is no role for mammography in detecting such relapse as there is limited breast tissue on the post mastectomy chest wall. Breast cancer follow up can be considered a screening programme for women in a high risk population group. It is generally well recognised that women who have suffered from breast cancer are at increased risk of developing a second contralateral primary. Studies have put this risk at between three and five times the risk of developing a primary breast cancer in the general

female population[Prior P and Waterhouse JAH, 1978;Chaudary MA et al., 2006;Adami HO et al., 1985]. Alternatively, some 15-20% of patients with breast cancer will eventually develop a contralateral breast cancer at some stage[Horn PL and Thompson WD, 1988;Fracchhia AA and Borgen PL, 1991] and the risk of a new contralateral primary is identical to the risk of ipsilateral relapse in women who have had breast conserving surgery[Freedman GM et al., 2005]. Mammography plays a part in screening previously unaffected patients, so there is surely a role for mammography in detecting metachronous contralateral new primary breast cancer in patients who have previously suffered malignant breast disease.

Also of relevance in the current era where many more women are treated with breast conserving surgery is the question of whether mammography can usefully be employed in detecting local relapse in the conserved breast. We have already seen that early detection of locoregional relapse can improve survival, as patients do better when such relapse is small or axillary node negative[Clark DH et al., 1985;Veronesi U et al., 1995;Voogd AC et al., 2005].

Interestingly, there have been few studies of the use of mammography in the routine follow up of breast cancer patients, and no prospective randomised controlled trials. In 1991, Mellink and colleagues published the results of an observational, retrospective cohort study on the benefits of mammography in routine breast cancer follow up[Mellink W et al., 1991]. The populations of two cities in Holland were compared. In the first, breast cancer follow up consisted of regular clinical examination coupled with routine mammography. This was

carried out annually. The second city made no use of routine mammography in the post breast cancer follow up programme.

In both cities, some 3% of patients developed metachronous contralateral breast cancers over 12 years of follow up. There were stark differences between the two groups when the grade and stage of cancers were looked at. In the group of patients who had mammograms, 35% of tumours were less than 10mm, while this was the case in only 7% of those who did not. The patients who underwent mammographic screening were also more likely to be node negative (75% against 57%). In the group screened mammographically, 17 cancers were detected by mammography (71%), compared with only 7 by clinical examination.

The authors conclude that the differences in stage of contralateral cancers between the two groups could be directly attributed to the mammography. Certainly, there are flaws in this study, not least of which is that two separate groups of doctors were used. Yet this was the first study to assess the use of mammography in routine breast cancer follow up and it certainly suggested that mammography has a role to play. If mammographically detected relapse is indeed smaller and more often node negative, one would expect such relapse to be associated with a relatively favourable prognosis[Clark DH et al., 1985;Veronesi U et al., 1995;Voogd AC et al., 2005]. The study makes no attempt to address the frequency with which mammography should be used.

Other authors have approached this subject in different ways, but with similar results. In a retrospective observational study looking at patients who presented to clinic with a primary

breast cancer, and who subsequently developed a metachronous contralateral primary during follow up, Samant and colleagues noted that the metachronous tumour was significantly more likely to be non-invasive than the primary breast cancer (11.4% v. 5.1%, $p < 0.02$) [Samant RS et al., 2001]. Among the invasive cancers, the mean size of the metachronous tumour was significantly less than that of the primary breast cancer (1.94 v. 2.55 cm, $p < 0.001$). Metachronous tumours were more than twice as likely to be diagnosed mammographically compared with the primary cancer (46.2 v. 19.9 $p < 0.001$), and tumours which were detected mammographically were smaller than those detected either by clinical examination by a doctor or self examination by a patient (1.30, 2.02 and 2.69 cm respectively).

While this is indirect evidence of a benefit of mammography at best, the authors nevertheless felt able to conclude that mammography results in earlier detection of metachronous contralateral breast cancer, and that this results in smaller cancers which are more likely to be node negative. If this is in fact the case, then mammographic screening of patients after treatment for breast cancer can be expected to improve the prognosis in patients who develop a second primary. Despite this indirect evidence from retrospective case note analysis of a benefit of mammography, there has been no attempt to prospectively establish how effective mammography is in detecting new contralateral cancers, or how frequently mammography should be undertaken.

There is similar evidence for the benefit of mammography in detecting ipsilateral breast cancer recurrence in the conserved breast. In 1999, Adri Voogd and colleagues reported that,

in women treated for breast cancer with breast conserving surgery and followed up by regular clinical examination and mammography, relapse detected only by mammography was smaller than that detected clinically, and patients with mammographically detected relapse had better disease free and overall survival than those with palpable relapse[Voogd AC et al., 1999]. Mammography detected a minority of relapses (24% in this series) and again there was no prospective assessment of how frequently mammography should be conducted.

There has been no attempt to establish prospectively the effect of regular follow up mammograms on overall survival after breast cancer. It has been suggested that there may be a reduced mortality rate in patients who have annual mammograms during follow up for breast cancer[Lash TL et al., 2006]. Among 65 year old women treated by breast conserving surgery for early stage breast cancer, there was a trend towards reduced all cause mortality with increasing numbers of mammograms performed routinely in the years after follow up. This group conclude that mammograms are therefore protective. These findings were far less robust than those of Voogd and colleagues as the authors did not analyse relapse or breast cancer specific mortality. Their outcome was based on all cause mortality and so it is very possible that they are witnessing either reduced likelihood of patients undergoing mammography if they have intercurrent health problems which later contribute to non-cancer death, or even a socio-economic effect whereby women less likely to attend for follow up mammograms are from lower socio-economic backgrounds with a lower life expectancy. This was an American study, were lack of health insurance in lower socio-economic groups is very likely to impact on the number of mammograms which are undertaken. These

explanations are speculative, but what is clear is that the study by Lash and colleagues does not prove a benefit to annual mammography in terms of breast cancer specific outcome.

Several authors have retrospectively analysed the way in which potentially treatable relapse is detected after breast conserving surgery, and reported varying proportions of relapses detected mammographically. One group stated that in a cohort of 643 patients treated between 1992 and 1998, there were 37 local relapses, none of which were detected by mammography[Donnelly J et al., 2001], while others have suggested that mammography detects up to 37.5% of local breast relapses[Jack WJL et al., 1998].

Several factors have limited the usefulness of these retrospective studies. For example, there has been no consistency in how local relapse is defined, with some authors including only true local relapse while others have included all new cancers in the ipsilateral breast, including those distant from the original operation. Additionally, small sample sizes and brief follow up have limited the usefulness of these studies. The study above which claimed that no relapses were detected by mammography followed only 643 patients for a median of 3 years and 11 months[Donnelly J et al., 2001]. As a consequence, the number of events is really too small to make sensible comment regarding the usefulness of mammography. A final issue affecting the usefulness of these studies is lack of clarity regarding the follow up schedule employed. A follow up schedule where mammography is conducted before a clinic visit will be more likely to detect relapses on mammograms than a schedule where mammography is carried out after clinical examination. It is uncertain what contribution more frequent mammography makes, as this has not been addressed.

The result of this lack of evidence regarding the place of mammography in follow up is that the recent guidelines which exist are vague in their recommendations regarding mammography. There is no evidence on which to base recommendations for an optimal schedule of mammography. As a result, the guidelines produced by the Royal College of Radiologists in 1995 which suggested that mammography should be performed one year after treatment to give a baseline for future comparison, and then every one to two years thereafter[Young JR and Wilson AR, 1995] are little more than best guess. The most recent guidelines produced by the National Institute for Clinical Excellence in England and Wales make no attempt to recommend a schedule, stating instead that the yield of mammography is low in follow up after breast cancer and that local networks should establish their own protocols[National Institute for Clinical Excellence, 2002]. They base their statements regarding the yield of mammography on just one paper, that by Donnelly and colleagues mentioned above in which no relapses were mammographically detected[Donnelly J et al., 2001]. The Canadian Steering committee state that mammography should be done, but acknowledge that the literature contains no high-level evidence to advocate any particular frequency[Grunfeld E et al., 2005].

As experience with the interpretation of post breast conserving surgery mammograms has increased, it is likely that the contribution of mammography to the detection of local relapse will have increased. It is also possible that these relapses will be detected earlier than relapse detected non-mammographically, given the evidence we have seen above regarding the early detection of contralateral relapse on mammography. It would be important to note if this has

happened, and if this translates to a survival advantage in patients whose relapse is detected mammographically compared with those clinically detected or symptomatic relapses. An investigation of this area should be central to any future recommendations regarding follow up in breast cancer.

Detection of Locoregional Relapse – Contribution of Clinical Examination

If the value of mammography in follow up is uncertain, then the need for routine clinical examination is even more so. Several authors have looked at the value of clinic visits in follow up for breast cancer. In 2004, Hiranamek published the findings of a retrospective review of relapse in 220 patients attending routine breast clinic follow up over the course of 1997[Hiranamek N, 2004]. This was a cross sectional analysis and so could not give any more than a snapshot of events. Moreover, due to the small number of patients analysed the event rate was very limited, with only 42 relapses in this group of patients. In four cases, data was incomplete, leaving just 38 patients for analysis. 25 of these patients had metastatic disease, leaving just 13 patients with potentially treatable locoregional relapse. Only one patient had a treatable relapse diagnosed by routine clinical examination while asymptomatic.

The current recommendations for follow up from NICE are based on a similarly small study, a retrospective analysis of a cohort of 643 patients operated on between 1992 and 1998 in the UK[Donnelly J et al., 2001]. Not only was the cohort relatively small, but median follow up duration was a very brief 3 years and 11 months. 108 patients suffered recurrent disease with full data available for 104 of these. The majority of these patients (67 patients, or 64% of the total) had metastatic disease. Only 37 patients had locoregional relapse. 77 (74%) were seen

at interval appointments having developed symptoms prior to their scheduled hospital visit. A further 18 (17.3%) drew attention to symptoms at a routine visit. Two cases of locoregional relapse were revealed by surveillance imaging. Unsuspected disease, locoregional in all cases, was detected on examination in 7 (6.7%) patients. The authors conclude that clinical examination adds very little in terms of diagnosing locoregional relapse. They go on to suggest that routine follow up after breast cancer should be continued to 2 years after treatment, and then delegated to general practice, the current recommendations from NICE[National Institute for Clinical Excellence, 2002].

Both of these analyses highlight some of the problems which have limited this area of investigation. Cross sectional analysis as used by Hiranmanek[Hiranmanek N, 2004] has been relatively uncommon as this type of study does not allow assessment of rate of relapse or an accurate reflection of the proportion of clinic visits at which relapse is detected. Retrospective cohort analysis is a better way to tackle the issue, as this allows an analysis of the rate of relapse and also allows the investigators to assess the rate of detection of relapse by each method. Most retrospective analyses, such as that carried out by Donnelly and colleagues[Donnelly J et al., 2001], have been in relatively small groups of patients and follow up has been brief. In addition, most have not included new contralateral breast cancers in their analysis, partly as a result of this being less common during the first few years after treatment.

This area of research has been hampered by fairly heterogeneous methodology among the various analyses which have been conducted, leading to studies which are really impossible

to compare and which add little in terms of evidence for what is required in follow up. Some investigators have set out to assess whether routine clinic visits are of value by establishing what proportion of relapses are detected at routine visits compared with interval visits, that is, visits where a patient develops problems and attends for review in between scheduled clinic visits. Many of these have included metastatic disease in their analysis. This will make interpretation difficult as metastatic disease is not routinely looked for and so usually will present when the patient develops symptoms[Donnelly J et al., 2001;Hiramanek N, 2004].

Even if metastatic disease is excluded from the analysis then routine clinic visits look inefficient. For example, Dewar and colleagues report that only 1% of routine clinic visits detected a treatable relapse, compared with relapse of some type being detected at 100% of interval visits in their analysis[Dewar JA and Kerr GR, 1985]. The overall number of events detected at each type of clinic was similar. They go on to analyse survival, and state that survival was greater in those whose relapse was detected at a routine appointment. Of course, there were more patients with metastatic disease detected at interval visits, which will skew this analysis considerably.

The analysis by Dewar and Kerr underlines the difficulty of conducting this type of research. Firstly, the low rate of treatable events which occur during follow up for breast cancer will inevitably make routine clinics look like an inefficient way to detect relapse. The findings by Dewar and Kerr in this regard are confirmed in a meta-analysis conducted in 2004[de Bock GH et al., 2004]. 5045 patients were studied. 378 relapses were detected. 58% of locoregional relapses were detected during routine follow up, the rest being detected at

interval appointments. The authors quote Morris and colleagues who point out that the incidence of locoregional breast cancer relapses after mastectomy or wide local excision is around 10% over ten years[Morris AD et al., 1997]. This will result in around 6 in every hundred women being diagnosed over the ten year period with locoregional relapse at routine scheduled appointments if the figures which form the meta-analysis are generally applicable. Furthermore, of the 58% who were diagnosed at routine clinic appointments, roughly one third had symptoms prior to their clinic appointment, but did not present early as they already had a clinic appointment arranged which they felt they should wait for. Thus, routine clinic appointments will, over ten years of follow up, detect only four asymptomatic cancers from a cohort of 100 women, but may lead to delayed diagnosis in 2 patients.

While routine clinics are an inefficient way to detect relapse, the analysis by Dewar and Kerr states that patients have a better outcome if their relapse is detected at a routine clinic visit. There are two reasons for this. Patients will usually present with symptoms of metastatic relapse between visits[Donnelly J et al., 2001;Dewar JA and Kerr GR, 1985;Hiramanek N, 2004]. If we include metastatic disease in an analysis of survival by type of clinic where relapse is detected, as was done in the analysis by Dewar and Kerr and the analyses of numerous other authors around this time[Scanlon EF et al., 1980;Valagussa P et al., 1981;Ormiston MC et al., 1985;Ojeda MB et al., 1987;Mansi JL et al., 1988;Muss HB et al., 1988], this will result in outcome appearing to be better for patients whose relapse is detected at routine clinics where less metastatic disease is detected. This artificial difference will be further enhanced by the fact that any relapse detected by routine mammography will be included in the routine clinic detection group. In other words, even removing distant relapse

from the analysis of outcome will not give a true reflection of the value of routine clinic visits.

This highlights the fact that these early studies were somewhat missing the point. The type of clinic at which relapse is detected is largely irrelevant. Even if all relapses are detected at routine clinics, if the patients who do relapse come to those clinics with symptoms of relapse or have palpable relapse detected by mammography then clinical examination, and by extension routine clinic visits are unnecessary. The only way to establish how valuable clinical examination is in the detection of relapse, and so whether routine clinics are necessary, is to analyse the proportion of relapses detected by routine clinical examination compared with mammography or symptoms and assess the impact, if any, that method of detection has on outcome.

Numerous groups have approached the issue in this manner[Churn M and Kelly V, 2001;Grogan M et al., 2002;Grinfeld E et al., 1996;Hussain ST et al., 1995;Jack WJL et al., 1998;Lees A et al., 1997;Mahoney L, 1986;Rutgers EJT et al., 1991;Snee M, 1996;Tate JT et al., 1989], but there have still been problems with these studies. In particular, short duration of follow up and inconsistency between the studies with regards to what types of relapse was included has led to widely varying results. Studies which are short in duration and which do not include new contralateral disease may have very different results than those with longer median follow up and where contralateral new breast cancers are included in the analysis. Only one study has made any attempt to establish what influence the method of detection has

on overall outcome, and in that study there were only 5 patients with locoregional relapse[Snee M, 1996].

The need for routine clinical examination in the follow up of breast cancer remains unclear. There is disagreement over the pattern of relapse, with some authors suggesting incidence peaks between three and five years[The ATAC Trialists' Group: 2005;Saphner T et al., 1996;Hussain ST et al., 1995], and other authors showing that potentially treatable relapse may in fact become more frequent with time[Freedman GM et al., 2005]. If the latter is the case, more frequent clinical examination in the first few years after treatment probably does nothing more than reduce the efficiency of routine clinics without actually improving patient outcome. This cannot be stated definitively as the issue of whether routine clinical examination has any impact on survival has been left unanswered in the literature. No-one has published survival statistics for patients with relapse which has taken into account the method by which relapse was detected.

Which components of follow up are useful and how frequently they should be conducted is not well dealt with in the literature. Most studies are retrospective, usually with short duration or small numbers, not including all types of potentially treatable relapse and confused by the inclusion of metastatic relapse. Many of the studies were conducted more than ten years ago. Since the publication of these studies, treatment has improved, risk of relapse has fallen, pattern of relapse may have changed as a result of better operative and adjuvant treatments and mammographic technology has moved on. In order to fully examine the contribution of the various components of follow up, either a retrospective analysis of a

large cohort with follow up to ten years or more or a randomised trial of similar size and duration is required. It would be important to fully explore the literature in a systematic way in order to establish whether either of these things has been done.

Additional aims of follow up

While early detection of relapse has long been held as a key aim of follow up, it is important not to forget that there are additional reasons for providing follow up after breast cancer. The National Institute for Clinical Excellence in England and Wales define these additional aims of follow up as being to detect and treat the adverse effects of therapy, particularly lymphoedema, and to provide psychological support[National Institute for Clinical Excellence, 2002].

There follows a discussion of the major adverse effects of breast cancer and its treatment which occur in women, both physical and psychological. In particular, the effectiveness of routine clinics at detecting and dealing with these effects is explored.

Adverse Effects of Treatment - Lymphoedema

Treatment for breast cancer will usually be surgical in the first instance, with or without neo-adjuvant chemotherapy, often followed by adjuvant chemotherapy and/or radiotherapy. During this period, women are in regular contact with clinicians involved in their treatment. Any problems which develop related to these treatments at this stage should be recognised and dealt with at the time, and there are very few problems which are likely to develop in a delayed manner. The exception to this would be lymphoedema, and in fact lymphoedema

alone is mentioned in the current NICE guidelines as being a side effect of treatment which should specifically be monitored for[National Institute for Clinical Excellence, 2002].

Post-operative arm swelling in the immediate period after axillary surgery is very common. This early swelling is felt to be due to disturbance to the lymphatic drainage and usually settles over time[Mortimer PS, 1998]. Lymphoedema may develop after this.

The incidence of lymphoedema after treatment for breast cancer varies greatly between studies. Historically, authors have reported that anywhere between 6.7% and 62.5% of patients suffer from lymphoedema[Pain SJ and Purushotham AD, 2000]. There is little doubt that decreasingly destructive primary breast surgery, coupled with advances in axillary procedures such as not performing radiotherapy after axillary dissection and the recent introduction of sentinel lymph node procedures, has reduced the incidence of this problem. Yet there is still wide variation in the reported incidence of lymphoedema, with two recent reviews reporting between 2% and 57.5% incidence in more modern studies in one review [Pain SJ and Purushotham AD, 2000] and between 6 and 30% in another [Petrek JA and Heelan MC, 1998].

This wide variation in incidence has mainly been as a result of methodological differences in the measurement of lymphoedema between studies, or of relying on subjective reports of arm swelling either from investigators or from patients. The difficulty of using a subjective impression to define the presence of lymphoedema is that patient impression of swelling is often not managed by a measurable increase in arm volume. 24% of patients in one study

reported feeling a swollen arm, only 11% had lymphoedema on objective measurement[Schunemann H and Willich N, 1998]. Gaining accurate figures has also been hampered by the fact that this condition can occur at some time after surgery, with examples being reported up to twenty years after surgery in some cases[Pain SJ and Purushotham AD, 2000].

The more consistent and objective assessments of incidence suggest that around one quarter of patients will develop lymphoedema at some stage using a definition of an increase in arm volume by 200ml from pre-operative values[Schunemann H and Willich N, 1998;Mortimer PS et al., 1996]. 75 % of these patients develop problems within the first year after surgery[Mozes M et al., 1982].

While the incidence is still uncertain, the psychological impact of the condition is not. Numerous authors have reported significant psychological problems developing as a result of lymphoedema[Maunsell E et al., 1993;Tobin MB et al., 1993;Fialka-Moser V et al., 2003]. Patients with lymphoedema experience more functional impairment, psychosocial maladjustment and increased psychological morbidity when compared with those experiencing no arm swelling in matched control studies[Tobin MB et al., 1993]. Generally, lymphoedema was considered more unsightly and distressing than the mastectomy scar by patients in Tobin and colleagues study. In fact, lymphoedema can be so distressing to individuals suffering this condition that one group described it as the most distressing and unpleasant sequelae of breast cancer[Fialka-Moser V et al., 2003].

There are numerous reasons why lymphoedema would cause major psychological problems. Many patients state simply that they have more problems hiding arm oedema than the mastectomy site, particularly with the current range of prostheses available for mastectomy patients, and that they find pressure garments to treat lymphoedema both unsightly and uncomfortable[Tobin MB et al., 1993]. They report being more distressed by the appearance of lymphoedema because, due to being less able to hide it, it has more impact on their body image to themselves and others.

There are also greater problems with sexuality in women who develop lymphoedema[Passik SD and Donald MV, 1998;Passik SD et al., 1995]. Women report feeling more unattractive as a result of arm swelling and are more uncomfortable with intimate relationships as a result. This can lead to increased social isolation, which is an independent risk factor for psychological distress[Bloom JR, 1982;Bloom JR and Spiegel D, 1984;Bloom JR et al., 2001].

Still others have problems relating to function. The lymphoedematous arm has been reported to feel heavy and less mobile[Passik SD et al., 1995;Hardy JR and Baum M, 1991], and that this leads to impaired ability to perform tasks such as dressing[Hardy JR and Baum M, 1991;Casley-Smith JR, 1985] or even going out to work[Tobin MB et al., 1993;Passik SD et al., 1995;Hardy JR and Baum M, 1991]. This in turn leads to feelings of increased social isolation, which again leads to more psychological problems[Bloom JR, 1982;Bloom JR and Spiegel D, 1984;Bloom JR et al., 2001]. Between 37.5 and 48% of women will complain of disabling pain[Haagensen DC, 1971].

Many patients see the development of lymphoedema as a secondary blow[Passik SD et al., 1995]. They have survived breast cancer, but the development of lymphoedema can bring back the emotions that they experienced at the time of diagnosis and act as a visual reminder of their malignancy.

The development of lymphoedema can give rise to very understandable concerns about relapse. In some 15% of patients, development of lymphoedema after the first year will represent axillary relapse of their breast cancer[Passik SD et al., 1995]. It is understandable that those worried about this possibility will have a marked rise in their anxiety levels until this possibility is ruled out.

Women who develop lymphoedema will require very specific physical and psychological help. They have physical impairment which limits what they can do, both in work and in leisure. Partly as a result of this physical limitation, and partly as a result of the cosmetic impact of lymphoedema, lymphoedema appears to be an independent risk factor for psychological problems above and beyond the diagnosis of breast cancer.

The value of routine clinical follow up for detection of lymphoedema is limited. The incidence of this condition varies from study to study as there is no standard way of measuring it and there is little correlation between a patient's perception of arm swelling and any objective measurement. As a result, there is no consistent way to diagnose this condition which will encompass all patients with troublesome symptoms and yet avoid diagnosing the

condition in patients with no symptoms or concerns. This is particularly important as there is no very good treatment for the condition, and patients' quality of life scores actually fall when treatment is initiated, as women realised that the symptom will require lifelong management[Howell D and Watson M, 2005]. Rather than have patients return to clinic routinely and have lymphoedema diagnosed with some objective measure, it may be better simply to allow open access back to clinics for people who develop troublesome symptoms, and to warn women of the possibility that these symptoms may occur. Treatment could then be offered only to women who have noticed this problem themselves and are troubled by it.

If we are to routinely screen for lymphoedema, this will require long term follow up, as this condition can arise long after initial treatment, as mentioned above. The yield of such follow up in detecting lymphoedema is likely to fall as new techniques for sampling the axilla, such as sentinel lymph node biopsy, are associated with much lower rates of lymphoedema[Schrenk P et al., 2000;Sener SF et al., 2001;Wilke LG et al., 2006].

Adverse Effects of Treatment – Non-Lymphoedematous arm dysfunction

Patients can suffer a number of arm symptoms after treatment for breast cancer unrelated to the development of lymphoedema. These symptoms include stiffness, weakness, limitation of movement, pain and numbness. It has been suggested that up to 80% of women can develop one or more of these after treatment for breast cancer, depending on the treatment received[Yap KPL et al., 2003;Maunsell E et al., 1993]. Axillary lymph node dissection and post operative radiotherapy, either to the axilla or to the breast, are risk factors for developing such problems[Yap KPL et al., 2003;Kwan W et al., 2002;Maunsell E et al., 1993], although

almost 11% of patients in one study who received only a lumpectomy with no radiation or axillary surgery experienced post operative arm symptoms[Yap KPL et al., 2003].

Like lymphoedema, arm dysfunction can cause both physical and psychological problems in patients. In fact, along with fear of recurrence, shoulder/arm morbidity is the most important source of distress after breast cancer treatment[Kuehn T et al., 2000]. Likelihood of suffering psychological morbidity at three months after treatment is directly related to the prevalence of arm symptoms, with more symptoms leading to a higher likelihood of morbidity[Maunsell E et al., 1993].

It is important that such arm symptoms are detected if they can have such an impact on function and on psychological well being. Arm symptoms develop in the immediate post-operative (or post treatment, if related to radiotherapy) period. Many of these symptoms can be expected to settle with time, perhaps with the intervention of physiotherapists and self directed exercises. In fact, spontaneous improvements can occur for two or more years after treatment[Segestrom K et al., 1991].. Yet these symptoms can be remarkably persistent in some patients, with almost 50% of respondents to a questionnaire reporting ongoing arm problems between 2 and 7 years after treatment[Kwan W et al., 2002]. It will be important in any follow up programme to ensure specific questions relating to arm symptoms are included, and the focus must not be simply on lymphoedema.

Adverse Effects of Treatment – Hormone Therapy

Around two thirds of breast cancers are oestrogen receptor positive, and as such are amenable to adjuvant treatment with hormonal therapy. This has, until recently, meant treatment with tamoxifen, which has been shown to reduce relapse by 47% and death by 26%[Early Breast Cancer Trialists' Collaborative Group, 1998]. More recently, a newer class of drugs, the aromatase inhibitors, has emerged as an alternative to tamoxifen for adjuvant therapy in post-menopausal women and the indications for this type of therapy are still being established.

Traditional hormonal therapy has involved treatment with tamoxifen for five years from completion of adjuvant chemo-radiotherapy. The aromatase inhibitors, in contrast, have been utilised in a variety of different ways in the various large trials. For example, the Anastrozole alone or in combination with tamoxifen trial (ATAC) has compared the use of anastrozole alone, tamoxifen alone or both drugs in combination for five years[Baum M et al., 2002]. Switch therapy (a change from tamoxifen to exemestane after two to three years of tamoxifen treatment) has been evaluated in the IES trial[Coombes RC et al., 2004], and extended adjuvant therapy (a further 2 years of letrozole therapy after completing five years of tamoxifen in post-menopausal women) has been evaluated in the MA-17 trial[Goss PE et al., 2003]. All of these have revealed reduced relapse rates and improved survival in women taking aromatase inhibitors compared to tamoxifen alone, so that the American Society of Clinical Oncology now recommends that post menopausal women with oestrogen receptor positive breast cancers should receive an aromatase inhibitor as adjuvant therapy, either alone or in some combination with tamoxifen[Winer EP et al., 2005]. As mentioned, the indications for these drugs are still to be established.

While all of these medications result in improved local control and outcome, they are not without side effects. 36% of patients develop some endometrial pathology while on tamoxifen[Cohen I, 2004]. Usually this will become evident as a result of inter-menstrual or post-menopausal bleeding. By far the most common complaints are simply endometrial hyperplasia or endometrial polyps, but tamoxifen has been shown to increase in the relative risk of developing uterine cancer by 6.4 times[Fornander T et al., 1989]. Tamoxifen can also cause a large range of persistent side effects, the most common and troubling of these being hot flushes, night sweats and other vasomotor symptoms, as well as gynaecological symptoms such as vaginal dryness[Couzi R et al., 1995;Day R et al., 1999]. The gynaecological symptoms in particular can result in significant changes in sexual functioning[Day R et al., 1999].

While the aromatase inhibitors have been reported to be better tolerated than tamoxifen in several large trials[Goss PE et al., 2003;Baum M et al., 2002;Coombes RC et al., 2004], tamoxifen and the aromatase inhibitors have similar side effect profiles and some authors have suggested that aromatase inhibitors are in fact no better tolerated than tamoxifen in the real world[Garreau JR et al., 2006]. Certainly, the aromatase inhibitors have been reported to cause significantly more problems with arthralgia[Garreau JR et al., 2006] and appear to have a more profound effect on bone health than does tamoxifen. In the ATAC trial, fractures were significantly less common in women on tamoxifen than on the aromatase inhibitors[Baum M et al., 2003]. This was as a result of increased bone turnover and reduced

bone density in patients on anastrozole compared with those on tamoxifen, who actually had an improvement in bone density[Eastell R et al., 2006].

From the point of view of follow up guidelines, there are a number of issues with regards to adjuvant hormonal therapy. The current NICE guidelines state simply that general practitioners should be responsible for looking after women on long term hormonal therapy and should stop treatment at five years[National Institute for Clinical Excellence, 2002]. This is an oversimplification. Guidelines as to what treatment to give and when are far from established and it is perhaps unfair to ask general practitioners to be responsible for keeping abreast of and interpreting the results of the changing recommendations. For this reason, keeping patients who are receiving adjuvant hormonal therapy under some type of review will become increasingly important, if only to change hormonal therapies in light of emerging new evidence and guidelines.

The side effects of these adjuvant therapies will also require some consideration when establishing follow up protocols. NICE states that follow up should detect and treat the adverse effects of therapy and should also detect and provide assistance with dealing with psychological concerns. Clearly, adjuvant hormonal therapy produces significant side effects. These have a physical impact, certainly, but they have also been shown to impact on a patients' psychological well being. Health related quality of life scores worsen from baseline to three months after starting a new hormonal therapy[Fallowfield L et al., 2004] and then stabilise. Historically there has been little work done to establish the prevalence or impact of side effects from hormonal therapy in women in routine breast cancer follow

up[Hopwood P, 1998]. Clinicians often fail to assess the side effect profile in their patients, and studies have revealed a disparity between doctors', nurses' and patients' perceptions of the impact of endocrine treatment[Leonard RCF et al., 1996;Denton S and on behalf of the EONS Advanced Working Group for Living With Advanced Breast Cancer Hormone Treatment, 1998]. The current approach of routine clinic visits may not be the most effective way of detecting these problems in patients, as these issues are often not asked about. It may be of more benefit to patients to simply warn them of potential side effects and provide open access to clinics. An alternative would be to focus more on these symptoms and their psychological impact than has been done previously. This could be done in a structured way, such as by questionnaire. This could even be done remotely, and only those with problems returned to clinic.

The final issues arising from the use of adjuvant hormonal therapy are bone health and endometrial pathology. Currently, screening for endometrial pathology is only done in those who develop abnormal bleeding. Therefore endometrial pathology is not a reason for bringing patients back to clinic routinely, but abnormal bleeding should certainly be considered if devising questionnaires to screen for problems remotely. In terms of bone health, it has been recommended that patients at low risk of osteoporosis who are being treated with aromatase inhibitors should be screened for deterioration in bone density at regular intervals by bone densitometry scans[Chien AJ and Goss PE, 2006]. Again, it will be important that any follow up system implemented can co-ordinate these and that an expert is available to interpret the result and institute any necessary treatment.

Quality Of Life

The final aim of follow up care in breast cancer, as defined by the current NICE guidelines, is detection and treatment of psychological distress[National Institute for Clinical Excellence, 2002]. There is some overlap between psychological distress and the adverse effects of treatment, as mentioned above. Lymphoedema, vasomotor and gynaecological side effects of adjuvant hormone treatment can all result in reduced quality of life and psychological distress in women after treatment for breast cancer. There is little doubt that the diagnosis and treatment of breast cancer per se can cause tremendous psychological distress. Feelings of isolation and difficulties with adapting to altered body image are all compounded by the fact that many women report subsequent problems with sexuality and intimacy. It has been reported that up to one third of breast cancer patients treated with surgery will develop anxiety or depression[Maguire P, 1994].

It is difficult to establish how effective clinicians are at detecting psychological distress during routine clinic visits. Morris and colleagues reported that women find breast clinics reassuring[Morris S et al., 1992]. A questionnaire was given to a cohort of 285 women after their attendance at breast clinic, subsequent to their being told they had no recurrent disease. It is understandable that the majority of women would find this experience reassuring, as 81% did. Over three quarters expressed a preference to continue coming to clinic for follow up when offered follow up by their general practitioner as an alternative, and 85% expressed a preference to continue coming back to clinic for follow up rather than just report if they had problems.

Others have approached the subject with a similar methodology, and obtained broadly similar results. In 127 patients surveyed after various cancer procedures, psychological stress was significantly higher among cancer patients one month before clinic and on the day of clinic visits than two weeks after[Kiebert GM et al., 1993]. Most psychological distress was suffered by breast cancer sufferers. In spite of this, most women expressed a preference for continuing to come to clinic, saying that it did provide reassurance. The possibility that the impending clinic visit itself was contributing to the anxiety was not addressed by that study.

Other investigators have replicated these findings of increased anxiety in the days or weeks before attending for appointments, and made some attempt to establish the cause. Pennery and Mallet conducted semi-structured interviews, and revealed that women had increased levels of anxiety before coming to clinic as they were concerned that they may be diagnosed with relapse[Pennery E and Mallet J, 2000].

These studies have all addressed the question of whether women derive some reassurance from a negative check up visit. Women find the idea of coming to clinic distressing. They suffer anxiety in the run up to attendance as there is concern that they may be diagnosed with relapse while at clinic. As a result, it has been reported that 71% of breast cancer patients suffer some anxiety during routine clinical follow up[Paradiso A et al., 1995]. The research by Kiebert and colleagues and Morris and colleagues simply highlights that women may be prepared to accept this increased anxiety attendant on clinic visits in exchange for the perceived reassurance they get from the visit itself. In fact, this was stated explicitly by several responders in the study by Pennery and Mallet[Pennery E and Mallet J, 2000], and is

implicit in the fact that 85% of women in the study by Morris and colleagues would like to continue coming back to clinic due to the reassurance they receive[Morris S et al., 1992].

These studies have established that women generally find it reassuring to attend routine breast clinics and be told they do not have recurrent disease. It has not been established in a large contemporary cohort that routine clinical examination is of value in the detection of recurrent disease. It may be that these women are receiving false reassurance. Because there has been no study in the recent literature which has clearly established the benefit of clinical examination in terms of relapse detection, it has been impossible until now to explore the question of whether women would still want to come to clinic for a clinical examination in the event that such examination was shown to be of no benefit. It has not been established whether the provision of routine clinical examination reduces women's willingness to perform self-examination. If this was the case, this could lead to a delay in detection of relapse in some women, which could have an impact on the survival of women in whom relapse is detected clinically.

Routine clinic visits have a generally reassuring effect on most women with regards to the detection of relapse. There will be a proportion of women who develop generalised anxiety and depression after diagnosis and treatment for breast cancer which is unrelated to the anxiety attendant on clinic visits. As mentioned above, up to one third of women are thought to suffer from generalised anxiety or depressive disorder at some point after diagnosis and treatment of breast cancer[Maguire P, 1994]. Most studies have focussed on the first two years after treatment, but other authors have reported that anxiety and depression will persist

in some patients for up to ten years after treatment[Hodgkinson K et al., 2007]. In addition, a proportion of women will develop side effects of treatment such as lymphoedema after axillary surgery or gynaecological problems while on adjuvant hormonal therapy and these side effects will impact on quality of life, possibly leading to symptoms of anxiety and depression. What is far more important than whether women find clinics generally reassuring is whether routine clinic visits are necessary for the detection and treatment of these problems.

Very little research exists which has specifically addressed this question. The evidence which does exist suggests that routine clinic visits are probably not the ideal setting for detecting psychological problems. Patients are often embarrassed or reluctant to report depressive symptoms[Valente SM et al., 1994]. In one study, 79% of women said they did not feel comfortable raising emotional/psychological concerns at a routine appointment[Pennerly E and Mallet J, 2000]. Reasons given by patients for this reluctance to express concerns were that the clinics were too physical, that the clinics were not oriented towards emotional needs and that they were conscious of taking up too much time. The concerns expressed by women in these studies are mirrored in other reports in the literature. It has been shown that women in cancer follow up often undergo a strictly medical consultation, with no time to discuss their psychological and social situation[Lampic C et al., 1994]. Others have highlighted that women often feel rushed during follow up visits for breast cancer[Adewuyi-Dalton R et al., 1998].

Follow up clinics are often rushed and concentrate on the detection of relapse. Even allowing for that, it may be that clinicians in the routine clinic setting are not ideally placed to detect psychological problems in their patients. Passik and colleagues point out that there has been little research into physicians' ability to recognise depression in cancer patients[Passik SD et al., 1998]. Junior doctors and nurses only recognised half of the depressed patients on a medical oncology service in one study[Hardman A et al., 1989]. The junior doctors recognised the majority of patients with significant anxiety, but only because they assumed that all patients were anxious and so greatly over diagnosed anxiety in this study. This over diagnosis is a result of the short term anxiety which we know is attendant on visiting clinic and has implications in the breast clinic, where the majority of patients report feeling anxious at attending clinics[Paradiso A et al., 1995]. Clinicians get only a short period of time with patients, and it will be difficult in this brief period to differentiate those with short term anxiety related to the clinic visit from those with more global psychological concerns.

More recently, work by Passik and colleagues suggested that oncologists frequently assessed their patients' levels of depressive symptoms inaccurately, and were especially ineffective at recognizing moderate or severe depression[Passik SD et al., 1998]. This was particularly concerning to the authors as the study was undertaken with the prior knowledge of the clinicians involved, who were aware that they would be expected to rate patients' level of depression immediately after the clinic visit. Thus the under-recognition of moderate to severe depression occurred when vigilance should have been somewhat higher than normal,

suggesting that the pick up rate for psychological problems at a standard follow up clinic will be particularly poor.

The question of how much psychological benefit women get from attending clinic is therefore a very complicated one. On the one hand, women feel reassured by attending clinic as they are told that they do not have relapse. This reassurance has not been demonstrated to be based on actual improved outcome related to clinic attendance and there is no great evidence that clinics are protective. On the other hand, women report feeling anxious about coming to clinic, and the relatively few women who have underlying psychological issues are probably not well served by the current clinic system, as the limited evidence which exists suggests that clinicians under-diagnose psychological problems.

The limited evidence which exists suggests that clinicians working in the routine clinic are not ideally placed to detect psychological problems. It may make sense to offer some alternative screening method, either at the clinic or remotely by postal or telephone questionnaire, which can separate the women who have concerns from those who don't.

Whether routine clinic attendance improves quality of life as a result of the reassurance provided is an important issue to clarify. It is unlikely that, particularly given the recent trend towards annual follow up visits after completion of treatment, this reassurance will last long enough to be of great benefit in terms of overall quality of life to many patients. The best way to answer this question is to compare quality of life in patients attending clinic to

those followed up by some alternative method. This could really only be done in the context of a randomised controlled trial of traditional follow up with a novel alternative.

Alternative follow up

While it has not yet been established whether routine clinic visits are of value, it is clear that women would benefit from better screening in terms of detecting psychological problems. This could be undertaken in addition to routine follow up if clinical examination was shown to be necessary. If, on the other hand, such examination was shown not to affect outcome, then it may be possible to do this screening remotely thereby establishing which patients need to attend clinic for extra psychological support. This would in addition reduce the number of patients coming to clinic thereby freeing up time to address the psychological problems detected.

There are a number of screening tools for detecting anxiety and depression. Many of the more general tools have been shown to be insensitive and non-specific when used on breast cancer patients[Hall A et al., 1999]. Questionnaires do exist which have been designed specifically to measure quality of life after breast cancer. These include the Functional Assessment of Cancer Therapy-Breast (FACT-B) quality of life instrument and the European Organization for Research and Treatment of Cancer Breast Cancer-specific quality of life questionnaire (EORTC-QLQ-C30/BR23)[Brady MJ et al., 1997;Sprangers MAG et al., 1996].

The FACT-B questionnaire was developed to provide a robust and reproducible way of measuring quality of life in patients during and after treatment for breast cancer[Brady MJ et al., 1997]. It is a 44 item self report questionnaire designed to measure general quality of life issues as well as more specific issues related to breast cancer, and it has been extensively used in both research and clinical practice. Since it was first introduced, additional subscales have been developed to measure arm and endocrine symptoms in patients who have had treatment for breast cancer.

The arm subscale consists of 4 additional items used in conjunction with the FACT-B and has been shown to be sensitive to deterioration in arm function associated with the development of lymphoedema, as well as being sensitive to improvement in function associated with effective treatment[Coster S et al., 2001].

The endocrine subscale consists of 18 items designed to detect endocrine symptoms related to adjuvant treatment with agents such as tamoxifen or the aromatase inhibitors. It is also useful in detecting symptoms in younger women who have been rendered menopausal either as a deliberate act during ovarian ablation or oophorectomy or as a consequence of premature menopause secondary to chemotherapy[Fallowfield LJ et al., 1999].

Both subscales have been developed to work in conjunction with FACT-B to provide a well integrated, simple to use questionnaire using a 5 point Lickert-type response scale. FACT-B with both subscales provides a comprehensive assessment of the physical and psychological health of women after a diagnosis of breast cancer, with particular emphasis on the areas

considered particularly important by NICE in their most recent guidelines. As such, FACT-B and the arm and endocrine subscales are ideally suited as a screening tool for physical and psychological problems after treatment for breast cancer.

The administration of such questionnaires could prove problematic. Completing them face to face with either clinicians or nursing staff would be time consuming, and would go no way to reducing the number of people coming to clinic. Moreover, it would still require someone to review all of these questionnaires and score them appropriately, which would also be time consuming. The same could be said of postal questionnaires.

Some authors have had success with telephone questionnaires. Wasson and colleagues reported on the use of telephone care as a substitute for routine clinic follow up in the setting of a primary care clinic[Wasson J et al., 1992]. A study group had the frequency of their follow up visits reduced by half, and had a telephone consultation in between clinic visits. A control group had normal scheduled clinic visits.

In the study group, there were fewer clinic visits, both scheduled and unscheduled, fewer admissions to hospital, shorter duration of stay when they were admitted and less medication use than the control group. Estimated total expenditure was 28% less for the study group. In those with poorer overall health at baseline, there was even a reduction in mortality in the study group as compared with controls. It was concluded that telephone follow up was not only safe, but could actually introduce improved outcome at a lower overall cost. Others have reported similar success with telephone follow up in the field of oncology. Telephone

follow-up by trained nurse practitioners resulted in good follow up provision both in terms of detecting medical problems and providing support with a reduction in outpatient workload of 30% in one study[James ND et al., 1994].

Telephone follow up of patients using a validated questionnaire to assess quality of life, as well as to detect symptoms which may either be related to relapse or amenable to treatment, certainly holds promise. This could even be automated, thereby reducing the work involved in scoring the questionnaires. This would allow the questionnaire to be done remotely with minimal involvement from clinical staff. It could then be used either in addition to routine clinical examination, or as an alternative method of establishing who needs to come to clinic.

Automated telephone systems have been utilised in other fields. One systematic review of the literature from 2003 highlights some of the uses to which such systems have been put. For example, attendance rates at clinic can be improved using automated phone systems, several studies have shown increased rates of childhood immunisations and such systems have been shown to improve various parameters in diabetes and hypertension management[Biem HJ et al., 2003]. The authors of that review point out that automated telephone systems have diverse potential applications, but conclude that the evidence is limited and that further studies are required. They do concede that, despite the relatively impersonal nature of such systems, the reliability and availability of such systems make them ideal tools for use in follow up of patients or for generating reminders.

While the authors of the systematic review described above see the relatively impersonal nature of these systems as a drawback, it is this impersonal aspect which has appealed in other settings. Millard and Carver utilised interactive voice recognition software to administer the SF-12 health status survey to a group of patients with low back pain. They compared the results with responses from another group to administration by a nurse over the telephone. While both methods produced similar results in the physical component scores, they differed in the mental component scores, with patients who used the voice recognition software acknowledging significantly greater mental interference, greater emotional concerns and poorer overall mood and health[Millard RW and Carver JR, 1999]. The authors conclude that, because automation eliminates the need to talk to a real person, patients are likely to be giving a more accurate representation of their feelings.

This increased reporting of emotional and/or sensitive issues is not an isolated finding. Computerised reporting systems have been widely used in the field of psychiatry, and in particular among patients with history of drug abuse. Several authors have reported that subjects may provide more honest responses, particularly with respect to substance abuse, when using interactive voice recognition systems compared with when questioned in a live interview setting, even by telephone[Turner CF et al., 1998;Kobak KA et al., 1997].

This aspect of automated telephone follow up is particularly attractive in the field of breast cancer. We have already described that, in the outpatient setting, patients are often embarrassed or reluctant to report depressive symptoms[Valente SM et al., 1994], and that a high proportion of women are uncomfortable raising emotional or psychological concerns at

a routine appointment[Pennery E and Mallet J, 2000]. It may be that introduction of an automated telephone screening service will increase diagnosis of psychological problems in women in follow up for breast cancer.

Summary

The aims of follow up are to detect relapse, provide psychological support and detect and ameliorate the side effects of treatment, in particular endocrine side effects and lymphoedema[National Institute for Clinical Excellence, 2002]. How well these aims are met by current routine follow up is unclear. Routine clinical examination appears to be an inefficient way of detecting relapse, but it is not certain whether relapse detected this way has a better outcome than symptomatic relapse, or relapse detected by mammography. The proportion of relapses detected by each of these three methods in a modern cohort has yet to be established. These questions would best be answered in the context of a randomised controlled trial. Comparing traditional follow up to a novel alternative where clinical examination was not performed or was performed in an alternative setting would establish definitively whether routine clinical examination by a breast surgeon or oncologist was necessary. If the information is unavailable from such trials, some information could be derived from retrospective analyses of how relapse presents and whether mode of presentation impacts on survival.

