

**Transcriptional Signatures as Early Predictors
of Response in Breast Cancers Treated with
Neoadjuvant Aromatase Inhibitors**

Miss Emma L. Murray

A Thesis Submitted for the Degree of MD to the
University of Edinburgh

2014

In Loving Memory of My Grandmother

Mrs Alexandrina Murray
(1922 - 2006)

who Lived and Died with Grace and Courage

Declaration

Transcriptional Signatures as Early Predictors of Response in Breast Cancers Treated with Neoadjuvant Aromatase Inhibitors

I hereby declare that this thesis was written by me. The work presented here is my own and relates to research carried out by me in the Edinburgh Breast Unit at the Western General Hospital between March 2006 and December 2007. Other members of the research group who contributed to the laboratory and clinical work embodied in this thesis are specifically acknowledged overleaf. This work has not been submitted in candidature for any other degree, award, or professional qualification.

Emma L. Murray, February 2014

Abstract

Introduction: Breast cancer is a leading cause of death in women of all ages in many parts of the world. The management of this disease is typically based upon its clinico-pathological features and usually involves a combination of surgery and systemic therapy. Many patients however, do not respond to these therapies as anticipated and may therefore suffer needlessly from medication side effects and delayed initiation of more effective treatment, all of which impose an enormous financial burden on healthcare systems.

Efforts are underway to improve the classification, prognostication and prediction of response to treatment for breast cancer patients, in the hope of providing better, individualised care. Advances in our understanding of the disease and its management are anticipated to come from investigations into the molecular pathways and gene expression underlying breast cancer development, growth and metastasis.

Microarray technology, used to simultaneously analyse the patterns of expression of tens of thousands of genes from tumour biopsies, has permitted the identification of new intrinsic subtypes of breast cancer based upon their transcriptional signatures, and of prognostic and predictive markers that are beginning to show clinical utility and promise to outperform standard clinico-pathological markers.

Methods: Microarray and quantitative polymerase chain reaction (qPCR) technologies were used to compare the efficacy of five multi-gene signatures for their ability to predict clinical response to three months of neoadjuvant treatment with the Aromatase inhibitor Letrozole, in a population of postmenopausal ER positive breast cancer patients. The levels of gene expression in biopsy samples acquired from each patient were measured prior to treatment and at week two of a three-month treatment regimen. Tumour response was assessed dynamically by means of three-dimensional ultrasound scanning. The predictive capacity of proliferation markers was further explored by focusing on the expression of several key genes involved in different stages of cell cycle progression, including Cyclins B1, A2, D1, CDKs 1, 2 and 4 and the NUSAP1 and Ki67 genes. In addition, changes in Ki67 after 3 months of neo-adjuvant treatment were related to long-term survival.

Results: A total of 394 women with large or locally advanced oestrogen receptor positive breast tumours were enrolled into the study population in two independent datasets.

A previously defined neoadjuvant predictive signature from our unit has been validated on both independent and extended datasets. In addition, two further proliferative-gene signatures were shown to have significant power to predict longer-term clinical response to Letrozole therapy, based upon their gene expression activity after two weeks of treatment. On the other hand, neither of the two stromal-gene signatures chosen for inclusion in this study was able to predict response based on their absolute expression profiles at two weeks. However, when considering the change in their gene expression profiles between baseline and two weeks, both stromal signatures became informative in one of the datasets. Only one of the proliferative signatures was predictive in this way.

Analysis focusing on the genes representing different phases of the cell cycle, demonstrated that two weeks of Letrozole therapy strongly decreased the expression of Cyclins B1, A2, D1, CDK1 and NUSAP1 but not CDK 2 and 4. Significant correlations between the change in Ki67 and changes in Cyclins B1/CDK1 and Cyclin A2/CDK2 were observed, demonstrating the importance of transcriptional regulation of S and G2-M phase Cyclins/CDKs for the anti-proliferative effect of neoadjuvant Letrozole.

Changes in the expression of NUSAP1 and Cyclin B1 significantly correlated with clinical response and Ki67 positivity, especially after two weeks of treatment.

The following factors were found to be significantly associated with breast cancer specific survival: T-stage at diagnosis; Nodal status at the time of surgery; and Ki67 at diagnosis and at three months and the corresponding change in Ki67 expression between these time-points.

Conclusions: The results presented here demonstrate that changes in the expression of multi-gene transcriptional signatures as well as single genes, can be measured in sequential tumour biopsies taken from patients during neoadjuvant treatment for breast cancer. The ability of several signatures and individual genes to predict the clinical response to neoadjuvant treatment has been compared and evaluated. The difference between the signatures composed of proliferative and stromal genes highlights the different contributions of these components to tumour development and response to treatment.

Acknowledgements

I would like to thank the following for their significant contributions to my research project.

Professor Michael Dixon, Professor of Breast Surgery at the Edinburgh Breast Unit, Clinical Lecturer at the University of Edinburgh, and also my supervisor; for enabling and overseeing the clinical aspects of this work and for his considered advice.

Dr Alexey Larionov, former Research Fellow at the Breakthrough Breast Cancer Research Unit in the University of Edinburgh and current Bioinformatician at the University of Cambridge; for training me in the required laboratory techniques and for supervising this work; for his insight and guidance in the planning, direction and execution of it; for providing me an understanding of the complexities of bio-informatics analysis and for unrelenting moral support. Dr Larionov was also the invaluable Bioinformatician for this project.

My predecessors in clinical research at the Edinburgh Breast Unit, in particular Miss Jane MacAskill for familiarising me with the various on-going studies with which I was involved and for showing me the ropes.

Miss Monica McGill, Mrs Jan Morrison, and Mrs Kirsty Birrell, Personal Assistants and Secretaries to Professor Dixon, along with the remainder of the clinical staff of the Edinburgh Breast Unit, for their support during my time there.

Novartis and the Breakthrough Breast Cancer Laboratories for their funding and for the use of their facilities

The University of Edinburgh, through which I was salaried.

The all-important patients who participated in this study and who continue to give generously of their time, to advance our understanding of breast cancer and to improve clinical care.

Patrick, my husband, for his love, support and tireless proofreading and as always, my parents – Thank you.