Whether clinics provide adequate psychological support can really only be answered in the context of a randomised controlled trial. Clinics are reassuring when patients are told they do not have recurrence. On the other hand, there is good evidence that attendance at clinic is

distressing to some patients. The impact of clinic visits on patients' overall psychological well being can only be established by objective comparisons of quality of life in women attending breast clinics and those being followed up in some novel manner.

Finally, while an important component of follow up is the detection and amelioration of side effects of treatment, a large discrepancy exists between patients' and clinicians' perceptions of side effects. A more robust and objective attempt must be made at detecting such problems among women. Perhaps the best way to provide follow up would be simply to ask patients to report when they develop side effects, and options exist for remotely screening women for such problems, such as by using structured questionnaire.

Aims of the Thesis

The overall aims of this thesis are:

- To establish the evidence base for the current recommendations for follow up in patients treated with curative intent for primary breast cancer.
- To evaluate patient opinion on follow up and develop an alternative follow up approach.

In order to achieve these aims, this thesis is split into 2 sections. In the first section, two systematic reviews and two retrospective analyses have been conducted to address the first of these aims. The value of routine follow up in meeting all of the aims of follow up as defined by NICE will be investigated. This will be achieved through establishing:

- The pattern of potentially treatable relapse, including timing of that relapse
- The contribution of routine clinic visits to relapse detection
- The value of routine clinic visits in meeting the additional aims of follow up.

Chapter 1 contains a systematic review of the literature analysing randomised controlled trials of alternative methods of follow up in breast cancer. This will establish the necessity of routine clinic visits for detection of relapse, and also highlight the value of routine clinic visits in meeting the additional aims of follow up. In Chapters 2 and 3, the pattern of potentially treatable relapse and the contribution of routine clinical examination to detection

of such relapse will be explored, firstly by analysis of local cohorts, and then by systematic review of the literature.

Section 2 addresses the second major aim of this thesis, which was to explore patients' expectations for follow up and to evaluate a remote, automated telephone based method of follow up for patients after breast cancer.

Patients' expectations for follow up will be investigated in a questionnaire based study in chapter 4, and a prospective cohort study of an alternative follow up method will be presented in chapter 5.

Section 1

Chapter 1:

The impact of alternative follow up methods on relapse detection, survival and quality of life

A Systematic Review of the Literature

Summary of the chapter

The randomised controlled trial is the gold standard for determining the value of one intervention over another. In this chapter, we systematically review the literature for randomised controlled trials of alternative follow up methods which have been used in patients after treatment for breast cancer, and for trials of alternative frequency of follow up visit. In this way, high level evidence confirming the value of traditional follow up clinics, either in terms of improved survival or detection of psychological problems or problems with symptoms of treatment, will be sought, along with evidence to recommend an ideal frequency of visit.

Introduction

Regular follow up of women after breast cancer is common practice. The aims of follow up have been discussed in detail above and include early detection of potentially treatable relapse, detection and amelioration of side effects of therapy and provision of psychological support in the aftermath of the diagnosis of breast cancer and subsequent treatment [National Institute for Clinical Excellence, 2002].

While it is generally believed that some form of follow up is required, it is not clear what form this should take. Multiple diagnostic interventions designed to detect metastatic disease are now no longer recommended[National Institute for Clinical Excellence, 2002;Khatcheressian JL et al., 2006;Grunfeld E et al., 2005]. These do not prolong survival and are considered to have a detrimental effect on quality of life[Del Turco MR et al., 1994;The GIVIO investigators, 1994;Palli D et al., 1999]. The only investigation now conducted routinely is mammography, yet the evidence base for this is limited. There is some indirect evidence that mammography may contribute to the earlier detection of contralateral disease[Mellink W et al., 1991;Samant RS et al., 2001], but there has been little work conducted on the efficacy of mammography in detecting locoregional recurrence in the conserved breast. One small British study concludes that the yield of mammography in this context is minimal[Donnelly J et al., 2001], and the Canadian Steering committee concedes that there is no strong evidence in the literature for the use of mammography in follow up[Grunfeld E et al., 2005]. The value of mammography in follow up after breast cancer will be explored in more detail in a subsequent section of this thesis.

In addition to mammography, most clinicians recommend that women attend routine clinic visits for history and clinical examination. Recommended frequency and duration of visits vary between guidelines. The American Society of Clinical Oncology (ASCO) recommends clinical examination three to six monthly for three years, six to twelve monthly for a further two years then annually[Khatcheressian JL et al., 2006]. In Canada, it is recommended that visits be tailored to the individual patient needs, but six monthly visits for clinical examination continued indefinitely is a considered a reasonable frequency and duration for

most patients[Grunfeld E et al., 2005]. In contrast, the National Institute for Clinical Excellence (NICE) in England advocates follow up for only two or three years, and gives no advice on how frequently visits should be conducted[National Institute for Clinical Excellence, 2002].

These marked variations in practice appear to stem from a lack of evidence that routine clinical examination is of benefit. The NICE guidelines cite only one study by Donnelly and colleagues[Donnelly J et al., 2001], a retrospective review of 643 patients with median follow up of 3 years and 11 months, as evidence for their recommendations for routine follow up.

A recent systematic review designed to investigate long term outcomes in the setting of routine clinical follow up failed to identify any studies comparing long term results (survival/disease free survival) in patients with clinically detected local relapse versus symptomatic presentation of local relapse[de Bock GH et al., 2004]. The authors were unable to comment on whether routine clinical examination during follow up conferred any survival benefit[de Bock GH et al., 2004]. They did report that routine clinics following the ASCO guidelines would detect only 4 asymptomatic relapses in every 100 women over ten years of follow up, and would lead to delayed diagnosis in 2 women who would notice problems but wait for their next clinic appointment[de Bock GH et al., 2004]. A recent simulation model has also suggested the survival benefit of follow up would be at best marginal[Jacobs HJM et al., 2001].

Most retrospective reviews of routine clinic visits have not been helpful in establishing either the best schedule for such visits, or indeed whether such visits are necessary[de Bock GH et al., 2004]. There has been wide variation in methodology among such studies, particularly in terms of what is counted as relapse. Despite this, the available evidence does suggest that the impact of routine follow up on survival will be at best marginal and may lead to delayed diagnosis of relapse. This will be discussed further in later chapters of this thesis.

Routine follow up should also be geared towards detection of psychological problems in women. Clinic visits themselves are not thought to be particularly useful at detecting psychological concerns and attendance at clinic itself is a stressful event for many. Women report increased levels of anxiety before coming to clinic as they are concerned that they may be diagnosed with relapse[Pennery E and Mallet J, 2000] and it has been reported that over 70% of breast cancer patients suffer distress during routine clinical follow up[Paradiso A et al., 1995]. Yet others suggest that breast clinic attendance is reassuring[Morris S et al., 1992].

It is unclear whether breast clinic attendance confers any survival advantage or contributes to improvement in quality of life. The optimal design to study these issues would be a randomised trial. Comparisons of different frequencies of visits to clinic or different durations of follow up would be of particular value in informing guidelines. Comparisons of traditional follow up with novel methods of care would help to establish whether traditional clinic visits were necessary.

In order to establish whether there is randomised evidence in the literature to suggest that routine clinic visits are either necessary or beneficial for the detection of potentially treatable relapse and/or psychological distress, a systematic review of the literature has been conducted.

Conducting systematic review

Systematic review of research evidence is an efficient and robust scientific method. In contrast with traditional narrative reviews, where the review is written by a recognised expert but without pre-defined methodology for identifying and bringing together all relevant studies, systematic review of the literature seeks to uncover all relevant evidence within the literature and compare the evidence in an unbiased way. Traditional narrative review has potential to introduce bias, particularly when the reviewer has pre-conceived opinions, whereas a systematic review aims to reduce the potential for introducing bias at every step in the process, from identifying relevant studies and selecting them for inclusion to collecting and combining their data, by using a predefined, explicit methodology. [Higgins JPT et al., 2006]. Studies should be sought regardless of their results.

Systematic review is an ideal way to establish the existing evidence base for routine clinical follow up after breast cancer. By searching multiple databases in a systematic way, the chance of uncovering published studies of breast cancer follow up will be maximised. By using a pre-defined methodology to select studies for inclusion and to analyse the results of these studies, potential sources of bias within the review itself will be minimised.

There follows a description of the methodology of conducting systematic review in general. The specific methodology used in this review will be presented in the methods section below, but the guidance described here was employed in full both in the review presented in this chapter and in the review presented in chapter 3 of this thesis.

Data Collection

Thorough data collection is of central importance in preparing a systematic review, yet there are challenges in finding all the relevant studies which exist. These challenges were highlighted in 1994 by Dickersin and colleagues who sought to establish the efficiency of searching the medical literature as part of the systematic review process [Dickersin K et al., 1994]. A systematic review of studies which had reported rates of article retrieval achieved by searching Medline was performed. The included studies varied in their methodology, but all reported the results of a comparison between a Medline search and a gold standard of all known studies conducted within an area of research. In some of the studies, the gold standard consisted of all studies within an area of medicine which had been published within journals cited in Medline, in the remaining studies, non-Medline indexed journals, books and proceedings of scientific meetings were included so that the gold standard was all studies known to exist anywhere in the literature.

The review concluded that a Medline search, even when conducted by a trained searcher, yielded on average only 51% (range 17 to 82%) of the studies known to exist when the gold standard was all articles known to exist in the literature [Dickersin K et al., 1994]. While there is a huge number of medical journals in print or available online (The United States

National Library of Medicine lists 10,767 titles indexed for online users alone), Medline indexes only a proportion of these; 4959 as of January 2006[United States National Library of Medicine, 2006]. Searching only within Medline will preclude finding articles in sources not indexed within Medline. If Medline is the only source searched, then rate of article retrieval will be lower than if multiple sources are searched. While multiple database searching will improve the yield, it is not the whole answer. While the sensitivity is higher if the gold standard consists only of known articles published in Medline indexed journals, it is nevertheless still a disappointing 63% (range 46-88%)[Dickersin K et al., 1994].

The review by Dickersin and colleagues highlights several important points. Firstly, it is necessary to use multiple sources when attempting to find studies for a systematic review. Each source of citations will include a different group of journals, and so multiple sources will maximise the number of journals searched. In addition, there is some overlap between electronic databases. For example, there is approximately 34% overlap between Medline and EMBASE, considered the European equivalent of Medline[Smith BJ et al., 1992]. Therefore, searching multiple databases will result not only in the search being expanded to cover more journals, but additionally, due to the overlap between databases, there will be increased chance of finding articles which are indexed in more than one.

The use of multiple search engines can contribute additional benefits as search engines do have a variety of differing features. Web of Sciences in particular is useful for conducting systematic reviews as it includes, along with each article indexed within the database, a list of articles which cite the source article. These citing articles can then be checked to establish

if they are eligible for inclusion in the review. Web of Science therefore provides a way of searching forwards in time from included or important articles.

Even when a thorough search was conducted of Medline, there was less than 100% retrieval of known studies cited within Medline. This is partly a result of inadequate indexing of studies[Dickersin K et al., 1994]. The emphasis traditionally was on developing MeSH terms (medical subject headings, a group of terms which are used to index scientific publications) which related to the subject matter, not the methodology. There has been a recent attempt to better classify studies according to their methodology, and more emphasis by editors of journals on ensuring that authors provide adequate description of the methods employed in studies in order that studies can be more accurately indexed. In addition, the indexing of previously published articles is being reviewed in order to better classify these older studies. This should lead to a higher sensitivity when searching online databases than would have been the case when Dickersin and colleagues published their report in 1994[Dickersin K et al., 1994].

Much work has been undertaken by researchers to develop highly sensitive search strings designed to identify certain types of studies, for example randomised controlled trials or systematic reviews, despite the limitations of indexing. These greatly facilitate searching for randomised controlled trials or pre-existing systematic reviews within an area of research, and are available in the Cochrane collaboration handbook[Higgins JPT et al., 2006]. These search strings have been made use of in designing the search strings used in this systematic review.

Medline searching has been shown to yield a fairly low median precision of 33% (precision being the proportion of articles which are examined for relevance compared to the number obtained which are relevant)[Dickersin K et al., 1994]. Sensitivity is highest at a precision of 35%, with reducing sensitivity as searchers attempt to improve precision. It is always necessary to strike a balance between comprehensiveness and precision when developing a search strategy. Increasing the comprehensiveness of a search entails reducing its precision and retrieving more non-relevant articles[Higgins JPT et al., 2006]. Particularly in non-expert hands, precision should be sacrificed for improved sensitivity[Higgins JPT et al., 2006]. Although this will result in searching through a larger number of irrelevant references, it will maximise retrieval of relevant articles. This is particularly the case if the search is conducted independently by more than one author, which is recommended by the Cochrane collaboration. This will not only reduce the chance of relevant studies being missed while searching through the titles of large numbers of irrelevant studies, but it also ensures that tasks such as selection of studies for inclusion and data extraction can be performed by at least two people independently, which will further increase the chance that errors are detected and that all relevant studies are identified[Higgins JPT et al., 2006].

There is advice available from several sources regarding the actual construction of a search strategy. In particular, the Cochrane collaboration offers guidance on the process drawn from the work of a number of experts[Higgins JPT et al., 2006]. As mentioned above, sensitive search strategies designed to detect specific types of articles, such as randomised controlled trials, exist and are reproduced in the Cochrane collaboration handbook[Higgins

JPT et al., 2006]. Additional search strategies will be required for two reasons. Firstly, the search strategies for randomised controlled trials will locate those randomised trials within the database being used, but the results will need to be further narrowed down to the subject area of interest. Secondly, as mentioned above, studies have often been indexed inaccurately in terms of their study design in the past. While these pre-prepared search strategies have been designed to be maximally sensitive to detect randomised controlled trials, they cannot be guaranteed to yield all of the studies which exist. For these reasons, it is important to develop a search strategy based on subject area in addition.

Authors of published trials can prove to be a valuable source of additional trial information. If a full text article obtained meets some of the inclusion criteria, or if some clarification is needed before the article can be selected for inclusion, the authors of the paper can be written to requesting further information. This maximises the chances of including articles which have relevant outcomes which will in turn maximise the data available for review and therefore the quality of the review. In turn, other authors are often a useful source of information on other studies in the relevant area, and may even have further data available which has not previously been presented or published. Again, this maximises the available data. It is also of relevance in so far as a weakness of systematic reviews in general is that they rely on published work. Negative or equivocal results are less likely to get published, so that the systematic review which includes meta-analysis can over-estimate the effect of an intervention simply by not including the more marginal studies.

Record Keeping

Using the search strategies to conduct a comprehensive and reproducible search for studies can be the most time-consuming and challenging task in preparing a systematic review. Yet it is also one of the most important. Identifying all relevant studies, and documenting the search for studies with sufficient detail so that it can be reproduced is what distinguishes a systematic review from a traditional narrative review[Higgins JPT et al., 2006]. The process of searching the literature and selecting studies for inclusion constitutes a large part of the methods whereby the review was conducted. This is akin to the methods used in a clinical trial and so should be published in sufficient detail to allow the process to be scrutinised by those who read the finished review, including editors of journals to whom the review is submitted.

Recently, there have been efforts to improve the general quality of reporting of systematic reviews. This refers particularly to the methods used for identifying and selecting studies for inclusion, as it is at this stage that significant scope for introducing bias exists. Guidelines for the reporting of systematic reviews/meta-analyses have previously been suggested[Cook DJ et al., 1995], but the Quality of Reporting of Meta-analyses (QUOROM) conference was convened in order to provide a cross-discipline consensus on the standards for improving the quality of reporting of meta-analyses[Moher D et al., 1999]. (A note on terminology: In Europe, a meta-analysis is a specific statistical method where data from more than one study is synthesised to produce an overall estimate of effect, so that all meta-analysis requires systematic review of the literature, yet a systematic review of the literature does not necessarily have to contain a meta-analysis. In contrast, in North America where the

QUOROM statement originated, the terms meta-analysis and systematic review are used interchangeably. Throughout this thesis, the European convention has been utilised).

The conclusions of this group were published in 1999 in the Lancet, and checklists were produced covering the methodology of conducting a review in order to ensure basic standards are met[Moher D et al., 1999]. It is recommended by the QUOROM group that these checklists are used by authors conducting systematic review, and several major journals will now only accept systematic reviews for publication which include a completed QUOROM statement.

Central to the completion of a QUOROM statement is the ability to report the number of titles uncovered by the original searches and the proportion of these titles assessed as being potentially sufficiently relevant to have the abstract retrieved. The total number of full text articles read and the number of these included should also be documented. Keeping meticulous records in this way will also ensure that work is not repeated unnecessarily. A QUOROM statement has been completed for this review and is included as appendix 5.

Selection Criteria

The methods by which articles are selected for inclusion or exclusion is a key component of conducting an unbiased systematic review. The selection criteria should be set at the beginning of the study, and clear inclusion and exclusion criteria presented.

Assessment of Methodological Quality

Assessing the quality of studies included in a systematic review is necessary not only to uncover sources of bias within the individual studies included in the review, but also to assess how comparable the included studies are and guide interpretation of findings. Several factors should be assessed, including those related to applicability of findings, validity of individual studies, and certain design characteristics that affect interpretation of results[Higgins JPT et al., 2006].

While an assessment of methodological quality is recommended, there is no generally accepted standard for doing this. One group of investigators identified 34 examples of scales and checklists which have been used throughout the literature to judge the methodological quality of randomised controlled trials alone[Moher D et al., 1995;Moher D et al., 1996]. Many more such check lists exist for judging the quality of other types of studies, such as retrospective reviews, which can be included in certain types of systematic review.

Because there is no ‘gold standard’ for the ‘true’ validity of a study, the possibility of validating any proposed scoring system is limited[Higgins JPT et al., 2006]. As such, the Cochrane collaboration advocates adopting a simple approach. Scales with multiple items and complex scoring systems take more time to complete than simple approaches, yet do not provide more reliable assessments of validity[Jüni P et al., 1999]. They may carry a greater risk of confusing the quality of reporting with the validity of the study. They are more likely to include criteria that do not directly measure internal validity, and they are less likely to be transparent to users of the review.

been carried out, as well as to whether the resultant cohorts are comparable and to whether there are significant differences between patients who participate and those who refuse.

It is crucial to ensure an adequate procedure is in place to conceal the allocation to intervention from the clinician making the decision about whether to include patients in the trial. If the clinician making the decision about inclusion knows in advance the intervention group to which the patient will be assigned, then this knowledge may influence the decision as to whether to include that patient. Research has shown that lack of adequate allocation concealment is associated with bias[Moher D et al., 1998;Schulz KF et al., 1995], and is more important than other components of allocation in preventing bias.

The Cochrane collaboration describes several approaches that can be used to ensure adequate concealment has occurred. The following are appropriate to concealment of allocation in trials of follow up (ie non-pharmaceutical trials)[Higgins JPT et al., 2006]:

- centralised (e.g. allocation by a central office unaware of subject characteristics)
- on-site computer system combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered
- sequentially numbered, sealed, opaque envelopes

Finally, there may be important differences between the characteristics of patients who refuse to participate and those who are recruited. This will often be related to personality or pre-existing anxiety levels, which are difficult to define, but are crucial if subjective outcome

measures such acceptability of alternative follow up methods or quality of life are used. Specifying differences between patients who participate and those who do not could be included as a criterion in assessing quality. In addition, the percentage of patients randomised relative to those asked to participate could be analysed.

Performance Bias

Performance bias refers to systematic differences in the care provided to the participants in the comparison groups other than the intervention under investigation. In the area of follow up, this is very difficult to assess. Patients in novel follow up schedules are free to present to a standard hospital clinic if they encounter problems, and similarly many patients attending for routine outpatient hospital visits will see their GP in between times.

There is evidence that blinding is important in preventing performance bias [Schulz KF et al., 1995]. The Cochrane collaboration suggests that adequate blinding of treatment from both the patient and the provider of treatment should be considered as criteria for the validity of the trial [Higgins JPT et al., 2006]. In a follow up trial, it is impossible to blind either the patient or the provider of follow up to the treatment allocation.

Performance bias is therefore potentially problematic. In the case of breast cancer follow up, main outcome measures often relate to survival, disease free survival and quality of life. While these first two outcomes are objectively measurable and should be free from the effects of performance bias, this is not necessarily the case with quality of life. Moreover, it would be easy to provide additional care to women in either the intervention group or the control group which may impact on quality of life. In fact, the very process of inquiring

about quality of life can allow patients to raise concerns which are then dealt with and which would not necessarily have been addressed were the patients not included in a trial. Any additional input the patients receive relative to either regular care out with the trial setting or to the follow up protocol specified is detailed in the assessment of methodological quality section. Another way to address this is to ensure that all outcome measures are sufficiently objective. This is dealt with in detection bias, outlined below.

Attrition Bias

Attrition bias refers to systematic differences between the comparison groups in the loss of participants from the study. This can have implications if significantly more women withdraw from one method of follow up than another, suggesting that one method of follow up is less agreeable. Reporting of loss to follow up should be looked at when assessing quality of trials.

Detection Bias

Detection bias refers to systematic differences between the comparison groups in outcome assessment. Certain end points or outcomes are inherently objective, such as death or developing recurrent disease. These outcomes are less prone to detection bias. For other more subjective outcome measures such as quality of life, method of assessment is important as knowledge of the intervention could affect assessment of quality of life related issues.

One way to ensure that outcome assessment is not affected by knowledge of the follow up procedure is to blind those measuring the outcomes, but this may not be possible within follow up trials. Blinding has been described as particularly important in research with such

subjective outcome measures[Schulz KF et al., 1995]. Several questionnaires are available to objectively measure quality of life. In trials where there are difficulties in blinding assessors to the patients' treatment arm, the use of such questionnaires to ensure objective measures of outcome should be considered.

Other Issues

Additional issues which are important in assessing the applicability of results, but do not necessarily affect the quality of the trial, include issues such as percent randomisation and length of follow up.

Finally, while a quality assessment rating is important for the reasons given above, trials should not be weighted in any way according to this assessment, as this is not recommended by any commentators. Trials should not be excluded on the basis of scores either, but any issues which are felt may be important to consider in interpreting the results should be included within the text of the discussion.

Methods

The methods adopted in this systematic review are guided by the general advice as detailed above.

MEDLINE, Embase, CancerLit, Web of Sciences, and EBM reviews including the Cochrane database were search for relevant studies in May 2006. All English language publications between 1966 and the time of the search were considered. A two stage search process was used. Relevant studies were identified using a less detailed search strategy. Several studies

and two systematic reviews already known to exist were obtained. By examining the MeSH terms and keywords of these studies, the search strategies employed within the two systematic reviews and the various pre-prepared search strategies for identifying certain types of studies as outlined in the Cochrane collaboration handbook[Higgins JPT et al., 2006], the final search strategies were produced. The reference lists of all relevant studies and pre-existing systematic reviews uncovered by the searching process were scrutinised for further relevant articles. Finally, all included articles were located in the Web of Science database in order to search for further articles which had been subsequently published and had cited the included studies.

The full search string used is reproduced as appendix 1.1. The full search string used for CancerLit is reproduced as appendix 1.2 and for Web of Sciences as appendix 1.3.

All searches were conducted independently by two authors in keeping with advice from the Cochrane collaboration[Higgins JPT et al., 2006]. The initial search was conducted independently by this author (DAM) and a second author (KK) and titles were studied to assess which abstracts should be obtained. Full details of all additional authors are given in appendix 1.4 of this chapter. The titles of all articles obtained from searching the included databases were examined independently and abstracts obtained for any articles likely to be relevant. All of the abstracts obtained were then read independently by both authors, and the full text of any articles considered relevant at this stage obtained. Consensus was not required in order to obtain the abstract or full text of any article. This was practical due to

the relatively small numbers of abstracts and articles involved in each of the reviews undertaken here.

All full text articles obtained were again read independently by both authors. Disagreement at this stage over whether an article should be included can be resolved by consensus, but the Cochrane collaboration suggests that a third author may be useful for arbitration[Higgins JPT et al., 2006]. In this review, any disagreement in whether to include a study in the final review was resolved by arbitration from a third author (TGC).

If a full text article obtained met some of the inclusion criteria, or if some clarification was needed before the article could be selected for inclusion, the authors of the paper were written to requesting further information. If additional information was available and the trial then had sufficient information to meet the inclusion criteria, then that trial was included.

All abstracts obtained were read and considered independently by both DAM and KK to assess whether the full text article should be obtained. Sources of disagreement at this stage resulted in the full text article being obtained. References of all full text articles obtained were also searched for further relevant studies.

Selection Criteria

Studies were included if they fulfilled the following criteria:

1. Patients included had been treated for primary operable breast cancer and were free of distant metastases outside the breast or axilla at time of initial treatment.
2. The study was a randomised controlled trial comparing routine clinical and mammographic follow up with an alternative, or comparing different frequency or duration of clinical follow up. Blinding of method of follow up to either clinicians or patients was considered unnecessary.

Data Extraction

Data were extracted by two investigators (DAM and KK) independently. Topics included number of patients, percentage agreeing to randomisation, age, primary surgical treatment, study comparison, follow up schedule and total duration of follow up.

Data was also extracted on the main outcome measures used and how these were determined. Outcome measures included patient satisfaction, quality of life, number of events (local relapse, distant relapse, significant adverse events and death) and economic analysis. Whether satisfaction and quality of life were measured in a validated manner was also recorded.

Assessment of Methodological Quality

Methodological quality was assessed independently by DAM and KK by means of a pre-defined form. This form was derived from a number of sources, guided by a summary by Greenhalgh [Greenhalgh T, 1997] on the measurement of methodological quality in scientific trials, and by the work of Altman et al [Altman DG and Lyman GH, 1998] and Laupacis et

al[Laupacis A et al., 1994]. The various issues which were considered important in terms of rating quality are described in the section above where the general advice on conducting systematic review is given. The quality checklist consisted of 13 points which were considered to be important, and so quality scores are out of a possible 13 points. As stated above, the quality rating was not used to weight the studies, or as exclusion criteria, but to give a general overview of any weaknesses in the trial, and these weaknesses were considered in the discussion section of this chapter.

Results

In all, 2248 titles were examined in MEDLINE, 944 in EMBASE, 225 in EBM reviews, 2882 in CancerLit and 331 in Web of Sciences. Five review articles were also obtained in order to examine the references for further relevant studies[de Bock GH et al., 2004;Collins RF et al., 2004;Anonymous, 1998;Heys SD et al., 2005;Rojas MP et al., 2005]. In total, 20 abstracts were obtained.

From these abstracts, 6 full text articles were obtained, all of which were eligible for inclusion in the analysis[Brown L et al., 2002;Grunfeld E et al., 1996;Grunfeld E et al., 2006;Gulliford T et al., 1997;Koinberg IL et al., 2004;Kokko R et al., 2005]. One abstract had been published as part of the proceedings of a scientific meeting, and the data had not yet been published in full[Baildam AD et al., 2004]. This abstract was also included. In total, we identified seven randomised controlled trials in the literature, all of which were eligible for inclusion.

Study Characteristics

The characteristics of all 7 included trials are presented in table 1.1.

| Study | Year | Patient Group | Percentage randomisation | Age | Primary Therapy | Study design | Main Outcome Measures | Follow up schedule | Duration of Study or median follow up |
|-----------|------|--|--------------------------|--|---|---|---|--|---|
| Grunfeld | 1996 | 296 primary operable patients attending routine follow up at all stages of follow up and free of metastatic disease | 66.5% | Mean 59.1 GP and 62.4 Hospital (sd 10.3 and 12) | 53% mastectomy, 47% Breast Conservation (BCT) | Randomized controlled trial of follow up in General practice versus hospital. | Quality of life as measured by several validated questionnaires. Number of recurrences. Number of deaths. Time to diagnose recurrence from onset of symptoms. NO data on survival | 3 monthly clinical exam and history during first year, 6 monthly for four years then annual in one hospital, with three monthly first year, four monthly second year, 6 monthly to five years and then annual in the other. GP as per hospital of diagnosis. One to two yearly mammograms. | 18 months |
| Gulliford | 1997 | 196 primary operable patients attending routine follow up over a 24 month period at all stages of follow up and free of metastatic disease | 93% | 56 <49 years, 96 =50-65 years and 41 >65 years | 68% BCT, 32% mastectomy | Comparison of frequent follow up versus annual follow up | Acceptability of randomisation and overall satisfaction. Interim use of telephone and GP. NO data on survival or quality of life | Three monthly clinical exam and history during first year, four monthly second year, 6 monthly to five years and then annual. One to two yearly mammogram, depending on whether mastectomy or BCT in the control group, annual clinical exam, history and mammogram in the trial group | Median 16 months |
| Brown | 2002 | 61 primary operable patients attending routine follow up at all stages of follow up and free of metastatic disease | 50% | Mean 63 standard clinic, 68 in patient initiated | 66% BCT, 34% mastectomy | Traditional clinic follow up versus patient initiated follow up | Quality of life using validated questionnaires and satisfaction using unvalidated structured interview. NO data on survival | 4 to 6 monthly clinical exam and history for first five years then annual in control group versus on request only in the study group. Annual mammograms in both | 12 months |
| Koinberg | 2004 | 264 primary operable patients attending routine follow up at all stages of follow up and free from metastatic disease | Not given | 58.8 (sd 10.1) in traditional group, 60 (sd 10.3) in on demand group | 84% BCT, 16% mastectomy | Traditional clinic follow up versus follow up on demand coordinated by breast care nurse | Quality of life and satisfaction using validated questionnaires. Number of contacts with health professionals. Number of events and survival | 3 monthly clinical exam and history for two years, 6 monthly for three years then annual for five years and annual mammogram in traditional follow up group, with appointments on demand only and annual mammograms in the nurse led follow up on demand group | 5 years |
| Baildam | 2004 | 525 primary operable patients attending routine follow up at all stages of follow up and free of metastatic disease | 78% | Not given | Not given | Comparison of standard follow up with hospital doctors or specially trained breast care sisters | Number of events Anxiety by validated questionnaire. Satisfaction using validated questionnaire. Economic comparison NO survival comparison | Not given, but identical for both arms | Not given |
| Kokko | 2005 | 472 primary operable patients attending routine follow up at all stages of follow up and free of metastatic disease | Not given | 56.9 in frequent follow up group, 60 in infrequent group | 54% mastectomy, 46% BCT | A comparison of three versus six monthly follow up (and of intense versus as required investigations) | Event detection and cost per event detected NO data on survival | Three versus six monthly clinical exam and history. | Median 4.2 years |
| Grunfeld | 2006 | 968 women between 9 and 15 months after diagnosis of early stage breast cancer who had completed treatment and were disease free | 55% | Mean 60.9 (in both groups) | 73% BCT, 20% mastectomy and 7% biopsy only | A comparison of follow up by GP versus hospital follow up | Quality of Life using validated questionnaires Significant clinical events (metastases related) Number of local recurrences and deaths | 3 to 6 monthly for three years, 6 monthly for two years then annual, with annual mammogram | Median 4.5 years from diagnosis. (3.5 from randomisation) |

Table 1.1: Characteristics of included studies.

Two trials compared follow up in hospital clinics with that provided by a general practitioner[Grinfeld E et al., 1996;Grinfeld E et al., 2006], two compared traditional follow up to follow up simply on demand by contacting a breast care nurse[Brown L et al., 2002;Koinberg IL et al., 2004] and one compared routine follow up by doctors with routine follow up by breast care nurses[Baildam AD et al., 2004]. Two trials compared different frequency of follow up within a traditional framework[Gulliford T et al., 1997;Kokko R et al., 2005].

Methodological Quality

As stated above, methodological quality was assessed by means of a form derived by DAM and KK prior to selecting trials. The form contained 13 points considered to be important, and so a score of 0 to 13 was possible. Methodological quality in general among these trials was good. One study [Baildam AD et al., 2004] scored poorly as a result of being published only in abstract form as part of the proceedings of an international meeting. The average score was 8.7 out of 13, or 10.7 if Baildam et al is excluded (see table 1.2)

| | Grunfeld 1996 | Gulliford 1997 | Brown 2002 | Koinberg 2004 | Baildam 2004 | Kokko 2005 | Grunfeld 2006 |
|---|------------------|-------------------|---------------------------------|------------------|-----------------|---------------|------------------|
| Is the population under study defined (with inclusion and exclusion criteria)? | Yes | Yes | No (exclusion criteria omitted) | Yes | No | Yes | Yes |
| Are the main prognostic factors defined (at least age of patient and stage of tumour)? | Yes | Yes | Yes | Yes | No | Yes | Yes |
| Is treatment of first tumour specified? | Yes | Yes | Yes | Yes | No | Yes | Yes |
| Has a power calculation been carried out to assess required cohort size? | Yes | No | No | Yes | No | No | Yes |
| Is cohort size sufficient for primary outcome measure (ie greater than any calculated sample size or of sufficient size to detect clinically significant difference)? | Yes | Yes | No | Yes | Yes | Yes | No |
| Is a comparison made of baseline characteristics (age and stage) between participants and those who refuse to participate? | Yes | Yes | Yes | No | No | No | No |
| Is mean or median follow up greater than five years? | No | No | No | Yes | Not given | No | No |
| Is loss during follow up specified? | Yes | No | Yes | No | No | No | No |
| Is the follow up schedule (including mammographic interval) specified? | Yes | Yes | No | Yes | No | Yes | Yes |
| Is mammographic frequency identical between follow up groups? | Yes | Yes | Yes | Yes | Not given | Yes | Yes |
| Are mostly objective or validated outcomes used? | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Were outcomes prospectively assessed? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Was article published in peer reviewed journal? | Yes | Yes | Yes | Yes | No | Yes | Yes |
| Total Score | 12 | 9 | 8 | 11 | 3 | 9 | 9 |

Table 1.2: Quality Rating of Included Studies

Rating of methodological quality took into account intended outcome measures when assessing adequacy of sample size and follow up duration. Many of the trials had either inadequate numbers of patients or insufficient follow up duration, or indeed both of these problems, for an analysis of long term survival to be conducted.

Detection of Relapse

Whether traditional clinical follow up is carried out three or six monthly makes no difference to the overall number of relapses detected[Kokko R et al., 2005]. In 239 women followed up

3 monthly, there were 66 relapses, compared with 57 relapses among 233 women followed up 6 monthly. Local relapse was not dealt with separately, and time to detection of relapse was not analysed. It is unclear whether further increasing follow up interval to annual would have an effect on detection of local relapse. One study attempted to address this issue, but recruited only 196 patients and analysed after a median of 16 months follow up, so had insufficient patient numbers and follow up duration to conduct such analysis[Gulliford T et al., 1997].

There was no difference in number of relapses between patients followed up by a physician compared with a breast care nurse (7 versus 6 respectively from 525 patients followed up for an undisclosed time)[Baildam AD et al., 2004]. Time to relapse was again not presented.

More local relapses were detected by nurses during on demand follow up than during routine visits in one trial[Koinberg IL et al., 2004]. Nurses detected 12 relapses from 133 patients compared with 8 local relapses detected by physicians from 131 patients. Again, these numbers are too small for this to be significant, and there was no difference in the number of patients diagnosed with metastatic disease. There were no significant differences between the groups in terms of time to event for any of the event types (time to loco-regional relapse, metastases or death) analysed (data not shown).

In the first of the trials dealing with hospital follow up compared with follow up in general practice 296 patients were recruited, 148 in each type of follow up. There were similar numbers of locoregional relapses in each group, with 4 in the general practice group and 3 in

the hospital group. There were more than twice as many metastatic relapses diagnosed in the hospital group (13 versus 6 in general practice, difference 4.7%, 95% c.i. -0.8 to 10.3%)[Grunfeld E et al., 1996]. Of interest, while all the relapses in the general practice group were detected by the GP, 44% of the relapses in the hospital follow up group were also diagnosed by the GP initially[Grunfeld E et al., 1996].

In the second of the hospital versus GP follow up trials, there was again no significant difference in the proportion of women presenting with local or distant relapses between the two groups (11.2% in GP group compared with 13.2% in the hospital group, difference of 2.02% 95% c.i. -2.13% to 6.16%). Time to relapse detection is not presented[Grunfeld E et al., 2006].

Adverse Clinical Events

One trial addressed the issue of relapse related serious clinical events(SCEs)[Grunfeld E et al., 2006]. The authors postulate that early diagnosis of relapse is likely to result in reduced rate of serious events related to relapse, such as spinal cord compression from spinal metastases etc. They found no difference in rate of SCEs between hospital and general practice follow up (3.7% of patients versus 3.5% respectively, difference 0.19% 95% c.i. -2.26 to 2.65).

Survival

There was no difference between physician follow up or follow up on demand by a breast care nurse in either number of deaths, with 14 deaths from 131 physician followed up

patients compared with 14 deaths from 133 patients followed up by the nurse). Overall survival on Kaplan Meier curves was similar for the two groups also (data not shown)[Koinberg IL et al., 2004].

One study showed a slight excess in mortality in hospital follow up compared with general practice[Grunfeld E et al., 1996]. The difference was not further analysed as the numbers were small, with only 2 deaths in the general practice group and 7 in the hospital group from 148 patients in each cohort. Moreover, this difference did not exist over longer follow up in a much larger cohort of 968 women in a subsequent trial of hospital follow up compared with general practice by the same authors[Grunfeld E et al., 2006]. In this later trial, there were 29 deaths in the general practice group and 30 in the hospital group.

Satisfaction

Willingness to participate (percentage randomisation) is displayed in table 2.

Two studies looked at satisfaction with an alternative method of follow up. One used a structured interview technique[Brown L et al., 2002] while the other employed a questionnaire previously validated by the authors[Koinberg IL et al., 2004]. Both reported high levels of satisfaction with follow up in both study and control groups, with no difference in satisfaction between the study and control groups in either trial. In the patient initiated follow up trial by Brown et al, there were differences in what patients perceived as advantages of different methods of follow up, with more women describing routine clinic

visits as reassuring ($X^2:27.63$, $p<0.000$, 1df), while more described on demand follow up as convenient ($X^2:17.354$, $p<0.000$, 1df)[Brown L et al., 2002].

Within a traditional follow up schedule, frequency can be reduced with no loss of satisfaction[Gulliford T et al., 1997]. Moreover, similar numbers in the high and low frequency groups will express a desire for lower or higher frequency visits than they received during the trial[Gulliford T et al., 1997].

In the trial by Baildam et al, the Fallowfield Satisfaction with Consultation Questionnaire revealed that women were significantly more satisfied with their consultation with a nurse than those seen by a doctor ($p<0.001$)[Baildam AD et al., 2004].

Quality of Life

Several trials looked at the issue of quality of life measured by validated tools such as the hospital anxiety and depression scale (HADS)[Brown L et al., 2002;Baildam AD et al., 2004;Grunfeld E et al., 1996;Grunfeld E et al., 2006;Koinberg IL et al., 2004], European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 and BR23[Brown L et al., 2002;Grunfeld E et al., 1996] and the Medical Outcomes Study Short Form 36-item General Health Survey (SF-36)[Grunfeld E et al., 1996;Grunfeld E et al., 2006]. None of these trials reported any differences in any of the scales between control and study groups at any stage in the respective trials.

While there were no differences in quality of life between control and study groups in any of the trials, one group did report that nurses were better at spotting psychological distress than doctors[Baildam AD et al., 2004]. From a cohort of 525 women, 86 patients had psychological distress indicated by HADS scores, 49 seen by a nurse and 37 by a doctor. While nurses failed to recognise this in 47%, doctors missed distress in 92% of these women.

Economic Analyses and Workload Concerns

Unsurprisingly, reducing frequency of clinic follow up is shown to reduce the cost of follow up, with a mean cost of follow up of 1656€ in the three monthly follow up group compared with 1050€ per patient in the six monthly follow up group over 4.2 years of follow up[Kokko R et al., 2005]. There was no increase in the number of phone calls or visits to General Practitioners as a result of reduced frequency of follow up visits[Gulliford T et al., 1997].

One other study revealed no advantage of breast care nurse follow up over follow up by clinicians in terms of cost (data not shown)[Baildam AD et al., 2004]. This finding was attributed to the fact that nurses spent longer with individuals than doctors, and that very senior nursing staff were required.

Grunfeld et al[Grunfeld E et al., 1996] subsequently published the results of an economic analysis of follow up in hospital compared with by general practitioner[Grunfeld E et al., 1999]. GP patients were seen significantly more frequently and each visit was significantly longer during the 18 months of the trial than were hospital patients. In addition, while overall numbers of tests were similar, GPs ordered significantly more blood tests,

mammograms and chest x-rays than did specialists. Overall, the cost to the health service was lower for general practice follow up than for follow up in hospital[Grunfeld E et al., 1999].

Follow up on demand by breast care nurses results in significantly fewer clinical contacts. There were 450 fewer visits to the physician and only 177 more phone calls and 88 more visits to the nurse in one study, so that there were 21% more primary contacts in the physician group[Koinberg IL et al., 2004]. There were more mammograms in the nurse led on demand group.

Discussion

Numerous guidelines exist for the follow up of breast cancer, with little concordance between them. In the UK, The National Institute for Health and Clinical Excellence (NICE) recommends two to three years of follow up then discharge to general practice[National Institute for Clinical Excellence, 2002]. The British Association of Surgical Oncology (BASO) suggests discharge at 5 years[The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005]. In sharp contrast, the American Society of Clinical Oncology recommend three to six monthly visits for five years followed by annual visits indefinitely[Khatcheressian JL et al., 2006], and the Canadian Steering Committee similarly advocate indefinite follow up visits for clinical examination[Grunfeld E et al., 2005]. But which of these organisations is correct?

Routine clinical examination after treatment for breast cancer is an inefficient way of detecting recurrent disease[de Bock GH et al., 2004]. Intuitively, more frequent clinic visits should lead to earlier detection of relapse and improved survival at the cost of reduced efficiency. There is no evidence in the literature that this is the case.

Only two randomised controlled trials have attempted to address frequency of follow up [Gulliford T et al., 1997;Kokko R et al., 2005]. Neither provided clear answers. Koko and colleagues demonstrated that three monthly and six monthly follow up are equivalent in terms of event detection, but made no comment on survival and did not separate locoregional from distant relapse[Kokko R et al., 2005]. Gulliford and colleagues' trial was insufficiently powered to comment on the safety of reducing follow up to annual visits[Gulliford T et al., 1997]. No randomised trials have looked at the issue of when patients should be discharged from clinic.

Five randomised controlled trials exist comparing traditional follow up to some alternative[Brown L et al., 2002;Baildam AD et al., 2004;Grunfeld E et al., 1996;Grunfeld E et al., 2006;Koinberg IL et al., 2004]. In these trials, the authors have largely sought to measure outcomes such as acceptability to patients, satisfaction, quality of life etc. Most of the trials have resultantly been underpowered or of inadequate duration to address the issue of survival.

Those investigators who have attempted to address the issue of survival have shown no disadvantage to follow up out with the hospital setting, either by general

practitioner[Grunfeld E et al., 2006] or on demand from a specialist nurse[Koinberg IL et al., 2004]. In one of these trials, almost 1000 women were followed until four and a half years after treatment and survival, locoregional and distant relapse detection and incidence of relapse related serious clinical events were identical in both arms[Grunfeld E et al., 2006]. Yet event rates will be small even in such a large cohort of women. In 1312 women in Edinburgh followed for a median duration of 10 years, only 110 treatable locoregional relapses occurred, with only 15 diagnosed by clinical examination, the rest being either symptomatic or mammographically detected (see chapter 2). Therefore despite the large recruitment and reasonable duration of follow up in the trial by Grunfeld and colleagues, even that trial was insufficiently powered to show a difference in outcome related to follow up.

Two trials have demonstrated a trend to more locoregional relapses being detected by alternative follow up methods[Koinberg IL et al., 2004;Grunfeld E et al., 1996] but both were of limited duration. With longer follow up, the question of whether this represents earlier detection of locoregional relapse during non-hospital based follow up, and whether this would result in improved long term survival with these alternative methods of follow up, may have been answered. Certainly, a large meta-analysis has demonstrated that routine clinical follow up results in delayed diagnosis of locoregional relapse in two thirds of women who recur[de Bock GH et al., 2004].

None of the trials revealed any reduction in quality of life or increase in levels of stress associated with alternative follow up [Brown L et al., 2002;Baildam AD et al.,

2004;Grunfeld E et al., 1996;Grunfeld E et al., 2006;Koinberg IL et al., 2004]. Moreover, alternative follow up methods appear to be acceptable to women, as evidenced by high randomisation rates in most of the trials. This suggests that, while many women report finding clinic visits stressful, this is a short term disadvantage and attendance at clinic does not negatively impact on quality of life. Unfortunately, attendance at clinic does not appear to improve quality of life. Attending clinic does not increase the likelihood of having psychological concerns detected and dealt with[Baildam AD et al., 2004], and the reassurance gained from attendance at clinic is also likely to be short term as it does not lead to any general improvement in quality of life compared to women not attending routine outpatient clinic visits.

Alternative methods of follow up can be more cost effective and reduce workload at busy specialist clinics[Gulliford T et al., 1997;Grunfeld E et al., 1999]. So why have so few investigators made any serious attempt to address this issue? In correspondence published in response to Grunfeld and colleagues' original trial[Grunfeld E et al., 1996], there was support from general practitioners for general practice follow up. On the other hand, there was an assertion from breast surgeons that clinical examination was central to the detection of locoregional relapse, and that those conducting such examination should have some degree of specialisation[Rainsbury D, 1996;Dixon M and Norman B, 1996]. They warn that non-specialists may be less well able to detect locoregional relapse, especially in those who have had breast conserving therapy, or to detect metastatic disease when it presents with non-specific symptoms[Rainsbury D, 1996].