Table of Contents

Declaration.....	i
Abstract.....	ii
Acknowledgements.....	iv

Breast Cancer

1.1 Epidemiology	1
1.1.1 Incidence	1
1.1.2 Mortality	3
1.1.3 Survival.....	5
1.1.4 Risk Factors.....	6
1.1.5 Reproductive Risks Factors.....	7
1.1.6 Non-Reproductive Risk Factors.....	9
1.2 Diagnosis.....	13
1.2.1 Triple Assessment	13
1.2.2 Radiological Imaging	14
1.2.3 Breast Biopsy.....	15
1.2.4 Accuracy of Investigations.....	15
1.3 Classification.....	16
1.3.1 An Overview	16
1.3.2 Histopathological Classification	16
1.3.3 Molecular Classification	19
1.4 Staging.....	20
1.4.1 TNM and UICC	20
1.4.2 Staging Terminology.....	21
1.4.3 Staging the Axilla.....	22
1.4.4 Clinical Terminology	23
1.5 Prognosis in Breast Cancer.....	24
1.5.1 Clinical Factors	24
1.5.2 Histological Factors	25
1.5.3 Molecular Biological Factors.....	26
1.5.4 Prognostic Indices	28
1.5.5 Predictive Indices.....	29
1.5.6 Future developments.....	30
1.5.7 Outcome Measures	31
1.6 Management Pathways – Local Treatments.....	32
1.6.1 Surgical Management of the Breast	32
1.6.2 Surgical Management of The Axilla	35
1.6.3 Radiotherapy	36
1.7 Management Pathways - Systemic Treatments	38
1.7.1 Chemotherapy	38
1.7.2 Biological Therapies	42

Breast Cancer Endocrinology

1.8 Endocrine Dependence and the Breast	44
1.8.1 Breast Development and Tumourigenesis	44
1.8.2 Oestrogen Production	45
1.8.3 The Oestrogen and Progesterone Receptors	46
1.8.4 The Biology of Oestrogen Receptor Signalling	47
1.8.5 Quantification of ER	49
1.9 Endocrine Therapy	50
1.9.1 Limiting Oestrogen Production	50
1.9.2 Blocking Oestrogen Function	52
1.9.3 Adjuvant Endocrine Therapy	54
1.9.4 Neoadjuvant Endocrine Therapy	56
1.10 Assessment of Response to Neoadjuvant Endocrine Therapy	60
1.10.1 Clinical Response	60
1.10.2 Pathological Response	61
1.10.3 Proliferative Response	62
1.10.4 Molecular Response	62
1.11 Molecular Mechanisms of Endocrine Resistance	63
1.11.1 Deregulation of Components of the Classical ER Pathway	64
1.11.2 Alterations in the Cyclins and Other Cell Cycle Regulators	64
1.11.3 Activation of Escape Pathways	65
1.11.4 Alteration of Cell Survival Proteins and Signalling Pathways	66

The Molecular Biology of Breast Cancer

1.12 The Molecular Basis of Tumourigenesis	67
1.12.1 Sustained Proliferative Signalling	68
1.12.2 Evasion of Growth Suppression	71
1.12.3 Resisted Cell Death	72
1.12.4 Replicative Immortality	73
1.12.5 Induction of Angiogenesis	73
1.12.6 Activation of Invasion and Metastasis	73
1.12.7 Involvement of the Tumour Microenvironment	74
1.12.8 Immune Destruction	76
1.12.9 Re-Programming of Energy Metabolism	77
1.13 The Molecular Classification of Breast Cancer	77
1.13.1 The Limitations of Current Classifications	77
1.13.2 The Genetic Basis of Cancer	78
1.13.3 Gene Expression Profiling and the Intrinsic Subtypes	79
1.13.4 Clinical Application of Microarray Technology	83
1.13.5 Prognostic Signatures	83
1.13.6 Second-Generation Prognostic Signatures	84
1.13.7 Predictive Signatures	84
1.13.8 Predictive Signatures in ER Positive Tumours	85
1.13.9 Stromal Gene Signatures	86

2. Aims and Objectives

2.1 Aims	88
2.2 Objectives	90

3. Methods and Results

3.1 Study Design	91
3.2 Patient Selection and Enrolment	93
3.3 Assessing Tumour Response and Clinical Follow-up	96
3.3.1 Clinical examination.....	96
3.3.2 Ultrasound	96
3.3.3 Mammography.....	96
3.3.4 Pathological assessment.....	97
3.3.5 Classification of Tumour Response.....	97
3.3.6 Outcome Monitoring	98
3.4 Tissue Collection and Storage	98
3.5 Tumour Processing and RNA Extraction	99
3.5.1 Extraction of Specimens from OCT and Stabilising RNA	100
3.5.2 Procedure for RNA extraction using Qiagen RNeasy Mini-kits.....	100
3.6 Microarray Analysis	102
3.6.1 Microarray Platforms.....	102
3.6.2 Microarray Data Processing	104
3.7 Reverse-Transcription and Quantitative Real-Time PCR for Selected Genes	107
3.7.1 Reverse Transcription Protocol	109
3.7.2 Selection and Design of Primers	110
3.7.3 Quantitative PCR Protocol.....	111
3.8 Datasets	113
3.8.1 Letma 2 (L2)	114
3.8.2 Letma 3 (L3)	115
3.8.3 Letma 2 and 3 (L23)	115
3.9 Tumour Response	116
3.10 Selection of Gene Signatures	119
3.10.1 Signature Translation.....	120
3.10.2 The Edinburgh Signature – Miller 2009 ^[510]	123
3.10.3 Proliferation Signature – Dai 2005 ^[471]	123
3.10.4 Proliferation Signature – Nagalla 2013 ^[524]	124
3.10.5 Stromal Signature – Farmer 2009 ^[496]	124
3.10.6 Stromal Signature – Beck 2008 ^[525]	125
3.11 Signature Interpretation	126
3.11.1 Incorporating Dynamic Measurements.....	126
3.11.2 Sign of Association	126
3.11.3 Heat Maps	126
3.11.4 Multi-Gene Expression Scores	127
3.11.5 Receiver Operating Characteristic (ROC) Curve Analysis	129
3.12 Validation of the Edinburgh Signature	130
3.13 Performance of Proliferative Signatures	132
3.14 Performance of Stromal Signatures	139

3.15 Signature Comparison	145
3.16 Quantitative PCR Validation of Microarray data	147
3.17 Changes in Cyclin and CDK expression during Neoadjuvant Letrozole Treatment.....	148
3.18 Ki67 and Other Factors Predicting Survival after Neoadjuvant Endocrine Therapy.....	151
3.18.1 Patients and Study Design	151
3.18.2 Statistical Analysis	152
3.18.3 Results.....	152
4. Discussion – PART 1	
The Present Study	
4.1 Findings of the Present Study	159
4.1.1 Validation of the Edinburgh Signature	159
4.1.2 Performance of the Proliferative Signatures.....	160
4.1.3 Performance of the Stromal Signatures.....	162
4.1.4 Expression of the Cyclins	163
4.1.5 Changes in Ki67 expression as a predictor of breast cancer-specific survival during neo-adjuvant AI treatment.....	163
4.2 Comments on Methodology.....	165
4.2.1 Use of the Neoadjuvant Treatment Setting.....	165
4.2.2 Timings of Treatment and Sample Acquisition	166
4.2.3 Tumour Response Assessment.....	167
4.2.4 Sample Collection.....	167
4.2.5 Signature selection	168
4.2.6 Microarray Analysis	168
4.2.7 Multi-Signature Comparisons	169
4. Discussion – PART 2	
The Clinical Utility of the Molecular Subtypes and Multi-Gene Signatures	
4.3 The Shortcomings of Current Classifications.....	169
4.4 The Potential and Limitations of Molecular Subtyping	170
4.5 The Promise of Gene Signatures in Prognostication and Prediction	172
4.5.1 Prognostic signatures	172
4.5.2 The Transition from Prognostic to Predictive Signatures.....	175
4.5.3 Predictive Signatures in the Adjuvant Endocrine Therapy Setting	177
4.5.4 Predictive signatures in the Neoadjuvant Chemotherapy Setting.....	179
4.5.5 Predictive Signatures in the Neoadjuvant Endocrine Therapy Setting	180
4.5.6 Combining Markers to Improve Power	181
5. Future Studies	182
5.1 Pertaining to the Present Study	182
5.2 General Considerations for Future Studies	183
6. Conclusions.....	185

Appendix

List of Figures

- Figure 1.1A:** The 10 Most Commonly Diagnosed Cancers in Females, Percentages of All Cancer Cases Excluding Non-Melanoma Skin Cancer, UK, 2009.
- Figure 1.1B:** Female Age Standardised Breast Cancer Trends, UK, 1975-2009.
- Figure 1.1C:** Average Number of new cases of Female Breast Cancer in the UK each year and Age Specific Incidence rates (2007–2009).
- Figure 1.1D:** Average Number of Deaths per Year and Age Specific Mortality Rates in UK Females (2008-10).
- Figure 1.1E:** Breast Cancer Mortality Rates, UK 1971-2010.
- Figure 1.1F:** Five-Year Survival by Age in UK Breast Cancer Patients.
- Figure 1.8A:** Functional Domains of the Oestrogen Receptor.
- Figure 1.8B:** Classical Oestrogen Signalling Pathway.
- Figure 1.8C:** Alternative Oestrogen Signalling Pathways.
- Figure 1.9A:** Mechanisms of Action of the Endocrine Therapies.
- Figure 1.11A:** Points in the ER Stimulated Pathway during which Resistance may Develop within a Breast Tumour.
- Figure 1.11B:** Mechanisms of Resistance to Endocrine Therapy.
- Figure 1.12A:** Mechanisms of Tumourigenesis.
- Figure 1.12B:** The Cell Cycle.
- Figure 1.12C:** Modulators of the Cell Cycle.
- Figure 3.1A-C:** Overview of Study Design
- Figure 3.2A:** Summary of Study Chronology
- Figure 3.5A:** OCT Embedded Tissue.
- Figure 3.6A:** Affymetrix Microarray Chip.
- Figure 3.6B:** Illumina Microarray System.
- Figure 3.7A:** Amplification of Specific DNA Sequences by PCR.
- Figure 3.8A:** Overview of Datasets.
- Figure 3.9A:** Tumour Responses by USS in L2 Series.
- Figure 3.9B:** Tumour Responses by USS in L3 Series.
- Figure 3.9C:** Tumour Responses by USS in L23 Series.
- Figure 3.10A:** Representation of Overlapping Genes within Chosen Signatures.
- Figure 3.10B:** Overview of Signature Translation.

- Figure 3.11A-B:** Heat Map and Multi-Gene Expression Score Chart for the Edinburgh Signature and Response for Samples from the L2 Series
- Figure 3.11C:** Model of Receiver Operating Characteristic Curve.
- Figure 3.12A:** Performance of Miller Signature on all Datasets.
- Figure 3.13A-C:** Performance of Nagalla Signature on each Dataset.
- Figure 3.13D-F:** Performance of Dai Signature on each Dataset.
- Figure 3.14A-C:** Performance of Beck Signature on each Dataset.
- Figure 3.14D-F:** Performance of Farmer Signature on each Dataset.
- Figure 3.15A-C:** Signature Comparison for each Dataset.
- Figure 3.17A:** Results of Cyclin/CDK Sub-Analysis.
- Figure 3.18A:** Cause Specific Survival by Node Status at Surgery.
- Figure 3.18B:** Cause Specific Survival by Ki67 at Diagnosis.
- Figure 3.18C:** Cause Specific Survival by Ki67 at 3 months.
- Figure 3.18D:** Cause Specific Survival by Change in Ki67.

List of Tables

- Table 1.5A:** Breast Cancer Survival According to TNM Stage.
- Table 1.5B:** Biological Factors Used For Prognostication.
- Table 1.5C:** Biological Markers of Uncertain Clinical Significance.
- Table 1.5D:** Breast Cancer Specific Survival According to NPI.
- Table 3.7A:** Primer sequences for all PCR Analyses.
- Table 3.10A:** Lists of Overlapping Genes within Paired Signatures.
- Table 3.16A:** Comparative Analysis of Microarray and qPCR on Selected Genes.
- Table 3.17A:** Changes in Selected Cell-Cycle Associated Genes during Treatment with Letrozole.
- Table 3.18A:** Patient Demographics and Tumour Pathology.
- Table 3.18B:** Tumour Response at 3 months by Ultrasound Measurement.
- Table 3.18C:** Ki67 at Diagnosis.
- Table 3.18D:** Ki67 at 3 Months.
- Table 3.18E:** Percentage Reduction in Ki67 at 3 Months.