Clinicians gain the experience they have in detecting recurrent disease by frequent examination of a range of normal. Reducing clinical follow up may result in de-skilling of clinicians, or at least reduced experience of those currently in training[Rodger A, 1997]. In addition, removing patients from routine follow up would have the effect that clinicians only see patients with problems. This may lead to reduced job satisfaction and the risk of subsequent burnout[Khatcheressian JL and Smith TJ, 2006]. While these are valid concerns, it does raise the question of whether we follow up well patients for their benefit or ours.

Whatever clinicians perceive as their reasons for continuing with clinical follow up, it is the perception we have of patients' expectations which has done most to reduce our exploration of alternative follow up methods. While the idea of coming back to specialist clinics less frequently is attended by a high degree of acceptance among women, as evidenced by a 93% randomisation to annual follow up in Gulliford's trial[Gulliford T et al., 1997], the idea of non specialist follow up is less acceptable. Randomisation rates for the two trials involving follow up by a non-specialist were 55% and 66.5%[Grunfeld E et al., 1996;Grunfeld E et al., 2006]. Morris and colleagues reported that 81% of women in routine breast cancer follow up were reassured by the clinic and 76% preferred hospital clinics to being followed up in general practice[Morris S et al., 1992]. Similar findings have been reported by others[Kiebert GM et al., 1993;Paradiso A et al., 1995;Renton JP et al., 2002].

When surveyed about their perceptions of cancer follow up, patients have expressed anger and distress about being discharged to their general practitioner[Maher J et al., 1995]. Patients viewed the hospital, and in particular diagnostic tests, specialist physicians and

breast care nurses, as their best defence against relapse. While they do report that there are negative points to coming to clinic – rushed consultations, long waiting times and lack of continuity of care – these are seen as an acceptable trade off for guarding against a relapse[Maher J et al., 1995].

Many women find follow up visits stressful. Almost half of patients report feeling worried, anxious or frightened before appointments[Renton JP et al., 2002] and psychological stress is significantly higher among cancer patients one month before clinic and on the day of clinic visits than two weeks after[Kiebert GM et al., 1993]. Up to 70% of breast cancer patients may suffer distress during routine clinical follow up[Paradiso A et al., 1995].

None of the trials of alternative methods of follow up included here set out to measure the psychological stress of attending clinic. As a result, it is not clear whether alternative methods of follow up have less attendant stress. Replacing routine clinical follow up with some alternative does not appear to have a negative impact on patient psychological well being in general. In the studies which reported on quality of life, there was no reduction in scores among patients in the trial group either compared with base-line or compared with the control patients[Brown L et al., 2002;Baildam AD et al., 2004;Grunfeld E et al., 1996;Grunfeld E et al., 2006;Koinberg IL et al., 2004]. There was no improvement in quality of life scores either.

Summary

There are no randomised trials in the literature with sufficient power to recommend an acceptable frequency or duration of follow up. Moreover, there are no randomised trials which can confirm the safety of alternative follow up methods. Those studies which have been conducted have not suggested that alternative methods are any less safe than routine clinical follow up [Grunfeld E et al., 1996; Koinberg IL et al., 2004; Grunfeld E et al., 2006], but recruitment and duration are such that the only conclusion which can be made is that the necessity for clinical examination and the safety of alternative follow up has not been proven.

Traditional routine clinic visits are an inefficient way of safeguarding against recurrent disease, and there is real doubt as to whether they are the ideal setting for providing psychological support to patients. Alternative follow up methods are acceptable to patients, are associated with no reduction in quality of life or increase in anxiety and may be conducted with significant economic and time savings.

More high quality, randomised controlled trials in this area are required in order to establish how best to provide effective psychosocial support to patients after treatment of breast cancer, while at the same time maintaining adequate surveillance to detect those few treatable relapses which occur.

Chapter 2:

Pattern and Timing of locoregional Relapse, Method of Detection and Impact on Survival

Retrospective analyses of two local cohorts

Summary of the chapter

It is unclear how frequently or for how long routine clinical examination should be provided in order to maximise detection of relapse. In the next two chapters, this issue is explored. In this chapter, the timing of relapse and the method of detection of that relapse is analysed in two local cohorts. This will help establish the incidence of relapse over time and provide more objective evidence for how long follow up should be provided. The analysis of method of detection will establish the contribution clinical examination makes to detection of relapse and thus outcome.

Introduction

Routine clinical examination remains a central component of follow up despite uncertainty over the contribution such examination makes to detection of potentially treatable relapse. As highlighted in chapter 1 above, there is no randomised evidence to suggest that patients do better when followed up by routine clinical examination compared with other methods of follow up, but this is simply because the few trials which have attempted to address this issue have been underpowered to do so.

A recent meta-analysis has reported that routine clinic visits are at best inefficient in the detection of potentially treatable locoregional relapse[de Bock GH et al., 2004]. Very few potentially treatable relapses are detected at routine clinic visits, and traditional follow up has involved very frequent clinic visits, at least within the first three to five years following diagnosis and treatment[American Society of Clinical Oncology, 1999]. As a result of this now recognised inefficiency, in the UK in particular the recent guidelines on follow up in breast cancer have recommended conducting less frequent or shorter duration follow up in an attempt to improve efficiency[National Institute for Clinical Excellence, 2002;The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005]. While reduced follow up has been shown to be acceptable to patients[Gulliford T et al., 1997] there has been little published on the impact of this strategy on loco-regional relapse detection and overall survival.

It is clear from chapter 1 and the general introduction above that there is still uncertainty over not only the need for regular clinical examination, but also the frequency with which it should be conducted and the duration of follow up required. There are no randomised controlled trials to advise the guidelines, and there is conflicting evidence from the various retrospective analyses which have been done.

In this chapter, the pattern and timing of potentially treatable relapse, along with the method of detection and the impact of method of detection on outcome, has been explored in two large contemporary cohorts.

Edinburgh Cohort

Patients and Methods

Between 1991 and 1998, 1312 patients were treated for early stage breast cancer by breast conservation surgery, axillary node sampling or clearance and postoperative radiotherapy to the breast +/- ipsilateral lymphatics. Systemic therapy was given according to local and national guidelines. Follow up was shared between the Edinburgh Breast Unit and the Department of Clinical Oncology, Western General Hospital, Edinburgh.

Although initial clinical follow up consisted of 3 to 4 monthly visits for the first two years, 6 monthly for three years then annual visits until the tenth anniversary, from 2000, all patients were changed to annual follow up visits only. Throughout, all patients were instructed in regular breast self-examination. Additional interval visits for assessment were arranged by patients, their general practitioner or other healthcare professionals as required. Patients were usually discharged to the national breast screening unit at ten years. Annual bilateral mammography was undertaken throughout. Median follow up was 10 years, range 1.5 to 15 years.

Patient characteristics

Characteristics and original pathology of the entire Edinburgh cohort are presented as table 2.1.

| | |
|--------------|-----------------|
| Age (years) | |
| Mean | 56 years |
| Range | 24 – 91 years |
| Tumour stage | No. of patients |
| T1 | 725 |
| T2 | 587 |
| Node status | No. of patients |
| Positive | 354 |
| Negative | 958 |

Table 2.1: Patient characteristics and original pathology of the entire Edinburgh Study population

Methods

Loco-regional relapse was defined as any relapse in the ipsilateral breast and axilla or the contralateral breast. Contralateral breast cancers are included here as the risk of subsequent contralateral breast cancer in patients who have had a previous breast cancer is 3-5 times greater than the population risk[Mellink W et al., 1991]. As such, detecting new contralateral cancers is an important component of breast cancer follow up programmes. Supraclavicular relapse was included with metastatic disease. Detection of metastatic disease within three months of a diagnosis of locoregional relapse was considered to constitute simultaneous local and distant relapse and these patients were excluded from the analysis.

A list of all patients known to have suffered an isolated locoregional relapse or new contralateral cancer was generated from our database in January 2006. A retrospective study of these case records was then undertaken. Time from diagnosis to relapse, method of detection of relapse (routine mammography, routine clinical examination in an asymptomatic patient or investigation of a patient who attended complaining of relevant symptoms) and type of clinic (routine or interval) was recorded. Time from relapse to development of metastatic disease or death was then calculated and survival was compared for each mode of

detection. For all relapses detected clinically in the first instance, we established whether the relapse was also visible on mammography.

Results

Between 1991 and 1998, 1312 patients were treated for primary operable breast cancer by breast conservation surgery. Figure 2.1 displays overall survival, with cause specific survival shown in Figure 2.2. The overall 5 year survival was 89.2%; and the ten year survival was 77.2%.

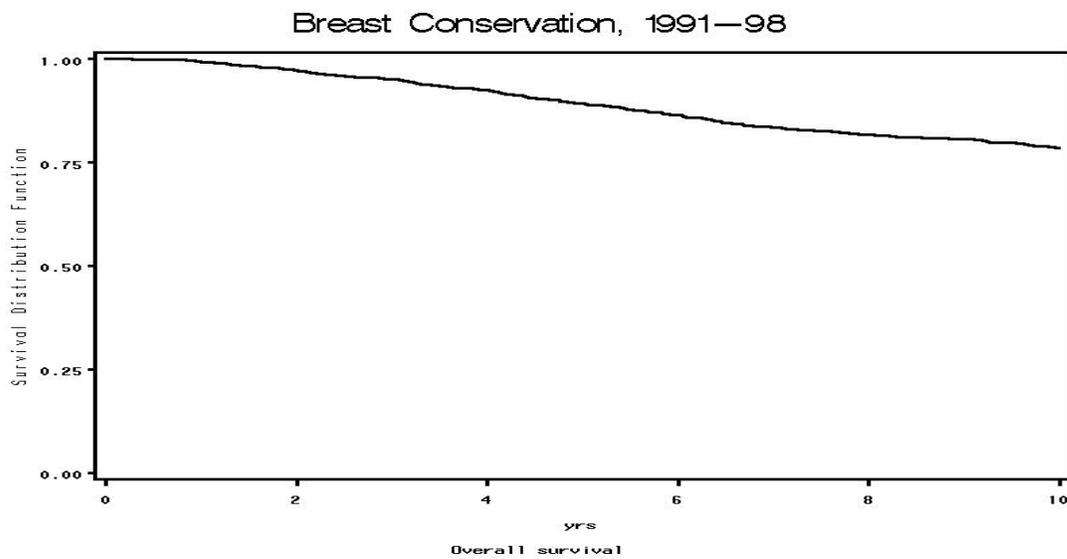


Figure 2.1: Overall survival for all 1312 patients in the Edinburgh cohort



Figure 2.2: Cause specific survival for all 1312 patients in the Edinburgh cohort

Local relapse

There were 31 patients who developed synchronous locoregional and metastatic disease and 11 patients who developed metastatic disease prior to developing locoregional relapse. These patients are not included in the analysis, as mentioned above. They are included as metastatic relapse in calculating the incidence of relapse in figure 2.3 (see below).

116 patients suffered an isolated locoregional relapse. Records were unavailable for four of these patients. This left 112 patients whose records were available. 4 patients have been lost to follow up. Two patients had moved away from the area before the diagnosis of their relapse and they were excluded from further analysis as the method of diagnosis was unknown. The other two had moved after the diagnosis and treatment of their relapse. They have been included in the analysis up until the time of their last clinic visit. This left 110 patients for analysis. Characteristics and the original pathology of all the patients in whom treatable relapse was detected are presented as table 2.2.

| | |
|--------------|-----------------|
| Age (years) | |
| Mean | 54 years |
| Range | 24 – 83 years |
| Tumour stage | No. of patients |
| T1 | 67 |
| T2 | 46 |
| Node status | No. of patients |
| Positive | 29 |
| Negative | 84 |

Table 2.2: Patient characteristics and original pathology of all patients from the Edinburgh Study population in whom relapse was diagnosed

There were 48 patients with ipsilateral breast relapse, 7 of whom developed concomitant axillary disease. 25 patients had isolated ipsilateral axillary relapse. 35 patients developed new contralateral cancers. One patient developed ipsilateral breast and axillary relapse and a new contralateral cancer simultaneously and one patient developed ipsilateral breast relapse and a new contralateral cancer simultaneously.

Timing of Relapse

The relapse rate in this cohort is presented as figure 2.3.

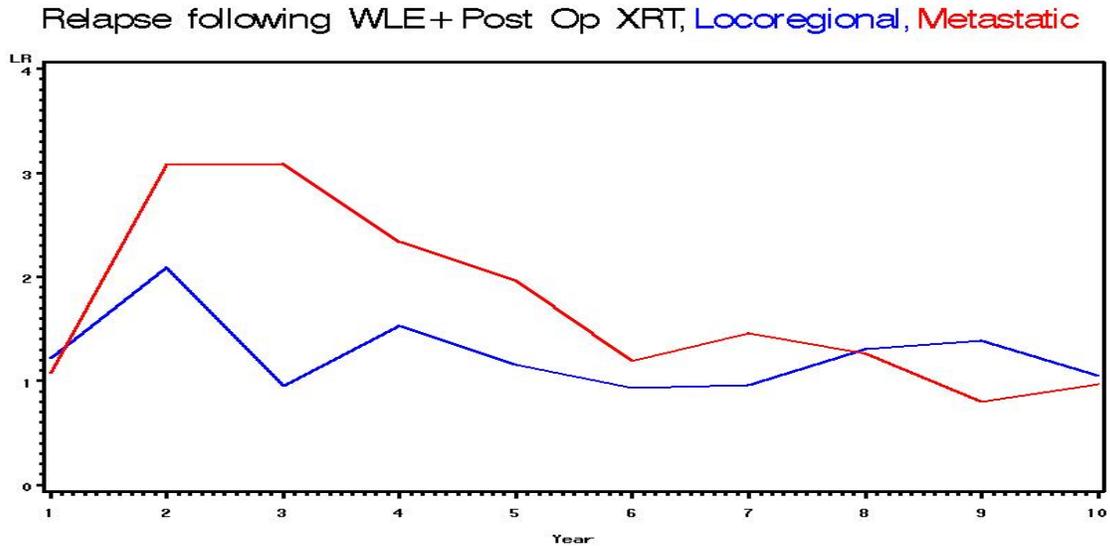


Figure 2.3: Annual incidence of relapse per hundred women at risk in the Edinburgh cohort. Metastatic disease is shown in red for comparison

The incidence of metastatic relapse peaks at just over 3% per annum at two to three years and remains above 2% per year for up to five years before falling off over the rest of the follow up period. In contrast, the incidence of locoregional relapse remains constant at 1 to 1.5% over the whole course of the follow up period.

Method of Detection

Site of relapse and the method by which relapse was detected is summarised for all 110 patients in table 2.3. The number of patients who subsequently died is included in brackets.

| Site of relapse | Symptoms | Clinical Finding | Mammogram | Interval clinic with symptoms | Total |
|--|---------------|------------------|----------------|-------------------------------|-----------------|
| Ipsilateral Breast | 4 (0) | 4 (4) | 20 (6) | 11 (4) | 39 (14) |
| Ipsilateral Axilla | 5 (5) | 9 (3) | 4 (1) | 7 (2) | 25 (11) |
| Ipsilateral Breast & Axilla | 1 (1) | 0 | 5 (0) | 1 (0) | 7 (1) |
| Contralateral breast | 3 (2) | 2 (0) | 25 (5) | 5 (1) | 35 (8) |
| Ipsilateral axilla & bilateral breast | 0 | 0 | 1 (0) | 0 | 1 (0) |
| Bilateral Breast | 0 | 0 | 1 (1) | 0 | 1 (1) |
| Total | 13 (8) | 15 (7) | 56 (13) | 24 (7) | 108 (35) |

Table 2.3: Site of relapse and method of detection. The number of patients who subsequently died is included in brackets. *2 patients, not included in this table, had ipsilateral breast relapse diagnosed incidentally during breast reshaping procedures. Both of these patients subsequently died.

Interval Appointments

23 of the 110 patients (21%) had relapse detected at an interval clinic visit. 22 of these patients complained of relevant symptoms, and the other had an abnormality noticed by carers. Of these 23 patients, 11 had ipsilateral breast relapse, 6 axillary relapse and 1 had synchronous breast and axillary relapse, with symptoms from both of these sites. Five patients discovered their new contralateral cancer.

Two patients (2%) were referred back to clinic having been discharged from further follow up at ten years post treatment. One complained of symptoms related to axillary relapse (included in table 2.3 as having been diagnosed at an interval clinic with symptoms) and one had abnormal routine breast screening mammograms which indicated breast and axillary relapse (included in table 2.3 as routine mammographically detected relapse).

Routine Appointments

The remaining 85 relapses (77%) were detected at routine clinics or by routine mammography. The pattern of relapse and method of detection for this group is also presented in table 2.3.

Of the four ipsilateral breast relapses diagnosed clinically, the most recent annual surveillance mammogram was negative in two cases and was not repeated. Mammography was repeated and was negative in the third. In the fourth, mammography did reveal some distortion, but this was reported as benign post surgical changes in their first routine follow up mammogram one year after surgery. Neither of the two contralateral cancers detected clinically were visible on mammography.

Of the 9 axillary relapses diagnosed clinically, one was also visible on mammography. It was included as a clinical diagnosis as it was diagnosed at a routine appointment at which mammography was not scheduled.

In total, 37 relapses (33.5%) were symptomatic, 56 (51%) mammographically detected, 15 (13.5%) clinically detected and 2 (2%) were diagnosed incidentally.

Rate of mammographic detection

10415 mammograms were undertaken during follow up, with 56 treatable relapses diagnosed. This equates to 5.37 treatable relapses diagnosed per 1000 mammograms undertaken.

Survival

In total, there were 37 deaths among the 110 patients during the follow up period. The number of patients who died for each site of relapse and method of detection is summarised as the bracketed figures in table 2.3.

Ipsilateral Breast

48 patients suffered ipsilateral breast relapse. Seven of these had additional axillary relapse. Five year survival for these patients overall was 87.5% from original operation, and 64% from diagnosis of relapse. Survival by method of detection is shown below (figure 2.4).

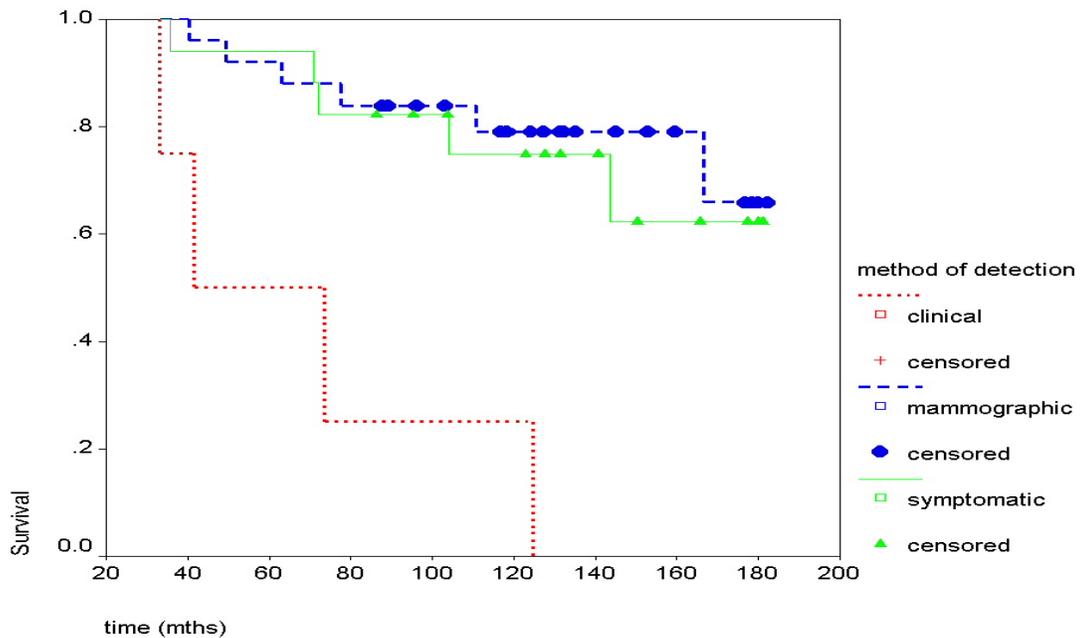


Figure 2.4: Survival from original operation in patients with ipsilateral breast relapse.

Overall survival was significantly reduced in the cohort in whom relapse was diagnosed clinically compared with either other method (log rank 2 df $p=0.0002$). This remains a highly significant association even if the seven patients with axillary disease are excluded ($p=0.0004$), and reflects a significantly longer survival from diagnosis of relapse (figure 2.5) (log rank 2 df $p=0.0014$) rather than a difference in time to relapse detection.

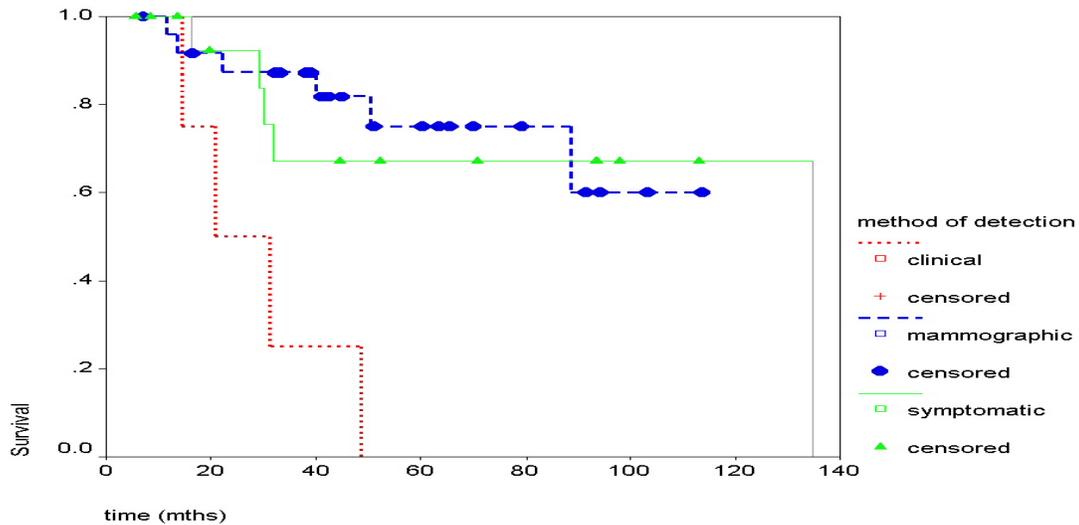


Figure 2.5: Survival from time of relapse in patients with ipsilateral breast relapse

The pathological characteristics of the recurrent disease are presented as table 2.4. Mean age at recurrence is also shown. Nodal status was available for all 48 patients. Size of the recurrent lesion was unavailable for ten women, with grade of the recurrence unavailable for 13 of the women. Receptor status was available for very few patients, and so is not presented.

| | Symptoms | Clinical Finding | Mammography |
|---|--|---|--|
| Age of patient at relapse | 57.7 (sd 14.7 yrs) | 62.77 (sd 17.7) | 59.74 (sd 11.8) |
| Mean Size of relapse (mm) | 25.85 (sd 26.38) | 17.67 (sd 10.97) | 12.71 (sd 8.53) |
| Grade of relapse | Grade 3: 9 Grade 2: 2 Grade 1: 2 Missing: 4 | Grade 3: 5 Grade 2: 10 Grade 1: 3 Missing: 7 | Grade 3: 2 Grade 2: 0 Grade 1: 1 Missing: 1 |
| Proportion of relapses with positive lymph nodes | 2 of 17 (8.5%) | 1 of 4 (25%) | 7 of 25 (28%) |
| Mean NPI* of recurrence | 4.36 (sd 1.33) | 4.02 (sd 1.26) | 3.87 (sd 1.16) |

Table 2.4: Clinicopathological features of recurrence by method of detection *NPI = Nottingham Prognostic Index

While there is a trend for patients with clinically detected relapse to be older than patients in the other two groups, this was not significant. There is also a trend for mammographically detected lesions to be smallest at time of detection, with clinically detected lesions being the larger and symptomatic being largest of all. Again this was not a significant difference. Overall, there was no significant difference in any of the clinicopathological features related to how relapse was detected and Nottingham Prognostic Index of the relapse was similar for all three groups.

Contralateral Breast

There was no association between method of detection of relapse and survival in patients who developed a new contralateral breast cancer. 25 out of 35 of these (71%) were diagnosed by mammography with 8 diagnosed by the patient. Only 2 were picked up on clinical examination. Both of the patients with new cancers detected clinically were alive and well at last follow up. Overall 5-year survival from time of relapse for all patients with contralateral breast relapse was 81%.

Ipsilateral Axilla

There was similarly no association between method of detection of relapse and survival in patients who had isolated ipsilateral axillary relapse, although the numbers overall were small. Overall 5-year survival from time of relapse for patients with axillary relapse was 61%.

Glasgow Cohort

Patients and Methods

Between October 1995 and September 2001, 198 patients were treated for early stage breast cancer by breast conserving surgery, axillary node sampling or clearance and postoperative radiotherapy to the breast +/- ipsilateral lymphatics. Systemic therapy was given according to local and national guidelines. Follow up was shared between the Glasgow Royal Infirmary Breast Surgery Unit and the Department of Clinical Oncology, Beatson Oncology Centre, Glasgow.

Although initial clinical follow up consisted of 3 to 4 monthly visits for the first two years, 6 monthly for three years then annual visits until the tenth anniversary, from 2000, all patients were changed to annual follow up visits only. Throughout, all patients were instructed in regular breast self-examination. Additional interval visits for assessment were arranged by patients, their general practitioner or other healthcare professionals as required. Patients were usually discharged to the national breast screening unit at ten years. Annual bilateral

mammography was undertaken throughout. The median follow up for the Glasgow cohort was 5.9 years (range 4 months to 10.5 years).

Methods

The breast specific case notes of all 198 patients were analysed in August and September 2006. For all patients found to have suffered locoregional relapse since the original operation, and for all patients whose breast case notes were missing or inadequate, the complete set of case records was obtained for a more detailed analysis.

Details collected were date of original operation, type of surgery, radiotherapy details and whether the patient had suffered relapse. For patients who had died, cause of death and presence of relapse was also recorded. Loco-regional relapse was defined as any relapse in the ipsilateral breast or axilla or the contralateral breast. Contralateral breast cancers are included here as the risk of subsequent contralateral breast cancer in patients who have had a previous breast cancer is 3-5 times greater than the population risk, as mentioned above[Mellink W et al., 1991]. Supraclavicular relapse was included with metastatic disease for the purposes of this study. Detection of metastatic disease within three months of a diagnosis of locoregional relapse was considered to constitute simultaneous local and distant relapse.

For those who had suffered relapse, time to relapse, method of detection of relapse (routine mammography, routine clinical examination in asymptomatic patient or investigation of a patient who attended complaining of relevant symptoms) and type of clinic (routine or

interval) was recorded. Time from relapse to development of metastatic disease or death was also recorded. For all those relapses detected clinically in the first instance, we recorded whether the relapse was also visible on mammography.

Results

198 patients were treated for primary operable breast cancer by wide local excision between October 1995 and September 2001. Type of axillary surgery and whether the patients had post-operative radiotherapy is given in table 2.5. The age range of the patients is also included in table 2.5.

| | |
|-------------------|------------------|
| Age (years): | |
| Mean | 58 (sd 13.6 yrs) |
| Range | 28 – 91 |
| Axillary surgery: | |
| Clearance | 175 |
| Sample | 12 |
| None | 11 |
| Radiotherapy: | |
| Yes | 168 |
| No | 15 |
| Attempted | 1 |
| Not recorded | 14 |

Table 2.5: Patient age and original treatment – Glasgow cohort

Only two patients have been lost to follow up. One has no case records available and the other was included in the analysis up until last clinic visit. Cause of death was available for all other patients known to have died. Figure 2.6 displays overall survival, with cause specific survival shown in Figure 2.7. The overall 5 year survival in our cohort was 83.5%.

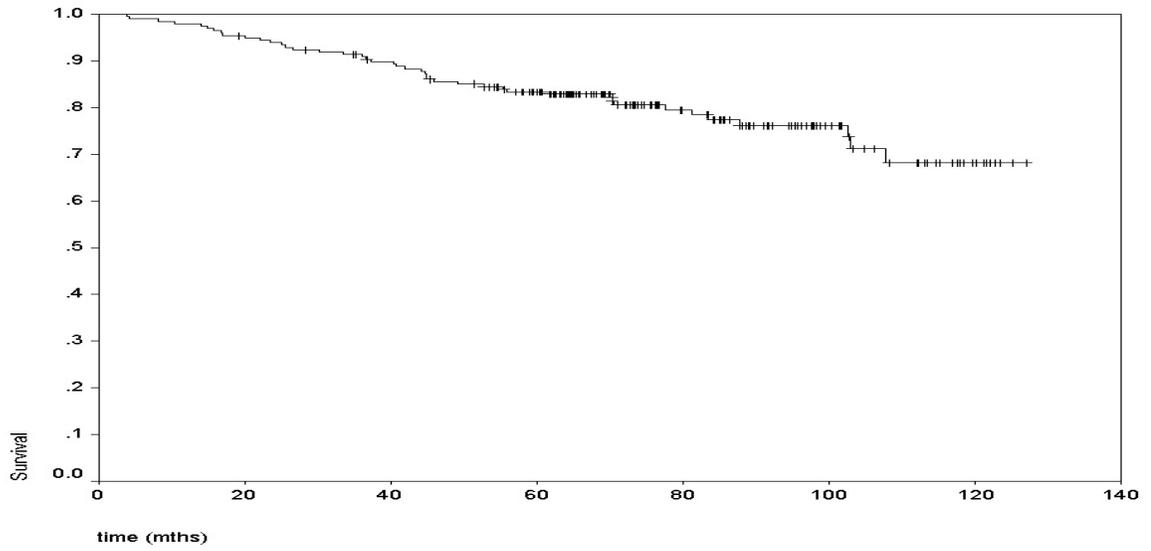


Figure 2.6: Overall survival for all 198 patients in the Glasgow cohort

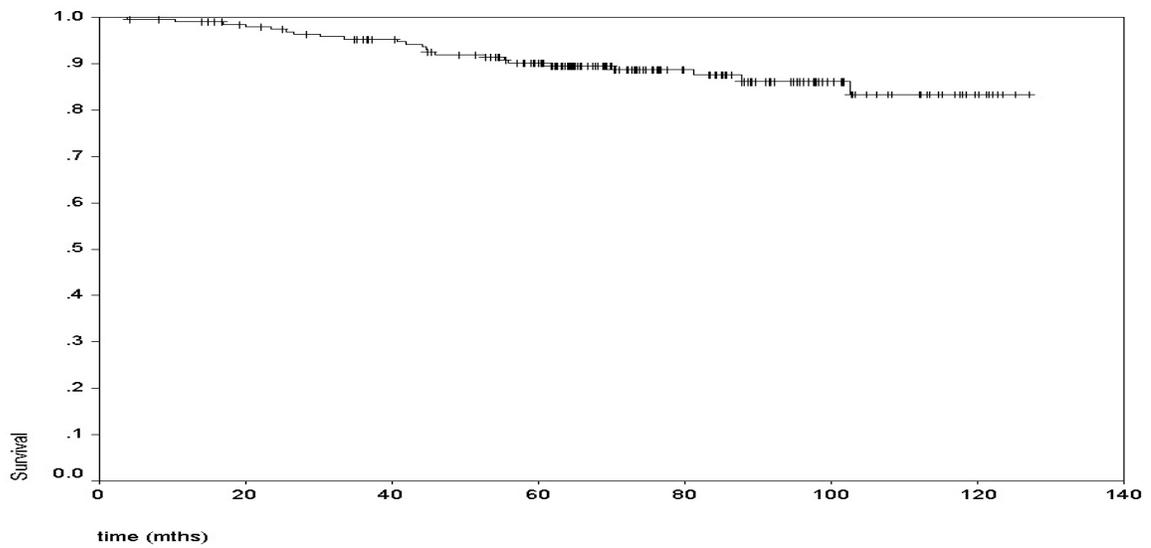


Figure 2.7: Cause specific survival for all 198 patients in the Glasgow cohort

Time to Relapse

The annual hazard rate of isolated locoregional relapse as a first event is shown in figure 2.8, with the numerical hazard rate for each year reproduced in table 2.6.

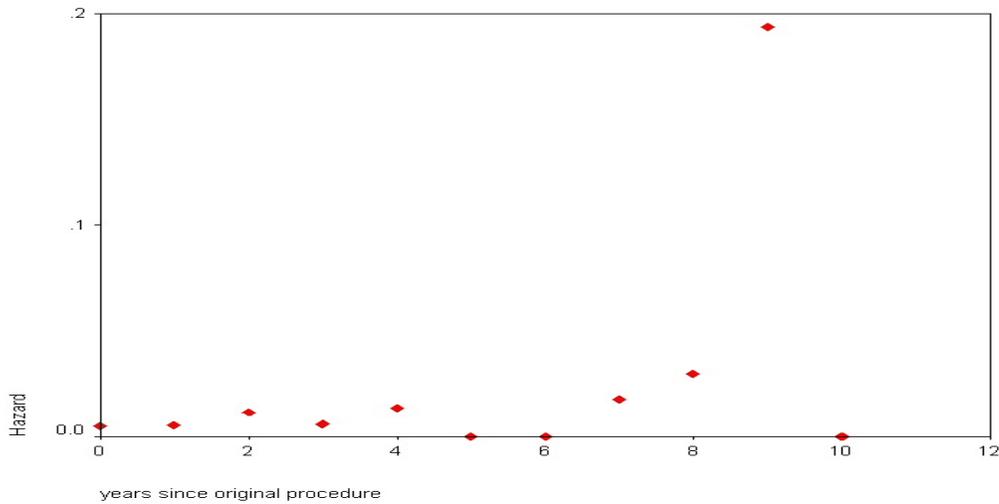


Figure 2.8: Annual incidence of relapse per hundred women at risk in the Glasgow cohort

| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------------------------------|-------|-------|-------|------|-------|-----|----|-------|-------|--------|----|
| Patients entering year | 198 | 193 | 183 | 173 | 160 | 140 | 96 | 71 | 44 | 23 | 8 |
| Number of relapses | 1 | 1 | 2 | 1 | 2 | 0 | 0 | 1 | 1 | 3 | 0 |
| Hazard rate | 0.51% | 0.53% | 1.12% | 0.6% | 1.32% | 0 | 0 | 1.72% | 2.94% | 17.65% | 0 |

Table 2.6: Annual hazard rate for isolated locoregional relapse – Glasgow cohort

Median time to relapse in these 12 patients was 64 months, with a range of 9.67 to 113.57 months. Only one patient had a second event within the first year, developing a new contralateral cancer. 3 patients had a second event within three years of their original operation. 1 developed ipsilateral axillary disease, one a true local relapse and one a new contralateral tumour. There were therefore 8 relapses (6 locoregional recurrences and 2 new contralateral cancers) in our cohort detected more than three years after initial diagnosis and treatment.

Method of Detection

12 patients suffered isolated locoregional relapse or new contralateral cancer as a first event, 6.1% of the cohort. Complete records were available for all of these patients.

Altogether, 5 of the 12 patients (41.66%) noticed symptoms of recurrent disease themselves. 2 of these presented early to an interval clinic, both with relapses in the same breast as the original cancer. Therefore 3 patients with locoregional relapse or new contralateral disease (25%) developed symptoms suggestive of this but did not attend early to have these investigated. Of these 3 patients, one had ipsilateral breast relapse, one ipsilateral axillary relapse and one contralateral breast relapse.

The other 7 patients had relapse detected by routine clinical examination or mammography. 3 were mammographically detected, all new contralateral primary breast cancers. 4 patients had their relapse detected during clinical examination. 2 had new contralateral primaries, 1 ipsilateral breast relapse and 1 ipsilateral axillary relapse.

Of the three clinically detected breast relapses, one of the contralateral cancers was seen on subsequent mammography, as was the ipsilateral breast relapse. The other contralateral cancer had not been seen on routine mammography six months previously, but the mammogram was not repeated.

Outcome of Relapse

10 of 12 patients had further surgery upon diagnosis of relapse. 7 patients had mastectomy, 2 had further conservation surgery and one had further axillary clearances. Of these 10

patients, two have subsequently developed metastatic disease. One was alive at last review having originally developed self-detected axillary relapse and the other has subsequently died (initial relapse for this patient was a clinically detected ipsilateral relapse).

Of the two patients who did not have salvage surgery, one was given letrozole due to other medical problems and was alive and well at last review, the other (whose initial relapse was self-detected axillary relapse) subsequently developed metastatic disease and has since died.

In this study, all three patients who have developed metastatic disease had locoregional relapse (two axillary and one local relapse) within three years of their original operation.

Survival from diagnosis of relapse is shown as figure 2.9, with survival from first diagnosis of breast cancer as figure 2.10.

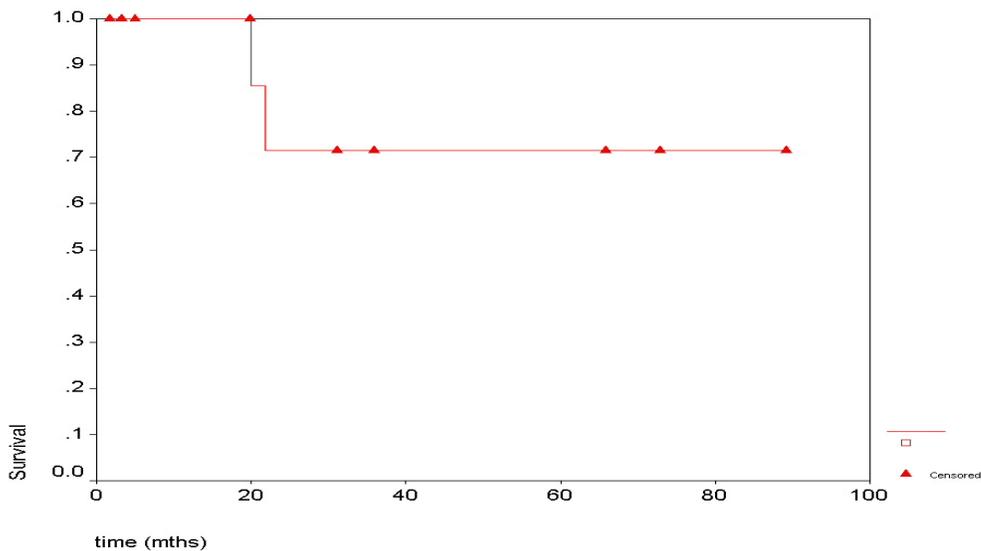


Figure 2.9: Survival from diagnosis of relapse in all women suffering relapse – Glasgow cohort

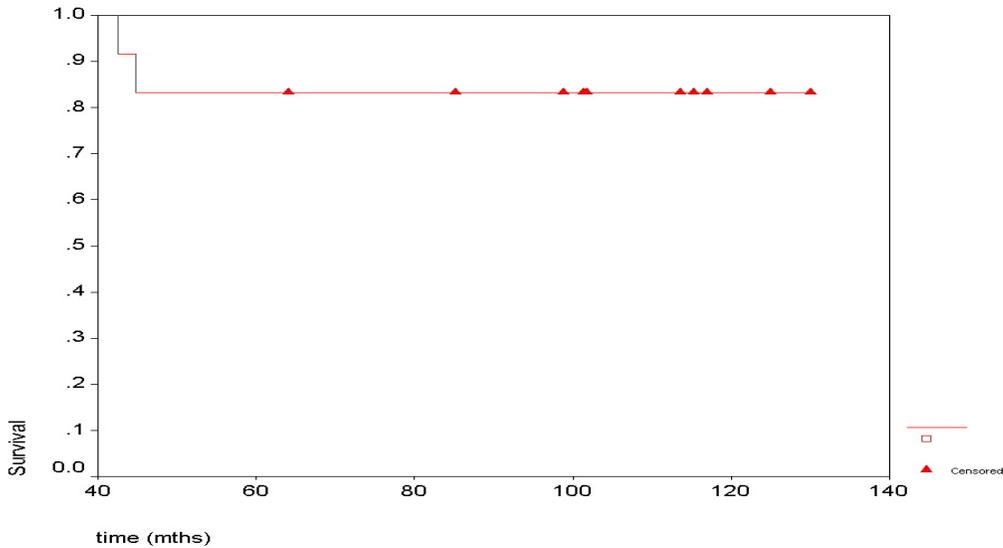


Figure 2.10: Survival from original operation in those with relapse – Glasgow cohort

Discussion

The various guidelines which exist for follow up in breast cancer have been discussed in detail above. While these guidelines appear to vary markedly in their recommendations, with 2-3 years of follow up recommended by NICE[National Institute for Clinical Excellence, 2002] and indefinite follow up recommended by the Canadian Steering committee[Grunfeld E et al., 2005], close scrutiny of the guidelines reveals that they are all underpinned by the assumption that relapse is most common in the first few years after treatment. In all the guidelines, follow up is concentrated on this initial period.

As discussed above, a greater hazard rate for relapse during the first few years after treatment has been reported by a number of authors. A higher hazard rate for three years after treatment was just evident in the ATAC trial[The ATAC Trialists' Group: 2005], but this reflects a higher rate of distant relapse in the first three years. Similarly, a large analysis of relapse among all patients recruited to the first seven Eastern Cooperative Oncology Group coordinated studies of post operative adjuvant therapy in breast cancer revealed a high initial hazard rate, falling off over from a peak between 1 and 3 years after treatment[Saphner T et al., 1996]. Again, locoregional relapse and metastatic relapse are not distinguished, despite being very different events in terms of prognosis. The rate of locoregional relapse is initially high but falls off over time, so that most treatable locoregional relapses occur in the first three to five years[Elder EE et al., 2006;Hussain ST et al., 1995]. This does not include new contralateral cancers and so does not accurately reflect the incidence of all treatable relapse. Often, analysis of relapse over the long term is not possible due to very short median follow up[Donnelly J et al., 2001].

The data presented in the retrospective analyses of the Edinburgh cohort confirm that there is an initial somewhat higher rate of distant relapse in the first five years, but reveal that this is not mirrored by the pattern of locoregional relapse, which remains constant at 1 to 1.5% for at least ten years. In the Glasgow cohort, while numbers were small, the majority of relapses occurred more than 3 years after diagnosis and treatment. The long term pattern is impossible to analyse in the Glasgow cohort due to the shorter median follow up.

The second assumption underpinning the various guidelines is that clinical examination is of more value than mammography. This is reflected in the strong emphasis in all the guidelines on providing more frequent clinical examinations during the perceived high risk period of the first few years. It is put most explicitly in the NICE guidelines which state that the yield of mammography during follow up after breast cancer is low[National Institute for Clinical Excellence, 2002]. Previous studies have consistently shown that multiple clinic visits do little to increase yield and only serve to reduce the “cost effectiveness” of follow up[de Bock GH et al., 2004]. Our data confirm this. In ten years of follow up in 1312 women, only 15 relapses were detected clinically. This is a very low yield. The yield within the Glasgow cohort does look more significant, with 4 potentially treatable relapses detected by routine clinical examination. Two of these women were due routine mammography on the same visit as their relapse was detected. In a unit where the mammography was carried out and reported before the clinic visit, these would both have been detected mammographically, reducing the yield of clinical examination to just 2 clinically detected relapses from 199 women, a rate similar to that seen in the Edinburgh cohort.

This distinction between how follow up is carried out between two units highlights one of the difficulties associated with this type of analysis. The schedule of activities at each clinic visit is rarely if ever described in retrospective analyses of method of detection of relapse, and differences in whether routine mammography is carried out before or after clinical examination could explain some of the differences between various studies with regards to the proportion of relapses detected by each of the various methods of detection.

Mammography, in contrast to clinical examination, makes a much larger and more significant contribution to relapse detection, not only in the area of contralateral new breast cancer detection but also in detecting ipsilateral breast relapse. In fact, a detection rate of 5.37 cancers per 1000 mammograms undertaken within the Edinburgh cohort is a higher rate of detection than that reported as recently as 2003 by the breast screening service in the UK who carry out mammography on a three yearly basis[NHS Cancer Screening Programmes, 2003], and is a rate equivalent to the most up to date report by this organisation published this year. Moreover, patients with ipsilateral breast cancer relapse which is detected clinically do significantly less well than patients whose relapse is detected in other ways – whether this is more biologically aggressive disease or cancer not detected by mammography is unclear.

Our data reveal that the basic assumptions behind the guidelines for follow up are incorrect. If one of the main aims of follow up is the detection of treatable relapse and not distant disease, as stated in the guidelines, there is no justification for focussing on the first two to three years after treatment. Treatable relapse occurs at a constant rate over at least ten years.

Clinical examination has a very low yield and the analysis of outcome carried out in the Edinburgh data suggests that such examinations do not improve outcome, particularly in ipsilateral breast relapse. Outcome was not analysed among the Glasgow patients by method of detection as the numbers were very small. This is, to our knowledge, the first description of outcome related to method of detection of relapse described anywhere in the literature, and it is therefore impossible to contrast our findings with other reports. It is also difficult to

establish why patients did badly when they had clinically detected relapse, as there did not appear to be any difference in the pathological stage of the relapse. This issue will be further investigated in a chapter 3.

While very few breast relapses are detected by clinical examination in either of the cohorts analysed above, axillary relapse may be somewhat different. In the Edinburgh cohort, axillary relapse is more often detected by clinicians than by patients with 36% of axillary relapses being detected by routine clinical examination. Comment is difficult for the Glasgow cohort, as only two relapses occurred within the axilla, one symptomatic and one clinically detected. Isolated axillary relapse is a relatively rare event in our patients, with just 25 events in 1312 patients over ten years, and only 9 clinically detected relapses in the Edinburgh group. There were only two axillary relapses in the Glasgow cohort from 198 patients followed for a median of 5.9 years.