List of Appendices

- | | |
|---------------------|--|
| Appendix I | The American Joint Committee on Cancer (AJCC) and the International Union Against Cancer (UICC) Staging Systems. |
| Appendix II | Patient Consent Forms |
| Appendix III | Patient Information Sheets |
| Appendix IV | Gene lists for each signature |
| Appendix V | Publications, Presentations and Supplementary files on CD-ROM |

Breast Cancer

1.1 Epidemiology

1.1.1 Incidence

Breast cancer is by far the most commonly diagnosed cancer in women worldwide. Over 1.38 million new cases are diagnosed globally each year, representing around 23% of all female cancers and 11% of all cancers regardless of gender ^[1].

The incidence of breast cancer is highest in Western Europe and lowest in Eastern and Middle Africa, with a nearly five-fold variation in age-adjusted incidence between various countries. In the United Kingdom, age standardised incidence and mortality for breast cancer remains amongst the highest in the world, ranking sixth amongst fellow European Union (EU) countries, with little variation between Scotland, England, Wales and Northern Ireland ^[2-4].

Breast cancer has been the most common malignancy in the UK since 1997, despite its relatively low incidence in men (<1 % of breast cancers). In 2009 more than 48,400 breast cancers were detected in women (317 in men), and an average of 133 new cancers were diagnosed each day. It has been estimated that breast cancer now constitutes over one third of all female cancers in the UK (Figure 1.1A) and affects 1 in 8 women ^[3, 5, 6].

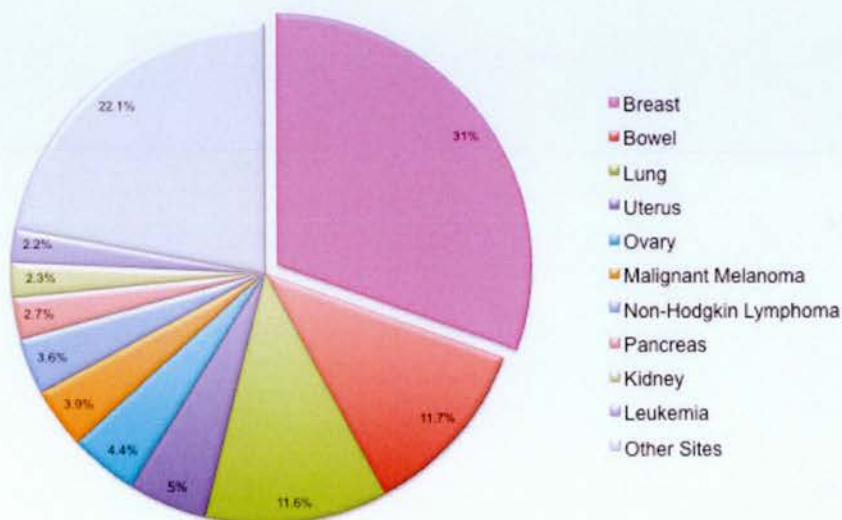


Figure 1.1A: The 10 Most Commonly Diagnosed Cancers in Females, Percentages of All Cancer Cases Excluding Non-Melanoma Skin Cancer, UK, 2009 ^[3, 5].

Furthermore, the UK has seen a rise in the age adjusted incidence of breast cancer of almost 70% since the mid 1970's and a 6% rise over the last decade (Figure 1.1B) ^[3, 5]. These figures are only partly explained by both a heightened public awareness of the disease and by the introduction of the National Health Service (NHS) breast-screening programme in 1988, which improved the detection rate of previously undiagnosed cancers.

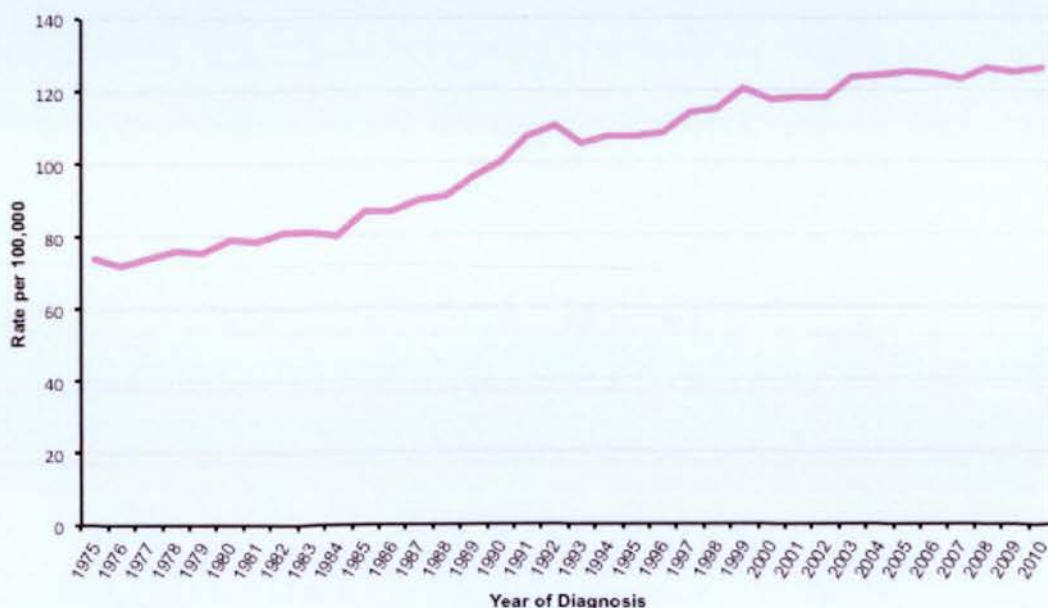


Figure 1.1B: Female Age Standardised Breast Cancer Trends for the UK, 1975-2009 ^[3, 5].

Mammographic screening is presently offered to women between the ages of 50-69 years, although a trial programme to include those aged 47 to 73 years is currently underway in England. Not surprisingly therefore, nearly half of all female breast cancer cases are diagnosed in the screened population and more than 80% are in women aged over 50 ^[3, 5].