While there is little from our analysis to recommend routine clinical examination for the detection of breast relapse, the argument is more complicated with regards to axillary relapse. There were too few patients to carry out formal analysis of outcome by method of detection of axillary relapse, but patients with axillary relapse overall had a 61% five year survival in the Edinburgh cohort, and clinically detected axillary relapse had a better prognosis, albeit not a statistically significant one. In the Glasgow cohort only the patient with clinically detected axillary relapse was disease free at the time of analysis, the patient who self diagnosed having died. Of course, our data also confirm that isolated axillary relapse is a relatively uncommon event after breast cancer treatment.

Summary

Our data highlight the flaws which exist in the recommendations for follow up after breast cancer, and provide evidence for developing alternative guidelines.

Treatable relapse occurs at a constant rate, with no early peak in incidence. Clinical examination appears to be of little value, particularly with regards to breast relapse, although even larger studies than reported here would help to clarify this issue still further. If such examination is to be provided for the detection of treatable relapse, then it must continue indefinitely, as recommended by the Canadian Steering Committee[The Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer, 1998], and should not be undertaken any more frequently than annually, even during the first three to five years after treatment, but with open access if necessary. This issue will be explored further in the next chapter.

Mammography is of central importance to the detection of treatable relapse. The yield when undertaken annually is 5.37 relapses detected for every 1000 mammograms performed, at least equivalent to the yield seen under a three yearly protocol in the general screening population. This fully justifies the use of annual mammography in this patient population. Given the rising incidence of new contralateral cancers seen after ten years in several studies, there is justification for continuing to provide annual mammography indefinitely.

Chapter 3:

Pattern of locoregional Relapse, Method of Detection and Impact on Survival

A Systematic review of the literature

Summary of the chapter

In the previous chapter, the pattern and timing of potentially treatable relapse, the relative contributions of mammography, patient self examination and routine clinical examination and the impact of detection method on survival in 2 local cohorts was reported. In this chapter, a systematic review of the literature is undertaken to search for studies, both retrospective cohort and randomised controlled trials, with similar data in an attempt to establish whether the findings reported in chapter 2 are seen more generally in the literature. From a summation of the data presented in the chapter above and any existing data in the literature, recommendations for follow up are made.

Introduction

The early detection of potentially treatable relapse remains a central purpose of follow-up after breast cancer[National Institute for Clinical Excellence, 2002]. Regular routine clinical examination is thought to be necessary if clinicians are to meet this aim, yet there is no agreement between the various published guidelines with regards to either how frequently routine clinical examination should be carried out, or for how long after the diagnosis and treatment of breast cancer that such examination is needed[Grunfeld E et al.,

2005;Khatcheressian JL et al., 2006;National Institute for Clinical Excellence, 2002;The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005].

Mammography is similarly considered a necessary component of follow up. It is believed that mammography may lead to earlier detection of contralateral new breast cancers[Mellink W et al., 1991] and local recurrence in the conserved breast[Voogd AC et al., 1999], and that this may lead to a survival advantage in women whose relapse is diagnosed mammographically[Voogd AC et al., 1999]. There is tremendous variation in the reported proportion of relapses diagnosed mammographically, with figures ranging from zero from 104 analysed relapses in a study by Donnelly and colleagues published in 2001[Donnelly J et al., 2001] to 51% of 110 analysed patients in a recent study by this author[Montgomery DA et al., 2007b], described in detail in chapter 2 above. Despite this tremendous variation, recent guidelines from NICE state that the yield of mammography in follow up is low[National Institute for Clinical Excellence, 2002], citing as their evidence the study conducted by Donnelly and colleagues where no relapses were diagnosed mammographically[Donnelly J et al., 2001].

A recent review of the evidence for mammography in follow up concedes that there are no randomised controlled trials, and that the evidence which does exist comes from a heterogenous collection of retrospective analyses, all of which are open to bias[Grunfeld E et al., 2002]. As a result, there is little agreement in how frequently mammography should be undertaken, and one set of guidelines states simply that there is very little high level evidence

in the literature to inform clinicians on how frequently mammography should be undertaken [Grunfeld E et al., 2005].

It has been shown by several investigators that early detection of locoregional relapse has a beneficial effect on survival [Clark DH et al., 1985; Kurtz JM et al., 1990; Dalberg K et al., 1998; Haffty BG et al., 1991; Fowble B et al., 1990; Recht A et al., 1989]. In patients who suffer relapse, cure is more likely in those with small, lymph node negative relapse than in those with more advanced disease at the time of diagnosis [Clark DH et al., 1985; Veronesi U et al., 1995; Voogd AC et al., 2005]. Mellink and colleagues have reported that in follow up programmes which include mammography, those relapses which do occur are more often small (35% vs 7%) and node negative (75% vs 57%) than relapses which occur in cohorts followed up with only clinical examination [Mellink W et al., 1991]. There were a number of flaws in Mellink and colleagues study. Most notably, the two compared populations were from different cities and were not matched in terms of either original diagnosis or treatment. A second group reported that, in patients who presented to clinic with a primary breast cancer and who subsequently developed a metachronous contralateral primary during follow up, the metachronous tumour was significantly more likely to be non-invasive than the primary breast cancer (11.4% v. 5.1%, $p < 0.02$) and that the mean size of the metachronous tumour was significantly less than that of the primary breast cancer also (1.94 v. 2.55 cm, $p < 0.001$) [Samant RS et al., 2001]. When method of detection was analysed, part of the reason for this became clear. Metachronous tumours were more than twice as likely to be diagnosed mammographically compared with the primary cancer (46.2 v. 19.9 $p < 0.001$), and tumours which were detected mammographically were smaller than those detected either by

clinical examination by a doctor or self examination by a patient (1.30, 2.02 and 2.69 cm respectively).

For ipsilateral breast cancer relapse, women whose relapse is detected only by mammography have better disease free and overall survival than women with palpable relapse[Voogd AC et al., 1999].

Mammography may contribute to detection of relapse when the disease is less advanced. Whether this is the case for routine clinical examination is not clear, and the relative contribution of each of these follow up tools in detecting relapse is uncertain, with conflicting evidence in the literature.

These issues are important for a number of reasons. Routine clinic visits are stressful for the women involved. Up to 70% of women reporting feelings of anxiety before such visits[Paradiso A et al., 1995]. While high levels of stress are attendant on most aspects of the patient journey through diagnosis and treatment of breast cancer, in the case of routine follow up it is uncertain whether this distress is offset by the benefits of routine clinical examination.

In the previous chapter, new retrospective analyses of local cohorts have been presented which detail the pattern and timing of relapse and method of detection. The aim of this chapter is to establish, through a systematic analysis of the literature, whether the findings reported in chapter 2 are reflected generally in the literature. The relative contributions of

clinical examination and mammography to the detection of treatable relapse after breast cancer will be analysed and the issue of whether any method of detection confers an advantage in terms of outcome will also be explored. From a summation of all the available evidence, recommendations for follow up will be made.

Methods

MEDLINE, Embase, CancerLit, Web of Sciences and EBM reviews were searched for relevant studies in May 2007. All English language publications between 1966 and May 2007 were considered. The search string used is reproduced as appendix 3.1. As mentioned in chapter 1 above, all searches should be carried out independently by more than one author. Details of additional authors used in the review process are included as appendix 3.2. This initial search was conducted independently by this author (DAM) and KK. Titles were studied to assess which abstracts should be obtained.

All abstracts were read and considered independently by DAM and KK to establish whether the full text article should be retrieved. If only one author thought the article relevant, the full text article was obtained. References of all full text articles obtained were also searched for further relevant studies.

Selection Criteria

Two separate analyses were conducted, one a comparison of methods used to detect relapse and a second analysis of the effect of method of detection on outcome. Studies were included in a comparison of methods used to detect relapse if:

- The study group comprised women with primary operable breast cancer without metastatic disease out with the breast and axilla at initial presentation.
- Data pertaining to salvageable locoregional relapse only were presented, or such data were presented separately from distant relapse data. Locoregional relapse was defined as relapse within the breast or axilla. New contralateral disease also could be included. Supraclavicular lymphadenopathy was considered to be distant disease for the purpose of this analysis.
- The type of relapse was presented separately (breast, axilla or contralateral breast). There may be differences in the pattern of detection of each of these types of relapse, and this must be fully explored.
- The method of detection (mammography, symptoms or clinical examination) of all types of relapse was included.

The authors of studies which contained some relevant information were written to for extra data, and the study included if the author could provide sufficient data to meet the inclusion criteria above. Studies were included in the further analysis of survival if there was adequate data on survival included in the initial paper, or if the authors were able to supply outcome data after correspondence. All authors who were included in the comparison of methods used to detect relapse were written to and asked if they were able to supply survival data if this was not available from the original publication.

Assessment of Methodological Quality of included studies

Methodological quality was assessed independently by two authors (DAM and KK) by means of a pre-defined form. There are no accepted criteria for measuring methodological quality in prognostic studies and so this form was a modified version of the form created by de Bock and colleagues[de Bock GH et al., 2004], derived from the work by Altman and colleagues[Altman DG and Lyman GH, 1998] and Laupacis and colleagues[Laupacis A et al., 1994]. The form is presented as table 3.1.

| | Mahoney 1986 | Tate et al 1989 | Rutgers et al 1991 | Snee 1994 | Hussain et al 1995 | Grunfeld et al 1996 | Lees et al 1997 | Jack et al 1998 | Churn et al 2001 | Grogan 2002 | van der Sangen et al 2006 | Montgomery et al 2007 | Glasgow cohort 2007 |
|--|-----------------|--------------------|-----------------------|--------------|-----------------------|------------------------|--------------------|--------------------|---------------------|----------------|------------------------------------|--------------------------|------------------------|
| Is the population under study defined (with inclusion and exclusion criteria)? | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Is the original cohort of patients from which those with relapse were drawn defined? | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | No | Yes | Yes |
| Were all those identified as having relapse analysed? | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes |
| Is loss during follow up specified? | No | No | No | No | No | Yes | No | Yes | No | Yes | No | No | Yes |
| Are the main prognostic factors defined (at least age of patient and stage of tumour)? | No | No | No | Yes | No | Yes | No | Yes | Yes | Yes | Yes (personal communication) | Yes | Yes |
| Is treatment of first tumour specified (including adjuvant)? | No | No | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Is mean or median follow up greater than five years? | Not given | No | Not given | Yes | Yes | No | Not given | Yes | Yes | Yes | Yes | Yes | Yes |
| Is the follow up schedule (including mammographic interval) specified? | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Were methods of diagnosis of relapse prospectively assessed? | Yes | Yes | No | No | No | Yes | No | No | No | No | No | No | No |
| Is all relapse, including axillary and new contralateral cancers, included? | Not given | Yes | No | Not given | Not given | Not given | yes | Yes | No | yes | No | Yes | Yes |
| Percentage of relapses not analysed due to inadequate information | 0 | 0 | 2% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3% | 0 |
| Total Score | 5 | 6 | 4 | 6 | 6 | 8 | 3 | 9 | 7 | 9 | 6 | 7 | 9 |

Table 3.1: Methodological Quality Rating of Included Studies

When analysing survival related to method of relapse, it is of particular concern if patients who are recognised as having relapse are not analysed. In retrospective analyses particularly, this may be because the patient has subsequently died and case notes have been destroyed. For all studies we have included the percentage of patients with recognised relapse not included in the final analysis due to lack of information.

Data Extraction

Two authors, DAM and KK, extracted data from included studies independently. Data collected were year of publication and year of initial operation or referral, population size, age, primary therapy, study design, follow up schedule including mammographic schedule, number of locoregional relapses and method of detection of locoregional relapse (scheduled versus interval clinic and whether detection was by patient, clinician or mammography in the first instance).

Relapse was recorded as clinically detected if it was first detected by a physician in a patient who had not complained of any relevant symptoms. Relapse was recorded as detected by the patient if the patient attended clinic with relevant symptoms, whether the patient waited for the next routine clinic visit or arranged an interval appointment. It was recorded as mammographically detected if an abnormal or suspicious mammogram was recorded before clinical examination revealed any abnormality.

Survival after locoregional relapse was recorded if that information was available.

Statistical Analysis

All data were analysed using SPSS version 11.01 (SPSS Inc, Chicago, IL). For survival calculations, individual data was available for each patient allowing analysis of all individual patients.

Results

In all, 4061 titles were studied in Medline, 4563 in EMBASE, 8906 in CancerLit and 3144 in Web of sciences. From all EBM reviews including the Cochrane database, 3 review articles were retrieved. From these titles, 188 abstracts were read and considered independently by DAM and KK. Nine review articles and four letters or editorials were also obtained from the 188 abstracts in order to examine the references of these articles for further relevant studies. In total, 68 full text articles were considered for inclusion.

From the 68 full text articles considered, 11 studies met the primary inclusion criteria for our analysis of method of detection of locoregional relapse[Mohoney L, 1986;Montgomery DA et al., 2007b;Tate JT et al., 1989;Rutgers EJT et al., 1991;Snee M, 1996;Hussain ST et al., 1995;Grunfeld E et al., 1996;Lees A et al., 1997;Jack WJL et al., 1998;Churn M and Kelly V, 2001;Grogan M et al., 2002]. One of these contained information on long term outcome after local relapse detection[Snee M, 1996]. The data from the Glasgow cohort presented in chapter 2 above was also included.

A further 25 studies contained some data appropriate to our meta-analysis and the authors of these were written to. One was able to provide us with further information, including method

of detection of local relapse and subsequent outcome for the complete study group and 4 additional patients, and has therefore been included in both analyses [van der Sangen MJC et al., 2006].

From the 12 published studies [Mohoney L, 1986; Montgomery DA et al., 2007b; Tate JT et al., 1989; Rutgers EJT et al., 1991; Snee M, 1996; Hussain ST et al., 1995; Grunfeld E et al., 1996; Lees A et al., 1997; Jack WJL et al., 1998; Churn M and Kelly V, 2001; Grogan M et al., 2002; van der Sangen MJC et al., 2006] and the retrospective analysis of relapses in the Glasgow cohort, data was available for 7815 patients with 552 relapses. Seven of these studies [Mohoney L, 1986; Montgomery DA et al., 2007b; Rutgers EJT et al., 1991; Hussain ST et al., 1995; Jack WJL et al., 1998; Grogan M et al., 2002; van der Sangen MJC et al., 2006] and the Glasgow cohort consisted of patients treated by conservation surgery, 2 studies a combination of mastectomy and conservation surgery [Grunfeld E et al., 1996; Churn M and Kelly V, 2001] and one study mastectomy alone [Snee M, 1996]. The final two studies [Lees A et al., 1997; Tate JT et al., 1989] did not specify original treatment.

There were only two published studies [Snee M, 1996; Montgomery DA et al., 2007b] which reported survival related to method of detection of relapse. One additional group had published outcome after locoregional relapse, and were able to supply us with the method of detection for all of these relapses [van der Sangen MJC et al., 2006]. The Glasgow cohort is also included in an analysis of outcome by relapse method. In total, there are 229 locoregional relapses or new contralateral cancers from a cohort of 4823 patients for whom we know the method of detection of relapse, site of relapse and outcome.

Characteristics of all the included studies are presented as table 3.2. In particular, table 3.2 describes the proportion of relapses detected in the same breast as the original tumour, the ipsilateral axilla and the contralateral breast for each included study.

| Study | Patient Group | Relapses included | Inclusion period | Age | Primary Therapy | Study design | Follow up schedule | Mammograms |
|---------------------------|---|--|--|--|--|--|--|---|
| Mahoney 1986 | 273 treated patients | 52 locoregional relapses. All ipsilateral breast | All patients treated by lumpectomy between July 1972 and October 1983. All relapses from July 1972 to December 1983 analysed | not given | lumpectomy | prospective cohort study of the use of thermography | 3 monthly for 1 year, 4 monthly for 1 year, 6 monthly for 3 years then annual | Biennial |
| Tate et al 1989 | 510 patients previously treated for early breast cancer attending a follow up review during the inclusion period. | 27 locoregional relapses. Ipsilateral breast (12), Ipsilateral axilla (12) and contralateral breast (3) | 6 month period (unspecified) neither original operation dates nor date of relapses specified | At follow up: mean 65 for interval attendees, 60 for symptomatic routine and 63 for asymptomatic routine | not specified | Non randomised, non controlled prospective cohort | 2 monthly for 1 year, 3 monthly for 1 year, 4 monthly for 1 year 6 monthly for 2 years then annual until 10 years | not specified |
| Rutgers et al 1991 | 44 patients with locoregional relapse presenting between 1982 and 1990 | 44 locoregional relapses All ipsilateral breast | All locoregional relapses diagnosed between 1982 and 1990 from a cohort of all patients treated between 1978 and 1990 | mean 47.1 (range 26-68) | Lumpectomy and axillary dissection | Non randomised, non controlled retrospective cohort | 3 monthly for two years, 6 monthly up to 5 years then annual | Annual |
| Snee 1994 | All 33 patients referred to regional centre for adjuvant treatment. | 5 locoregional relapses chest wall (3) and axilla (2) | All referrals: jan-feb 1982. No information given on original operation date or period of follow up scrutinised. | At referral: mean 57 (range 34-78) | mastectomy | Non randomised, non controlled, prospective cross-sectional | not detailed, mean of two visits each per year. | Not specified |
| Hussain et al 1995 | 354 treated patients | 33 locoregional relapses. Ipsilateral breast (24), Ipsilateral axilla (3) ipsilateral breast and axilla (6). Did not include new contralateral disease | All patients treated between October 1980 and December 1991. Date of analysis not given | not given | WLE + radiotherapy + at least axillary sample | Non randomised, non controlled retrospective cohort | 3 monthly for 2 years, 6 monthly for 3 years then annual until 10 years | 6 months then annual |
| Grunfeld et al 1996 | 296 patients randomised to GP v. hospital follow up. | 7 locoregional relapses ipsilateral breast/chest wall (5) and ipsilateral axilla (2) | All patients treated between 1988 and 1992 were randomised to the trial at the end of this period and followed for 18 months from that point | GP follow up mean 55.6. Hospital follow up mean 59 | 153 mastectomy and 138 WLE | prospective randomised comparison of GP v hospital follow up | three monthly for one year and 6 monthly for four in one group, 3, 4 and 6 monthly years 1, 2 and 3 for the other then annual both groups. | Year one then every 1 to 3 years |
| Lees et al 1997 | A selected group of 458 treated patients. Selection criteria not given | 83 locoregional relapses. All ipsilateral breast | All patients were treated between 1980 and 1985. follow up complete until December 1991 | not given | Mastectomy or conservation surgery | Non randomised, non controlled retrospective cohort | three monthly for 2 years then 6 monthly to 5 years then annual | Annual |
| Jack et al 1998 | 341 treated patients | 39 locoregional relapses. Ipsilateral breast (24), ipsilateral axilla (11) and contralateral breast (4) | All patients treated between 1986 and 1990 and followed for ten years. Date of analysis not given | mean 52.2 (range 24-82) | Wide Local Excision (WLE) + radiotherapy | Non randomised, non controlled retrospective cohort | 3-4 monthly for 3 years, then 6 monthly until 10 years | Annual |
| Churn et al 2001 | All 612 patients with early breast cancer referred to regional oncology centre for adjuvant therapy in 1993 | 34 locoregional relapses. 25 in wle group and 9 in mastectomy group. Ipsilateral breast, axilla or chest wall (not separated, but did not include new contralateral disease) | All referrals received in 1993 for adjuvant therapy were analysed during 1996 | 189 patients <50, 423 patients >50 | 105 mastectomy, 511 conservation, 3 radiotherapy after neo adjuvant chemotherapy. Variable LN dissection | Non randomised, non controlled retrospective cohort | 3 to 4 monthly for 2 to 3 years, 6 monthly to 5 years then annual | less than annual, according to clinician preference |
| Grogan et al 2002 | 104 treated patients. | 4 salvageable locoregional relapses. Ipsilateral breast (3), Ipsilateral axilla (1) and contralateral breast (0) | Patients treated between January 1988 and June 1991. follow up was for five years from end of treatment in all patients | mean 53, (range 28-81) | WLE + radiotherapy | Non randomised, non controlled retrospective cohort | 3 monthly for 2 years, 4 monthly for 1 year, 6 monthly thereafter | Annual |
| Van der Sangen et al 2006 | 3280 treated patients. All patients from cohort with locoregional relapse > 5 years after original procedure | 102 relapses. All ipsilateral breast | All patients treated between 1982 and 1997. All relapses were between 31 October 1988 and 15 March 2003 | Mean 51 (range 32-85) | WLE + radiotherapy | Non randomised, non controlled retrospective cohort | 3 monthly for 2 years, 6 monthly for three years then annual | Annual (referenced) |
| Montgomery et al 2007 | 1312 treated patients. | 110 locoregional relapses ipsilateral breast (48), Ipsilateral Axilla (25), Bilateral breast (2), and contralateral breast (35) | All patients treated between 1991 and 1998. follow up complete until January 2006 | 54 (range 24-83) | WLE and either sample or clearance of axilla | Non randomised, non controlled retrospective cohort | 3 – 4 monthly for 3 years, 6 monthly to five years then annual. Annual for all patients from 2000 onwards | Annual |
| Glasgow cohort 2007 | 198 treated patients. | 12 locoregional relapses. Ipsilateral breast (4), Ipsilateral axilla (2) and contralateral breast (6) | All patients treated between October 1995 and September 2001. Follow up complete until August 2006 | 58 (range 28 – 91) | WLE +/- axillary surgery +/- radiotherapy (see table 1) | Non randomised, non controlled retrospective cohort | 3 – 4 monthly for 3 years, 6 monthly to five years then annual. Annual for all patients from 2003 onwards | Annual |

Table 3.2: Characteristics of included studies

Quality rating of studies

The median quality score was 6.5 with a range from 3 to 9. Six of the 13 studies included new contralateral breast cancers in the analysis. Eleven of the 13 included studies analysed all of the locoregional relapses which they were aware of within their cohort. In the two remaining studies, one failed to analyse 2% of relapses[Rutgers EJT et al., 1991] and one failed to analyse 3% of relapses[Montgomery DA et al., 2007b].

Method of detection of relapse

The proportion of relapses detected by patient symptoms, mammography and clinical examination for each of the studies is presented in table 3.3.

| | Initial surgery | Patient detected | Mammographic | Clinical examination | Unknown or incidental | Total |
|---------------------------|---------------------------|------------------|--------------|----------------------|-----------------------|-------|
| Tate et al 1989 | not given | 16 (59%) | 0 | 11 (41%) | 0 | 27 |
| Churn et al 2001 | mastectomy | 1 (12%) | 0 | 6 (66%) | 2 (22%) | 9 |
| Snee 1994 | mastectomy | 2 (40%) | n/a | 3 (60%) | 0 | 5 |
| Lees et al 1997 | mastectomy & conservation | 46 (55%) | 15 (18%) | 22 (27%) | 0 | 83 |
| Grinfeld et al 1996 | mastectomy & conservation | 2 (28.66%) | 2 (28.66%) | 2 (28.66%) | 1 (14%) | 7 |
| Mahoney 1986 | conservation | 20 (38%) | 1 (2%) | 31 (60%) | 0 | 52 |
| Rutgers et al 1991 | conservation | 26 (59%) | 8 (18%) | 10 (23%) | 0 | 44 |
| Hussain et al 1995 | conservation | 4 (12%) | 5 (15%) | 24 (73%) | 0 | 33 |
| Jack et al 1998 | conservation | 15 (38%) | 12 (31%) | 12 (31%) | 0 | 39 |
| Churn et al 2001 | conservation | 9 (36%) | 7 (28%) | 8 (32%) | 1 | 25 |
| Grogan et al 2002 | conservation | 2 (50%) | 2 (50%) | 0 | 0 | 4 |
| Van der Sangen et al 2006 | conservation | 41 (41%) | 32 (32%) | 13 (13%) | 16 (16%) | 102 |
| Montgomery et al 2007 | conservation | 37 (33.5%) | 56 (51%) | 15 (13.5%) | 2 (2%) | 110 |
| Glasgow cohort 2006 | conservation | 5 (41.667%) | 3 (25%) | 4 (33.33%) | 0 | 12 |
| Total | | 227 (41%) | 138 (25%) | 166 (30%) | 22 (4%) | 552 |

Table 3.3: Method of detection of all locoregional relapses for all studies

For Churn et al[Churn M and Kelly V, 2001], relapses were presented separately for mastectomy and for conservation surgery and so these results are presented separately in the table. Most of the studies were retrospective analyses and in some cases it was not certain

how the relapse was detected. These are included in the table as unknown. Relapse detected during further surgery for cosmetic reasons are described as incidental relapses.

Few studies addressed the issue of relapse after mastectomy alone. It is not clear what initial surgery was employed in the study by Tate et al [Tate JT et al., 1989], but it is likely to have been mastectomy given the date of publication. Clinical examination was an important method of relapse detection in patients after mastectomy, with between 41% and 66% of relapses detected this way. Mammography played no role in the detection of relapse in these studies, as new contralateral disease was not included in the analysis.

Three studies reported on locoregional relapse in a mixed treatment population [Grunfeld E et al., 1996; Lees A et al., 1997; Churn M and Kelly V, 2001]. In two studies, it was not possible to separate patients treated by mastectomy from those treated by wide local excision [Grunfeld E et al., 1996; Lees A et al., 1997]. Discerning a pattern of relapse detection in these studies is difficult as mammography plays a much smaller role in the follow up of patients treated with mastectomy than in those treated with breast conservation. Clinical examination detected a smaller proportion of relapses in these studies than in the mastectomy studies, with less than one third of relapses detected this way.

There were 8 studies which looked at the issue of relapse after breast conserving surgery [Mohoney L, 1986; Montgomery DA et al., 2007b; Churn M and Kelly V, 2001; Rutgers EJT et al., 1991; Hussain ST et al., 1995; Jack WJL et al., 1998; Grogan M et al., 2002; van der Sangen MJC et al., 2006]. In addition, the Glasgow cohort consisted of

patients who had undergone breast conserving surgery, giving 9 studies of breast conserving surgery in total.

In these 9 studies, 38% of relapses were detected by the patient, 30% by mammograms and 28% by clinical examination. In 4%, method of detection was unknown.

There was some overlap between the studies both with regards to when the included patients had been treated and when relapse was diagnosed. Moreover, the date of relapse of included patients is not always clear in the included studies. The information which does exist is included in table 3.2. An attempt was made to assess whether mammography has made a changing contribution to relapse detection over time. Comparison was made of the proportion of relapses detected by each method in studies published prior to 2000, compared with those published after 2000. Date of publication was chosen as a surrogate for date of diagnosis of relapse as it was the only consistent date available for all the published studies (see table 3.2). It is likely to underestimate any increase in the importance of mammography as the paper by Churn et al (2001) was published in 2001, yet the analysis included relapse diagnosed only until 1996 and has a pattern of relapse similar to that seen in studies published before 2000. The impact of the study by Grogan et al (2002) is limited as so few patients are included.

While the proportion of relapses detected by the patient remains fairly constant (39% in studies from before 2000, 37% from those published after), the proportions detected by mammography and clinical examination reverse. Before 2000, 15% of relapse was

mammographically detected with 46% detected by routine clinical examination. After 2000, 40% is mammographically detected with 16% detected on routine clinical examination.

Three of the studies reported only on ipsilateral breast relapse after wide local excision[Mahoney L, 1986;Rutgers EJT et al., 1991;van der Sangen MJC et al., 2006]. In three other studies[Grogan M et al., 2002;Montgomery DA et al., 2007b;Jack WJL et al., 1998] and the Glasgow cohort, it was possible to extract individual data on method of detection for each area of relapse. Table 3.4 displays method of detection of ipsilateral breast relapse, table 3.5 axillary relapses and table 3.6 new contralateral breast cancers.

| | Patient detected | Mammographic | Clinical examination | Unknown or incidental | Total |
|---------------------------|------------------|--------------|----------------------|-----------------------|-------|
| Mahoney 1986 | 20 (38%) | 1 (2%) | 31 (60%) | 0 | 52 |
| Rutgers et al 1991 | 26 (59%) | 8 (18%) | 10 (23%) | 0 | 44 |
| Jack et al 1998 | 9 (37.5%) | 9 (37.5%) | 6 (25%) | 0 | 24 |
| Grogan et al 2002 | 1 (33%) | 2 (66%) | 0 | 0 | 3 |
| Van der Sangen et al 2006 | 41 (41%) | 32 (32%) | 13 (13%) | 16 (16%) | 102 |
| Montgomery et al 2007 | 17 (36%) | 25 (52%) | 4 (8%) | 2 (4%) | 48 |
| Glasgow cohort 2006 | 3 (75%) | 0 | 1 (25%) | 0 | 4 |

Table 3.4: Method of detection of all ipsilateral breast relapses in conservation surgery studies

| | Patient detected | Mammographic | Clinical examination | Total |
|-----------------------|------------------|----------------|----------------------|-----------|
| Jack et al 1998 | 4 (36%) | 1 (9%) | 6 (55%) | 11 |
| Grogan et al 2002 | 1 | 0 | 0 | 1 |
| Montgomery et al 2007 | 12 (48%) | 4 (16%) | 9 (36%) | 25 |
| Glasgow cohort 2006 | 1 | 0 | 1 | 2 |
| Total | 18 (46%) | 5 (13%) | 16 (41%) | 39 |

Table 3.5: Method of detection of axillary relapse

| | Patient detected | Mammographic | Clinical examination | Total |
|-----------------------|-------------------|-------------------|----------------------|-----------|
| Jack et al 1998 | 2 | 2 | 0 | 4 |
| Grogan et al 2002 | 0 | 0 | 0 | 0 |
| Montgomery et al 2007 | 8 | 25 | 2 | 35 |
| Glasgow cohort 2006 | 1 | 3 | 2 | 6 |
| Total | 11 (24.4%) | 30 (66.6%) | 4 (9%) | 45 |

Table 3.6: Method of detection of new contralateral primaries

Two patients from Montgomery et al 2007[Montgomery DA et al., 2007b] had simultaneous bilateral breast relapses and so are not included in these tables. Both had their relapses detected by mammography.

As can be seen in table 3.4, there is a trend towards increasing proportions of ipsilateral breast relapses being detected by mammography the more recently published the study. In contrast, clinically detected ipsilateral breast relapse becomes less common with time.

Very little pattern can be made of axillary relapse, shown in table 3.5, as numbers are small. The two larger studies are from the same unit and show again a trend towards fewer relapses being detected by clinical examination[Jack WJL et al., 1998;Montgomery DA et al., 2007b], this time with more patient detected relapses.

Very few contralateral breast cancers are detected by clinical examination. 66.6% are detected by mammography and 24.4% by the patients themselves.

Survival

Outcome data were available from the Glasgow cohort and three other studies within our analysis[Snee M, 1996;Montgomery DA et al., 2007b;van der Sangen MJC et al., 2006]. In one study, primary surgery was mastectomy in all cases[Snee M, 1996]. In the local retrospective analysis and other two studies[Montgomery DA et al., 2007b;van der Sangen MJC et al., 2006], patients were initially treated by wide local excision.

Mastectomy

Snee reported on 5 patients who had suffered locoregional relapse after mastectomy[Snee M, 1996]. 2 patients with ipsilateral axillary relapse were diagnosed by clinical examination, one in the breast clinic and one incidentally at another surgical clinic. 3 patients had chest wall relapse, 2 with symptoms and one diagnosed clinically. There was no difference in survival related to method of detection or area of relapse, although three patients died within 2 years of relapse. There were very few relapses in this study[Snee M, 1996].

Conservation Surgery

van der Sangen et al provided outcome data and data on method of detection of relapse for 86 patients who developed ipsilateral breast relapse more than 5 years after wide local excision, method of detection of relapse being unknown in 16 of their cohort[van der Sangen MJC et al., 2006]. In Montgomery et al, method of relapse detection and outcome was known for 108 of their cohort of 110 patients, the other two patients having had relapse detected incidentally[Montgomery DA et al., 2007b]. Outcome and method of relapse detection were known for all 12 patients in the Glasgow cohort. In total, therefore, method of detection of relapse and long term outcome is known for 206 patients.

Conservation surgery – ipsilateral breast relapse

Figure 3.1 shows survival from time of original operation for all patients with ipsilateral breast relapse only.

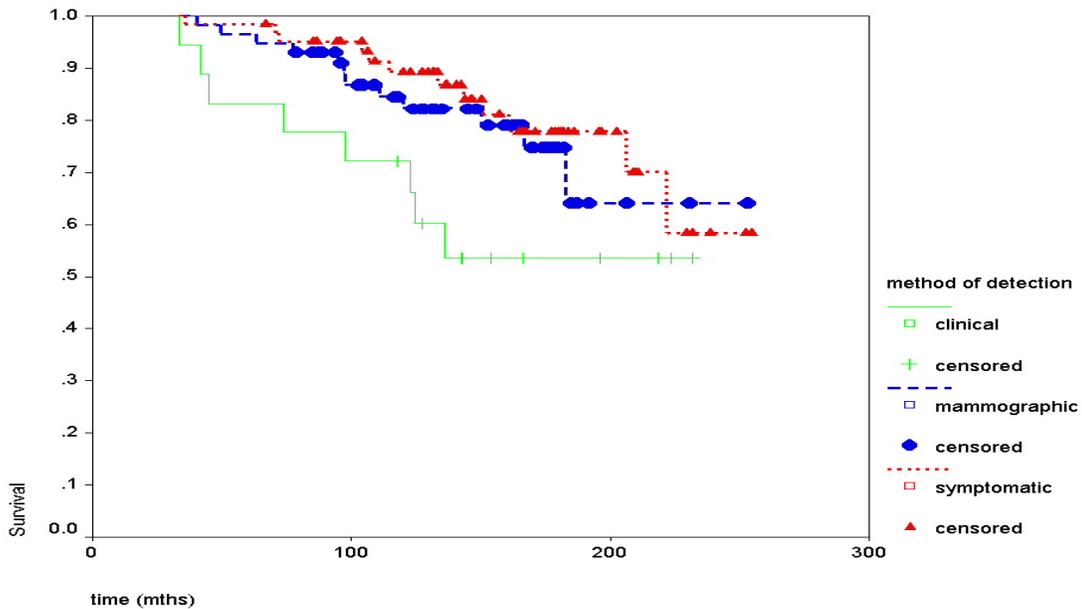


Figure 3.1: Survival from original operation in patients suffering ipsilateral breast relapse

In the cohorts we have analysed, patients with ipsilateral same breast relapse either noticed themselves or detected by mammography have a trend towards improved survival both from original operation and from detection of relapse. This was not significant on Kaplan Meier analysis.

Mean age at diagnosis of relapse was similar for all three methods of relapse detection, being 59.5 years (std deviation 13.66) for symptomatic, 60.68 (std deviation 10.46) for mammographic and 60.95 (std deviation 12.28) for clinically detected relapse.

Conservation surgery – ipsilateral axillary relapse

There were 27 patients who relapsed in the ipsilateral axilla alone. Survival from original diagnosis is shown in figure 3.2. There was no significant difference in overall survival related to method of detection of relapse. There was no difference in time to diagnosis of axillary relapse or in time from relapse to death among the methods of diagnosis either. Numbers were very small in each group.

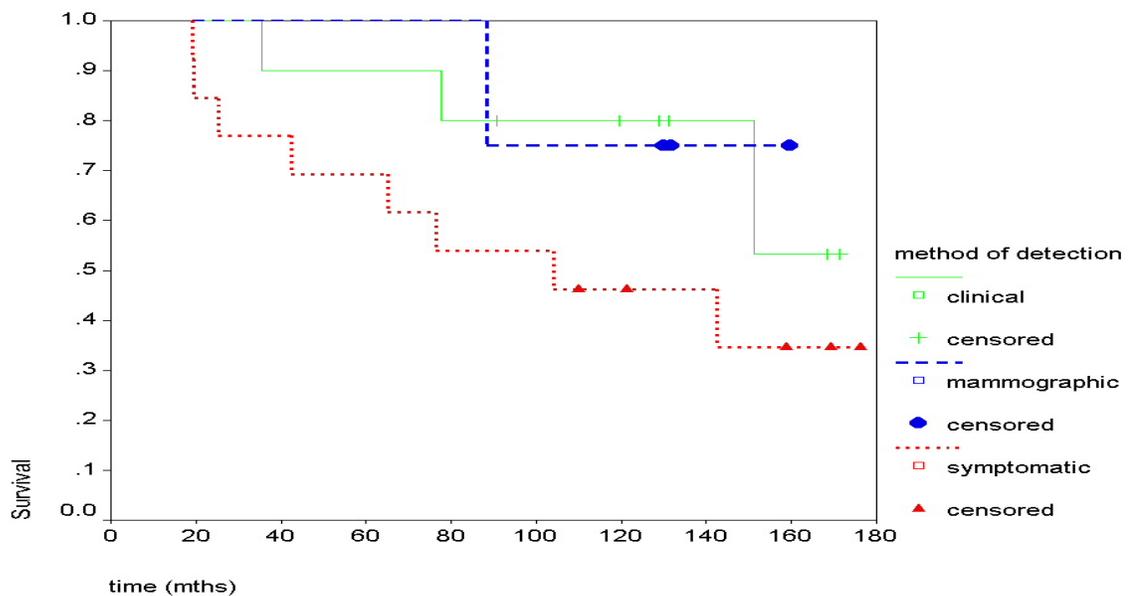


Figure 3.2: Survival from original operation in patients suffering axillary relapse.

Conservation surgery – contralateral breast relapse

There were 41 patients with a new contralateral breast cancer in whom method of detection was known. Overall 5 year survival from diagnosis of contralateral breast primary was 82.65%.

28 of these new contralateral cancers were detected mammographically and 5 year survival from diagnosis of new cancer in these patients was 86.5%. There were 9 new cancers detected by the patient, with a 5 year survival of 62.5%. All of the 4 women whose new cancer was detected first by clinical examination were alive and well at the time of analysis. These results are shown as figure 3.3. There was no significant difference in overall survival or survival from time of diagnosis of new contralateral cancer by method of diagnosis. The number of new cancers diagnosed by means other than mammography was small.

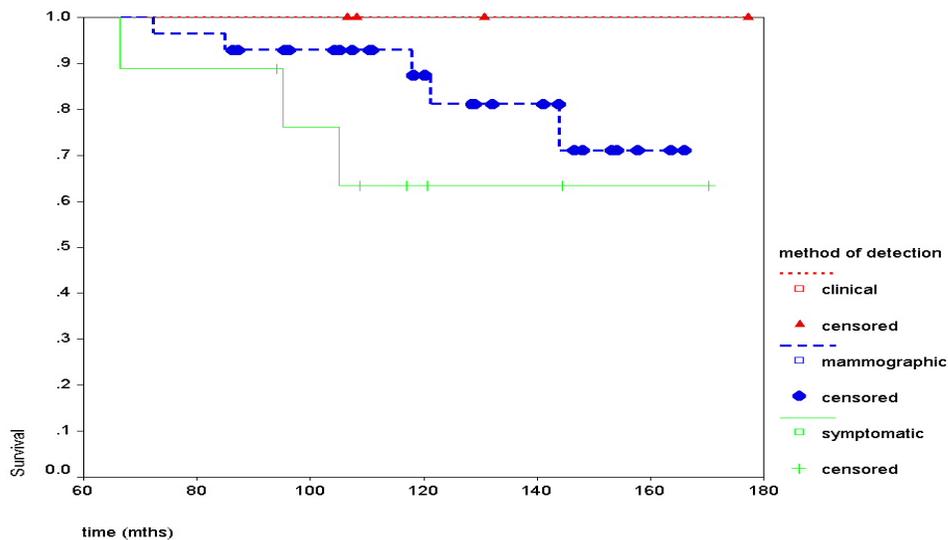


Figure 3.3: Survival from original operation in patients suffering contralateral breast relapse

Discussion

Follow up in breast cancer is an area of importance, both clinically and economically. Yet despite this, there is very little evidence within the literature to advise clinicians on how best to provide follow up care. Consequently, the multiple guidelines which have been published have very little concordance. Current guidelines published by the American Society of Clinical Oncology recommend frequent visits for routine clinical examination and

mammography for up to ten years after treatment [Khatcheressian JL et al., 2006]. It is recommended that patients attend for clinical examination three to four monthly for three years, four to six monthly for two further years and then annually for five more years. There is particular emphasis on follow up in the first three years after treatment as the rate of relapse has been reported to be particularly high at this time [The ATAC Trialists' Group: 2005; Hussain ST et al., 1995; Saphner T et al., 1996].

This schedule of routine clinic visits remains very labour intensive. The National Institute for Clinical Excellence (NICE) in England and Wales estimates the cost savings to the NHS to be around £3.7 million if follow up were limited to 5 years compared with ten years of follow up, and £9.3 million total savings if follow up were limited to just three years [National Institute for Clinical Excellence, 2002]. It is recommended in their most recent guidelines that follow up should be limited to the first two to three years after treatment followed by discharge to general practice [National Institute for Clinical Excellence, 2002]. NICE claim that recurrent breast cancer will usually cause symptoms and that these are almost always noticed by the patient themselves. The organisation claims further that the yield of mammography in follow up is low.

Most of the basic assumptions underpinning the NICE guidelines are in disagreement with the findings of this review. There certainly is a slight increase in the hazard rate for relapse in the first three years after treatment in several published studies [Hussain ST et al., 1995; Saphner T et al., 1996; The ATAC Trialists' Group: 2005]. It is likely that this represents partly a peak in distant relapse and partly a failure on behalf of other investigators

to include contralateral new breast cancers as a type of relapse in their analysis. In the analysis of the Edinburgh cohort outlined in chapter 2 of this thesis [Montgomery DA et al., 2007b], there is an initial peak in rate of distant relapse at around 3% of patients per year. This falls to around 2% per year where it remains constant for almost ten years. In contrast, potentially treatable relapse occurs at a constant rate of 1 – 1.5% per year for at least ten years [Montgomery DA et al., 2007b].

Studies which include metastatic disease in the analysis will report a higher initial rate of relapse which falls off with time. Metastatic disease does peak early before falling off [Montgomery DA et al., 2007b], and is twice as common as locoregional relapse in the Edinburgh analysis, at least in the first two to three years [Montgomery DA et al., 2007b]. If metastatic disease is included in an analysis of relapse then overall relapse will occur at a higher rate in the first two to three years before falling off. Yet the guidelines do not recommend routine screening for metastatic disease. If the aims of follow up as stated by NICE include the detection of potentially treatable locoregional disease only and not distant metastases [National Institute for Clinical Excellence, 2002], then the pattern of relapse should not be complicated by the inclusion of metastatic disease.

This issue will be complicated further if new contralateral cancers are excluded from the analysis or if analyses are conducted in cohorts with brief median follow up. The rate of new contralateral cancer rises with time [Freedman GM et al., 2005]. New contralateral cancer is often not included in analysis of relapse as technically it is not relapse. If follow up in breast cancer is intended to detect new breast cancers as early as possible, then new contralateral

cancers must be a focus. Patients are as likely to develop a new contralateral cancer as suffer from locoregional relapse over 15 years of follow up[Freedman GM et al., 2005;Veronesi U et al., 2001;Veronesi U et al., 2002]. So while studies which report on locoregional relapse only may show an early peak in disease, studies which include new contralateral cancers will more accurately reflect the true pattern of relapse.

The reasons for the discrepancy between the recommendations in the NICE guidelines and the findings of this analysis in terms of recommended duration of follow up are clear when the study on which NICE base their guidelines is examined. NICE base their recommendations for follow up on the findings of just one retrospective review[Donnelly J et al., 2001], a review which was not included in our analysis due to lack of clarity of the data. The authors of this review report method of detection of relapse in 67 patients who developed metastatic disease and 41 locoregional relapses or new contralateral cancers. Follow up was for a median of 3 years and 11 months. The authors did include new contralateral cancers, but only 4 patients relapsed this way. This study is therefore both too brief in terms of follow up, has too few patients and includes metastatic disease (and in fact a third more patients with metastases than potentially treatable relapse). The pattern of relapse in this cohort will not reflect the true pattern of potentially treatable relapse over the long term.

This review casts doubt over the value of clinical examination. Routine clinical examination is responsible for detection of only 14% of the relapses overall in the three most contemporary data sets analysed[van der Sangen MJC et al., 2006;Montgomery DA et al.,

2007b],Glasgow data]. The proportion was higher in the Glasgow cohort, although the numbers in that cohort are small and half of the relapses described as clinically detected were in fact also seen on mammography which was due at the same time as the clinic visit. This highlights an important issue when analysing pattern of relapse, which is that in centres where mammography is carried out before the clinic visit, more relapses will be detected that way. If the mammography is carried out after, as happened in Glasgow at that time, the reported rate of clinically detected relapse will be higher.

Relapse detected by clinical examination does not appear to be associated with improved outcome and, for ipsilateral breast relapse, may be associated with particularly poor outcome compared with relapse detected by other means. This analysis is clear cut in the Edinburgh cohort[Montgomery DA et al., 2007b], where survival after ipsilateral breast relapse is significantly poorer in those in whom the relapse is detected clinically. In this review, the issue was less clear cut. There certainly was a trend towards relapse detected clinically doing less well than that detected by other means, although this difference was not significant. In the study by van der Sangen and colleagues, all patients analysed had relapsed more than five years after the original diagnosis and treatment[van der Sangen MJC et al., 2006]. These patients made up a significant number of those whose survival after ipsilateral breast relapse was analysed. The authors of that study themselves report survival as being significantly better in patients whose relapse occurs after five years compared with those who relapse before five years. It is likely that late relapse has such a good prognosis that the method of detection of that relapse is irrelevant, so that inclusion of the patients from van der Sangen

and colleagues' study has simply confused the issue. Further work should be done to clarify this.