Indeed, there is a strong positive correlation between age and breast cancer incidence, which doubles every 10 years until the menopause and then slows considerably. A further rise in incidence occurs beyond the age of 65, dropping between 70-74, and rising again steadily thereafter (Figure 1.1C) ^[2-5, 7].

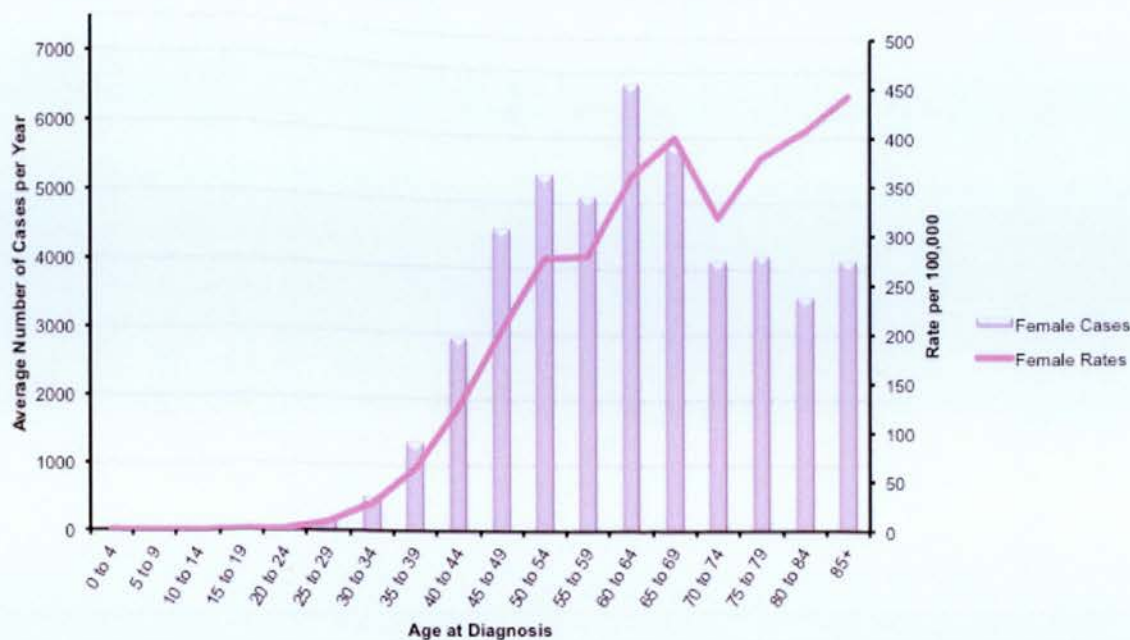


Figure 1.1C: Average Number of new cases of Female Breast Cancer in the UK each year and Age Specific Incidence rates (2007–2009) ^[3, 5].

A subtle downward trend is notable in breast cancer incidence in the over 50 population since the mid 2000's ^[5, 7]. This is likely to be attributable to a decline in the use of hormone replacement therapy (HRT), a risk factor for breast cancer development, since the late 1990's ^[8-14].

1.1.2 Mortality

Each year, breast cancer is responsible for the death of over 458,000 women worldwide. Europe accounts for around 89,000 of the total annual mortality and the UK, ranking 7th among its EU counterparts, has around 11,550 deaths ^[15, 16]. In fact, with an estimated average daily mortality rate of 32, breast cancer is now second only to lung cancer as a leading cause of cancer death in British women, representing 15% of all UK female cancer deaths ^[1, 3].

Female breast cancer mortality is also strongly related to age, with rates increasing steadily from the age of 30 onwards (Figure 1.1D). In women over the age of 74, a rapid acceleration in mortality is seen ^[15, 16].

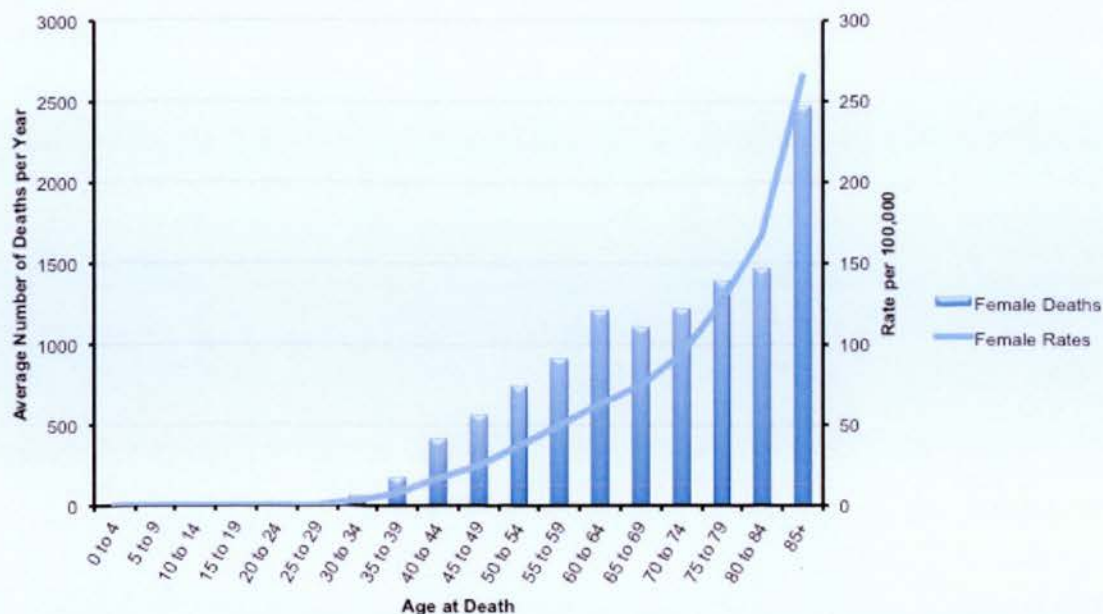


Figure 1.1D: Average Number of Deaths per Year and Age Specific Mortality Rates in UK Females (2008-10) ^[3].

On a more optimistic note, the overall breast cancer death rates in the UK have fallen by almost 40% since peaking in the late 1980's, and by as much as 20% over the past decade (Figure 1.1E) ^[15, 16].

Advances in early diagnosis as well as improvements in early detection and treatment have no doubt driven this reduction in mortality ^[17, 18]. The mortality rates have fallen most dramatically for women under the age of 70 years. The slower change in mortality rates in the older group may be explained by the fact that women aged over 71 are not eligible for screening and are less likely to receive surgery and/or radiotherapy for their cancer and may die from other conditions ^[19-21].

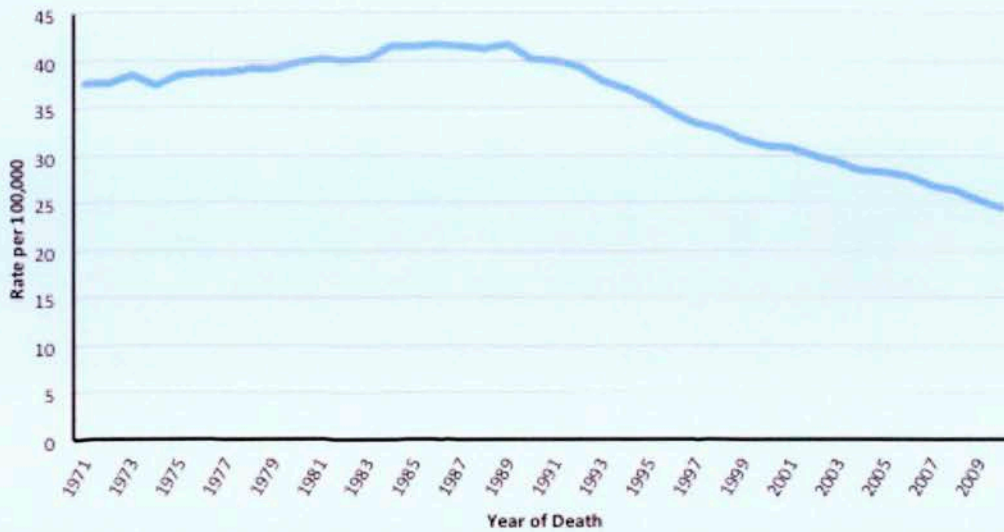


Figure 1.1E: Breast Cancer Mortality Rates, UK 1971-2010 ^[3].

1.1.3 Survival

Conversely, breast cancer survival rates have been improving over the past four decades. In the 1970's around half of women survived their disease beyond five years, that number has now risen to more than 80% ^[22]. Three-quarters of those currently diagnosed will survive beyond 10 years and two-thirds survive more than 20 years following their original diagnosis ^[22].

For most other cancers, survival rates decrease with age whilst in breast cancer, women diagnosed in their 50's and 60's have consistently higher survival rates than younger or older women (Figure 1.1F) ^[23, 24].

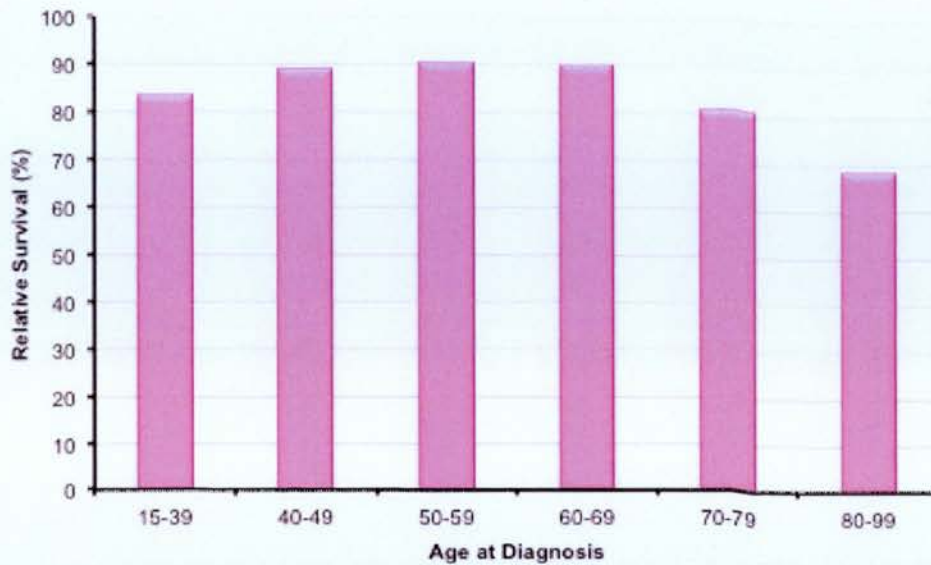


Figure 1.1F: Five Year Survival by Age in UK Breast Cancer Patients ^[22].

Clearly, survival rates are generally better the earlier the cancer is diagnosed and treated. It is estimated that whilst 90% of women diagnosed with early (Stage I) disease will survive beyond 5 years, this falls to 10% for advanced (stage IV) disease ^[22].

In terms of socio-economic effect, the chance of survival is significantly higher for women living in affluent areas compared to their more deprived counterparts. When survival by cancer stage is compared internationally, Scotland in general and Glasgow in particular, fare poorly with respect to many other parts of the UK, after adjustment for known prognostic factors ^[25].

1.1.4 Risk Factors

The strongest risk factor for breast cancer, after gender, is age ^[26]. Although an interplay of both genetic and environmental factors also contribute to the risk of developing the disease, the incidence patterns amongst migrant populations suggest that environmental factors are of greater importance ^[27]. In developed countries, it has been estimated that genetic factors contribute around 25% of the differences in individual susceptibility, whilst environmental and lifestyle factors contribute the remaining three quarters ^[28].

1.1.5 Reproductive Risks Factors

Age at Menarche

An early age at menarche has been consistently associated with an increased risk of breast cancer and each five-year delay is estimated to reduce this risk by 22% ^[29]. The average age of menarche in developed countries has fallen since the mid 19th century by around 4-5 years ^[30], largely as a consequence of improved nutrition ^[31].

Parity

Child bearing reduces the risk of breast cancer, with a higher number of pregnancies conferring greater protection. Mothers have a 30% lower risk of developing breast cancer compared to nulliparous women. This equates to a 7% risk reduction for each pregnancy ^[32-34].