Although the NICE guidelines claim that mammography has a low yield during routine follow up, this review indicates that in both the conserved breast and the contralateral breast, the contribution of mammography appears not only to be important, but in fact may be of increasing importance. While early studies such as those by Mahoney and colleagues[Mahoney L, 1986] reported very few relapse detections using mammography, the proportion of relapses detected this way has increased so that in the studies published since 2000, up to 50% of all treatable breast relapses have been diagnosed first on mammography[Churn M and Kelly V, 2001;Montgomery DA et al., 2007b;Grogan M et al., 2002;van der Sangen MJC et al., 2006]. This was mirrored in the Glasgow cohort and, as mentioned above, 2 out of the 4 clinically detected relapses in our series had mammographically apparent relapse and were scheduled to have routine mammography as part of the visit where their relapse was detected.

This change in the impact of mammography is highlighted when comparing two cohorts from the same unit. The number of treatable relapses diagnosed by mammography in Edinburgh increased from 31% among patients diagnosed with relapse between 1986 and 1998[Jack WJL et al., 1998] to 46% in the more recent cohort diagnosed with relapse between 1991 and 2006[Montgomery DA et al., 2007b]. This is a consequence both of technical improvements in mammography and increasing experience in interpretation of post breast conserving surgery mammograms. As a result, mammography detected 5.37 new

cancers per thousand routine mammograms undertaken during follow up in the Edinburgh cohort[Montgomery DA et al., 2007b], which compares very favourably with the observed rate at incidence screening within the NHS breast screening programme in the UK[NHS Cancer Screening Programmes, 2003].

The proportion of relapses detected by patients, particularly in the treated breast and ipsilateral axilla, is fairly constant at 30 to 40% in all the studies analysed here. This applies less to contralateral breast relapse, where mammography has a much larger impact. This may reflect high pick up from mammography, less aggressive contralateral disease or simply a lack of patient awareness of the risk of contralateral breast relapse.

The difficulty faced in compiling recommendations for follow up is obvious. Much of the data in this review, and almost all of the data relating to survival by method of detection of relapse, has come from retrospective analysis of a local cohort which has not previously been published, unpublished data from other authors[van der Sangen MJC et al., 2006] or very recently published data analysed as part of this thesis[Montgomery DA et al., 2007b]. There are only 11 studies in the literature which fully present the pattern of relapse after breast cancer with regards to how that relapse is detected. Only one of these[Lees A et al., 1997] reports on more than 50 patients and only one[Snee M, 1996] has any outcome data. In this review, large datasets are analysed which give a more accurate representation of pattern of potentially treatable relapse and the impact of the various follow up strategies on detecting such relapse.

Summary

Very few analyses of pattern of relapse in large contemporary cohorts have been presented. From the data presented here, largely derived from analysis of local cohorts or unpublished data, we can make several observations and recommendations for follow up based on this analysis. Potentially treatable relapse is not common. Around 1-1.5% of women per year will suffer treatable relapse[Montgomery DA et al., 2007b]. Relapse occurs at a constant rate after treatment, so that the majority of relapses occur more than three years after treatment. Patients with later relapses can expect to do particularly well, and so effort should be made to diagnose later relapse at a treatable stage. If any follow up for the detection of treatable relapse is to be offered, this cannot stop at three years.

The need for clinical follow up for the detection of relapse is uncertain. The majority of relapses are now detected by patients or mammography. There are few relapses detected by clinical examination and, certainly in the case of ipsilateral breast relapse, those which are diagnosed clinically may do less well.

Mammography is less important in axillary relapse. Patients notice a similar proportion of axillary relapses themselves as they do relapse overall, although clinical examination picks up a large proportion of axillary relapse. Isolated axillary relapse is very uncommon, and it may be that better patient education could increase the proportion of such relapse detected by the patient. While most women are well schooled in breast self examination, it may be that the importance of axillary examination is less well appreciated.

Future guidelines should take these facts into account, but should also try to address the additional needs of patients during follow up for breast cancer. These include both detection of psychosocial problems and side effects of treatment which are central to maintaining patients well being. We are aware that these are currently poorly dealt with in the outpatient clinic.

Novel methods of follow up, with less emphasis on routine clinic visits for examination and more emphasis on detecting psychosocial problems, should be explored as a matter of urgency.

Section 2

Chapter 4:

Patients' Expectations for follow up

In Breast Cancer

Summary of the chapter

Section 1 of this thesis ends with the conclusion that routine clinical examination may be unnecessary for the detection of local relapse. This suggests that alternative methods of follow up could be explored and that routine attendance at the outpatient clinic could be avoided in many patients. Prior to the introduction of an alternative method of follow up, patients' expectations of follow up have been sought prospectively before their first routine clinic visit, and these expectations are presented here.

Introduction

Regular clinic visits for follow up after breast cancer remains a routine component of care. It is widely held that women find such follow up visits reassuring[Morris S et al., 1992], and expect to receive some form of follow up after breast cancer. 85% of women questioned by one group of investigators at the end of a clinic visit reported that they would prefer to continue coming to routine clinics rather than simply return if they noticed problems[Morris S et al., 1992].

On the other hand, some women find routine clinic visits stressful. Many women report feelings of anxiety in the week before attending clinic as they are worried that they may be diagnosed with relapse[Pennery E and Mallet J, 2000]. Other investigators have quantified

the proportion of women who suffer such anxiety, reporting that 71% of women suffer some distress during routine follow up[Paradiso A et al., 1995].

A careful balance needs to be struck between improving quality of life through regular feedback, support and reassurance after treatment for breast cancer and temporarily reducing quality of life through bringing patients back to what many see as a stressful routine clinic appointment. From the very limited number of randomised controlled trials which have been performed, it is difficult to be sure that balance is successfully struck by the current practice of providing routine follow up care. Certainly, reducing or removing routine clinics altogether does not appear to alter significantly patients' quality of life either for the better or worse (see chapter 1). It is likely that the reassurance which routine clinics provide is every bit as short lived as the stress which comes before them.

Routine clinics are not an effective method of detecting psychological problems in women, partly due to unwillingness on the part of women to discuss problems in the outpatient setting[Valente SM et al., 1994;Pennery E and Mallet J, 2000] and partly as a result of physicians' inability to spot problems in such a setting[Baildam AD et al., 2004]. Any improvement in quality of life, temporary or otherwise, which comes as a result of routine clinic visits must be due to the reassurance that a clinic visit where no relapse is detected brings. Yet the preceding chapters of this thesis reveal that such reassurance is largely false, as clinic visits contribute little to the detection of relapse[Montgomery DA et al., 2007b].

There has been little work published on what women expect in terms of follow up after breast cancer. Most of the studies regarding expectations for follow up have been conducted in women already undergoing routine breast cancer follow up, who therefore have set ideas about what they expect in terms of frequency and duration of visits. It is less clear whether women have pre-conceived ideas regarding the length and type of follow up they will receive before embarking on their routine follow up.

It is also unclear what women expect their clinician to achieve at routine clinic visits. As has been described above, follow up after breast cancer has several aims, including relapse detection, provision of psychological support and detection and amelioration of side effects of treatment[National Institute for Clinical Excellence, 2002]. It is unclear which of these aims women see as components of the process, and whether women are aware of the limitations of routine clinic visits in terms of meeting these various aims.

There is pressure to reduce the total amount of follow up provided for women after breast cancer in the UK[National Institute for Clinical Excellence, 2002]. The preceding chapters of this thesis confirm that reduced follow up is likely to be safe in terms of long term outcome, as routine follow up contributes little to detection of relapse. It is less clear how acceptable this limited follow up would be to patients.

In order to explore this issue, a survey of the expectations of women for their follow up provision was undertaken.

Methods

Study Design

A questionnaire based survey of the expectations for follow up was undertaken. The study was approved by the Glasgow Royal Infirmary Local Research Ethics committee.

Section Criteria

Women were considered eligible for this questionnaire based study if they met the following criteria:

- Breast cancer treated with curative intent
- Disease free at the time of enrolment
- No history of previously treated breast cancer
- Less than one year from original treatment
- Not yet attended for routine follow up clinic visit
- Ability to consent to completing the questionnaire

Methods

In our unit, patients are seen by the surgical staff shortly after their original operative procedure to discuss the pathology findings and prepare them to meet the oncology staff. Patients are then referred to the oncologists for provision of adjuvant chemo-radiotherapy, if appropriate. A routine follow up appointment is arranged for the following year, and patients are not routinely seen by the surgical team again until this first annual review appointment. Information on the follow up process is not normally given until the first annual review appointment, where the whole process of follow up is explained to the patient.

In December 2005, a complete list of all patients operated on with curative intent for breast cancer in the preceding twelve months was generated. For the following 7 months a list of all women undergoing curative resection of a breast cancer was maintained prospectively. Therefore all patients attending for their first annual review appointment between January 2006 and July 2007 were invited to participate in this study.

Letters were sent to all patients due to attend for their first annual review appointment between two and three weeks before that appointment. Along with the questionnaire, a letter outlining the purpose of the questionnaire was sent, as was a consent form approved by the local research ethics committee. Stamped addressed envelopes were also included to facilitate response. The questionnaire was designed by this author in consultation with several senior consultant breast surgeons in Glasgow. The full questionnaire is reproduced as appendix 4.1.

Method of detection of the original tumour was asked as this may influence patients' confidence in their ability to self monitor for relapse. Expectations for frequency and duration of follow up were asked and opinion on whether more frequent follow up would lead to greater anxiety or greater reassurance was sought. Traditional follow up regimes, such as advocated by the American Society for Clinical Oncology [Khatcheressian JL et al., 2006], have reducing frequency of visits over ten years as the perceived threat of relapse reduces. We were keen to establish whether reducing the frequency of clinic visits would worry patients, or reassure them as they perceive their clinicians level of anxiety over their

risk of relapse to be reducing. The final questions on page one of the questionnaire were designed to establish what the patients felt the main purpose of follow up clinics were and how useful they thought routine clinic visits were for detecting relapse compared with their own self-examination.

It was felt that providing information on the success of clinics in terms of relapse detection may influence patients' opinions. An information sheet was produced which explained in some detail that the likelihood of suffering relapse for most women was low. In addition, the common methods by which relapses are detected in our unit were outlined. To establish whether the provision of this information altered patients' opinions about clinic attendance, the patients were asked whether they were still keen to come to clinic despite all they had read. Finally, they were asked to suggest how they would like to be followed up from a number of choices including telephone follow up and traditional clinic visits.

Results

Over the 18 months of the study, 102 patients were eligible for participation. Letters were sent to all of these women, and responses received from 79, a 77% response rate. Mean age of the respondents was 59 years (SD 14 yrs, range 26 – 86 yrs).

Method of detection of original cancer

62 women (78.5%) were diagnosed with breast cancer after originally finding a lump themselves during breast self examination. 4 women (5.1%) had been diagnosed by their General Practitioner (GP) after attending with other breast related symptoms. 3 women

(3.8%) had been to their GP with other non-breast related complaints and had their breast cancer diagnosed incidentally. The remaining 9 women (11.4%) had their cancer detected within the breast screening service.

Expectations for frequency and duration of follow up

74 of the 79 women (94%) expected to be seen back at the clinic for more follow up after completion of their adjuvant chemo-radiotherapy. One woman did not answer the question, leaving only 4 women who expected not to have to come back to clinic for further follow up visits beyond their first annual review. Three of these four women detected their own cancer originally, one was discovered in the breast screening service.

The frequency with which patients expected to be seen back at clinic is shown in table 4.1.

| Frequency of clinic visits | No of patients (%) |
|-----------------------------------|---------------------------|
| Three monthly | 13 (16.5%) |
| Six monthly | 30 (38%) |
| Annual | 29 (36.7%) |
| Less than annual | 1 (1.3%) |
| Not answered | 6 (7.6%) |

Table 4.1: Expected frequency of clinic visits

Of the four women who did not expect to return to clinic for further follow up, three did not answer the question regarding how frequently this should occur. One patient stated that they expected annual follow up. The patient who did not answer the question of whether she expected to come back for more clinics did not respond to this question on frequency of visit either.

There was a difference in expectations for follow up between women over 50 years old compared with those under 50. There was a very even distribution in patients under the age of 50 with regards to how frequently they expected to be brought back, with 7 patients expecting three monthly visits, 6 patients expecting six monthly visits and 7 expecting 12 monthly visits to the clinic. This was not the case in patients over 50, where only 6 expected to return 3 monthly compared with 24 expecting 6 monthly visits and 22 expecting annual visits (χ^2 3 df $p=0.048$).

In question four, patients were asked when they expected to be discharged from clinic. Expectations regarding the duration of follow up varied remarkably between patients. Of the four women who stated in question 2 that they did not expect to return to the clinic for more follow up, one patient stated that she thought she would be discharged at her next clinic visit (consistent with not expecting any more follow up) and one stated that she would be discharged “when she no longer requires follow up”. The other two patients did not answer this question.

Of the 74 patients who did expect more follow up, 12 (16%) did not answer how long they expected this to continue for and 17 (23%) state that they did not know when they would be discharged. 3 patients (4%) suggested that they would be discharged after all of their treatment including adjuvant hormone therapy was completed, but gave no indication as to how long they felt this would take.

8 patients (11%) expected between 1 and 3 more years of follow up, 6 patients (8%) between 3 and 5 more years and 17 (23%) expected exactly 5 more years of follow up, with 3 of these stating that this would coincide with the end of their tamoxifen treatment. Only 4 patients (5%) expected ten more years of follow up, the current duration of follow up provided in our unit, although one other patient thought she would receive follow up visits indefinitely. Finally, two patients thought they would be discharged when the surgeon was happy with their progress and two when they were considered by the surgeon to be “cured” or “in remission”.

Expectations of what clinics are designed to achieve

All 79 women answered the question of whether being brought back to clinic more regularly would increase or decrease their anxiety. 13 women (16%) stated that more regular clinic visits would increase their anxiety. Of these women, 8 expected only annual review visits, 4 expected six monthly visits and one had not answered the question above on how frequently they expected to be brought back to clinic. 51 of the 79 women (65%) felt that they would be more reassured by more regular clinic visits, and 15 (19%) stated that they would be neither more anxious nor more reassured by more regular clinic attendance. There was no significant association between the age of the patient and how this question was answered. Similarly, there was no difference between patients who had discovered their own cancer and patients whose original cancer was detected by other means in terms of whether more clinic visits would provide more or less reassurance.

The women were asked the main reason/s they thought they were brought back to clinic. 8 patients selected more than one option, and one patient did not answer the question. 46 women (58%) thought that the main reason they attended clinic regularly was to catch the cancer early when it came back, with 8 others (10%) reporting that it was one of the main reasons. Therefore 54 women (68%) in total saw detection of relapse as a main aim of routine clinic visits. Only 10 patients (13%) saw detection of side effects of treatment as one of the main reasons for clinic visits. 24 patients (30%) felt that provision of reassurance was one of the main aims of regular clinic visits. Again, method of detection of original cancer did not affect the way this question was answered.

In keeping with the feeling that regular clinic visits were focussed mainly on the detection of relapse, 43 women (54%) thought that the clinician was more likely to detect relapse than they were. 33 patients (42%) believed that they were more likely to detect relapse than their doctor at the clinic. The remaining 3 patients did not answer the question. When this was analysed by method of detection of original relapse, women who had not detected their original cancer themselves were significantly more likely to expect a clinician to detect any subsequent relapse than those women who had detected their own original cancer (χ^2 1 df $p=0.004$).

Effect of information on expectations

Patients were then shown the enclosed information sheet before completing the final two questions. The information sheet stressed that routine clinics had a very low yield in terms of detecting recurrent cancer, and that regular self examination and mammography were the

best safeguards against relapse. Patients were then asked whether they would still like to come to clinic, knowing what they had just read. They were assured that they would still have regular mammograms whether they elected to come to clinic or not.

The majority of patients (49 of the 79, or 62%) reported that they would still prefer to come to clinic regularly. Only 27 of the 79 patients (34%) reported that they would be happy not to come to clinic regularly. 3 patients did not answer the question. Interestingly, there was no relationship between whether the patient wanted to continue coming back to clinic and who they thought was most likely to detect relapse if it occurred, nor was there a relationship with how the original cancer had been diagnosed.

When asked what method of follow up they should have instead of routine clinic visits, 20 of the patients who were happy not to come back routinely suggested that they could come and see us only if they had a problem, one suggested GP or nurse follow up and 5 replied that they would be happy to receive telephone follow up.

Discussion

Recommendations for follow up have changed dramatically in the past few years in the United Kingdom. From multiple tests and investigations carried out at frequent clinic visits for a prolonged number of years, we have now reached a situation where follow up is recommended for only 2 to 3 years after treatment for breast cancer in the most recent guidance from the National Institute for Clinical Excellence (NICE)[National Institute for Clinical Excellence, 2002].

Yet the majority of clinicians involved in follow up in the UK do not adhere to these guidelines. A survey of 562 clinicians registered in the Cancer Research UK Clinical Trials Unit database of clinicians involved in breast cancer follow up revealed that, while a protocol existed for follow up of breast cancer patients in the units of 84% of respondents, in only 9% of cases did this protocol conform to the NICE guidelines[Donnelly P et al., 2007].

There are many reasons why clinicians would be reluctant to discharge patients from their care at 3 years. In the study by Donnelly and colleagues, detection of side effects of treatment, detection of new abnormalities in either breast and detection of psychological morbidity, in that order, were the three most commonly cited reasons for providing follow up[Donnelly P et al., 2007]. We know that psychological problems after breast cancer can persist for up to ten years[Hodgkinson K et al., 2006] and we know from the evidence in chapter 2 of this thesis that the rate of potentially treatable relapse in either breast is constant to at least ten years and probably beyond. Yet routine clinic visits may not be necessary for the detection of such abnormalities, with mammography and patient self examination being much more central to the detection of locoregional relapse[Montgomery DA et al., 2007b]. Such visits are probably not effective at detecting psychological problems or symptoms in patients as women are reluctant to discuss such issues and doctors have been shown to be poor at detecting these problems within the clinic setting[Baildam AD et al., 2004;Pennery E and Mallet J, 2000;Valente SM et al., 1994].

While several investigators have asked women already undergoing follow up whether they find routine clinic visits reassuring, very little work has been done on perceptions of women prior to embarking on their follow up. It is often assumed that women find clinic visits reassuring and want to come for such visits, but this research has usually been conducted in women who are part of routine follow up programmes[Kiebert GM et al., 1993;Morris S et al., 1992]. This is the first time research has been conducted on perceptions of follow up in women prior to commencing routine follow up, or being informed of the protocol.

The majority of women in this study did expect some further follow up, but there was tremendous variation in what was expected. The majority of respondents expected either 6 or 12 monthly clinic visits, but a small proportion thought they would be brought back more frequently than this. Over half of the patients questioned expected to be brought back to clinic more frequently than annually, the current frequency adopted in our unit.

Interestingly, the majority of women felt that more frequent clinic visits would provide more reassurance. This is somewhat contrary to our own experience with reducing frequency of follow up in practice. In the past, patients have often associated a reduction in the frequency of their visits as some indication that our anxiety over their risk of relapse is lessening, and they often report feeling that reducing the frequency with which they attend clinic is a positive step towards the “all clear”. It is therefore surprising to see such a high proportion of women keen to receive higher frequency visits.

Women were equally uncertain about how long they would remain in follow up. 39% of women either did not know or did not answer, and only 5% of women expected to be followed up for ten years, the duration of follow up currently provided in our unit. So while women usually have an expectation of the amount of follow up they will receive, their expectation is that the visits will be more frequent than is currently provided but for a much briefer duration.

Women in this study consider detection of relapse to be the most important reason for coming back for routine follow up visits, with 68% stating this is the most important reason for follow up. Given the expectations that women have for frequency and duration of follow up, it is possible to surmise that women have a perception that their risk of relapse is high in the first few years after treatment of their original cancer but that the risk falls off. While we know that to be the pattern of relapse for metastatic disease (see chapter 2), routine clinic visits are not usually focussed on the detection of distant disease. We know that the risk of potentially treatable relapse does not fall off after five years.

In this study, we did not highlight the pattern of relapse to our patients. It would be interesting to see whether being told that the risk of potentially treatable relapse remains constant for at least ten years and probably beyond altered any of the findings in this study.

The most commonly cited reason for being brought back to clinic was detection of relapse, with 68% of women seeing this as a main reason for providing follow up. While this was the most frequently cited reason for providing follow up, a smaller proportion of women cited it

as an aim of follow up than did clinicians in the study by Donnelly and colleagues, where more than 80% of respondents felt that detecting an new breast abnormality early was a reason for bringing patients back to clinic[Donnelly P et al., 2007]. Perhaps one of the reasons for this disparity is the high proportion of women (42%) who felt that they would be likely to detect any relapse before their doctor.

While the majority of clinicians in Donnelly and colleagues' study felt that detection of side effects of treatment and provision of psychological support were central aims of follow up, only 13% of our patients saw side effect detection as important and only 30% felt that provision of reassurance was an aim of the clinic visits. It is difficult to know how much this view is related to the experiences women in our study have had. At the time of completing the questionnaire, none of our patients had been back to a routine surgical follow up clinic. Many will have attended a post-operative results clinic, and will have attended oncology clinics for planning and provision of therapy, and perhaps for a routine oncology follow up clinic. It would be disappointing if the perceptions of patients in this study about our interest in any side effects of therapy they were suffering were based on their experiences of coming to the few results/treatment clinics that they had attended prior to completing this questionnaire. It is equally possible that women simply expect to monitor themselves for side effects, and feel they would come back to tell us if they had problems rather than require diagnosis of those problems in routine outpatient clinics.

Patients appear to have very different expectations from clinicians regarding what we are trying to achieve with follow up. If patients do not see the clinic as being geared towards

detecting side effects of treatment and/or psychological problems, then they may be less willing to divulge such problems. This would explain why routine clinic visits have been so poor at detecting such problems in the past. And if such a high proportion of patients expect to find their own relapse, then it may be that these patients would be open to alternative methods of follow up where they self monitor for relapse and for other problems such as side effects of treatment, perhaps receiving some remote screening to prompt self evaluation, and attend outpatient clinics only if they detect problems.

And yet, when provided with confirmatory evidence of the relatively ineffective nature of clinics in terms of relapse detection, the majority of patients (62%) stated that they would still prefer to come to clinic, even 19 patients who believed that they were more likely to detect their own relapse than a clinician.

This study highlights a number of important issues, and raises several questions. Women expect some follow up, but there is no consistency regarding how frequently or for how long they expect such follow up to take place. Even though the majority of women see relapse detection as important, provision of the information that routine clinic visits are not effective at detecting such relapse does not deter all women from wanting to come back to clinic, even among those who feel that they are more likely to detect their own relapse than the clinician. What is important to note from this study is that more than one third of women at the end of their first year of treatment would be happy not coming back to clinic when provided with the information that such clinics are not useful for detecting relapse. In these women, there is certainly scope for introducing alternative follow up methods more geared towards

psychosocial concerns and side effects of therapy. It may even be that more women would be willing to accept an alternative follow up method if a more structured option was presented to them. In the present study, no structured alternative was offered and many women may have been concerned at being discharged from follow up altogether with no opportunity to get back in touch quickly should the need arise. This was not explored directly in the present study.

There is some discrepancy between what patients expect follow up to achieve and what clinicians believe are the central purposes of such follow up. Patients are much less likely than clinicians to cite detection of psychological problems or side effects of therapy as being reasons for follow up. In order to help achieve these additional aims of follow up, such as provision of psychological support and detection of side effects of treatment, then patients will need more education regarding the purpose of follow up, or clinicians will require to search for problems in a more pro-active way than is the case at present.

Summary

This study highlights that women are uncertain about the structure and goals of follow up prior to their first annual review appointment. Their expectations are often at odds with those of clinicians, and this may hinder clinicians in achieving some of the central aims of follow up.

A large proportion of women would be open to alternative follow up methods even at such an early period after diagnosis and treatment, particularly when educated regarding the limited

use of clinic visits in terms of relapse detection, and this proportion may increase if a well structured alternative was presented to them.

Chapter 5:

Automated Telephone Follow-up After Breast Cancer

Design and introduction of a new system

Summary of the chapter

The preceding chapters of this thesis cast doubt on the necessity of routine follow up after breast cancer for the detection of potentially treatable relapse. In addition, such routine visits do not seem to be the best way of detecting symptoms and side effects of adjuvant treatment or psychological problems in women.

In this chapter, a prospective cohort based feasibility/acceptability study of a novel method of follow up is presented.

Introduction

Routine clinical examination is an inefficient way of detecting potentially treatable relapse after breast cancer[de Bock GH et al., 2004]. Local control after breast conserving surgery and modern radiotherapy is excellent, so event rates are low. As reported in chapter 2, locoregional relapse, including new contralateral cancers, occurs with a rate of 1-1.5% of patients per year in a contemporary cohort[Montgomery DA et al., 2007b]. Such low event rates will inevitably result in any follow up method being considered inefficient if detecting potentially treatable relapse is considered the main outcome of follow up.

Older studies looking how relapse was detected after treatment for breast cancer reported that clinical examination was responsible for detecting up to 73% of potentially treatable relapses[Hussain ST et al., 1995]. It was felt that clinical examination would result in earlier detection of relapse and potentially lead to survival benefit in women who suffer relapse. And while this has never before been addressed formally in the literature (see chapter 3), in women who develop a metachronous contralateral tumour during follow up, at least one study has shown that that tumour will be smaller at diagnosis when detected by a clinician compared with when diagnosed by the patient themselves[Samant RS et al., 2001]. This mirrors the findings reported in chapter 2 above, where, in the Edinburgh cohort, relapse was smaller when clinically detected than when detected by the patient (see chapter 2, table 2.4). Consequently, despite routine clinical examination being an inefficient way to provide follow up, such examination was believed to be the main method by which potentially treatable relapse was diagnosed and so examination remained a central component of follow up regimes [Khatcheressian JL et al., 2006;Grunfeld E et al., 2005].

Our analyses presented in the preceding chapters have revealed a number of new and important considerations. Clinical examination is becoming less relevant in detecting potentially treatable relapse. In recent studies, as few as 13% of all relapses are detected by the clinician during routine examination[van der Sangen MJC et al., 2006] (see chapter 3). Improvements in mammographic technology make this investigation far more important than routine clinical examination, and patients themselves are well able to self-survey for relapse.

In this thesis, an analysis of outcome by method of detection of relapse has been performed. This has never before been conducted in the literature (see chapter 3). While it has been confirmed again that recurrence detected by routine clinical examination may be smaller than that which is detected by the patient, this does not appear to translate into a survival advantage (see chapter 2)[Montgomery DA et al., 2007b]. In fact, despite a slightly higher average Nottingham Prognostic Index of the relapse in patients who detect their own relapse compared with those which are diagnosed clinically, patients with clinically detected ipsilateral relapse do less well than relapse detected either by mammography or by the patient[Montgomery DA et al., 2007b]. From our analyses, routine clinical examination is an inefficient and probably unnecessary component of follow up after breast cancer. Such routine examinations detect very small numbers of relapses, and those which are detected appear to do worse than those relapses detected by the patient or by mammography.

Detection of potentially treatable relapse is not the sole purpose of follow up provision. Detecting psychological problems and side effects of treatment are both important considerations[National Institute for Clinical Excellence, 2002]. Yet there is little evidence that routine clinical follow up meets these additional aims. As described in chapter one, few randomised trials of follow up exist. Those trials which have compared traditional follow up with some novel alternative have revealed no reduction in quality of life when an alternative method of follow up is provided. None of these trials has reported improvement in quality of life with alternative follow up either. Notably, removing routine follow up altogether and replacing it with clinic visits on demand does not appear to be associated with any detrimental effect on patients' psychological well-being[Gulliford T et al., 1997]. From this,

it is possible to suggest that the clinic visits are having little impact on patients' psychological well-being.

An alternative follow up method which has beneficial effects in terms of improved detection of psychological problems and which is more cost effective than traditional follow up has not yet been found (see chapter 1). Follow up by breast care nurses may detect more psychological problems, but there is no reduction in cost due to the seniority of the nursing staff involved[Baildam AD et al., 2004]. Moreover, the patients still come up to clinic so the same levels of hospital resources are needed.

Follow up by telephone has been tried in other disciplines. James and colleagues reported good outcomes in terms of detecting medical problems and providing support to oncology patients using a nurse led telephone system[James ND et al., 1994]. While these authors did report a 30% reduction in outpatient workload as a consequence of this method of follow up, training of the nursing staff was required and nurses were needed to undertake the telephone call to the patients. This method of follow up was acceptable to the patients and provided good levels of support, yet there was no comment made regarding the overall economic impact of this type of follow up. As we have seen before, follow up based on experienced nursing staff may be no cheaper than traditional methods[Baildam AD et al., 2004].

Automated telephone follow up has potential benefits in terms of reducing clinic visits, reducing workload for clinical staff and improving detection of psychological problems in women after breast cancer. Such follow up could be used as a screening tool to detect

women with psychosocial problems or side effects of therapy after treatment for breast cancer, and clinic visits could then be targeted to those who require such visits. This could be done using an automated computer telephone system to further save staff resources.

A computer telephone system combines the computer and telephone in order to receive process and store or transmit information. There is some overlap between computer telephone systems and other telephone based healthcare such as telemedicine or NHS 24. In computer telephone systems, automation can replace person-to-person telephone contact, thus removing the need for a trained staff member to call the patient.

Follow up using an automated telephone system has been used in several fields. Millard and Carver utilised voice recognition software to administer the SF-12 health status survey to patients with back pain, and concluded that, compared with live telephone administration by a nurse, the voice recognition software resulted in more accurate reporting of negative emotional issues due to the more impersonal nature of the automated system, and could reduce the cost of data entry and of labour[Millard RW and Carver JR, 1999]. A systematic review of uses of automated telephone computer systems concluded that the reliability and availability of such systems made them attractive for use in follow up, but that little evidence for their use in this setting existed and that further trials were needed[Biem HJ et al., 2003].

In this chapter, an account of the design and development of an electronic patient record incorporating an automated telephone follow up system, and its use in an initial feasibility

study in a group of women currently receiving follow up after treatment for breast cancer, is presented.

Methods

An electronic case record (Excelicare™) with linked telephone system (Excelicare™ Direct) is available from AxSys Technology, Glasgow, UK. Excelicare™ is a toolset based application that allows the creation of tailored clinical systems. Excelicare™ therefore provides a platform for developing an electronic patient record relevant to breast cancer patients but which can later stand alongside and interact with clinical modules developed by other specialties, potentially resulting in the long term in a fully electronic, multi-disciplinary patient record. This company have recently won the contract to provide a generic clinical system to the NHS in Scotland, the ultimate goal of which is to construct a complete electronic case record.

In addition to the Excelicare™ patient record, AxSys technology produces a linked telephone system, Excelicare™ Direct, which can communicate with the Excelicare™, allowing data to be entered into screens within the Excelicare™ patient record over the telephone. A patient phoning in to the Excelicare™ Direct system will hear a pre-recorded greeting before being asked to enter their unique pin number. This gives them access to a pre-recorded questionnaire. This questionnaire is answered using digit capture, where the telephone keypad (keys 0 to 9, * and #) can be used to navigate through menus or provide answers which can then be stored to the patient electronic record and accessed at any time by the patients clinicians. In addition, the computer system can be programmed to calculate or derive scores according to the answers given in the questionnaire and undertake specific

actions in the event of certain answers being given or scores being achieved. For example, an acceptable score can result in a reassuring letter being sent to the patient, and a request for a routine mammogram being sent to radiology. Poor scores or deterioration from the last recorded score can result in an email being sent to a clinician.

Excelicare™ and Excelicare™ Direct were chosen by this author to develop a fully integrated electronic patient record and telephone system which would be used not only to provide a complete record of the patients diagnosis and treatment, but which would also be used for much of the follow up for each patient and automate some of the clerical work associated with follow up in patients. The development of the generic clinical system is at a relatively early stage, with isolated applications being used in various specialties across Scotland. This system had never been utilized in breast cancer prior to this report. As a result, all components of the breast cancer patient record and follow up system required to be designed by this author prior to patient recruitment. As described elsewhere, programmers develop applications for computer telephone systems by first assessing the needs of the client then flow charting, voice file scripting, programming (in this case both the phone system and the electronic patient record) and piloting the new system before implementing and assessing the system in real patients. Throughout the development of this project, this author was involved in all of these processes as described below.

Designing the Electronic Patient Record

Initially, meetings with an AxSys project manager were arranged to assess how the electronic patient record and telephone system had been developed for use in other specialties. In particular, the system is currently utilised in obstetrics and gynaecology at Glasgow Royal

Infirmery for conducting multi-disciplinary meetings. The Excelicare™ Direct telephone software is also utilised by the rheumatologists for increasing drug compliance and reducing side effects of therapy. These initial meetings allowed this author to assess the functionality of this system and its appropriateness for use in the follow up of breast cancer patients. Additionally, these meetings generated ideas for how to design the patient record and the telephone system in order that they would be applicable to breast cancer patients.

Needs assessment for this project was complicated by the multi-disciplinary nature of breast cancer care. In order to maximise the usefulness of the new system, and ensure that a fully integrated system was introduced which could follow the patient through all aspects of their breast cancer care and provide clinical details of each step in the diagnosis and treatment to all authorised clinicians, it was felt that the record should contain data pertaining to more than just the immediate diagnosis and treatment of the breast cancer. The system was designed to store data from all aspects of the patients' breast cancer diagnosis and treatment.

Meetings were arranged with heads of department in oncology, pathology, radiology and plastic surgery to discuss the needs of these departments with regards to such a system. Aspects that each department considered important were integrated. More specific aspects of care were given separate screens within the system. Input was sought from several senior breast consultants across the city as to content of the system before this author finally designed screens to cover all aspects of the breast cancer diagnosis and treatment pathway. Each screen constituted a 'page' in the electronic case record. Separate screens were designed for each of the following areas:

- Patient demographic and contact details, along with initial referral details.
- Initial clinical investigations conducted to reach diagnosis or to stage disease, if done.
- Radiological investigations for initial diagnosis, including staging investigations if done.
- Waiting times and planning data.
- Nursing audit, particularly to record contact with breast care nurses.
- Initial surgery, including options for further excisions or additional surgical procedures as necessary.
- Plastic surgery procedures including reconstruction or contralateral reshaping.
- Subsequent revision/contralateral plastic surgery procedures.
- Pathology records.
- Radiotherapy record.
- Chemotherapy record.
- Additional therapy such as hormonal.
- Screens to record the diagnosis and subsequent treatment of local recurrence or metastatic disease.
- Details on death of the patient.

In addition to these screens, several screens were designed to compliment the telephone component. On each patient's record, copies of the complete telephone script were stored. Telephoning in populated the data fields in one of these scripts and generated a summary page (see below). These screens also required to be designed and tested. Finally, in order to facilitate automation of the system, a final screen was designed with links to standard letters.

These standard letters could be accessed and printed off by secretaries or clinicians in response to an event, such as the patient telephoning in and completing the questionnaire.

Subsequent to designing these screens, the finished items were then shown to the original contributors in the various specialties. Further amendments as necessary were made before the design templates were given to AxSys technology to construct the electronic case record.

Further meetings were then held over a three month period to ensure that the content of our electronic case record complied with that required by the various organisations involved in data collection after breast cancer. In particular, the requirements of the managed clinical network (MCN) in breast cancer were considered. Several other individuals were involved at this stage for their expertise in the requirements of these external organisations. These individuals were Paul Burton, clinical effectiveness co-ordinator at Glasgow Royal Infirmary, Eileen Kesson, data manager, Flora MacInnes, an NHS IT project manager at Glasgow Royal Infirmary, and the AxSys technology project manager who provided technical advice. During these meetings, the final version of the electronic case record was agreed in such a way that they would be both useful clinically and able to provide all the data required by MCN and other external auditors.

During the building of the system, links were added to various existing software systems to ensure easy transfer of data into the new patient records. These included links to Patient Administration systems (PAS) allowing patients demographic and contact details to be pulled into the system along with details on general practitioners etc as well as links to the

Chemocare and Varis systems allowing chemotherapy and radiotherapy details to be automatically pulled across for all patients without this data needing to be re-typed.

Finally, the completed electronic case record was delivered by AxSys technology. Full testing of the system was carried out by this author to ensure that the content was correct and that there was complete functionality of all aspects including the links to external software. Problems at this stage were corrected by AxSys technology and again another round of complete testing was carried out. Complete testing of the final patient record was completed by this author, Paul Burton, Eileen Kesson and Flora McInnes independently and a final meeting arranged to sign off on the project and accept delivery of the system.

Developing the Telephone Questionnaire

As stated above, the aims of follow up as described by the National Institute for Clinical Excellence (NICE) are detection of potentially treatable relapse, detection and amelioration of side effects of therapy and psychological support in the aftermath of diagnosis and treatment[National Institute for Clinical Excellence, 2002]. NICE stress the importance in particular of identifying lymphoedema as this symptom can have a marked psychological impact, as well as being very troubling physically.

This first aim, of detecting locoregional relapse, can be adequately achieved by regular mammography and patient education on self examination in our patient group. The need for regular self examination should be highlighted to patients by the questionnaire, thus reinforcing the need to carry out this examination. It is also important that an opportunity to

report changes within the breasts or axillae should be given within the questionnaire. In order to facilitate both of these requirements, seven questions were designed purely to encourage self examination and allow the reporting of symptoms which may be associated with relapse, either locoregional or metastatic. This part of the questionnaire had not previously been validated, and is reproduced as appendix 5.1 below.

The main objective of this telephone follow up system is to improve detection of psychosocial concerns among our patient, as well as to detect treatment related side effects. In selecting a questionnaire for use in this study, a number of characteristics were sought. Detection of psychosocial concerns requires a detailed questionnaire capable of detecting not only anxiety and/or depression in patients, but also of detecting more specific breast cancer related problems and concerns. The more general anxiety and depression questionnaires which exist, such as the hospital anxiety and depression scale, are insufficiently specific to detect breast cancer related anxiety in our patients. These more general questionnaires have been shown to be non-specific in relation to breast cancer and often insensitive to problems in breast cancer patients[Hall A et al., 1999].

In addition to being fairly specific for breast cancer related psychosocial problems, the questionnaire should be able to detect women who are having problems with side effects from treatment. In particular, lymphoedema, other arm symptoms such as stiffness or poor mobility secondary to axillary procedures and side effects of endocrine therapy are considered to be important areas which the questionnaire should be sensitive to. These are all areas highlighted as important and these should be addressed when providing follow up

for breast cancer patients[National Institute for Clinical Excellence, 2002]. The final questionnaire should have a very strong focus on detecting problems in these areas specifically.

Finally, the questionnaire should be as concise as possible. This will ensure that it can be completed by patients in a reasonable time, thus hopefully maximising compliance and minimising patient inconvenience.

As described in the general introduction to this thesis, questionnaires do exist which have been designed specifically to measure quality of life after breast cancer. After a brief review of the literature, it was considered that two of the more widely used and better validated questionnaires were the Functional Assessment of Cancer Therapy-Breast (FACT-B) quality of life instrument and the European Organization for Research and Treatment of Cancer Breast Cancer-specific quality of life questionnaire (EORTC-QLQ-C30/BR23)[Brady MJ et al., 1997;Sprangers MAG et al., 1996]. These two separate questionnaires have been used for assessing quality of life in a number of breast cancer trials in recent years.

The benefits of the FACT-B questionnaire have been described in the introduction, but will be briefly reiterated. The FACT-B questionnaire was developed to provide a robust and reproducible way of measuring quality of life in patients during and after treatment for breast cancer[Brady MJ et al., 1997]. It is a 44 item self report questionnaire designed to measure general quality of life issues as well as more specific issues related to breast cancer, and it has been extensively used in both research and clinical practice. Since it was first

introduced, additional subscales have been developed to measure arm and endocrine symptoms in patients who have had treatment for breast cancer.

The arm subscale consists of 4 additional items used in conjunction with the FACT-B and has been shown to be sensitive to deterioration in arm function associated with the development of lymphoedema, as well as being sensitive to improvement in function associated with effective treatment[Coster S et al., 2001].

The endocrine subscale consists of 18 items designed to detect endocrine symptoms related to adjuvant treatment with agents such as tamoxifen or the aromatase inhibitors. It is also useful in detecting symptoms in particularly younger women who have been rendered menopausal either as a deliberate act during ovarian ablation or oophorectomy or as a consequence of premature menopause secondary to chemotherapy[Fallowfield LJ et al., 1999].

Both subscales have been developed to work in conjunction with FACT-B to provide a well integrated, simple to use questionnaire using a 5 point Lickert-type response scale. FACT-B with both subscales provides a comprehensive assessment of the physical and psychological health of women after a diagnosis of breast cancer, with particular emphasis on the areas considered particularly important by NICE in their most recent guidelines. Moreover, it was designed with extensive consultation with women previously treated for breast cancer and as such should accurately represent the concerns of this patient group. As such, FACT-B and the arm and endocrine subscales should be ideally suited as a screening tool for physical and

psychological problems after treatment for breast cancer, and were chosen for use in this study for these reasons. The complete questionnaire, including both the arm and endocrine subscales, as it was presented to the patient is shown in appendix 5.2.

The FACT-B questionnaire is usually administered on paper. This author is unaware of any previous attempt to use FACT-B with computer telephone technology. The layout of FACT-B makes it simple to convert to use on the telephone. The Lickert scale of the original questionnaire means that questions are answered by circling a number from 0 to 5, depending on how much the patient agrees with a given statement. For the automated system introduced in this study, the questionnaire can instead be answered by pressing the corresponding number on the telephone keypad of an ordinary touch-tone keypad.

Prior to the questionnaire script being programmed as a voice file on the Excelicare™ Direct system, the length of time taken by the patient to listen to this author read the questions and then respond to the question by pressing a button on the keypad was established using a small focus group of around ten routine follow up patients. This confirmed that the complete FACT-B questionnaire with arm and endocrine subscales, along with the seven initial questions to encourage self examination, could be completed in less than ten minutes. This was considered acceptable by this small group. The FACT-B questionnaire with arm and endocrine subscales was therefore considered to meet all of the requirements for the questionnaire.

The scoring of the FACT-B questionnaire is well structured. The questionnaire consists of a number of statements, grouped into five areas: Physical well being, Social/family well being, Emotional well being, Functional well being and the breast cancer subscale. In addition, as mentioned above, the arm and endocrine subscales were also used in this study, giving seven sections in total in this study. For each statement, the patient decides how closely that applies to them over the past seven days, answering from not at all (0 points) to very much (4 points). All positive statement scores are added, negative statement scores are reversed and then added. The higher the score overall the better the patient is doing. In addition to scores for each section, several derived scores can be calculated by the addition of different sections of the questionnaire.

Using the Excelicare™ and Excelicare™ Direct system in this study, every time a patient telephones in and completes a questionnaire, the answers are stored in the patient's record and the questionnaire script and answers can then be viewed by accessing that patient's record. The computer telephone system was designed to calculate the scores for each section as well as overall and derived scores as mentioned above. These summary scores are displayed in the patient record in a file along with the complete questionnaire answers. For the seven initial questions (appendix 5.1), answering any of them 'yes' results in this being flagged up on the summary page also. Therefore clinicians can look at the summary page to assess scores, or enter beyond this to analyse individual answers.

The function of all of these components of the telephone system, including the scoring system, again required to be tested by this author. Problems at this stage again were

corrected by AxSys technology before further testing was carried out until the complete application was functioning.

Study Design

The final step in the development of the computed telephone follow up system is to implement the system and assess the acceptability and feasibility of this type of follow up. A prospective cohort study of automated telephone follow up was undertaken. The study was approved by the local research ethics committee at Glasgow Royal Infirmary.

In addition, a smaller validation study was undertaken to ensure equivalence of telephone administration of the FACT-B questionnaire and traditional paper administration of the questionnaire.

Inclusion and Exclusion Criteria

All patients with a previous history of breast cancer treated with curative intent who had completed their adjuvant chemo-radiotherapy and who attended a routine follow up clinic between 1st May and 31st August 2006 were asked to participate in the main feasibility study. For the smaller validation study, all patients with a previous history of breast cancer treated with curative intent attending a routine follow up clinic between 15th March and 7th June 2007 were asked to participate.

The following exclusion criteria were applied:

- Presence of local or distant disease at time of clinic. (Successfully treated previous locoregional recurrence was not an exclusion criterion).
- Inability to complete the questionnaire on paper independently.
- No access to a telephone.
- Inability to consent to participation.

Several patients felt that, due to hearing difficulties or other physical disabilities, they would be unable to complete telephone based follow up themselves. Several of these women were keen to participate in this form of follow up and suggested that they be sent written versions of the questionnaire and have a relative telephone their answers in for them after completion on paper. As such, inability to complete the questionnaire by phone due to physical difficulties was not an exclusion criterion if a paper version could be completed and the patient had assistance from a willing family member to complete the questionnaire over the phone.

Study Protocol – Acceptability and feasibility Study

Follow up is conducted on an annual basis within our unit. All patients attending a routine clinic appointment over a four-month period between 1st May and 31st August 2006 were invited to participate in the acceptability/feasibility study. At the end of a routine scheduled clinic visit, the nature of the study was explained and patients were offered to complete their next follow up visit one year later using the remote telephone system.

Patients who agreed to participate in the acceptability/feasibility study completed the full FACT-B questionnaire including arm and endocrine subscales on paper before leaving the clinic in order to obtain a baseline score. They were then informed that they would be written to ten months later and asked to complete the same questionnaire over the telephone. The automated nature of the system was explained to them in detail. All participants were informed that they could withdraw consent and come for a standard clinic visit at any time, and were informed that this author would telephone them after they had completed the telephone follow up to ensure they had no problems and to obtain feedback on the system.

Ten months after recruitment these women were written to, inviting them to telephone into the system. Full instructions on how to use the telephone system are included on the telephone system, but these instructions were backed up by full written instructions on the letter. The complete questionnaire was sent to all women so they could go over their answers beforehand if they wished. Finally, the telephone number of the unit secretary and this author were included for women who had questions or wanted to now withdraw consent and return for a standard clinic appointment.

A reminder letter was generated for patients who had not called within two weeks of the original letter being sent. Failure to respond after this second letter resulted in the woman being telephoned to ensure that the letters had been received and to remind them to call, or to offer a regular clinic appointment if they preferred.

All patients were telephoned by the author within one week of completing the telephone questionnaire. A semi-structured interview was used, with the following questions being asked to establish how acceptable this type of follow up was:

- Did you find the system easy to use?
- What, if anything, did you like about this type of follow up, compared with coming up to the clinic?
- What, if anything, did you not like about this type of follow up?
- Do you have any ideas regarding how it could be improved?
- Do you feel sufficiently reassured by this type of follow up? (Remembering that you would get feedback on how you scored and will also get a mammogram in the next few weeks)
- Do you feel that the telephone questionnaire gave you more, less or just about the same opportunity to tell us about things that are worrying you compared with coming to clinic?
- Would you be happy to use this as your only method of follow up, along with the mammograms each year?

Patients were sent a copy of the full FACT-B questionnaire on paper along with the letter asking them to phone in, as stated above. During the semi-structured interview, they were asked whether they had completed this beforehand and used it as a prompt when phoning in their answers. They were also asked if they had completed the questionnaire on paper prior to having a relative telephone in their answers.

Any areas of low score or significant deterioration from the baseline score in any part of the questionnaire were discussed. Previous studies have suggested that the minimally important change in score in any section of the questionnaire which is associated with a true deterioration in symptoms important to the patient is 5% of the total available score for that section[Cella D et al., 2002;Eton DT et al., 2004]. This was taken to be the case for this initial study. If the patient answered yes to any of the initial seven questions, this was also discussed. Problems which could not be resolved over the telephone resulted in an appointment being sent for the patient to come to clinic. Patients were also brought back to clinic if they requested to come for clinical examination. Finally, all patients were sent appointments to attend for mammography, the results of which were communicated to the women by letter subsequent to the mammography being conducted.

Study Protocol – Validation Study

The FACT-B questionnaire has been extensively validated in the past, and the various components have been shown to be sensitive to change over time, yet stable over short time periods with good test-retest reliability[Fallowfield LJ et al., 1999;Coster S et al., 2001;Brady MJ et al., 1997]. All of the previous validation work has been carried out using paper versions of the questionnaire. This is the first time of which we are aware that the FACT-B and the additional subscales has been administered using a computer telephone system.

The computer telephone method of administration differs from the paper version filled out in clinic in three important ways. Firstly, the computer telephone version is completed by the patient at home, away from the potentially stressful environment of the hospital clinic. This

may alter how the patient is feeling and may lead to a difference in score even over such a short time period. Secondly, in our computer telephone version there is no scope to miss out questions, unlike when completing the questionnaire on paper. The paper version of FACT-B and the additional subscales allows scores to be calculated for each section by multiplying the score for each section by the number of questions and then dividing this by the number of questions answered. It is likely that the fewer questions which are answered then the less accurate will be the score. There may be discrepancy between the scores on the telephone version and on the paper version in patients who miss out a high proportion of questions when filling out the paper version. Finally, there is some evidence that patients using voice recognition software to complete questionnaires on health related issues are more likely to acknowledge mental interference, emotional concerns and poor overall mood and health when using the voice recognition software than when completing questionnaires on paper in clinic[Millard RW and Carver JR, 1999]. It has been proposed that, because automation eliminates the need to talk to a real person, patients are likely to be giving a more accurate representation of their feelings. This phenomenon has been reported in psychiatry, especially in the field of drug addiction, where subjects may provide more honest responses when using interactive voice recognition systems compared with when questioned in a live interview setting, even by telephone[Turner CF et al., 1998;Kobak KA et al., 1997].

This telephone computer follow up system will be used as a general screening test and overall scores on the questionnaires used will not be taken as diagnostic of problems in the live setting. The FACT-B questionnaire and the arm and endocrine subscales being used in this system have usually been completed on paper. Even within this study, all of the patients

recruited completed the questionnaire on paper initially to generate baseline scores. Due to the potential differences between telephone and paper administration as highlighted above, it was considered important to establish the correlation between paper and computer telephone administration of the FACT-B questionnaire including the arm and endocrine subscales.

In the past, test-retest reliability has been conducted in small groups of around 29 to 32 patients in previous studies[Coster S et al., 2001;Brady MJ et al., 1997], although the reliability of the endocrine subscale was tested on a slightly larger cohort of 56 women[Fallowfield LJ et al., 1999]. These small cohort studies showed the FACT-B questionnaire, the arm subscale and the endocrine subscale to be very stable over short periods of time, with high test-retest correlation.

To test correlation between paper administration and computer telephone administration, a second cohort of patients was recruited between 15th March and 7th June 2007. These patients completed the FACT-B questionnaire with arm and endocrine subscales on paper at the end of a clinic visit in an identical manner to the patients recruited to the acceptability/feasibility study. They were given written and verbal instructions on the use of the telephone system, as well as a copy of the questionnaire on paper, again just as had been the women recruited to the acceptability/feasibility study. They were asked to telephone in and complete the questionnaire between 3 and 5 days later. Non-respondents at this stage were telephoned on the evening of day four after the clinic visit to remind them to telephone. Non-respondents at that stage were discounted from the validation analysis.

All patients in the smaller validation study were telephoned by the author within one week of completing the telephone questionnaire in the same way as for the acceptability/feasibility study cohort. The same semi-structured interview was used, with the questions as described above. These patients were given a copy of the full FACT-B questionnaire on paper when they left the clinic visit. During the semi-structured interview, they were also asked whether they had completed this beforehand and used it as a prompt when phoning in their answers. They were also asked if they had completed the questionnaire on paper prior to having a relative telephone in their answers.

Results

Acceptability and feasibility of telephone follow up

Over the four month recruitment period from 1st May until 31st August 2006, 121 patients who were suitable for inclusion attended clinic visits. 110 of these 121 patients (91%) agreed to participate and completed a baseline questionnaire. The mean age of participants was 62 years (SD 11.8 years), with those refusing to participate having a mean age of 68 years (SD 7.4 years), (p=NS). The characteristics of the included patients are shown in table 5.1. For patients with more than one procedure to the same breast, the most recent procedure is included. For patients with bilateral disease, the poorest prognosis pathology is included. One patient, not included in the table, had a single scalp lesion resected which was pathologically an oestrogen receptor positive breast cancer. Following treatment she has now been disease free for 5 years in routine follow up and was keen to give feedback on the new telephone system.

| | | |
|---------------------------|---|--|
| Age (years) | Mean SD Range | 62 years 11.8 years 35 – 87 years |
| Tumour grade | DCIS High grade phylloides 1 2 3 unknown | No. of patients 2 3 21 38 40 6 |
| Node status | Positive Negative unknown | No. of patients 40 62 8 |
| Procedure | Conservation surgery Mastectomy Mastectomy/reconstruction Bilateral procedures | No. of patients 63 32 11 4 |
| Adjuvant hormonal therapy | Never Previously Current Stopped between paper and phone versions | No. of patients 37 33 37 3 |
| Time in follow up | Median Range | Time (yrs) 5.16 yrs 1 to 20 |

Table 5.1: Characteristics of the patients enrolled into the feasibility/acceptability of automated telephone follow up study

Mean scores with standard deviations for the baseline questionnaire recorded on paper are shown in table 5.2.

| | PWB | SWB | EWB | FWB | BCS | ESS | ARM | Fact B | Fact G | Fact ES |
|--|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|--------------------|-------------------|--------------------|
| No Hormone therapy (n=37) | 22.45 sd 5.67 | 24.00 sd 4.99 | 18.43 sd 5.09 | 20.70 sd 5.58 | 22.39 sd 7.81 | 55.72 sd 11.35 | 15.35 sd 4.73 | 107.97 sd 22.84 | 85.58 sd 16.05 | 141.56 sd 24.71 |
| Previous Hormone therapy (n=33) | 25.28 sd 2.66 | 24.41 sd 3.67 | 20.28 sd 2.76 | 22.34 sd 4.39 | 26.48 sd 4.13 | 61.35 sd 8.40 | 16.68 sd 3.86 | 116.97 sd 13.75 | 90.65 sd 11.74 | 152.02 sd 14.46 |
| Current Hormone therapy (n=37) | 23.61 sd 4.66 | 24.22 sd 4.69 | 19.91 sd 2.83 | 22.03 sd 5.38 | 25.91 sd 6.91 | 55.38 sd 11.08 | 16.82 sd 4.02 | 112.80 sd 25.36 | 87.54 sd 19.53 | 141.61 sd 31.28 |
| Overall (n=3) | 23.71 sd 4.68 | 24.21 sd 4.48 | 19.50 sd 3.80 | 21.66 sd 5.19 | 24.89 sd 6.74 | 57.34 sd 10.68 | 16.27 sd 4.25 | 112.34 sd 21.81 | 87.76 sd 16.38 | 144.64 sd 25.30 |
| One way Anova | P=0.041 | P=0.933 | P=0.095 | P=0.373 | P=0.018 | P=0.032 | P=0.264 | P=0.236 | P=0.447 | P=0.155 |

Table 5.2: Mean scores for baseline questionnaire delivered on paper for patients in the feasibility/acceptability study

PWB = Physical Well being (7 items); SWB = Social Well Being (7 items); EWB = Emotional Well Being (6 items); FWB = Functional Well Being (7 items); BCS = Breast Cancer Subscale (9 items); ESS = Endocrine SubScale (18 items); ARM = arm subscale (5 items) Fact B = sum of PWB, SWB, EWB, FWB, BCS; Fact G = sum of PWB, SWB, EWB, FWB; Fact ES = sum of PWB, SWB, EWB, FWB, BCS, ESS.

In March 2007, the 110 patients who had agreed to participate were cross referenced with our recurrence database. 5 patients had died since enrolment. This left 105 patients who were written to over the following two months with instructions to telephone in and complete a follow up questionnaire over the telephone.

Willingness to participate

Of the 105 patients written to, 75 patients (71%) telephoned in and completed a follow up questionnaire. Of the patients who did not call in to complete the questionnaire over the telephone, 3 (3%) were uncontactable by either letter or telephone, and so were considered lost to follow up and 5 (5%) did not complete the questionnaire even after telephone reminder to do so but gave no reason for non-participation.

Of the remaining 21 patients who did not complete the questionnaire over the telephone, 8 (8%) could not manage to use the telephone system for technical reasons, with 4 of them stating that the system was engaged or simply rang out. 3 patients (3%) stated that they were “too old” for the system and did not want to try it. 2 (2%) felt that their hearing had deteriorated too much since agreeing to participate and that they would be unable to hear the

questions, 3 (3%) had inter-current illness that they felt would interfere with the interpretation of the results, one patient claimed that she did not have the time, one patient felt that it was too detailed with too many questions and one patient stated that she does not want to think about her breast cancer in such detail. Only three patients did not want to complete the questionnaire because they preferred to come to clinic for examination.

Patients who decided not to telephone in to the system were on average significantly older than those who did try the system. The mean age of patients who called and completed the follow up questionnaire using the computer telephone system was 59 years (sd 11 yrs) while those who chose not to call the system had a mean age of 67 years (sd 11 yrs) ($p=0.002$). This issue was analysed further in an attempt to establish a cut off point where women would usually be successful in using the system. 66.66% of women born between 1930 and 1939 successfully completed the telephone version, with only 30% of patients born between 1920 and 1929 successfully completing the system (χ^2 1 df $p=0.046$).

Ease of use and acceptability of the telephone system

Mean scores with standard deviations are shown in table 5.3 for the 75 questionnaires completed using the computer telephone system.

| | PWB | SWB | EWB | FWB | BCS | ESS | ARM | Fact B | Fact G | Fact ES |
|--|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|----------------------|-------------------|--------------------|
| No Hormone therapy (n=24) | 23.83 sd 2.86 | 22.56 sd 4.28 | 17.54 sd 4.51 | 19.75 sd 6.17 | 22.09 sd 5.56 | 56.05 sd 9.97 | 14.39 sd 4.83 | 105.15 sd 19.35 | 83.69 sd 14.78 | 138.88 sd 23.86 |
| Previous Hormone therapy (n=26) | 24.69 sd 4.92 | 23.95 sd 4.71 | 19.35 sd 3.65 | 20.46 sd 5.31 | 23.8 sd 6.12 | 59.65 sd 13.02 | 15.88 sd 5.40 | 111.64 sd 15.85 | 88.45 sd 12.92 | 146.96 sd 19.84 |
| Current Hormone therapy (n=25) | 24.35 sd 2.86 | 22.89 sd 5.73 | 19.88 sd 3.06 | 21.80 sd 4.80 | 24.28 sd 6.57 | 53.26 sd 11.87 | 46.88 sd 3.43 | 113.1987 sd 16.92 | 88.92 sd 12.33 | 142.72 sd 21.97 |
| Overall (n=75) | 24.30 sd 4.14 | 23.15 sd 4.92 | 18.94 sd 3.85 | 20.68 sd 5.44 | 23.42 sd 6.10 | 56.34 sd 11.92 | 15.75 sd 4.67 | 110.13 sd 14.48 | 87.08 sd 13.39 | 143.10 sd 21.73 |
| One way Anova | 0.768 | 0.583 | 0.083 | 0.411 | 0.435 | 0.192 | 0.181 | 0.247 | 0.323 | 0.491 |

Table 5.3: Mean scores for follow up questionnaire completed using the computer telephone system for patients in the feasibility/acceptability study

PWB = Physical Well being (7 items); SWB = Social Well Being (7 items); EWB = Emotional Well Being (6 items); FWB = Functional Well Being (7 items); BCS = Breast Cancer Subscale (9 items); ESS = Endocrine SubScale (18 items); ARM = arm subscale (5 items); Fact B = sum of PWB, SWB, EWB, FWB, BCS; Fact G = sum of PWB, SWB, EWB, FWB; Fact ES = sum of PWB, SWB, EWB, FWB, BCS, ESS.

71 of the 75 patients who completed the questionnaire found the computer telephone system easy to use and had no trouble with the technology whatsoever. Two patients had telephones with the keypad on the handset, and this complicated the process. Both managed to complete the questionnaire successfully despite this. Two further patients were anxious about using this technology, so completed the questionnaire on paper and had a member of their family call in the results. All patients stated that having the questionnaire on paper in front of them when using the computer telephone service reduced the anxiety associated with this new technology and made the process far easier.

49 of the 75 patients (65.33%) who completed the computer telephone follow up questionnaire liked the system and stated that they would be happy to use it instead of routine clinic visits as their method of follow up along with regular mammography. A further 9 patients (12%) liked many aspects of the computer telephone system and stated that they would be happy to use it as a component of their follow up, although all 9 of these patients stated that they appreciated the reassurance of regular clinical examination in addition. These 9 patients made several suggestions for the use of this new system. These included completing the questionnaire before they came to clinic in order that their clinic visit could

be more focussed on problems they had raised in the questionnaire and alternating between telephone follow-up and clinic visits each year to reduce the frequency of clinic visits. One patient suggested coupling this new system with clinical examination by the general practitioner.

Patients cited a number of benefits of this follow up method. In particular, time saved by the patient and the patients' families was considered greatly beneficial. One patient stated that she was able to complete the questionnaire with her daughter and have her daughter call it in during the evening, thus saving her daughter from taking a day off work to take her to clinic. She additionally stated that coming to the hospital and trying to find a parking space added greatly to the stress of an already stressful visit, a concern mirrored by several others. Several patients stated that not having to take time off work themselves was of benefit to them.

A number of patients reported that they found the telephone follow up method much less stressful than coming to clinic. They were able to complete the questions in their home environment, which they found less threatening, more relaxing and less likely to drag up memories of their treatment. As a result of feeling less stressed, they felt more able to think about how they were feeling in detail, particularly as they completed the paper version of the questionnaire over a longer period of time than they usually got to see their doctor in clinic before phoning in to record their answers on to the computer system. A number of patients cited additional factors which limited their communication of problems within the clinic. Reasons for not being able to communicate problems in clinic included feeling embarrassed,

feeling rushed or being so anxious that they forgot to mention things that had been concerning them. As a result of alleviation of all of these factors, 29 of the 49 patients who liked the system felt it gave them much more opportunity to divulge problems than was the case in clinic.

Most patients described the questionnaire as being very detailed. This allowed them not only to report all the problems which they had encountered, but several women commented that, as a result of participating in this study, they felt much more aware of the type of problems which they may encounter in the future and they were now more confident that they would notice if something went wrong. These patients actually felt that they had been given back control over their condition and felt empowered by this form of follow up.

12 patients (16%) stated that they would not be happy to use this type of follow up. 2 of these had medical problems ongoing at the time of completing the telephone questionnaire and suggested that their anxieties over these problems may have affected their view of the system. 4 patients had technical problems with the system or were nervous about the new technology, including 2 patients who completed the questionnaire on paper and got their daughter to call in. These patients actually felt less in control than they did when attending clinic as they had to rely on family members to convey problems. Interestingly, 2 patients stated that the questionnaire had too much detail and they did not want to think about their breast cancer treatment in the amount of detail required to complete this type of follow up. The 4 remaining patients in this group stated that they would prefer to come to clinic for an examination.

3 patients (4%) were undecided about whether they would like to use this type of follow up. 2 of these patients were at any early stage in their follow up and felt that they would like to keep coming to clinic for a few more years until things were a bit more stable. One did not feel that the questionnaire asked the right kind of questions, but felt that clinic was not of much benefit to her either. In fact, this patient had a very detailed conversation with this investigator over poor scores in a number of areas of the questionnaire and it is likely that her level of anxiety required additional input. By the end of the conversation she did concede that perhaps longer clinic appointments to discuss her problems were needed, and that it had been the new telephone system which had uncovered this fact.

Finally, one patient completed the telephone questionnaire after several attempts to contact her by post and by phone. She was uncontactable after completing the questionnaire either by post or by telephone and subsequently did not attend a clinic appointment. Her thoughts about the system have not been recorded.

Correlation between paper and telephone administration of the questionnaires

Between the 15th March and 7th June 2007, 40 patients were recruited to this smaller validation section of the study to explore the degree of correlation between scores obtained using the telephone system and those obtained when the FACT-B questionnaire with arm and endocrine subscales is completed on paper.

All 40 patients filled in questionnaires on paper at baseline. 26 patients (65%) completed the telephone version of the questionnaire between 3 and 5 days after completing the paper version. The 14 non-responders were excluded from the analysis.

Prior to calculating the degree of correlation between the methods of administration of the questionnaire, it is important to ensure that any relationship between the two methods is linear and that there is no systematic bias, such as the telephone giving consistently lower or higher scores. A perfectly linear relationship with one of the methods scoring consistently lower would result in a high degree of correlation, but high correlation in such a situation would not confirm equivalence of the two methods of administration.

To explore the relationship prior to calculating the degree of correlation, each individual's score from the questionnaire completed on paper while in the clinic was plotted against their score from the questionnaire completed using the telephone system. These scatter plots are shown below. The line of equality ($x = y$) is super-imposed on the plot. The line of equality represents perfect correlation between the two methods of administration of the questionnaire. If the two methods of administration are equivalent, there should be a linear relationship with an equal spread of points above and below the line of equality.

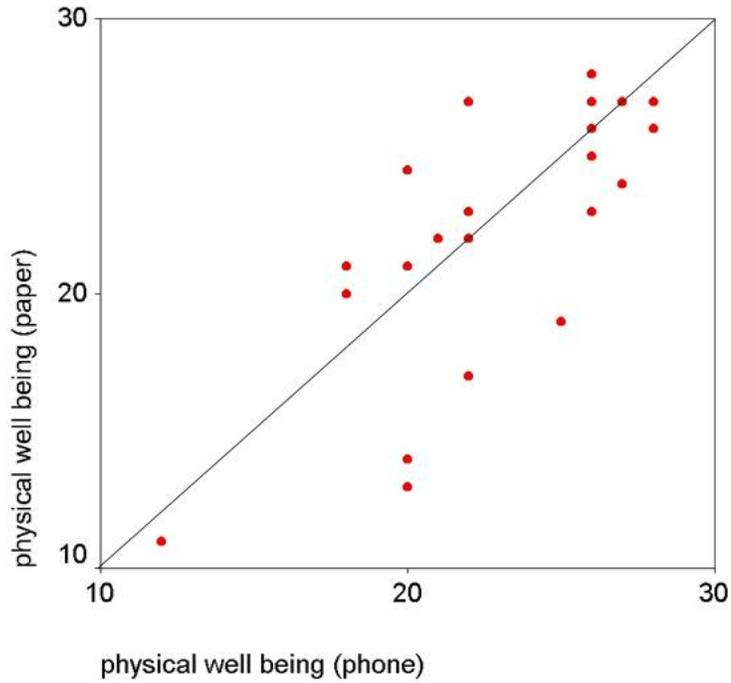


Figure 5.1: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – physical well being. The line of equality super-imposed represents a line of perfect correlation.

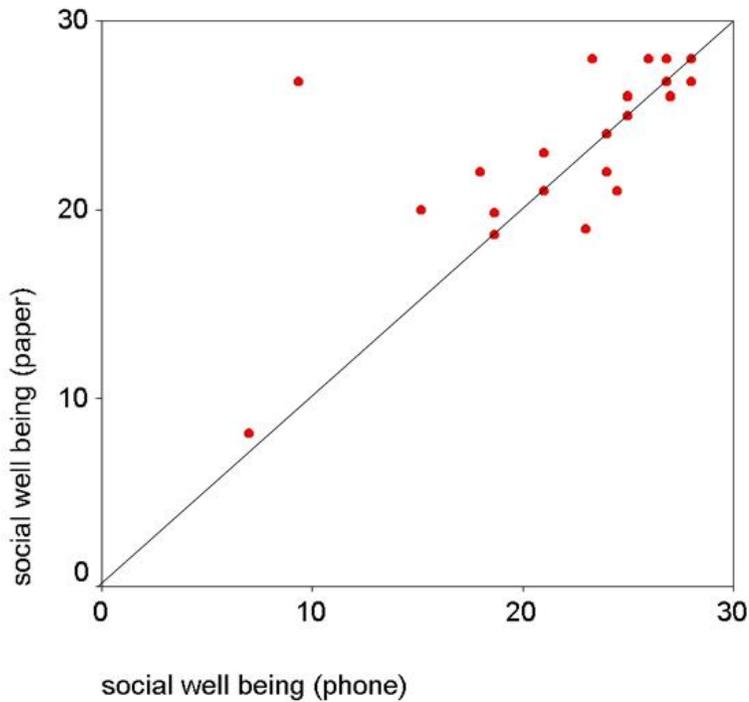


Figure 5.2: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – social well being. The line of equality super-imposed represents a line of perfect correlation.

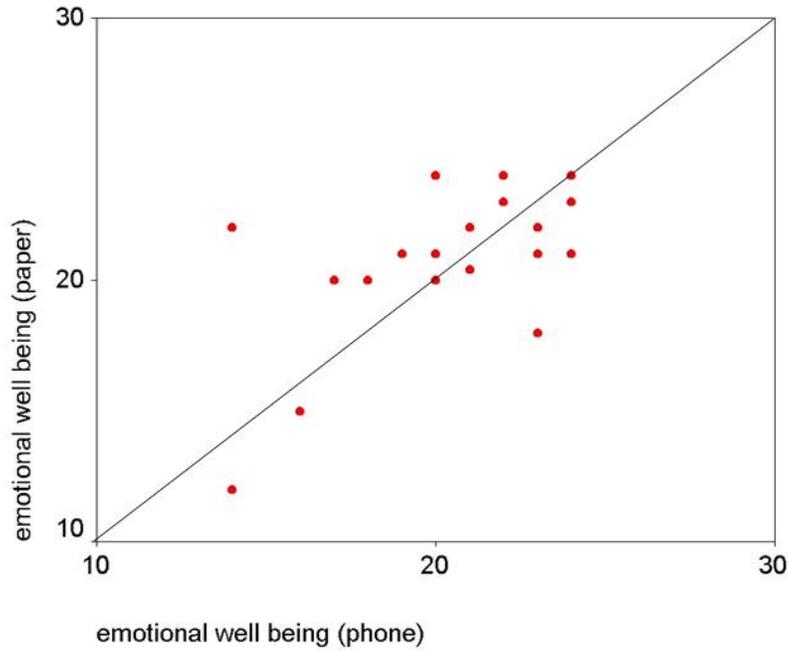


Figure 5.3: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – emotional well being. The line of equality super-imposed represents a line of perfect correlation.

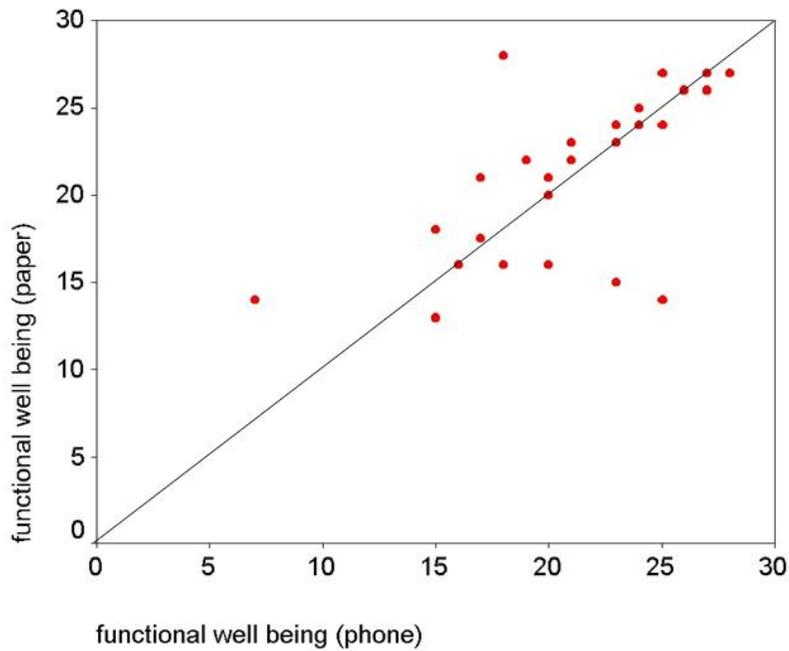


Figure 5.4: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – functional well being. The line of equality super-imposed represents a line of perfect correlation.

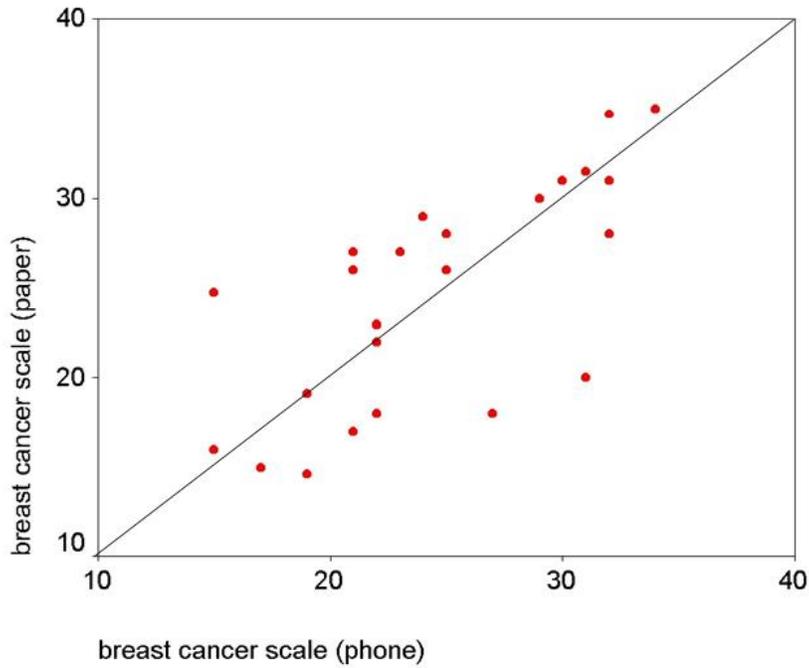


Figure 5.5: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – breast cancer subscale. The line of equality super-imposed represents a line of perfect correlation.

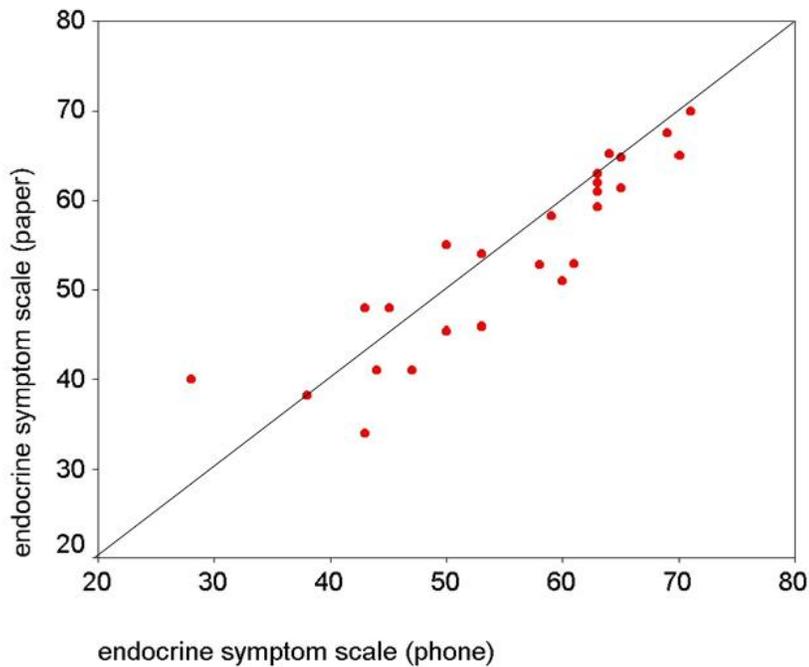


Figure 5.6: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – endocrine subscale. The line of equality super-imposed represents a line of perfect correlation.

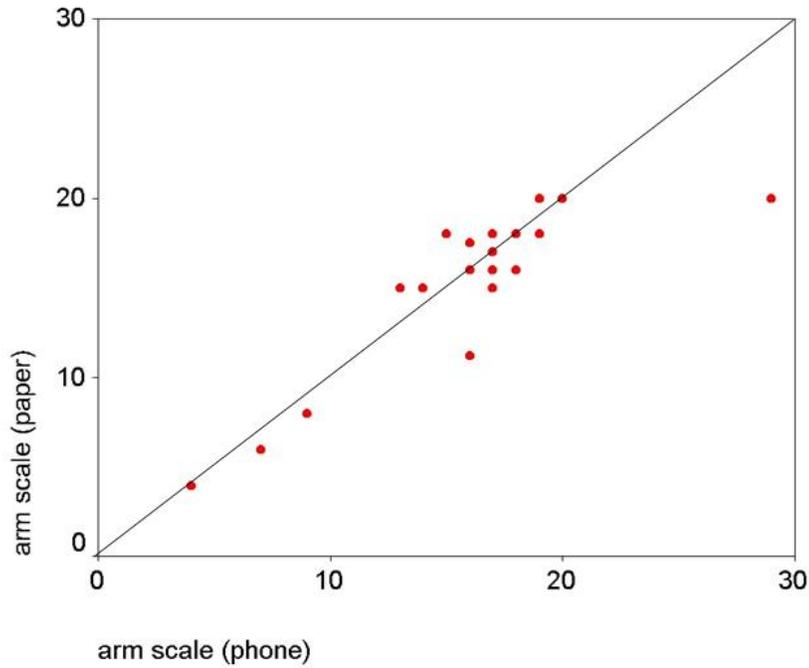


Figure 5.7: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – arm scale. The line of equality super-imposed represents a line of perfect correlation.

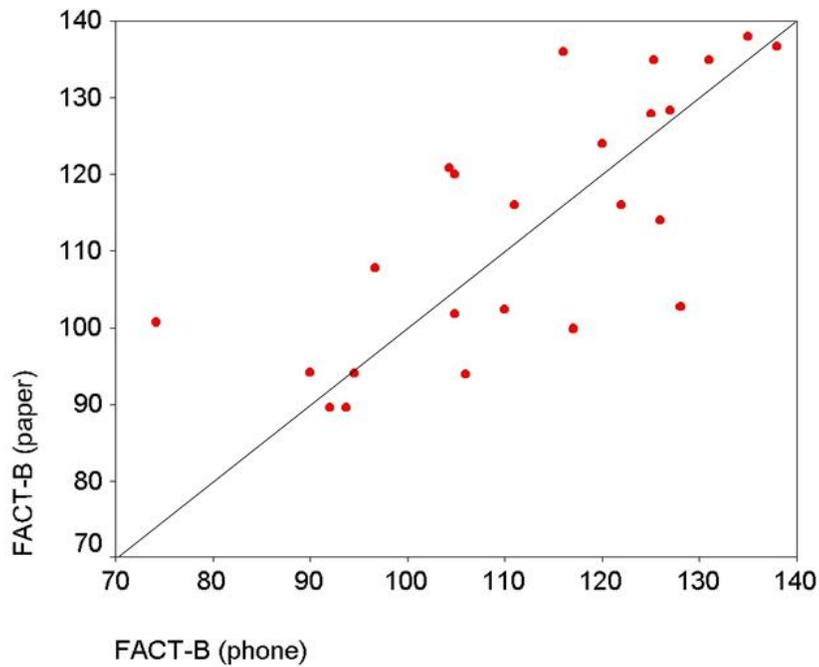


Figure 5.8: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – FACT-B overall. The line of equality super-imposed represents a line of perfect correlation.

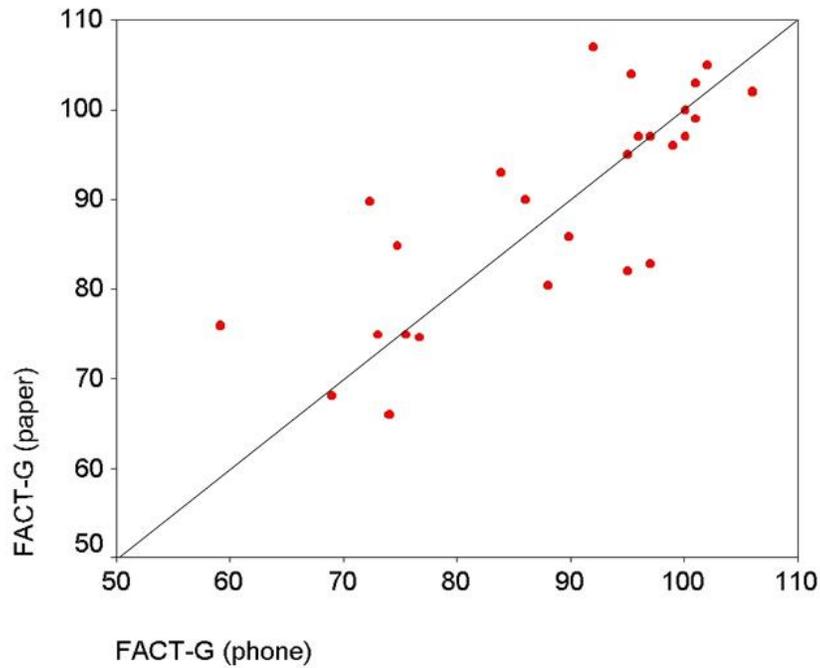


Figure 5.9: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – FACT-G overall. The line of equality super-imposed represents a line of perfect correlation.

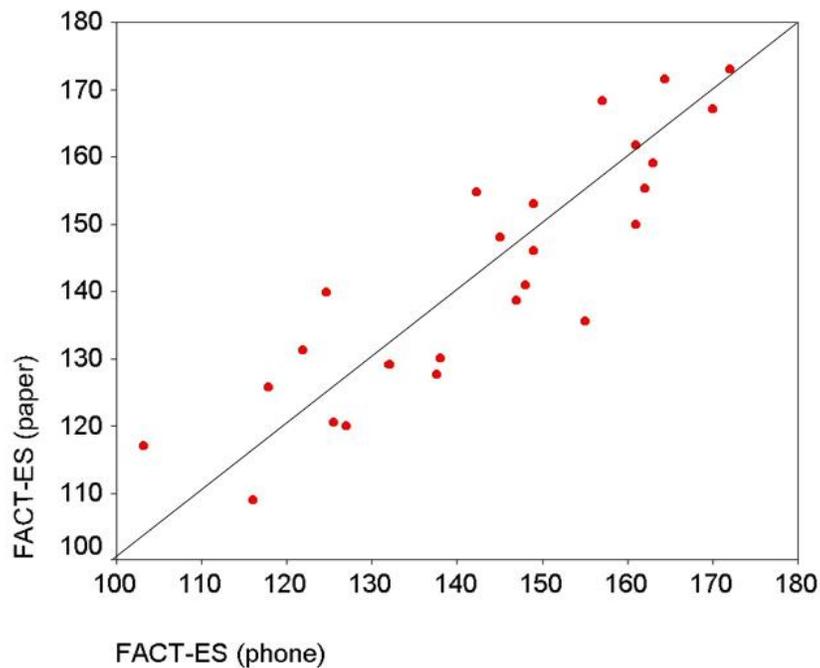


Figure 5.10: Scatter plot of the correlation between scores when administered by telephone and on paper for each individual patient – FACT-ES overall. The line of equality super-imposed represents a line of perfect correlation.

As can be seen from the graphs, there is no evidence of a systematic bias in any of the subscales of FACT-B, the arm or endocrine subscales or the overall scores (FACT-B, G and ES). There is a generally linear relationship in the scatter plots and an even spread of points above and below the line of equality in all graphs. The relationship was further explored by conducting one sample t-tests of the differences between the telephone score and the paper score for each subscale. If the results are spread evenly above and below the line of equality and there is no systematic bias between the methods, one would expect a mean difference between the methods of administration to approach zero, and for the 95% confidence interval to cross zero. The results of this analysis are shown in table 5.4 below.

| subscale | Number of patients | Mean difference (paper vs phone) | 95% confidence interval | T | P |
|-----------------|---------------------------|---|--------------------------------|----------|----------|
| PWB | 26 | -0.5577 | -1.7806 – 0.6652 | -0.939 | 0.357 |
| SWB | 26 | 1.0666 | -0.5173 – 2.6505 | 1.387 | 0.178 |
| EWB | 26 | 0.3231 | -0.6880 – 1.3341 | 0.658 | 0.516 |
| FWB | 26 | 0.2115 | -1.4163 – 1.8394 | 0.268 | 0.791 |
| BCS | 24 | 0.1129 | -1.8431 – 2.0689 | 0.119 | 0.906 |
| ESS | 25 | -1.7204 | -3.7147 – 0.2740 | -1.780 | 0.088 |
| ARM | 24 | -0.5104 | -1.5049 – 0.4840 | -1.062 | 0.299 |
| FACT-B | 24 | 1.3687 | -3.6230 – 6.3603 | 0.567 | 0.576 |
| FACT-G | 26 | 1.0438 | -2.2208 – 4.3083 | 0.658 | 0.516 |
| FACT-ES | 25 | -0.6349 | -4.3239 – 3.0542 | -0.355 | 0.726 |

Table 5.4: Mean difference between paper administered questionnaire score and computer telephone administered score for each of the subscales.

PWB = Physical Well being (7 items); SWB = Social Well Being (7 items); EWB = Emotional Well Being (6 items); FWB = Functional Well Being (7 items); BCS = Breast Cancer Subscale (9 items); ESS = Endocrine SubScale (18 items); ARM = arm subscale (5 items); Fact B = sum of PWB, SWB, EWB, FWB, BCS; Fact G = sum of PWB, SWB, EWB, FWB; Fact ES = sum of PWB, SWB, EWB, FWB, BCS, ESS.

As can be seen, confidence intervals for the mean difference between paper and phone versions of each subscale cross zero for each subscale and there are no significant results in the one sample t-tests. This confirms the graphical impression from the scatter plots above that there are no systematic differences between the paper administration and the phone administration of these questionnaires.

In the Endocrine scale, the confidence interval almost fails to include zero. Analysis of the scatter plot for the endocrine score (figure 5.6 above) shows a trend towards patients having lower scores in the endocrine subscale when completing the questionnaire by telephone than when completing it on paper. This did not reach significance ($p=0.088$).

The analysis above confirms that there is no systematic difference between one method of administration or the other and that the relationship is generally linear. The degree of correlation between the two methods of administration of the questionnaire was assessed by computing Pearson's correlation coefficient. This analysis is shown in table 5.5 below.

| Subscale | PWB | SWB | EWB | FWB | BCS | ESS | ARM | FACT-B | FACT-G | FACT-ES |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| Pearson Correlation | 0.774* | 0.732* | 0.642* | 0.640* | 0.704* | 0.900* | 0.876* | 0.741* | 0.786* | 0.883* |

Table 5.5: Pearson correlation coefficients for the degree of correlation between paper and telephone scores for each of the questionnaire subscales, and for the calculated overall scores.

PWB = Physical Well being (7 items); SWB = Social Well Being (7 items); EWB = Emotional Well Being (6 items); FWB = Functional Well Being (7 items); BCS = Breast Cancer Subscale (9 items); ESS = Endocrine SubScale (18 items); ARM = arm subscale (5 items); Fact B = sum of PWB, SWB, EWB, FWB, BCS; Fact G = sum of PWB, SWB, EWB, FWB; Fact ES = sum of PWB, SWB, EWB, FWB, BCS, ESS. * $p=0.01$

There is generally a high degree of correlation between the scores for the questionnaire completed on paper and completed using the telephone system. The exception to this is in the emotional well being score and in the functional well being score where correlation was less strong.

Ease of use and acceptability among the small validation/correlation cohort

43 patients were asked to participate in the smaller validation study designed to assess the degree of correlation between paper and telephone administration of the questionnaire. 40

patients agreed to participate, and acceptability of 93%. 26 patients subsequently completed the telephone questionnaire, a 65% response rate.

All of the patients who telephoned in to this section of the study found the system easy to use. 22 patients liked the system and were keen to use it as their main method of follow up. 1 patient was not sure, but felt she may be happy a few years down the line as she was only two years post diagnosis. 3 patients did not feel sufficiently reassured, and all three stated that talking to a machine was less personal, less reassuring and more difficult. None mentioned a desire to come to clinic for examination and when asked did not rate clinical examination as a priority.

Detecting problems using computer telephone follow up

The potential impact of introducing the new computer telephone follow up system was assessed by establishing how many patients would require to be telephoned or brought back to clinic in light of their scores on the questionnaire. As mentioned above, a change in score of 5% or more of the total score available for each section is considered significant[Cella D et al., 2002;Eton DT et al., 2004]. Patients recording deterioration in score of 5% or more would be telephoned to discuss this. Patients with issues unable to be resolved through telephone consultation would be brought back to clinic for further assessment, investigation or treatment. Patients answering yes to any of the initial seven questions outlined in appendix 5.1 and designed to allow the patients to highlight easily any new problems they had developed in their breast or arm pit were also telephoned to discuss the problem which had led to them answering yes, and to arrange for them to be brought back to clinic if these

problems required further investigation. Table 5.6 outlines the changes in scores for each of the sections for the 75 patients who responded.

| | PWB | SWB | EWB | FWB | BCS | ESS | ARM | Fact B | Fact G | Fact ES |
|----------------------------|-----|-----|-----|-----|-----|-----|-----|--------|--------|---------|
| Improved score | 21 | 11 | 20 | 12 | 18 | 14 | 18 | 7 | 9 | 7 |
| Static score | 42 | 36 | 26 | 33 | 20 | 33 | 36 | 44 | 42 | 38 |
| Deteriorating score | 11 | 27 | 29 | 30 | 35 | 16 | 19 | 21 | 22 | 16 |
| Not available* | 1 | 1 | 0 | 0 | 2 | 12 | 2 | 3 | 2 | 14 |

Table 5.6: Change in scores between the baseline paper version and the telephone questionnaire completed at ten months. Scores were considered static if the telephone score was within 5% of baseline score.

PWB = Physical Well being (7 items); SWB = Social Well Being (7 items); EWB = Emotional Well Being (6 items); FWB = Functional Well Being (7 items); BCS = Breast Cancer Subscale (9 items); ESS = Endocrine SubScale (18 items); ARM = arm subscale (5 items); Fact B = sum of PWB, SWB, EWB, FWB, BCS; Fact G = sum of PWB, SWB, EWB, FWB; Fact ES = sum of PWB, SWB, EWB, FWB, BCS, ESS.

*several patients did not complete all sections of both the baseline and telephone questionnaires and so a difference between the scores was not calculable.

The number of patients with improved scores was similar for all subsections of the questionnaire at between 11 patients for the social well being score and 21 for the physical well being (14.66% and 28%). There was much greater variety in the number whose score deteriorated with 11 patients (14.66%) recording a deterioration in score in the physical well being scale and 35 patients (46.66%) doing so in the breast cancer subscale.

In total, 66 of the patients (88%) had deterioration in score in at least one section of the questionnaire of more than 5% of the total score available for that section. 43 patients answered yes to at least one of the initial questions shown in appendix 5.1. 40 of these had also recorded deterioration in some aspect of their questionnaire and 3 had not. In total, 69 patients (92%) either answered yes to one of the initial 7 questions or had deterioration in their score from baseline of more than 5% in at least one section of the questionnaire.

The majority of patients with deterioration in score stated that they felt no different at the time of the telephone questionnaire compared to when filling out the baseline questionnaire the year before, or felt that their change in score was related to other problems in their life. In particular, a number of patients with arthritis attributed much of their deterioration in scores to joint pain which had deteriorated over the course of the year and several patients with shortness of breath or cough were currently suffering upper respiratory tract infections or had asthma or COPD being looked after by their general practitioner. One patient had suffered deterioration in her functional and emotional scores, but had already been investigated by her GP and was seeing a psychologist. All patients with deterioration in endocrine scores were given advice on dose splitting, one was sent to her GP for venlafaxine prescription and one patient was brought back to clinic to further discuss treatment options.