The risk is also affected by a woman's age at first full term pregnancy. The older a woman is when she begins childbearing, the higher her risk of developing breast cancer. The relative risk is thought to increase by 3% for each year of delay, though it is possible that this may only be applicable to oestrogen receptor positive tumours. The highest risk appears to be for women who have a first child beyond 35 years of age. These women appear to be at even higher risk than nulliparous women ^[32, 33].

Breast Feeding

Women who breastfeed reduce their relative risk of developing breast cancer compared to those who do not, by an order of 4.3% for every 12 months of breast feeding ^[32].

Age at Menopause

Early menopause decreases the risk of breast cancer. In fact, women who have undergone the menopause have a lower risk of developing breast cancer than pre-menopausal women of the same age. This risk is thought to increase by 3% with each year of delay in onset of the menopause ^[35]. Women who experience a natural menopause after the age of 55 years are twice as likely to develop breast cancer as those who do so before the age of 45 ^[35].

Exogenous Hormones

Diethylstilboestrol

Women who took the synthetic oestrogen, Diethylstilboestrol during pregnancy in the 1940-60's, to decrease their risk of miscarriage, were exposed to high levels of oestrogen in early pregnancy and have been shown to have a 27% increase in their risk of breast cancer. Whilst more commonly used in the USA than the UK, this drug is no longer prescribed ^[36].

Oral Contraceptives (OC)

The use of the oral contraceptive pill (OCP) is associated with a 25% increase in breast cancer risk for current and recent users, but with no excess risk ten or more years after discontinuation ^[37]. OC users are generally young women whose breast cancer risk is comparatively low, so this equates to a relatively small increase in cases each year of around 1% ^[37].

Hormone Replacement Therapy (HRT)

Women currently using HRT have a 66% increased risk of breast cancer. Again this risk is temporary, returning to that of a non-user within five years of discontinuation ^[38]. The risk is larger for oestrogen-progestogen therapy compared to oestrogen only and larger still for women with a low BMI ^[35, 38, 39]. Evidence also suggests that breast cancers in women on HRT tend to be larger and are more likely to be node positive, possibly because HRT renders mammographic detection more difficult ^[14]. The use of HRT in recent years has fallen. It is estimated that 3% (around 1530 cases per year) of British breast cancer cases are linked to HRT use ^[40].

Endogenous Hormones

Intrinsic Levels

Higher levels of endogenous hormones have long been hypothesised to increase breast cancer risk. It would appear that post-menopausal women with particularly high levels of oestrogen and testosterone are indeed at much greater risk than the normal population ^[41]. Similarly, high levels of prolactin, insulin and insulin-like growth factor are associated with a higher incidence of oestrogen receptor positive post-menopausal cancers, possibly explaining the slightly higher relative risk seen in diabetic patients ^[42-45]. The link between these hormones and pre menopausal cancers is however less clear ^[46, 47].

In-Utero Exposure

It has been suggested that in-utero exposure to high levels of oestrogen as may occur with older mothers who have offspring with high birth weights, may significantly increase the subsequent risk of breast cancer for the offspring. Conversely, studies suggest that breast cancer risk may be reduced in the offspring of mothers who suffered from pre-eclampsia or eclampsia, where maternal oestrogens are often low ^[48].

1.1.6 Non-Reproductive Risk Factors

A Study published in December 2011 which excluded reproductive factors, estimated that 27% of female breast cancers in the UK are linked to largely modifiable lifestyle and environmental risk factors ^[49].

Lifestyle Factors

Obesity

The correlation between breast cancer risk and body weight is likely to be hormonally driven. Obesity increases the risk of post-menopausal breast cancer by up to 30% and is one of the modifiable risk factors for breast cancer. The main source of endogenous oestrogen in post-menopausal women is from the peripheral conversion in adipose tissue ^[50]. Studies have shown that breast cancer incidence is higher if the extra fat is around the waist and if significant weight is gained between the ages of 20-40 years ^[51]. In contrast, obese pre-menopausal women have a 20% reduction in risk ^[51]. In this case, the risk reduction may well be due to their greater proportion of anovulatory menstrual cycles ^[52].

Diet

The role of the diet in breast cancer development is controversial and may not be nearly as important as was initially thought. The strongest evidence suggests that saturated fat intake significantly increases risk ^[53-55]. Phyto-oestrogens (plant compounds containing oestrogens) have been studied extensively and implicated in breast cancer development ^[56, 57]. Alcohol intake also appears to increase the risk ^[58-63].

Physical Activity

An active lifestyle appears to reduce breast cancer risk, particularly for post-menopausal women ^[64, 65]. Once again, this is probably due to the effect of exercise on hormone levels, with a reduction in oestrogen and testosterone seen in women who are more physically active ^[66]. A significant number of breast cancers in the UK are linked to an indolent lifestyle ^[67].

Sleep

It is thought that a lack of sleep suppresses the production of melatonin, which has anti-carcinogenic properties. Sleep deprivation may therefore increase breast cancer risk. Melatonin also suppresses the production of other hormones with similar anti-carcinogenic effects ^[68-71].

Medication

Regular use of Aspirin or other non-steroidal anti-inflammatory drugs (NSAIDs) is associated with a risk reduction of up to 25%, possibly due to a lowering of Oestradiol ^[72-75]. Conversely, the use of anti-hypertensive medication for more than five years is associated with an increased risk of breast cancer development ^[76].

Smoking

The evidence linking tobacco smoking to breast cancer is limited ^[77]. The increased risk reported in the relevant studies was of the order of 10-20% in heavy smokers ^[78, 79]. However, the evidence remains inconsistent and inconclusive for both pre and post-menopausal groups and more research is required ^[79].

Environmental Factors

Ionising Radiation

Ionising radiation is an established risk factor for breast cancer, especially at younger ages of exposure ^[80, 81]. Studies have shown up to 25 fold increases in secondary breast cancer development in women treated with mantle radiation therapy to the chest for Hodgkin's lymphoma before the age of 30 ^[82-85]. A much lower relative risk increase of 9% has been shown in women treated with radiotherapy for a previous breast cancer ^[86, 87]. Overall, it is estimated that around 1% of UK breast cancers in 2010 were linked to radiation. Almost half of these were linked to medical radiation and the rest attributable to background radiation ^[88].