4 patients complained of new bone pain which could not be dealt with satisfactorily over the telephone and all four of these patients were brought back to clinic for further treatment or investigation. One patient had had recent further axillary surgery for cosmesis and had developed some troublesome lymphoedema. She was referred to the lymphoedema clinic. 3 patients brought back to investigate new lumps in the breast or axilla (all were benign) and one patient had been back for investigation of a new lump one week before completing the telephone questionnaire. One patient had suffered a marked deterioration in endocrine scores and complained of vaginal bleeding, but had just been diagnosed with endometrial cancer having been referred to gynaecology by her general practitioner.

From all 75 patients who responded, only 8 (10.66%) required to be brought back to clinic for further investigation or treatment of problems raised during the telephone follow up.

Discussion

It is still routine practice to provide regular clinic visits for women after diagnosis and treatment of breast cancer. Both the UK national guidelines and most clinicians in a recent large UK survey of breast follow up practice among breast cancer specialists agree that detection of potentially treatable relapse, detection and amelioration of the side effects of therapy and provision of psychological support are the main reasons for continuing to see women in the long term[Donnelly P et al., 2007;National Institute for Clinical Excellence, 2002].

Yet there is little evidence regarding how well routine clinic visits meet these aims of follow up. A large systematic review conducted by de Bock described routine breast cancer follow up as an inefficient way of detecting potentially treatable relapse[de Bock GH et al., 2004]. Very few relapses are detected by clinical examination during routine clinic visits. The analyses outlined in chapters 2 and 3 of this thesis confirm that very few relapses are detected by routine clinic visits and highlight that those relapses which are detected clinically do not appear to do any better than those picked up by other methods such as mammography. In fact, the Edinburgh patients with ipsilateral breast relapse analysed in chapter 2 do less well when such relapse is detected clinically[Montgomery DA et al., 2007b].

Two large trials by Grunfeld and colleagues failed to show any difference in survival when patients do not come back to breast clinic routinely and are followed instead by their GP[Grunfeld E et al., 1996;Grunfeld E et al., 2006]. While both of the trials by Grunfeld were ultimately underpowered to show a difference in survival, this is largely a component of the event rate in breast cancer being low (1 to 1.5% locoregional relapse per year) and the low rate of detection of such relapse by clinical examination (see chapter 2), again highlighting how relatively uncommon an event is diagnosing relapse through routine clinical examination.

There is little evidence that routine clinic visits are useful at detecting side effects of treatment or psychological problems. Clinicians are not good at detecting psychological problems in women[Baildam AD et al., 2004], and women are not keen to discuss psychological concerns in clinic for a variety of reasons including embarrassment, feeling rushed or not wanting to waste the doctors time[Pennery E and Mallet J, 2000;Valente SM et al., 1994]. Removing clinic visits does not appear to reduce quality of life, as would be expected if such visits contributed to the detection and amelioration of physical or psychosocial problems in the aftermath of breast cancer (see chapter 1). In the study presented in this chapter, women who preferred the new computer telephone method of follow up over traditional clinic visits were asked what it was about routine clinics that they did not like. A number of women stated explicitly that they felt rushed in clinic, embarrassed about talking about some of the problems they were having or anxious at the thought that they may have a recurrence diagnosed, all factors which they felt made it less easy to discuss their concerns with the doctor they saw in clinic.

Despite the suggestion that routine outpatient clinics may not be the most effective or efficient way to provide follow up after breast cancer, few alternative methods of providing follow up have been evaluated. As outlined in chapter 1 of this thesis, only 5 randomised controlled trials of alternative follow up methods have been conducted. Most of these have been too small or too brief to show a convincing benefit of the alternative under scrutiny, and none has been able to comment on the safety of abandoning clinics from the point of view of recurrence. More crucially, none of these alternative methods has been introduced with a proven reduction in cost or clinical commitment, and only one, the study of routine breast care nurse follow up, has suggested that alternative follow up methods may be introduced with resultant improvements in quality of life amongst patients compared with those followed in a traditional manner.

This is the first study of which we are aware where the use of specific quality of life tools to more objectively uncover psychosocial concerns and side effects of treatment in women in routine follow up has been evaluated. Quality of life tools have been used to assess the impact of follow up methods in the past, as described in some detail in chapter one, but never specifically to screen for problems. Using such tools in this way should be done with care, as the FACT-B questionnaire was not developed as a diagnostic tool in its own right. This questionnaire, particularly when coupled with the additional arm and endocrine subscales as was the case here, does allow a very detailed assessment of patients' functioning across a wide spectrum of areas, covering both psychosocial concerns as well as the physical side effects of treatment, and is sensitive to change over a period of time. There are well defined

criteria as to what is likely to constitute a significant change in score, usually considered to be a change in 5% or more of the total score available for each section[Cella D et al., 2002;Eton DT et al., 2004], and there are extensive data pertaining to the use of this questionnaire and the range of scores normally expected, although much of this data comes from clinical trials and may not reflect the range of scores seen in non-trial patients.

Because FACT-B was not developed as a diagnostic tool, the overall scores and changes in score for an individual patient must be interpreted with caution. As mentioned by Cella and colleagues, the derived change in score in an individual may reflect true change, although the degree of measurement error in individual assessment precludes the use of these scores for individual diagnosis[Cella D et al., 2002]. Rather, the scores derived for any patient can be interpreted in much the same way as any screening test, as long as one keeps in mind the consequences of false positive and false negative results. This was the approach adopted here in the large feasibility cohort, where poor scores on any section or deterioration in score between the paper version done at baseline and the telephone version done between 10 and 12 months later of more than 5% of the total score available in any section resulted in a telephone call to the patient to assess whether a clinic visit was required. In the majority of cases, some advice was needed. Only a minority required a clinic visit.

The small validity section of this study carried out to establish correlation between paper administration and computer telephone administration suggests that correlation between the two methods of administration is not perfect. High correlation coefficients were seen for some sections, especially the endocrine and arm subscales and the overall FACT-ES scores,

but many of the other sections, particularly scores in the emotional well being and functional well being sections, showed only moderate correlation between the paper and telephone scores. The endocrine subscale showed the greatest degree of correlation, but there a very strong trend towards scores being systematically lower when the questionnaire was completed by telephone, although this trend did not reach significance.

That the correlation is not perfect is not surprising given the differences between the two methods of administration. Patients completed the paper version in clinic and the phone version at home several days later. It might be expected that patients would be more relaxed and less anxious in their own environment, which could contribute to better scores in some patients, especially in some of the psychological or emotional sections. There was variability in whether family members were present when completing the different versions of the questionnaire, with some patients completing the baseline paper version of the questionnaire while in the company of family members in clinic and some completing it while alone in clinic, and similarly some patients reported discussing the questions with family members at home before calling the computer telephone line while others did not involve family. Any discrepancy in the level of family involvement between the two methods of completing the questionnaire for an individual patient may affect the answers given and resultantly the correlation between their two questionnaires. Several studies have shown that patients are often more honest when reporting symptoms or other issues to an automated computer telephone system than when completing questionnaires in other ways[Turner CF et al., 1998;Millard RW and Carver JR, 1999;Kobak KA et al., 1997], presumably as the system

feels more anonymous. This may also contribute to differences in score between the two methods of administration.

There were also technical differences between the methods of administration which might contribute to slight discrepancy in score between the two methods of administration. The telephone version did not allow questions to be missed out, while this was possible in the paper version. Most patients missed out at least one question at some stage during completing the paper version, and so scores for the paper version were very often derived scores based on the calculations available to account for missed questions. In contrast, the scores from the telephone questionnaires were based on complete data, and are therefore likely to be more accurate.

Correlation between the paper and the telephone versions of the questionnaire is not sufficient to allow these methods to be used interchangeably. It will be important to do further work in the future to ensure there is good test-retest reliability using the telephone administration of this questionnaire. This should be done by inviting a cohort of patients to complete the questionnaire by telephone on two occasions a few days apart. For any future patients recruited to the system, baseline questionnaires should also be completed on the telephone. This will hopefully result in a higher degree of correlation between baseline and follow up scores, which may also result in fewer patients recording a deterioration in score over the year, in turn further reducing the number of patients who need to be telephoned each year to discuss scores or reported symptoms.

The seven initial screening questions used in this study (see appendix 5.1) were designed to allow patients to report new findings in the breast, scar or axilla or to report significant new bone pain, breathlessness or persistent cough, any of which may be indicative of recurrent disease. These seven questions had not been previously validated. In practice, these questions were not selective of patients with true problems. The questions regarding bone pain, shortness of breath and persistent cough in particular were very non-specific and were answered positively by most women in what is a relatively elderly population, often with intercurrent medical conditions such as arthritis or respiratory problems. The evidence from early studies is that most patients with metastatic disease will present between annual reviews having developed symptoms of their metastases, such as intractable bone pain, breathlessness etc. In future studies of this computer telephone follow up system, these seven additional questions as outlined in appendix 5.1 will be removed and replaced with general advice to the patient to get in touch if they have any new concerns.

A large proportion of the women approached to take part in this study were happy to do so, with a randomisation rate of 91% to the study on acceptability. There was a relatively high drop-out rate subsequently, with some 30 patients (29% of patients who agreed to take part) not completing the telephone questionnaire 10 to 12 months after recruitment. This drop out rate is probably not completely reflective of patients' willingness to try the system. Certainly, 3 patients were uncontactable and have subsequently failed to attend scheduled review appointments also. 5 patients who were contactable said they intended to call in and take part but subsequently did not, and at least three of these claimed not to have received the letters asking them to call in. A number of women failed to complete the telephone

questionnaire due to technical reasons and while some of these were simply unable to operate the technology, half of them blamed the system itself, stating that there was no answer when they called or that the system cut out while they were trying it. Several of the patients who did not complete the small validation study complained that this was the case also. There were occasional technical problems with the system, and further work on the part of the software suppliers will be undertaken to ensure the system operates consistently before patients are transferred to this method of follow up. Finally, 3 patients said that intercurrent illness precluded their participation. We would anticipate that in real clinical practice, when there is no back up clinic available unless specifically asked for and when the initial technical problems have been completely rectified, a number of these women may have taken part. This remains speculative.

One consistent finding in this pilot study was that women born before 1930 usually did not manage to use this system, with less than one third of these women completing the questionnaire on telephone successfully. In contrast, two thirds of women born in the decade from 1930 called in and completed the questionnaire successfully using the computer telephone system. This does not preclude this system being offered to older women as a form of follow up, but it certainly highlights the need for provision of an alternative for women who are likely to not manage. A selective approach will be needed when offering this system to older women, with strong reassurance that they can come back to clinic if they do not manage and perhaps even a trial run using the phone system at recruitment while the patient is in clinic. This was not possible in this study as the telephone system was not in place during the recruitment process.

Among participants, this novel method of follow up was very acceptable, with almost 80% agreeing that there were benefits to this type of follow up over traditional clinic visits and stating that they would be keen to use the system in future. Most of these patients felt that the follow up process was considerably more detailed than was usually the case, with a number of women commenting that the depth of questioning was considerably greater than is the case during routine clinic visits. In this initial feasibility study, there was no formal assessment of the impact of this system on quality of life as there was no control group for comparison. We would hope this more detailed questioning would lead to greater detection of problems, a higher level of input to attempt to resolve these problems and ultimately would translate to an improved quality of life among participants in this type of follow up process. This is an issue which could be explored with a randomised controlled trial in the future.

The final issue which remains to be explored is the economic impact of this follow up system. Most women needed some form of clinical contact after completing the phone system, but in most cases this was simply a telephone call. In the future, patients will complete their baseline questionnaire over the telephone also, rather than on paper as was the case in this study. This should lead to a greater degree of correlation between the baseline score and further scores generated during follow up, with any change recorded over the course of follow up being more likely to be a true change rather than variation between the two methods of completing the questionnaire. This should in turn result in fewer patients recording a deteriorating score during follow up.

Only a very few patients required to be brought back to clinic for investigation or further treatment. In most of these cases, they could have requisite investigations carried out first, or they could be referred directly to another specialty clinic, such as a lymphoedema clinic, thus saving them at least one visit to the breast clinic.

It is likely that this system will dramatically reduce the need for routine clinic visits, but this will be offset in part by the need for a specialist nurse to oversee the telephone system, analyse patients' scores and telephone those patients who raise concerns with their telephone follow up. Again, full evaluation should take place within the context of a clinical trial to establish the economic benefits.

Summary

This novel follow up method has definite advantages over the traditional clinic system, and some potential benefits which remain to be quantified. Patients are largely accepting of the new technology, and feel that it has the potential to identify problems more effectively than the current clinic system. Coupled with this, patients and their families appreciate the time and expense saved compared with the traditional approach of coming to clinic.

The system has potential to reduce the number of patients returning to clinic and to make the visits of those who do come more focussed and more organised. Potential problems identified before clinic visits can lead to necessary investigation taking place before the

patient attends the clinic, obviating the need for an initial assessment visit then a review visit to discuss results of tests.

Offset against this will be the work required to look at patients' scores and then telephone patients with low or deteriorating scores to assess the need for a clinic visit. Much of this process can be automated and carried out using the computer software, but a senior breast nurse will need to be responsible for the running of the system and the day to day telephoning of patients with problems. Working practices could be much more flexible as a result, so it is possible that there will be cost savings and space may be saved in clinics, allowing new patients to be seen more quickly. A full analysis as part of a randomised trial will be needed to fully quantify any cost or time savings.

Automated computer telephone follow up after breast cancer is a promising alternative to the relatively un-focussed routine follow up clinics which women currently attend. It has the potential to save patients time and effort, reduce the burden on outpatient clinics and at the same time improve the detection of particularly psychosocial concerns and side effects of treatment, hopefully to the benefit of patients' well being.

Discussion

And

Conclusions

Discussion

The expressed aims of breast cancer follow up are consistent across several sets of guidelines, and are explicitly that follow up should be provided to detect relapse, diagnose and treat complications of therapy and ensure detection of, and support for, psychosocial problems which develop in the face of diagnosis and treatment of breast cancer. Despite a consistency of aims, there is marked variation in what is recommended by the various guidelines.

In the UK, very limited follow up is advocated, with the National Institute for Clinical Excellence (NICE) recommending 2-3 years of follow up [National Institute for Clinical Excellence, 2002] and the British Association of Surgical Oncology (BASO) supporting a 5 year follow up duration [The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005]. The American Society of Clinical Oncology (ASCO) recommends a very traditional follow up scheme, with frequent initial visits reducing to annual visits after five years and continued indefinitely [Khatcheressian JL et al., 2006]. At the opposite end of the spectrum altogether from the UK guidelines are those issued by the Canadian Steering Committee which recommends a pragmatic approach to frequency of visits, stating that frequency be tailored to the individual, but suggests that follow up should be indefinite [Grunfeld E et al., 2005].

Much of the discrepancy between the NICE recommendations for two to three years of follow up [National Institute for Clinical Excellence, 2002] and at the other extreme the Canadian Steering Committee guidelines for indefinite follow up [Grunfeld E et al., 2005]

results from a lack of high quality evidence in the literature to support any particular approach. As revealed by the systematic review conducted in chapter 1 of this thesis, only two randomised trials have been performed to help establish how frequently routine clinic visits should be provided[Kokko R et al., 2005;Gulliford T et al., 1997]. There was no difference between three and six monthly follow up in the study by Kokko and colleagues in terms of the number of relapses detected overall, but locoregional relapse and metastatic disease were not analysed separately and there was no comparison of time to detection of relapse[Kokko R et al., 2005]. In the study by Gulliford and colleagues there was neither sufficient numbers recruited nor sufficient duration of follow up to make comment on the safety of further reducing frequency of visits to annual[Gulliford T et al., 1997]. All that can be said, then, is that in terms of relapse detection there is no evidence from randomised controlled trials to support any particular frequency of visit.

There is similar uncertainty with regards to the duration of follow up required, with periods ranging from 3 years[National Institute for Clinical Excellence, 2002] to indefinite follow up[Grunfeld E et al., 2005] recommended by the various guidelines. No randomised controlled trials have been conducted which have addressed this issue (see chapter 1). An acceptable approach to establishing the necessary duration of follow up would be to perform a cohort study analysing the frequency of relapse over time. If a time point where relapse becomes very infrequent or even ceases altogether could be identified, then an ideal duration of follow up could be recommended.

Such cohort studies have been performed by a number of investigators, but with varying methodologies and definitions of what constitutes relapse. The NICE guidelines for duration of follow up[National Institute for Clinical Excellence, 2002] are based on one study which analysed all relapses (locoregional, contralateral and metastatic) in a cohort of 643 patients followed for a median of 3 years and 11 months[Donnelly J et al., 2001]. There was no analysis of frequency of relapse in the longer term in this study. Others have shown relapse rate to peak in the first 2 to 5 years after treatment and decline after this, but again there are problems with these studies. For example, Hussain and colleagues showed that 93% of locoregional relapses occurred within five years of diagnosis and treatment, but did not include new contralateral breast cancers in their analysis[Hussain ST et al., 1995]. Saphner and colleagues reported that relapse peaked in the first 2 years and that rate of relapse declined after this, but in this study metastatic disease was included in the analysis[Saphner T et al., 1996]. In neither of these studies did the hazard rate for relapse reach zero at any stage, despite follow up for 10 years by Hussain and colleagues and for 12 years in the study by Saphner and colleagues.

The analysis of hazard rate for relapse in the Edinburgh cohort presented in chapter 2 explores this issue in more detail than has been the case previously. The rate of metastatic relapse peaks initially at around 3% of patients at risk per year and then falls off with time but never reaches zero. In contrast, the rate of potentially treatable relapse (locoregional relapse and new contralateral cancers) remains essentially constant at around 1-1.5% of those at risk per year for at least ten years[Montgomery DA et al., 2007b]. There was no analysis of rate of relapse for individual sites in the Edinburgh data presented in chapter 2 due to the

relatively small numbers involved when the data was subdivided in this manner. It has been highlighted by other investigators that ipsilateral breast relapse is more common initially and that new contralateral cancers become more common over time so that by 15 years after treatment the incidence of each is identical[Freedman GM et al., 2005;Veronesi U et al., 2001;Veronesi U et al., 2002]. From this, it can be seen that authors who have included metastatic disease in, or excluded new contralateral cancers from, their analysis of relapse rate will demonstrate an initial peak and then fall off in incidence. It is now clear that this is not an accurate picture of the rate of potentially treatable relapse.

The data presented in chapter 2 of this thesis does not support the contention that return clinic visits should be more frequent initially, as advocated by the ASCO guidelines[Khatcheressian JL et al., 2006] or that patients can be safely discharged after this time, as recommended by NICE[National Institute for Clinical Excellence, 2002] and BASO[The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005]. None of these guidelines recommends screening for metastatic disease, which most commonly presents with symptoms and for which regular visits or intensive monitoring for metastatic disease does not, from the evidence available in the literature, improve survival[Del Turco MR et al., 1994;Palli D et al., 1999;The GIVIO investigators, 1994]. The guidelines do say that potentially treatable relapse should be the focus of follow up. This must include new contralateral cancers. New contralateral cancers have been shown to be as common as true ipsilateral breast relapse over prolonged follow up periods [Freedman GM et al., 2005;Veronesi U et al., 2001;Veronesi U et al., 2002] and are likely to be just as treatable as, if not more treatable than, local relapse if detected early.

It is the conclusion of this author that if routine clinic visits are to be provided for follow up of patients after breast cancer with the aim of detecting potentially treatable relapse, then such visits must be continued indefinitely. The rate of potentially treatable relapse does not fall off over time, and patients continue to be at risk of relapse long after they would normally be discharged under the NICE guidelines[National Institute for Clinical Excellence, 2002]. In this regard, the Canadian steering committee guidelines are the closest to recommending ideal practice[Grunfeld E et al., 2005].

There remains considerable uncertainty over the value of routine clinic visits in meeting any of the aims of follow up. This is an important consideration as there is evidence that routine clinic visits are an inefficient way to provide follow up for the detection of relapse[de Bock GH et al., 2004], that they cause anxiety in women who attend them[Paradiso A et al., 1995] and that they may not be an effective way of providing psychosocial support or detecting side effects of adjuvant therapy.

The only definitive way to establish the safety of abandoning traditional follow up is to conduct randomised controlled trials comparing traditional follow up to some alternative. As described in the systematic review conducted in chapter 1 of this thesis, 5 such trials exist. None has been sufficiently powered to definitively suggest that alternative follow up methods are safe, but at least one very large study, of general practice follow up versus traditional follow up in hospital, failed to show any difference between these two methods in terms of relapse detection or survival[Grunfeld E et al., 2006]. None of the five trials of alternative

follow up has shown traditional clinics to be better at detecting recurrence or improving survival (see chapter 1).

Numerous retrospective analyses of method of detection of relapse have been conducted. These have often focused on irrelevant outcomes. For example, several studies have analysed the proportion of relapses detected at routine clinics versus interval clinics and others have compared the proportion of relapses detected by routine follow up (either clinical examination or mammography) compared with those which present with symptoms. Neither of these issues is important in establishing the value of routine clinic visits. It is intuitive that interval clinics will be more effective than routine clinics as all those presenting to an interval clinic will have detected something which concerns them enough to present again early. The rate of relapse detection in a symptomatic population will be higher than in an asymptomatic population, and this difference will be potentially more dramatic when event rates are as low as the 1-1.5% per year suggested by this author's data (chapter 2). Similarly, bundling all relapse detected clinically and mammographically in together in an analysis hides valuable information. Patients do not need to come up to a clinic to have a mammogram performed, and vice versa. What is important is the proportion of relapses detected by routine clinical examination compared with mammography compared with patient self examination. These are the three main methods by which relapse is likely to be detected, given that the guidelines recommend no additional investigations [Grunfeld E et al., 2005; Khatcheressian JL et al., 2006; National Institute for Clinical Excellence, 2002; The Association of Breast Surgery @ BASO and Royal College of Surgeons of England, 2005].

Few retrospective analyses have looked at this issue in so much detail. As described in chapter 3, prior to the work conducted here only ten studies have presented method of detection of relapse separated into the three components described above[Mohoney L, 1986;Tate JT et al., 1989;Rutgers EJT et al., 1991;Snee M, 1996;Hussain ST et al., 1995;Grinfeld E et al., 1996;Lees A et al., 1997;Jack WJL et al., 1998;Churn M and Kelly V, 2001;Grogan M et al., 2002]. Only one of these has presented outcome data, and then only on 5 patients[Snee M, 1996]. Many of these studies have been based on what are now historical cohorts and most have been very small. As part of this work, three new datasets have been presented for this first time. Two of these have been retrospective reviews of local cohorts, analysed in detail in chapter 2. The third is an analysis of previously unpublished data from a large Dutch series[van der Sangen MJC et al., 2006], analysed as part of the systematic review conducted in chapter 3. Outcome data is available for all three of these cohorts, making this the first time that outcome after relapse related to method of detection has been presented (with the exception of the study by Snee et al where such data was presented for 5 women[Snee M, 1996]).

In these analyses, mammography is more important in modern follow up than has historically been the case. Over half of relapses are detected this way in the modern cohorts analysed in chapter 3, and mammography detects 5.37 new cancers per 1000 mammograms undertaken in the Edinburgh cohort[Montgomery DA et al., 2007b] analysed in chapter 2, comparing very favourably with the rate of detection in breast screening during the equivalent period in time. This confirms that mammography should be a central component of follow up, and, according to the evidence presented here, should be provided on an annual basis. This is

contrary to the advice in the NICE guidelines based on far smaller patient numbers[National Institute for Clinical Excellence, 2002].

The systematic review conducted in chapter 3 also confirmed the importance of patient self examination. Across all of the retrospective reviews analysed, patients detect around one third of their own relapses. The value of patient self examination should not be underestimated and patients should be encouraged to conduct such examinations. Particular attention should be given to axillary self-examination, as a far larger proportion of axillary relapse was detected by a clinician in the Edinburgh series presented in chapter 2 than either ipsilateral or contralateral breast relapse.

The value of routine clinical examination appears to be limited. In the Edinburgh cohort analysed in chapter 2, routine clinical examination detected few relapses. The annual hazard rate for potentially treatable relapse was constant at 1-1.5% for the first ten years. 13% of these relapses were detected by clinical examination. Thus less than 0.2% of women at risk per year have relapse diagnosed by clinical examination. Moreover, those few ipsilateral breast relapses which were detected in this manner appear to have a poorer outcome. Certainly, patients with clinically detected relapses did not do better than relapses detected by other means. This issue becomes more complicated when the series from Holland is included[van der Sangen MJC et al., 2006]. In an analysis of survival by method of detection as part of the systematic review conducted in chapter 3, there was no difference in outcome between any of the methods of relapse detection, but again clinically detected relapses were not associated with improved survival. One possible explanation is that all of the relapses in

the Dutch group occurred more than five years after treatment and were associated with particularly good prognosis. It may well be that the prognosis of late relapse is so good that method of detection becomes irrelevant in these patients. This issue will require further study.

Whether routine clinic visits are necessary for achieving the additional aims of follow up, namely detection of psychosocial problems and of side effects of therapy, is possibly easier to answer. Particularly in the first few years after treatment, incidence of psychological distress will be higher than will incidence of potentially treatable relapse, with estimates that up to one third of women will suffer generalised anxiety and depression at some stage after surgery for breast cancer[Maguire P, 1994]. Clinic visits themselves are responsible for causing short term anxiety in the weeks leading up to the clinic and during the actual visit itself in 71% of women[Paradiso A et al., 1995]. The event rate for anxiety and depression is therefore much higher and so it should be easier to detect a difference between patients followed up in routine clinics and patients who receive a novel alternative.

Again, the only way to answer this question definitively is by randomised controlled trial. Various questionnaire or interview based studies of cohorts of women undergoing follow up in breast cancer have been undertaken to attempt to answer this question. This approach is complicated by the fact that women find it reassuring to be told they do not have relapse, and so are likely to report satisfaction from attending routine follow up when no relapse is detected. Questionnaire based studies of routine follow up have shown high levels of satisfaction with follow up under these circumstances[Morris S et al., 1992;Kiebert GM et

al., 1993]. What is less clear is whether attendance at routine clinics provides short term reassurance only, as might be implied from the fact that women report high levels of anxiety before attending clinic as they are worried about being diagnosed with relapse [Pennery E and Mallet J, 2000], or whether there are longer term quality of life gains as a result of detection and amelioration of psychosocial concerns or side effects of treatment in routine follow up.

Unlike for issues such as relapse detection and survival, the randomised controlled trials which exist in the literature and which were analysed in the systematic review conducted in chapter 1 of this thesis go some way to answering this question. Most of the randomised controlled trials of alternative follow up methods which have looked at issues such as quality of life or patient satisfaction have shown no difference between traditional clinic visits and alternatives such as coming to clinic only if problems occur [Brown L et al., 2002; Koinberg IL et al., 2004], or having follow up conducted by the general practitioner [Grunfeld E et al., 1996; Grunfeld E et al., 2006]. Reducing frequency of follow up does not impact negatively on patient satisfaction with follow up either [Gulliford T et al., 1997]. Women report being more satisfied with breast care nurse follow up than with traditional follow up by a doctor, and nurses were also more likely to recognise anxiety/depression in the patients they saw than were doctors [Baildam AD et al., 2004]. These findings were attributed to the length of time spent by nurses with each individual patient, which was far longer than was spent by doctors. The evidence suggests that routine follow up by clinicians is not the best way to detect psychological problems.

Routine clinic visits are not necessarily the best place to detect side effects of treatment either. It has been shown that the number of side effects a patient suffers is significantly correlated with quality of life score[Macquart-Moulin G et al., 1997] and detection of side effects is a key aim of the follow up guidelines[National Institute for Clinical Excellence, 2002]. It has also been reported that clinicians underestimate side effects in their patients in routine oncology follow up, even in the setting of a clinical study where physicians perceptions of patients symptoms is being specifically addressed[Macquart-Moulin G et al., 1997;Fromme EK et al., 2004].

None of the randomised trials reviewed in chapter 1 has analysed detection of side effect as an outcome, but if side effects of treatment are related to quality of life, then quality of life could perhaps be used as a surrogate marker. It is certainly the case that alternative follow up methods do not reduce quality of life, but neither have an alternative follow up methods been shown to improve quality of life (see chapter 1).

Most patients do expect some form of follow up, with 94% of women expecting to be seen again in clinic in the study reported in chapter 4. Most have very little idea of what to expect in terms of follow up with only 36.7% predicting the correct frequency of follow up visits and only 5% of patients predicting the correct duration. Few patients had realistic expectations of what clinics were likely to achieve. In our practice, most local screening detected cancers are treated elsewhere. The result is that 78.5% of patients questioned about their expectations for follow up in the study presented in chapter 4 had self detected cancers originally. Despite that, 68% of women thought that early detection of relapse was a main

reason for coming back to clinic and 54% of women thought the clinician was more likely to detect relapse than they were.

Providing information of the inefficiency of clinics did alter some patients' feelings, with 34% stating that they would be happy not to come back to clinic having been informed of how infrequently routine clinical examination leads to detection of relapse. This does mean that almost two thirds of women in the study reported in chapter 4 state that they would still like to come back to clinic regularly, even when asymptomatic, despite being made aware of the very low yield from clinical examination and despite the majority of them having detected their own cancers initially.

The reasons for this were not fully explored here. There was no obvious relationship between either method of detection of original cancer or perception of who was more likely to diagnose any subsequent cancers and desire to come back to clinic. It is possible that, even though many women accepted that they had diagnosed their original cancer themselves and were most likely to diagnose any relapse, they still appreciate the reassurance they get from a negative clinic visit. It is also possible that the proportion who want to stay in follow up was high as these women were all very early on in the process, having been diagnosed and treated less than a year before taking part in the study. Certainly, women recruited to the computer telephone follow up system in the first two or three years after their treatment did occasionally report being unwilling to change to this new method of follow up at such an early stage.

The other possibility is that patients were unaware of the alternatives available. Discharge from follow up at one year is unlikely to be a popular option for women, especially as many are still having treatment such as tamoxifen or aromatase inhibitors. More explanation of the alternatives to routine follow up may have resulted in a higher proportion of women being accepting of these alternative follow up methods. This theory is borne out by the much higher acceptance of alternative follow up plans when the structure of these plans is laid out in detail, as in chapter 5.

In chapter 5, an acceptability study of a unique method of follow up using an automated computer telephone system was presented. With the computer telephone follow up studied here, for the first time a follow up method concentrating specifically on maximising detection of psychosocial concerns and side effects of therapy was introduced. As mentioned above, structured quality of life assessment is more commonly used to measure outcomes in trials, but is rarely, if ever, a component of routine scheduled follow up plans.

There was a high level of acceptability of this type of follow up. Most participants felt that it provided a greater opportunity to relate problems than they are given in clinic. This should lead to earlier and more comprehensive diagnosis of psychosocial problems and side effects of therapy, which will in turn allow better targeting of resources towards those who have more problems. Whether this leads to improvements in quality of life remains to be seen.

While there was a high level of acceptance of this new system, it was not suitable for all patients. Elderly patients in particular often struggled to use the technology. A small

proportion of the patients who tried the system did not feel sufficiently reassured by it. Many of them could recognise the potential benefits, but felt that face to face contact was more reassuring, or that they would gain most reassurance from regular clinical examination. Alternative provision will be required for these patients. For those who like the idea of this system, ideas such as telephone interview with a health professional should be explored further. For others, it may be possible to use this system in addition to routine clinic visits, thereby drawing attention to areas of concern that can be subsequently addressed during the routine visit.

The questionnaires used in this study have been well validated in the past and were considered by most of the patients to be comprehensive and address all areas of concern. For all elements of the questionnaire, there was reasonable correlation between scores when completed on paper compared with those completed using the telephone system. For some sections the level of correlation was inadequate to allow comparison between a paper version completed at baseline and a telephone version completed at some point in the future. Further exploration of the test-retest reliability will be needed with the questionnaire filled out using the telephone system on two occasions a few days apart. If greater correlation is seen in that situation, it would be expected that fewer people would record deterioration in score over the course of the year, so that in practice fewer people would need to be phoned to discuss their scores than was the case in the study here.

There are still questions regarding this new system which remain to be answered. A fuller study of the economic impact is required now that the acceptability and feasibility of this

type of follow up has been established. The impact of this follow up method on quality of life also remains to be seen. Certainly, patients did feel that more areas of concern were addressed with this system, but whether this leads to greater input from clinical staff and resultant resolution of problems remains to be confirmed. Both of these issues, along with an assessment of the impact of this type of follow up on long term outcome, should be further explored in a randomised controlled trial. This will require multi-centre input and long follow up to avoid the problem which has plagued similar trials in the past; that of failing to answer the questions posed due to lack of power.

Overall Conclusions

In this thesis, the value of routine clinical follow up after breast cancer has been investigated. At the outset, it was considered that at least some of the aims of follow up were poorly addressed by routine clinic visits, and so an alternative method of follow up was proposed and subjected to a prospective cohort study in order to assess the acceptability and feasibility of this alternative follow up method. In addition, the expectations which women have for follow up before embarking on that follow up was sought. The initial aims laid out at the beginning of this thesis have been addressed, and the overall conclusions of this work are as follows:

Routine clinical examination in follow up of breast cancer contributes little to detection of potentially treatable relapse in a modern cohort.

There is no good evidence that routine clinical examination improves outcome, either from modern retrospective analyses or from randomised controlled trials.

There is good evidence from the cohorts presented for the first time as part of this work that, for certain types of relapse, clinically detected relapses may do less well than relapse detected in other ways.

Mammography and patient self examination are of great value in the detection of potentially treatable relapse, and should form the core of any future follow up regimes.

The additional needs of patients and aims of follow up, those of detecting psychological problems and side effects of treatment, are not well met in the routine outpatient setting.

Despite the apparent shortcomings of routine outpatient clinics in meeting the aims of follow up, there have been few attempts to evaluate alternative methods of follow up.

Alternative methods of follow up are acceptable to patients, and follow up using an automated computer telephone system in particular was acceptable for a large proportion of women in routine follow up, with multiple benefits including convenience and a more thorough assessment of psychosocial concerns and side effects of treatment.

Further work is required to fully develop and assess this new and unique method of follow up

Further work

There are three areas where further work related to this project should be considered.

Firstly, much of the evidence in assessing the benefit of clinical examination in routine follow up has come from patients whose original breast cancer was treated by wide local excision. Further exploration of the benefits of routine clinical examination should be undertaken in patients post mastectomy. Mammography will have a role to play in detecting contralateral cancers in these patients, but not in detecting ipsilateral chest wall relapse. It was not possible to analyse this issue in a local cohort to any satisfactory degree due to inadequate patient numbers.

With regards the technical aspects of the computer telephone follow up system, an assessment of the test-retest reliability of the phone system should be undertaken, and this will be done as part of ongoing work by this author.

The telephone system should be subjected to randomised controlled trial in order to definitively establish the economic impact and the impact on patients' quality of life of this method of follow up. These questions could be answered in a local cohort. A larger multi-centre trial would be needed to establish the impact of this method of detection on survival, and this should be considered.

Reference List

- Adami HO, Bergstrom R, Hansen J (1985) Age at first primary as a determinant of the incidence of bilateral breast cancer. *Cancer* 55: 643-647
- Adeuyi-Dalton R, Ziebland S, Grunfeld E, Hall A (1998) Patients' views of routine hospital follow-up: A qualitative study of women with breast cancer in remission. *Psycho-oncology* 7: 436-439
- Altman DG, Lyman GH (1998) Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Research and Treatment* 52: 298-303
- American Society of Clinical Oncology (1999) Update of recommended breast cancer surveillance guidelines. *Journal of Clinical Oncology* 17: 1080-1082
- Anonymous (1998) Follow-up after treatment for breast cancer. The Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. *Canadian Medical Association Journal* 158: S65-S70
- Baildam AD, Keeling F, Thompson L, Bundred N, Hopwood P (2004) Nurse-led surgical follow up clinics for women treated for breast cancer - a randomised controlled trial. *Breast Cancer Research and Treatment* 88: S136-S137
- Baum M, Budzar AU, Cuzick J, Forbes J, Houghton JH, Klijn JG, Sahmoud T, ATAC trialists' Group (2002) Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. *Lancet* 359: 2131-2139
- Baum M, Buzdar A, Cuzick J, Forbes J, Houghton J, Howell A, Sahmoud T, The ATAC Trialists' Group: (2003) Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early-stage breast cancer: results of the ATAC (Arimidex, Tamoxifen Alone or in Combination) trial efficacy and safety update analyses. *Cancer* 98: 1802-1810
- Biem HJ, Turnell RW, D'Arcy C (2003) Computer telephony: automated calls for medical care. *Clinical and Investigative Medicine* 26: 259-268
- Bloom JR (1982) Social support, accommodation to stress and adjustment to breast cancer. *Social Science and Medicine* 16: 1329-1338
- Bloom JR, Spiegel D (1984) The relationship of two dimensions of social support to the psychological well-being and social functioning of women with advanced breast cancer. *Social Science and Medicine* 19: 831-837
- Bloom JR, Stewart SL, Johnston M, Banks P, Fobair P (2001) Sources of support and the physical and mental well-being of young women with breast cancer. *Social Science and Medicine* 53: 1513-1524

Brady MJ, Cella DF, Mo F, Bonomi AE, Tulsy DF, Lloyd SR, Deasy S, Cobleigh M, Shiimoto G (1997) Reliability and validity of the Functional Assessment of Cancer Therapy-Breast Quality-of-Life instrument. *Journal of Clinical Oncology* 15: 974-986

Brown L, Payne S, Royle G (2002) Patient initiated follow up of breast cancer. *Psychology* 11: 346-355

Casley-Smith JR (1985) Discussion of the definition, diagnosis and treatment of lymphoedema. In *Progress in lymphology. Proceedings of the Tenth International Congress on Lymphology.*, Casley-Smith JR, Piller NB (eds) pp 1-16. University of Adelaide press: Adelaide

Cella D, Eton DT, Fairclough DL, Bonomi P, Heyes AE, Silberman C, Wolf MK, Johnson DH (2002) What is a clinically meaningful change on the Functional Assessment of Cancer Therapy - Lung (FACT-L) questionnaire? Results from Eastern Cooperative Oncology Group (ECOG) study 5592. *Journal of Clinical Epidemiology* 55: 285-295

Chaudary MA, Millis RR, Hoskins EO, Halder M, Bullbrook RD, Cuzick J, Hayward JL (2006) Bilateral primary breast cancer: a prospective study of disease incidence. *British Journal of Surgery* 71: 711-714

Chien AJ, Goss PE (2006) Aromatase inhibitors and bone health in women with breast cancer. *Journal of Clinical Oncology* 24: 5305-5312

Churn M, Kelly V (2001) Outpatient follow-up after treatment for early breast cancer: Updated results after 5 years. *Clinical Oncology* 13: 187-194

Clark DH, Le MG, Sarrazin D, Lacombe MJ, Fontaine F, Travagli JP, May-Levin F, Contesso G, Arriagada R (1985) Analysis of loco-regional relapses in patients with early breast cancers treated by excision and radiotherapy: experience of the institut Gustave-Roussy. *International Journal of Radiation Oncology, Biology and Physics* 11: 137-145

Cohen I (2004) Endometrial pathologies associated with postmenopausal tamoxifen treatment. *Gynaecologic Oncology* 94: 256-266

Collins RF, Bekker HL, Dodwell DJ (2004) Follow-up care of patients treated for breast cancer: a structured review. *Cancer Treatment Reviews* 30: 19-35

Cook DJ, Sackett DL, Spitzer W (1995) Methodologic guidelines for systematic reviews of randomized controlled trials in health care from the Potsdam consultation on meta-analysis. *Journal of Clinical Epidemiology* 48: 167-171

Coombes RC, Hall E, Gibson LJ, Paridaens R, Jassem J, Delozier T, Jones SE, Alvarez I, Bertelli G, Ortmann O, Coates AS, Bajetta E, Dodwell D, Coleman RE, Fallowfield LJ, Mickiewicz E, Andresen J, Lonning PE, Cocconi G, Stewart A, Stuart N, Snowdon CF, Carpentieri M, Massimini G, Bilss JM, van de Velde C, Intergroup Exemestane Study (2004) A randomized trial of exemestane after two to three years of tamoxifen therapy in

postmenopausal women with primary breast cancer. *New England Journal of Medicine* 350: 1081-1092

Coster S, Poole K, Fallowfield L (2001) The validation of a quality of life scale to assess the impact of arm morbidity in breast cancer patients post-operatively. *Breast Cancer Research and Treatment* 68: 273-282

Couzi R, Helzlsouer KJ, Fetting JH (1995) Prevalence of menopausal symptoms among women with a history of breast cancer and attitudes toward oestrogen replacement therapy. *Journal of Clinical Oncology* 13: 2737-2744

Dalberg K, Mattsson A, Sandelin K, Rutqvist LE (1998) Outcome of treatment for ipsilateral breast tumour recurrence in early-stage breast cancer. *Breast Cancer Research and Treatment* 49: 69-78

Day R, Ganz PA, Constantino JP, Cronin WM, Wickerham DL, Fisher B (1999) Health-related quality of life and tamoxifen in breast cancer prevention: A report from the National Surgical Adjuvant Breast and Bowel Project P-1 study. *Journal of Clinical Oncology* 7: 2659-2669

de Bock GH, Bonnema J, van Der Hage J, Kievit J, van de Velde CJH (2004) Effectiveness of routine visits and routine tests in detecting isolated locoregional recurrences after treatment for early stage invasive breast cancer: a meta analysis and systematic review. *Journal of Clinical Oncology* 22: 4010-4018

Del Turco MR, Palli D, Carridi A, Ciatto S, Pacini P, Distante V (1994) Intensive diagnostic follow-up after treatment of primary breast cancer. A randomized trial. National Research Council Project on Breast Cancer Follow-up. *Journal of the American Medical Association* 271: 1593-1597

Denton S, on behalf of the EONS Advanced Working Group for Living With Advanced Breast Cancer Hormone Treatment (1998) Living with hormone therapy for advanced breast cancer: a European survey of oncologists' and nurses' perceptions. *European Journal of Cancer* 31A:

Dewar JA, Kerr GR (1985) Value of routine follow up of women treated for early carcinoma of the breast. *British Medical Journal* 291: 1464-1467

Dickersin K, Scherer R, Lefebvre C (1994) Systematic Reviews: Identifying relevant studies for systematic reviews. *British Medical Journal* 309: 1286-1291

Dixon M, Norman B (1996) Most recurrences after breast conservation are detected by regular hospital visits. *British Medical Journal* 313: 1548

Donnelly J, Mack P, Donaldson LA (2001) Follow-up of breast cancer: time for a new approach? *International Journal of Clinical Practice* 55: 431-433

Donnelly P, Hiller L, Bathers S, Bowden S, Coleman R (2007) Questioning specialists' attitudes to breast cancer follow-up in primary care. *Annals of Oncology*

Early Breast Cancer Trialists' Collaborative Group (1998) Tamoxifen for early breast cancer: An overview of the randomized trials. *Lancet* 351: 1451-1467

Eastell R, hannon RA, Cuzick J, Dowsett M, Clack G, Adams JE, ATAC trialists' Group (2006) Effect of an aromatase inhibitor on bmd and bone turnover markers: 2-year results of the Anastrozole, Tamoxifen, Alone or in Combination (ATAC) trial. *Journal of Bone and Mineral Research* 21: 1215-1223

Elder EE, Kennedy CW, Gluch L, Carmalt HL, Janu NC, Jospeh MG, Donellan MJ, Molland JG, Gillett DJ (2006) Patterns of breast cancer relapse. *European Journal of Surgical Oncology* 32: 922-927

Elkhuizen PHM, van de Vijver MJ, Hermans J, Zonderland HM, van de Velde CJH, Leer JH (1998) Local recurrence after breast conserving therapy for invasive breast cancer: high incidence in young patients and association with poor survival. *International Journal of Radiation Oncology, Biology and Physics* 40: 859-867

Eton DT, Cella D, Yost KJ, Yount SE, Peterman Ah, Neuberg DS, Sledge GW, Wood WC (2004) A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *Journal of Clinical Epidemiology* 57: 898-910

Fallowfield L, Cella D, Cuzick J, Francis S, Locker G, Howell A (2004) Quality of life of postmenopausal women in the Arimidex, Tamoxifen, Alone or in Combination (ATAC) Adjuvant Breast Cancer Trial. *Journal of Clinical Oncology* 22: 4261-4271

Fallowfield LJ, Leaity SK, Howell A, Benson S, Cella D (1999) Assessment of quality of life in women undergoing hormonal therapy for breast cancer: Validation of an endocrine symptom subscale for the FACT-B. *Breast Cancer Research and Treatment* 55: 189-199

Fentiman IS, Matthews PN, Davison OW, Millis RR, Hayward JL (1985) Survival following local skin recurrence after mastectomy. *British Journal of Surgery* 72: 14-16

Fialka-Moser V, Crevenna R, Korpan M, Quittan M (2003) Cancer rehabilitation: particularly with aspects on physical impairment. *Journal of Rehabilitation Medicine* 35: 153-162

Fornander T, Rutqvist LE, Cedermark B, Glas U, Mattsson A, Silfversward C, Skoog L, Somell L, Theve T, Wilking N, et al (1989) Adjuvant tamoxifen in early breast cancer: occurrence of new primary cancers. *Lancet* 1: 117-120

Fortin A, Larochelle M, Laverdiere J, Lavertu S, Tremblay D (1999) Local failure is responsible for the decrease in survival for patients with breast cancer treated with conservative surgery and postoperative radiotherapy. *Journal of Clinical Oncology* 17: 101-109

Fowble B, Solin LJ, Schultz DJ, Rubenstein J, Goodman RL (1990) Breast recurrence following conservative surgery and radiation: patterns of failure, prognosis and pathological findings in mastectomy specimens with implications for treatment. *International Journal of Radiation Oncology, Biology and Physics* 19: 833-842

Fracchhia AA, Borgen PL (1991) Bilateral breast cancer. *Seminars in Surgical Oncology* 7: 305

Freedman GM, Anderson PR, Hanlon AL, Eisenberg DF, Nicolaou N (2005) Pattern of local recurrence after conservative surgery and whole-breast irradiation. *International Journal of Radiation Oncology, Biology and Physics* 61: 1328-1336

Fromme EK, Eilers KM, Mori M, Hsieh Y, Beer TM (2004) How accurate is clinician reporting of chemotherapy adverse events? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *Journal of Clinical Oncology* 22: 3485-3490

Garreau JR, DeLaMelena T, Walts D, Karamlou K, Johnson N (2006) Side effects of aromatase inhibitors versus tamoxifen: the patients' perspective. *The American Journal of Surgery* 192: 496-498

Goss PE, Ingle JN, Martino S, Robert NJ, Muss HB, Piccart MJ, Catiglione M, Tu D, Shepher LE, Pritchard KL, Livingston RB, Davidson NE, Norton L, Perez EA, Abrams JS, Therasse P, Palmer MJ, Pater JL (2003) A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *New England Journal of Medicine* 349: 1793-1802

Greenhalgh T (1997) Chapter 4: Assessing methodological quality. In *How to read a paper*, BMJ publishing group: London

Grogan M, Rangan A, Gebiski V, Boyages J (2002) The value of follow-up of patients with early breast cancer treated with conservative surgery and radiation therapy. *The Breast* 11: 163-169

Grunfeld E, Dhesy-Thind S, Levine M, for the Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer (2005) Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment for breast cancer (summary of the 2005 update). *Canadian Medical Association Journal* 172: 1319-1320

Grunfeld E, Gray A, Mant D, Yudkin P, Adewuyi-Dalton R, Coyle D, Stewart J, Fitzpatrick R, Vessey M (1999) Follow-up of breast cancer in primary care vs specialist care: results of an economic evaluation. *British Journal of Cancer* 79: 1227-1233

Grunfeld E, Levine MN, Julian JA, Coyle D, Szechtman B, Mirsky D, Verma S, Dent S, Sawka C, Pritchard KI, Ginsburg D, Wood M, Whelan T (2006) Randomized trial of long-term follow-up for early-stage breast cancer: a comparison of family physician versus specialist care. *Journal of Clinical Oncology* 24: 848-854

Grunfeld E, Mant D, Yudkin P, Adewuyi-Dalton R, Cole D, Stewart J, Fitzpatrick R, Vessey M (1996) Routine follow up of breast cancer in primary care: randomised trial. *British Medical Journal* 313: 665-669

Grunfeld E, Noorani H, McGahan I, Paszat L, Coyle D, van Walraven C, Joyne J, Sawka C (2002) Surveillance mammography after treatment of primary breast cancer: a systematic review. *The Breast* 11: 228-235

Gulliford T, Opomu M, Wilson E, Hanham I, Epstein R (1997) Popularity of less frequent follow up for breast cancer in randomised study: initial findings from the hotline study. *British Medical Journal* 314: 174-177

Haagensen DC (1971) In *Diseases of the breast, 2nd Ed*, pp 718-724. W B Saunders: Philadelphia, Pennsylvania

Haffty BG, Fischer D, Beinfield M, McKhann C (1991) Prognosis following local recurrence in the conservatively treated breast cancer patient. *International Journal of Radiation Oncology, Biology and Physics* 21: 293-298

Hall A, A'hern R, Fallowfield L (1999) Are we using appropriate self-report questionnaires for detecting anxiety and depression in women with early breast cancer? *European Journal of Cancer* 35: 79-85

Hardman A, Maguire P, Crowther D (1989) The recognition of psychiatric morbidity on a medical oncology ward. *Journal of Psychosomatic research* 33: 235-239

Hardy JR, Baum M (1991) Lymphoedema - prevention rather than cure. *Annals of Oncology* 2: 532-533

Heys SD, Chaturvedi S, Hutcheon AW, Sarkar TK (2005) Guidelines, guidelines and more guidelines: and we still do not know how to follow-up patients with breast cancer. *World Journal of Surgical Oncology* 3: 54

Higgins JPT, Green S, Editors (2006) *Cochrane Handbook for Systematic Reviews of Interventions* 4.2.6. John Wiley & Sons Ltd: Chichester, UK

Hiramanek N (2004) Breast cancer recurrence: follow up after treatment for primary breast cancer. *Postgraduate Medical Journal* 80: 172-176

Hodgkinson K, Butow P, Hunt GE, Pendlebury S, Hobbs KM, Wain G (2006) Breast cancer survivors' supportive care needs 2-10 years after diagnosis. *Support Care Cancer* Epub. ahead of print:

Hodgkinson K, Butow P, Hunt GE, Pendlebury S, Hobbs KM, Wain G (2007) Breast cancer survivors' supportive care needs 2-10 years after diagnosis. *Supportive Care in Cancer* 15: 515-523

- Hopwood P (1998) Living with advanced breast cancer: development and application of a clinical checklist for patients on endocrine therapy. *The Breast* 7: 14-21
- Horn PL, Thompson WD (1988) Risk of contralateral breast cancer. *Cancer* 62: 412-424
- Howell D, Watson M (2005) Evaluation of a pilot nurse-led, community-based treatment programme for lymphoedema. *International Journal of Palliative Nursing* 11: 62-69
- Hussain ST, Gui GPH, Lee KS, Plowman PN, Gilmore OJA, Allum WH (1995) Detection of loco-regional recurrence after breast-conserving surgery and radiotherapy. *Journal of the Royal College of Surgeons of Edinburgh* 40: 163-166
- Jack WJL, Kerr GR, Kunkler IH (1998) Long-term follow-up after breast conservation: The Edinburgh experience. *The Breast* 7: 80-84
- Jacobs HJM, van Dijck JAAM, de Kleijn EMHA, Kiemeny LALM, Verbeek ALM (2001) Routine follow-up examinations in breast cancer patients have minimal impact on life expectancy: A simulation study. *Annals of Oncology* 12: 1107-1113
- James ND, Guerrero D, Brada M (1994) Who should follow up cancer patients? Nurse specialist based outpatient care and the introduction of a phone clinic system. *Clinical Oncology* 6: 283-287
- Jüni P, Witschi A, Bloch R, Egger M (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association* 282: 1054-1060
- Kemperman H, Borger J, Hart A, Peterse H, Bartelink H, van Dongen J (1995) Prognostic factors for survival after breast conserving therapy for stage I and II breast cancer. The role of local recurrence. *European Journal of Cancer* 31A: 690-698
- Khatcheressian JL, Smith TJ (2006) Randomized trial of long-term follow-up for early-stage breast cancer: A comparison of family physician versus specialist care. *Journal of Clinical Oncology* 6: 835-837
- Khatcheressian JL, Wolff AC, Smith TJ, Grunfeld E, Muss HB, Vogel VG, Halberg F, Somerfield MR, Davidson NE (2006) American Society of Clinical Oncology 2006 update of the breast cancer follow-up and management guidelines in the adjuvant setting. *Journal of Clinical Oncology* 24: 5091-5097
- Kiebert GM, Welvaart K, Kievit J (1993) Psychological effects of routine follow-up on cancer patients after surgery. *European Journal of Surgery* 159: 601-607
- Kobak KA, Taylor LH, Dottl SL, Greist JH, Jefferson JW, Burroughs D, Mantle JM, Katzelnick DJ, Norton R, Henk HJ, Serlin RC (1997) A computer-administered telephone interview to identify mental disorders. *Journal of the American Medical Association* 278: 905-910

- Koinberg IL, Fridlund B, Engholm GB, Holmberg L (2004) Nurse-led follow-up on demand or by a physician after breast cancer surgery: a randomised study. *European Journal of Oncology Nursing* 8: 109-117
- Kokko R, Hakama M, Holli K (2005) Follow-up cost of breast cancer patients with localized disease after primary treatment: a randomised trial. *Breast Cancer Research and Treatment* 93: 255-260
- Kuehn T, Klauss W, Darsow M, Regele S, Flock F, Maiterth C, Dahlbender R, Wendt I, Kreienberg R (2000) Long-term morbidity following axillary dissection in breast cancer patients: Clinical assessment, significance for life quality and the impact of demographic, oncologic and therapeutic factors. *Breast Cancer Research and Treatment* 64: 286
- Kurtz JM, Almaric R, Brandone H, Ayme Y, Jacquemier J, Pietra J-C, Hans D, Pollet J-F, Bressac C, Spitalier J-M (1989) Local recurrence after breast conserving therapy: frequency, time course and prognosis. *Cancer* 63: 1912-1917
- Kurtz JM, Spitalier J-M, Almaric R, Brandone H, Ayme Y, Jacquemier J, Hans D, Bressac C (1990) The prognostic significance of late local recurrence after breast conserving therapy. *International Journal of Radiation Oncology, Biology and Physics* 18: 87-93
- Kwan W, Jackson J, Weir LM, Dingee C, McGregor G, Olivotto IA (2002) Chronic arm morbidity after curative breast cancer treatment: Prevalence and impact on quality of life. *Journal of Clinical Oncology* 20: 4242-4248
- Lampic C, Wennberg A, Schill JE, Brodin O, Glimelius B, Sjöden P-O (1994) Anxiety and cancer-related worry of cancer patients at routine follow-up visits. *Acta Oncologica* 33: 119-125
- Lash TL, Fox MP, Silliman RA (2006) Reduced mortality rate associated with annual mammograms after breast cancer therapy. *The Breast Journal* 12: 2-6
- Laupacis A, Wells G, Richardson WS, Tugwell P (1994) Users' guide to the medical literature: How to use an article on prognosis. *Journal of the American Medical Association* 272: 234-237
- Lees A, Gabos A, Jenkins H (1997) Investigations for staging and follow-up of breast cancer patients. *Cancer Prevention and Control* 1: 157-160
- Leonard RCF, Lee L, Harrison ME (1996) Impact of side-effects associated with endocrine treatments for advanced breast cancer: clinicians' and patients' perceptions. *Breast* 5: 259-264
- Loomer L, Brockschmidt JK, Muss HB, Saylor G (1991) Postoperative follow-up of patients with early breast cancer. *Cancer* 67: 55-60
- Macquart-Moulin G, Viens P, Bouscary ML, Genre d, Resbeut M, Gravis G, Camerlo J, Maraninchi D, Moatti JP (1997) Discordance between physician's estimations and breast

cancer patients' self-assessment of side-effects of chemotherapy: an issue for quality of care. *British Journal of Cancer* 76: 1640-1645

Maguire P (1994) Psychological aspects. In *ABC of breast diseases*, Dixon JM (ed) BMJ publishing group: London

Maher J, Bradburn J, Adewuyi-Dalton R (1995) Follow up in breast cancer. *British Medical Journal* 311: 54

Mahoney L (1986) Methods for detecting locally recurrent and contralateral second primary breast cancer. *The Canadian Journal of Surgery* 29: 372-373

Mansi JL, Earl HM, Powles TJ, Coombes RC (1988) Tests for detecting recurrent disease in the follow up of patients with breast cancer. *Breast Cancer Research and Treatment* 11: 249-254

Maunsell E, Brisson J, Deschenes L (1993) Arm problems and psychological distress after surgery for breast cancer. *Canadian Journal of Surgery* 36: 315-320

Mellink W, Holland R, Hendriks J, Peeters P, Rutgers E, Van Daal W (1991) The contribution of routine follow-up mammography to an early detection of asynchronous contralateral breast cancer. *Cancer* 67: 1844-1848

Millard RW, Carver JR (1999) Cross-sectional comparison of live and interactive voice recognition administration of the SF-12 health status survey. *The American Journal of Managed Care* 5: 153-159

Moher D, Cook DJ, Eastwood S, Olkin I, Drummond R, Stroup DF et al (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 354: 1896-1900

Moher D, Jadad A, Nichol G, Penman M, Tugwell T, Walsh S (1995) Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Controlled Clinical Trials* 16: 62-73

Moher D, Jadad AR, Tugwell P (1996) Assessing the quality of randomized controlled trials: current issues and future directions. *International Journal of Technical Assessments in Health Care* 12: 195-208

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 352: 609-613

Mohoney L (1986) Methods for detecting locally recurrent and contralateral second primary breast cancer. *The Canadian Journal of Surgery* 29: 372-373

Montgomery DA, Krupa K, Cooke TG (2007a) Alternative Methods of Follow Up in Breast Cancer: A Systematic Review of the Literature. *British Journal of Cancer* 96: 1625-1632

Montgomery DA, Krupa K, Jack WJL, Kerr GR, Kunkler IH, Thomas J, Dixon JM (2007b) Changing Pattern of the Detection of Loco-regional Relapse in Breast Cancer: The Edinburgh Experience. *British Journal of Cancer* 96: 1802-1807

Morris AD, Morris RD, Wilson JF, White J, Steinberg S, Okunieff P, Arriagada R, Le MG, Blichert-Toft M, van Dongen JA (1997) Breast-conserving therapy vs mastectomy in early-stage breast cancer: A meta-analysis of 10-year survival. *Cancer Journal from Scientific American* 3: 6-12

Morris S, Corder AP, Taylor I (1992) What are the benefits of routine breast cancer follow up? *Postgraduate Medical Journal* 68: 904-907

Mortimer PS (1998) The pathophysiology of lymphoedema. *Cancer* 83: 2798-2802

Mortimer PS, Bates DO, Brassington HD, Stanton AWB, Strachan DP, Levick JR (1996) The prevalence of arm oedema following treatment for breast cancer. *Quarterly Journal of Medicine* 89: 377-380

Mozes M, Papa MZ, Karasik A, Reshef A, Adar R (1982) The role of infection in post-mastectomy lymphedema. *Annals of Surgery* 14: 73-83

Muss HB, McNamara MJC, Connelly RA (1988) Follow-up after stage II breast cancer: a comparative study of relapsed versus nonrelapsed patients. *American Journal of Clinical Oncology* 11: 451-455

National Institute for Clinical Excellence. Guidance on Cancer Services: Improving Outcomes in Breast Cancer. Manual Update.

<http://www.nice.org.uk/page.aspx?o=csgbcguidance>. 2002.

Ref Type: Internet Communication

NHS Cancer Screening Programmes. NHS Breast Screening Review (2003) NHS Breast Screening Programme. 2003. England.

Ref Type: Pamphlet

Nishimura R, Akizuki M, Tashima R, Ootao R (2006) Investigation of factors related to periods to ipsilateral breast tumor recurrence after breast-conserving surgery and measures for preventing recurrence in early breast cancer. *Breast Cancer* 13: 152-158

Ojeda MB, Alonson MC, Bastus R, Alba E, Piera JM, Lopez Lopez JJ (1987) Follow-up of breast cancer stages I and II. An analysis of some common methods. *European Journal of Cancer and Clinical Oncology* 23: 419-423

Ormiston MC, Timoney AG, Quereshi AR (1985) Is follow up of patients after breast cancer worthwhile? *Journal of the Royal Society of Medicine* 78: 920-921

Pain SJ, Purushotham AD (2000) Lymphoedema following surgery for breast cancer. *British Journal of Surgery* 87: 1128-1141

- Palli D, Russo A, Saieva C, Ciatto S, Del Turco MR, Distante V, Pacini P (1999) Intensive vs clinical follow-up after treatment of primary breast cancer: 10-year update of a randomized trial. *Journal of the American Medical Association* 281: 1586
- Paradiso A, Nitti P, Frezza P, Scorpiglione N, on behalf of G.S.Bio.Ca.M. (1995) A survey in Puglia: the attitudes and opinions of specialists, general physicians and patients on follow up practice. *Annals of Oncology* 6: s53-s56
- Passik SD, Donald MV (1998) Psychosocial aspects of upper extremity lymphedema in women treated for breast cancer. *Cancer* 83: 2817-2820
- Passik SD, Dugan W, McDonald MV, Rosenfeld B, Theobald DE, Edgerton S (1998) Oncologists' recognition of depression in their patients with cancer. *Journal of Clinical Oncology* 16: 1594-1600
- Passik SD, Newman M, Brennan M, Tunkell R (1995) Predictors of psychological distress, sexual dysfunction and physical functioning among women with upper extremity lymphoedema related to breast cancer. *Psycho-oncology* 4: 255-263
- Pennery E, Mallet J (2000) A preliminary study of patients' perceptions of routine follow up after treatment for breast cancer. *European Journal of Oncology Nursing* 4: 138-145
- Petrek JA, Heelan MC (1998) Incidence of breast carcinoma-related lymphedema. *Cancer* 83: 2776-2781
- Prior P, Waterhouse JAH (1978) Incidence of bilateral tumours in a population-based series of breast cancer patients: I. Two approaches to an epidemiological analysis. *British Journal of Cancer* 37: 620-634
- Rainsbury D (1996) Routine follow up of breast cancer in primary care: Follow up by non-specialists should not be encouraged. *British Medical Journal* 313: 1547
- Recht A, Schnitt SJ, Connolly JL, Rose MA, Silver B, Come S, Henderson IC, Slavin S, Harris JR (1989) Prognosis following local or regional recurrence after conservative surgery and radiotherapy for early stage breast carcinoma. *International Journal of Radiation Oncology, Biology and Physics* 16: 3-9
- Renton JP, Twelves CJ, Yuille FAP (2002) Follow-up in women with breast cancer: the patients' perspective. *Breast* 11: 257-261
- Rodger A (1997) Routine follow up of breast cancer in primary care. *British Medical Journal* 314: 1129-1130
- Rojas MP, Telaro E, Russo A, Moschetti I, Coe L, Fossati R, Palli D, del Roselli TM, Liberati A (2005) Follow-up strategies for women treated for early breast cancer. *Cochrane Database of Systematic Reviews* 25: CD001768
- Rose G, Barker DJP (1994) *Epidemiology for the uninitiated*. BMJ Publications: London

Rutgers EJT, van Rossum AB, Peterse JL, Cohen P, Borger JH, Kemperman HWPM (1991) Breast-conserving surgery for invasive carcinoma: diagnosis of local recurrence. *The Netherlands Journal of Surgery* 43: 110-113

Samant RS, Olivotto IA, Jackson JSH, Mates D (2001) Diagnosis of metachronous contralateral breast cancer. *The Breast Journal* 7: 405-410

Saphner T, Tormey DC, Gray R (1996) Annual hazard rates of recurrence for breast cancer after primary therapy. *Journal of Clinical Oncology* 14: 2738-2746

Scanlon EF, Oviedo MA, Cunningham MP, Caprini JA, Khandekar JD, Cohen E, Robinson B, Stein E (1980) Preoperative and follow up procedures on patients with breast cancer. *Cancer* 46: 977-979

Schrenk P, Rieger R, Shamiyeh A, Wayand W (2000) Morbidity following sentinel lymph node biopsy versus axillary lymph node dissection for patients with breast carcinoma. *Cancer* 88: 608-614

Schulz KF, Chalmers I, Hayes RJ, Altman D (1995) Empirical evidence of bias. *Journal of the American Medical Association* 273: 408-412

Schunemann H, Willich N (1998) Lymphoedema of the arm after primary treatment of breast cancer. *Anticancer Research* 18: 2235-2236

Scottish cancer registry, ISD.

www.isdscotland.org/isd/cancer_definition.jsp?pContentID=1421&p_applic=CCC&p_service=Content.show&. 2005.

Ref Type: Generic

Segestrom K, Bjerle P, Nystrom A (1991) Importance of time in assessing arm and hand function after treatment of breast cancer. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery* 25: 241-244

Seidman H. Cancer of the breast: statistical and epidemiological data. 1972.

Ref Type: Generic

Sener SF, Winchester DJ, Martz CH, Feldman JL, Cavanaugh JA, Winchester DP, Weigel B, Bonnefoi K, Kirby K, Morehead C (2001) Lymphedema after sentinel lymphadenectomy for breast carcinoma. *Cancer* 92: 748-752

Smith BJ, Darzins PJ, Quinn M, Heller RF (1992) Modern methods of searching the medical literature. *Medical Journal of Australia* 157: 603-611

Snee M (1996) Follow-up of women treated for breast cancer. *Clinical Oncology* 8: 85-89

Sprangers MAG, Groenvold M, Arraras JI, Franklin J, te Velde A, Muller M, Franzini L, Williams A, de Haes HCJM, Hopwood P, Cull A, Aaronson NK (1996) The European Organization for Research and Treatment of Cancer breast cancer - specific Quality-of Life

questionnaire module: First results from a three country field study. *Journal of Clinical Oncology* 14: 2756-2768

Tate JT, Clifford PC, Royle GT, Buchanan RB (1989) After surgery for breast cancer: Which mode of surveillance? *Journal of the Royal Society of Medicine* 82: 737-738

The Association of Breast Surgery @ BASO, Royal College of Surgeons of England (2005) Guidelines for the management of symptomatic breast disease. *European Journal of Surgical Oncology* 31: S1-S21

The ATAC Trialists' Group: (2005) Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* 365: 60-62

The GIVIO investigators (1994) Impact of follow-up testing on survival and health-related quality of life in breast cancer patients. A multicenter randomized controlled trial. *Journal of the American Medical Association* 271: 1587-1592

The Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer (1998) Follow-up after treatment for breast cancer. *Canadian Medical Association Journal* 158(suppl 3): S65-S70

Tobin MB, Lacey HJ, Meyer L, Mortimer P (1993) The psychological morbidity of breast cancer related arm swelling. *Cancer* 72: 3248-3252

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL (1998) Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 280: 867-873

United States National Library of Medicine. List of journals indexed for Medline. www.nlm.nih.gov/tsd/serials/lji.html . 2006.

Ref Type: Electronic Citation

Valagussa P, Tess JD, Rossi A, Tancini G, Banfi A, Bonadonna G (1981) Adjuvant CMF effect on site of first recurrence, and appropriate follow up intervals, in operable breast cancer with positive axillary nodes. *Breast Cancer Research and Treatment* 1: 349-356

Valente SM, Saunders JM, Zichi Cohen M (1994) Evaluating depression among patients with cancer. *Cancer Practice* 2: 65-71

van der Sangen MJC, van de Poll-Franse LV, Roumen RMH, Rutten HJT, Coebergh JWW, Vreugdenhil G, Voogd AC (2006) The prognosis of patients with local recurrence more than five years after breast conservation therapy for invasive breast carcinoma. *European Journal of Surgical Oncology* 32: 34-38

Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A, Aguilar M, Marubini E (2002) Twenty-year follow-up of a randomized study comparing breast-conserving surgery

with radical mastectomy for early breast cancer. *New England Journal of Medicine* 347: 1227-1232

Veronesi U, Marubini E, Del Vecchio M, Manzari A, Andreola S, Greco M, Luini A, Merson M, Saccozzi R, Rilke F, et al (1995) Local recurrences and distant metastases after conservative breast cancer treatments: partly independent events. *Journal of the National Cancer Institute* 87: 19-22

Veronesi U, Marubini E, Mariani L, Galimberti V, Luini A, Veronesi P, Salvadori B, Zucali R (2001) Radiotherapy after breast-conserving surgery in small breast carcinoma: long-term results of a randomized trial. *Annals of Oncology* 12: 997-1003

Voogd AC, Cranenbroek S, de Boer R, Roumen RMH, Rutten HJT, van der Sangen MJC (2005) Long-term prognosis of patients with axillary recurrence after axillary dissection for invasive breast cancer. *European Journal of Surgical Oncology* 31: 485-489

Voogd AC, van Tienhoven G, Peterse HL, Crommelin MA, Rutgers EJT, van de Velde CJ, van Geel BN, Slot A, Rodrigus PT, Jobsen JJ, von Meyenfeldt MF, Coebergh JW, for the Dutch study group on local recurrence after breast conservation (BORST) (1999) Local recurrence after breast conservation therapy for early stage breast carcinoma: Detection, treatment and outcome in 266 patients. *Cancer* 85: 437-446

Wasson J, Gaudette C, Whaley F, Suavigne A, Baribeau P, Welch HG (1992) Telephone care as a substitute for routine clinic follow-up. *Journal of the American Medical Association* 267: 1788-1793

Whelan T, Clark R, Roberts R, Levine M, Foster G (1994) Ipsilateral breast tumour recurrence post lumpectomy is predictive of subsequent mortality: results from a randomised trial. Investigators of the Ontario Clinical Oncology Group. *International Journal of Radiation Oncology, Biology and Physics* 30: 11-16

Wilke LG, McCall LM, Posther KE, Whitworth PW, Reintgen DS, Leitch AM, Gabram SGA, Lucci A, Cox CE, Hunt KK, Herndon JE, Giuliano AE (2006) Surgical complications associated with sentinel lymph node biopsy: results from a prospective international cooperative group trial. *Annals of Surgical Oncology* 13: 491-500

Winchester DP, Sener SF, Khandekar JD, Oviedo MA, Cunningham MP, Caprini JA, Burkett FE, Scanlon EF (1979) Symptomatology as an indicator of recurrent or metastatic breast cancer. *Cancer* 43: 956-960

Winer EP, Hudis C, Burstein HJ, Wolff AC, Pritchard KI, Ingle JN, Chlebowski RT, Gelber R, Edge SB, Gralow J, Cobleigh MA, Mamounas EP, Goldstein LJ, Whelan TJ, Powles TJ, Bryant J, Perkins C, Perotti J, Braun S, Langer AS, Browman GP, Somerfield MR (2005) American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for postmenopausal women with hormone receptor-positive breast cancer: status report 2004. *Journal of Clinical Oncology* 23: 619-629

Yap KPL, McCreedy DR, Narod S, Manchul LA, Trudeau M, Fyles A (2003) Factors influencing arm and axillary symptoms after treatment for node negative breast carcinoma. *Cancer* 97: 1369-1375

Young JR, Wilson AR (1995) The use of imaging in the follow-up of patients with breast cancer. Royal College of Radiologists Breast Group. *Clinical Oncology* 7: 239-240

Appendices

Appendix 1.1:

Search string used to identify relevant studies

- 1 Adult/
- 2 "Aged, 80 and over"/ or Aged/ or Middle Aged/
(neoplasm\$ or cancer\$ or tumour\$ or tumor\$ or
- 3 malign\$ or carcinoma\$ or adenocarcinoma\$ or
sarcoma\$ or dcis or ductal or infiltrat\$ or intraduct\$
or lobular\$ or medullary\$).mp.
- 4 breast.mp.
- 5 3 and 4
- 6 (follow up or follow-up or relapse).mp. [mp=ti, ab,
tx, kw, ct, ot, sh, hw]
- 7 1 and 2 and 5 and 6
- 8 quality of life.mp. or "Quality of Life"/
- 9 patient satisfaction.mp. or Patient Satisfaction/
- 10 survival rate.mp. or Survival Rate/
- 11 8 or 9 or 10
- 12 7 and 11

Appendix 1.2:

Search string used to identify relevant studies in CancerLit

- 1 Adult/
- 2 "Aged, 80 and over"/ or Aged/ or Middle Aged/
(neoplasm\$ or cancer\$ or tumour\$ or tumor\$ or
- 3 malign\$ or carcinoma\$ or adenocarcinoma\$ or
sarcoma\$ or dcis or ductal or infiltrat\$ or intraduct\$
or lobular\$ or medullary\$).mp.
- 4 breast.mp.
- 5 3 and 4
- 6 (follow up or follow-up or relapse).mp. [mp=ti, ab,
tx, kw, ct, ot, sh, hw]
- 7 1 and 2 and 5 and 6
- 8 quality of life.mp. or "Quality of Life"/
- 9 patient satisfaction.mp. or Patient Satisfaction/
- 10 survival rate.mp. or Survival Rate/
- 11 8 or 9 or 10
- 12 7 and 11
- 13 11 and randomised

Appendix 1.3:

Search string used to identify relevant studies in Web of Sciences

1 TS=(neoplasm* or cancer* or tumour* or tumor* or
malign* or carcinoma* or adenocarcinoma* or
sarcoma* or dcis or ductal or infiltrat* or intraduct* or
lobular* or medullary*)
2 TS=breast
3 1 and 2
4 TS=(follow up or follow-up or relapse)
5 1 and 2 and 4
6 1 and 2 and 4 and randomised

Appendix 1.4:

Details of additional authors involved in the review process:

Author 2 (KK):

Katherine Krupa

Consultant surgeon in Breast and Endocrine surgery

Paisley Royal Alexandra Hospital

Corsebar Road

Paisley

PA2 9PN

Author 3 (TGC):

Professor T G Cooke

St Mungo Professor of Surgery

Department of Surgery

Glasgow Royal Infirmary

Alexandra Parade

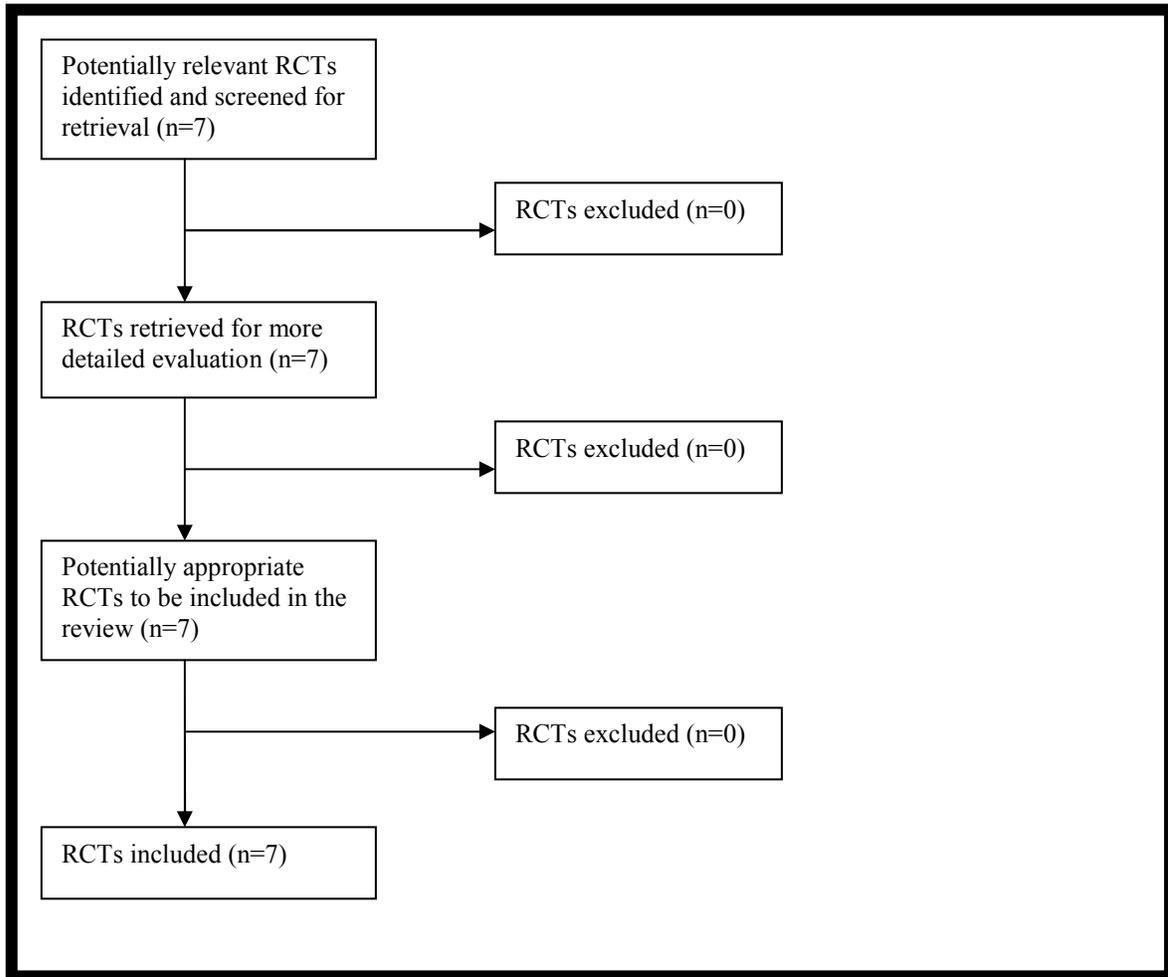
Glasgow

G31 2ER

Appendix 1.5:
QUOROM statement for systematic review

Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement checklist

| Heading | Subheading | Descriptor | Reported? (Y/N) | Page number |
|---------------------|-----------------------------|---|---|------------------|
| Title | | Identify the report as a meta-analysis [or systematic review] of RCTs | Y | |
| Abstract | | Use a structured format | Not Applicable in the context of thesis, Although completed in published paper (see ref[Montgomery DA et al., 2007a]) | |
| | | Describe | | |
| | Objectives | The clinical question explicitly | N/A | |
| | Data sources | The databases (ie, list) and other information sources | N/A | |
| | Review methods | The selection criteria (ie, population, intervention, outcome, and study design); methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis in sufficient detail to permit replication | N/A | |
| | Results | Characteristics of the RCTs included and excluded; qualitative and quantitative findings (ie, point estimates and confidence intervals); and subgroup analyses | N/A | |
| | Conclusion | The main results | N/A | |
| | | Describe | | |
| Introduction | | The explicit clinical problem, biological rationale for the intervention, and rationale for review | Y | 77 |
| Methods | Searching | The information sources, in detail (eg, databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status, language of publication) | Y | 94 |
| | Selection | The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design) | Y | 95 |
| | Validity assessment | The criteria and process used (eg, masked conditions, quality assessment, and their findings) | Y | 97 |
| | Data abstraction | The process or processes used (eg, completed independently, in duplicate) | Y | 97 |
| | Study characteristics | The type of study design, participants' characteristics, details of intervention, outcome definitions, &c, and how clinical heterogeneity was assessed | Y | 100 |
| | Quantitative data synthesis | The principal measures of effect (eg, relative risk), method of combining results (statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed; a rationale for any a-priori sensitivity and subgroup analyses; and any assessment of publication bias | N/A | |
| Results | Trial flow | Provide a meta-analysis profile summarising trial flow (see figure) | Y | 283 (appendix) |
| | Study Characteristics | Present descriptive data for each trial (eg, age, sample size, intervention, dose, duration, follow-up period) | Y | table 1.1 (p100) |
| | Quantitative data synthesis | Report agreement on the selection and validity assessment; present simple summary results (for each treatment group in each trial, for each primary outcome); present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses (eg 232 tables of counts, means and SDs, proportions) | N/A | |
| Discussion | | Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (eg, publication bias); and suggest a future research agenda | Y | 108 |



Appendix 3.1:

Search string used to identify relevant studies

```
1 BREAST/  
  (neoplasm$ or cancer$ or tumour$ or tumor$ or malign$ or carcinoma$ or  
2 adenocarcinoma$ or sarcoma$ or dcis or ductal or infiltrat$ or intraduct$ or lobular$ or  
  medullary$).mp. [mp=title, abstract, subject headings, heading word, drug trade name,  
  original title, device manufacturer, drug manufacturer name]  
3 1 and 2  
4 Relapse or recurrence  
5 3 and 4  
6 Diagnosis or mammogr$ or Exam$  
7 5 and 6
```

Appendix 3.2:

Details of additional authors involved in the review process:

Author 2 (KK):

Katherine Krupa

Consultant surgeon in Breast and Endocrine surgery

Paisley Royal Alexandra Hospital

Corsebar Road

Paisley

PA2 9PN

Author 3 (TGC):

Professor T G Cooke

St Mungo Professor of Surgery

Department of Surgery

Glasgow Royal Infirmary

Alexandra Parade

Glasgow

G31 2ER

Appendix 4.1

Complete questionnaire and information sheet

Clinic expectation questionnaire

1. How was your cancer diagnosed?
 - You discovered a lump in your breast
 - You went to the GP with other breast symptoms
 - You went to GP with other unrelated problem
 - Discovered by screening service
 - other(please specify)

2. Now that you have completed all your treatment, do you think we will want to see you back at clinic again for more follow up?
 - Yes
 - No

3. If yes, how often would you expect us to want to see you?
 - Monthly
 - Every three months
 - Every six months
 - Every year
 - Less than every year?
 - Other.....

4. When do you expect to be discharged from clinic?

5. If we asked you to come back to clinic **more** regularly, would this make you feel?
 - More anxious
 - More reassured
 - Neither?

6. If you would like to come back to clinic, what do you think is the main reason for that?
 - To catch the cancer early if it comes back
 - To detect and treat any side effects you may have from treatment
 - Just for reassurance
 - Other.....

7. If you are unlucky enough that your cancer does come back, who do you think is most likely to discover it?
 - Us
 - Yourself

Please read information sheet now, and then answer the questions overleaf.

Patient Follow Up Information Sheet

The way we look after our patients following treatment for breast cancer has changed over the years. It was once felt that the main reason to keep seeing patients back after breast cancer was to make sure the cancer had not come back and to catch it early if it did. We now understand this disease better and in particular the way patients respond afterwards. We now know that:

- While breast cancer is very common (about 3500 people develop the disease in Scotland every year), in most people it is treated successfully and they do not develop problems afterwards.
- Seeing people frequently at clinic has been shown not to make any difference to how well women do after treatment for breast cancer.
- Seeing people frequently at clinic can cause distress and worry among many women.
- It has been found that most of the people who do suffer a relapse discover it themselves in between clinic visits and come back to clinic early.
- In a few people, their cancer comes back early but is so small that they don't notice it. Very often, though, we can't feel it either. In these people, mammogram is the best way to detect it.

So we now know that if your cancer does come back, it is more likely to do so in between clinic visits and that you are far more likely to notice than we are. If it is so small that you do not notice it, it is much more likely to be seen on mammogram than noticed by us in clinic.

The main purpose of bringing you back to clinic now is to keep an eye on you and make sure you do not develop any problems related to treatment.

Because of all these things, we wonder if there might be a better way to look after people after their breast cancer.

Please turn over now and complete the last two questions

8. Given what you have just read, do you still feel you want to come back to clinic routinely, even when everything is going well? **(You would still get a mammogram every year either way)**

Yes

No

9. We think that clinics may not be the best way of keeping an eye on you. There are other ways that we could make sure you are doing well without bringing you up to the hospital. Which of the following three ways would you prefer us to look after you from now on?

You could phone us to tell us how you are doing, instead of coming in to clinic

You could come back only if you have a problem

Or you could continue to come and see us at regular intervals, even if you feel fine

Have you any ideas of your own?.....

Thank you for taking the time to complete this questionnaire!

Appendix 5.1:

Initial seven questions:

1. **How often have you examined your breasts since we last saw you? (please circle one)**
not at all, every day every week every month when and if I remember

2. **Have you noticed any changes in your breast? Yes No**

3. **Have you noticed any changes in your operation scar? Yes No**

4. **Have you noticed any changes in your arm pit? Yes No**

5. **Have you had any NEW problems with breathlessness Yes No**

6. **Have you had any NEW problems with persistent cough Yes No**

7. **Have you developed any NEW aches and pains in your bones recently? Yes No**

Appendix 5.2:

FACT-B with arm and endocrine subscales as used in this project:

PHYSICAL WELL-BEING

| | | Not at all | A little bit | Some- what | Quite a bit | Very much |
|-----|---|-----------------------|-------------------------|-----------------------|------------------------|----------------------|
| GP1 | I have a lack of energy..... | 0 | 1 | 2 | 3 | 4 |
| GP2 | I have nausea | 0 | 1 | 2 | 3 | 4 |
| GP3 | Because of my physical condition, I have trouble meeting the needs of my family | 0 | 1 | 2 | 3 | 4 |
| GP4 | I have pain | 0 | 1 | 2 | 3 | 4 |
| GP5 | I am bothered by side effects of treatment | 0 | 1 | 2 | 3 | 4 |
| GP6 | I feel ill | 0 | 1 | 2 | 3 | 4 |
| GP7 | I am forced to spend time in bed..... | 0 | 1 | 2 | 3 | 4 |

SOCIAL/FAMILY WELL-BEING

| | | Not at all | A little bit | Some- what | Quite a bit | Very much |
|-----|---|-----------------------|-------------------------|-----------------------|------------------------|----------------------|
| GS1 | I feel close to my friends | 0 | 1 | 2 | 3 | 4 |
| GS2 | I get emotional support from my family | 0 | 1 | 2 | 3 | 4 |
| GS3 | I get support from my friends..... | 0 | 1 | 2 | 3 | 4 |
| GS4 | My family has accepted my illness | 0 | 1 | 2 | 3 | 4 |
| GS5 | I am satisfied with family communication about my illness.. | 0 | 1 | 2 | 3 | 4 |
| GS6 | I feel close to my partner (or the person who is my main support) | 0 | 1 | 2 | 3 | 4 |
| Q1 | Regardless of your current level of sexual activity, please answer the following question. If you prefer not to answer it, please check this box <input type="checkbox"/> and go to the next section. | | | | | |
| GS7 | I am satisfied with my sex life..... | 0 | 1 | 2 | 3 | 4 |

EMOTIONAL WELL-BEING

| | | Not at all | A little bit | Some- what | Quite a bit | Very much |
|-----|--|-----------------------|-------------------------|-----------------------|------------------------|----------------------|
| GE1 | I feel sad..... | 0 | 1 | 2 | 3 | 4 |
| GE2 | I am satisfied with how I am coping with my illness..... | 0 | 1 | 2 | 3 | 4 |
| GE3 | I am losing hope in the fight against my illness..... | 0 | 1 | 2 | 3 | 4 |
| GE4 | I feel nervous..... | 0 | 1 | 2 | 3 | 4 |
| GE5 | I worry about dying..... | 0 | 1 | 2 | 3 | 4 |
| GE6 | I worry that my condition will get worse..... | 0 | 1 | 2 | 3 | 4 |

FUNCTIONAL WELL-BEING

| | | Not at all | A little bit | Some- what | Quite a bit | Very much |
|-----|---|-----------------------|-------------------------|-----------------------|------------------------|----------------------|
| GF1 | I am able to work (include work at home)..... | 0 | 1 | 2 | 3 | 4 |
| GF2 | My work (include work at home) is fulfilling..... | 0 | 1 | 2 | 3 | 4 |
| GF3 | I am able to enjoy life..... | 0 | 1 | 2 | 3 | 4 |
| GF4 | I have accepted my illness..... | 0 | 1 | 2 | 3 | 4 |
| GF5 | I am sleeping well..... | 0 | 1 | 2 | 3 | 4 |
| GF6 | I am enjoying the things I usually do for fun..... | 0 | 1 | 2 | 3 | 4 |
| GF7 | I am content with the quality of my life right now..... | 0 | 1 | 2 | 3 | 4 |

BREAST CANCER SUBSCALE

| | | Not at All | A little bit | Some- what | Quite a bit | Very much |
|-----|---|-----------------------|-------------------------|-----------------------|------------------------|----------------------|
| B1 | I have been short of breath | 0 | 1 | 2 | 3 | 4 |
| B2 | I am self-conscious about the way I dress..... | 0 | 1 | 2 | 3 | 4 |
| B3 | One or both of my arms are swollen or tender | 0 | 1 | 2 | 3 | 4 |
| B4 | I feel sexually attractive | 0 | 1 | 2 | 3 | 4 |
| B5 | I am bothered by hair loss..... | 0 | 1 | 2 | 3 | 4 |
| B6 | I worry that other members of my family might someday get the same illness I have..... | 0 | 1 | 2 | 3 | 4 |
| B7 | I worry about the effect of stress on my illness..... | 0 | 1 | 2 | 3 | 4 |
| B8 | I am bothered by a change in weight..... | 0 | 1 | 2 | 3 | 4 |
| B9 | I am able to feel like a woman | 0 | 1 | 2 | 3 | 4 |
| P2 | I have certain parts of my body where I experience significant pain | 0 | 1 | 2 | 3 | 4 |
| Q6 | On which side was your breast operation? | Left | Right | | | |
| B10 | Movement of my arm on this side is painful | 0 | 1 | 2 | 3 | 4 |
| B11 | I have a poor range of arm movements on this side | 0 | 1 | 2 | 3 | 4 |
| B12 | My arm on this side feels numb | 0 | 1 | 2 | 3 | 4 |
| B13 | I have stiffness of my arm on this side | 0 | 1 | 2 | 3 | 4 |

ENDOCRINE CONCERNS

| | | Not at all | A little bit | Some- what | Quite a bit | Very much |
|------|---|---------------|-----------------|---------------|----------------|--------------|
| ES1 | I have hot flashes | 0 | 1 | 2 | 3 | 4 |
| ES2 | I have cold sweats..... | 0 | 1 | 2 | 3 | 4 |
| ES3 | I have night sweats | 0 | 1 | 2 | 3 | 4 |
| ES4 | I have vaginal discharge | 0 | 1 | 2 | 3 | 4 |
| ES5 | I have vaginal itching/irritation | 0 | 1 | 2 | 3 | 4 |
| ES6 | I have vaginal bleeding or spotting | 0 | 1 | 2 | 3 | 4 |
| ES7 | I have vaginal dryness | 0 | 1 | 2 | 3 | 4 |
| ES8 | I have pain or discomfort with intercourse..... | 0 | 1 | 2 | 3 | 4 |
| ES9 | I have lost interest in sex | 0 | 1 | 2 | 3 | 4 |
| ES10 | I have gained weight..... | 0 | 1 | 2 | 3 | 4 |
| An9 | I feel light-headed (dizzy)..... | 0 | 1 | 2 | 3 | 4 |
| O2 | I have been vomiting | 0 | 1 | 2 | 3 | 4 |
| C5 | I have diarrhoea | 0 | 1 | 2 | 3 | 4 |
| An10 | I get headaches..... | 0 | 1 | 2 | 3 | 4 |
| Tax1 | I feel bloated..... | 0 | 1 | 2 | 3 | 4 |
| ES11 | I have breast sensitivity/tenderness | 0 | 1 | 2 | 3 | 4 |
| ES12 | I have mood swings | 0 | 1 | 2 | 3 | 4 |
| ES13 | I am irritable | 0 | 1 | 2 | 3 | 4 |
| BRM | I have pain in my joints | 0 | 1 | 2 | 3 | 4 |