Stature

The Million women study confirmed a 17% increase in the risk of breast cancer for every 10cm increase in a woman's height. The underlying mechanism is unclear but it seems likely that height is a marker for other environmental or lifestyle factors that influence breast cancer development ^[89, 90]. It may also be the case that the hormones that affect a woman's height, cause an increase in the volume of breast parenchymal tissue, and therefore susceptibility to breast cancer.

Family History

Approximately 5% of breast cancer cases in Western countries appear to have a strong genetic or familial predisposition. In general, individuals with a mother, sister or daughter with breast cancer have an almost two-fold increased risk of developing the disease themselves. This risk increases with the number of affected first-degree relatives and is higher if the relative is diagnosed under the age of 50 years ^[91, 92].

The Breast Cancer Genes

A number of genes may predispose to breast cancer and are usually inherited in an autosomal dominant manner, with limited penetrance. Thus they can be transmitted through the maternal or paternal lines and an individual can transmit the gene without him or herself developing cancer. Inheritance of a predisposing gene may result in breast cancer developing at a young age, often involving both breasts.

a) High Risk Genes

There are very few families in whom it is possible to be certain of the dominant inheritance of a cancer predisposing gene. A small proportion of women have a particularly strong family history, with four or more affected family members, and are, not surprisingly, at very high risk. Rare mutations in two well-documented cancer susceptibility genes, BRCA1 and BRCA2, located on the long arms of chromosomes 17 and 13 respectively, account for a substantial proportion of these high-risk families ^[93]. The estimated prevalence of BRCA1 and 2 mutations in the general population is only 0.11% and 0.12% respectively but carriers have a 60-85% chance of developing breast cancer over their lifetime, due to the high penetrance of the genes ^[94, 95], along with a risk of ovarian cancer.

Mutations of the tumour suppressing genes TP53 (chromosome 17p) and PTEN (phosphate and tensin homolog), associated with Li Fraumeni (sarcoma, brain malignancy and other tumours) and Cowden's syndromes (skin lesions and skull abnormalities) respectively, are also associated with a very small proportion of familial breast cancers ^[96].

b) Moderate Risk Genes

Intermediate penetrance gene variants that confer a 2-3 fold increase in risk, have been found in the genes that regulate the identification of DNA damage leading to cell cycle arrest, DNA repair or apoptosis including CHEK2, ATM (ataxic telangiectasia), BRIP1 and PALB2 genes. These genes are relatively rare, with a population frequency of 0.1-0.2%.

c) Low Risk Common Alleles

Common single nucleotide polymorphisms (SNPs) have now been identified by genome-wide association studies. Around 20 SNPs with population frequencies of >5% are associated with increases in the relative risk of breast cancer varying from 1.05-1.24. Other low penetrance gene variants have yet to be identified ^[96].

Whilst 85% of breast cancers still occur in women with a negative family history ^[91], it may be that hereditary factors still play a part in a proportion of these cases. Testing of genes, other than BRCA1/2 and, rarely, TP53, has however not been shown to have clinical utility.

Previous Breast Disease

Women with benign proliferative breast changes without atypia have a two-fold increase in risk while those with atypical epithelial hyperplasia have four to five times the risk of developing breast cancer ^[97]. On the other hand, women with non-proliferative, benign pathologies including palpable cysts, complex fibroadenomata, ductal papillomata and sclerosing adenosis only carry a significantly higher risk if they have a particularly strong family history of breast cancer ^[97].

Ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS) and, atypical lobular hyperplasia (ALH), combined now as lobular intra-epithelial neoplasia (LIN), are non-invasive conditions of the breast which can progress to invasive disease. Overall, women with LIN are four to five times more likely to develop an invasive breast cancer compared to the general population. Whilst it is not possible to predict which women will progress to invasive cancer, the risk is higher with high grade compared to low-grade in-situ disease ^[98].

A previous diagnosis of breast cancer also increases the risk of developing a second primary breast cancer. Whilst reported risk ratios vary, this is of the order of 3-4 fold ^[99-102].

Breast Density

The breast is composed of fat, connective tissue and epithelium. Breasts with a high epithelial content are described as dense in comparison with those with a higher fat content. Breast density is an independent risk factor for breast cancer ^[103, 104] associated with an almost five times greater risk of developing the disease ^[105].

1.2 Diagnosis

1.2.1 Triple Assessment

In the UK, women may present to a breast clinic with symptoms of disease, or they may be detected through the NHS breast screening programmes. All patients with a suspicious breast lesion should undergo triple assessment. This comprises:

- (i) Clinical examination of both breasts and the regional lymph nodes in the axilla and supraclavicular fossae.
- (ii) Radiological imaging of the breast by mammography or ultrasound, depending on the patient's age.
- (iii) Biopsy by fine needle aspiration cytology (FNAC), Trucut core biopsy or both.

1.2.2 Radiological Imaging

Mammography

Outwith breast cancer screening programmes, mammography is performed on symptomatic women over forty years of age. Below this age the normal breast tissues are too radio-dense and mammography is of limited value and should only be performed if a strong clinical suspicion of a cancer exists. Two views, oblique and cranio-caudal are obtained to allow detection of mass lesions, parenchymal distortion and, breast micro-calcifications, which may warrant further investigation. Cancers usually appear as solid, spiculated lesions on the radiographs.

Screening mammography is performed every 2-3 years from the age of 50 and has been shown known to reduce mortality through early diagnosis. In addition, all patients with a proven cancer should undergo mammography, with or without magnification views and whole breast ultrasound, prior to treatment to help assess the extent of their disease ^[106].

Ultrasound

In women under the age of forty years, breast ultrasonography (US) is the most useful imaging modality, whilst in older women it is used to further define localised palpable and mammographically identified lesions. Cancers usually have indistinct outlines on US and are hypoechoic. Patients with invasive cancer should also undergo axillary ultrasound with guided biopsy of any suspicious nodes.

Ultrasound is also a useful modality in monitoring variations in tumour size in response to treatment [106].

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is not currently in routine clinical use for breast cancer assessment in part because of its relatively high cost and time taken, although it can provide the best image detail of any currently available breast imaging technique. MRI sensitivity is high and it is valuable in demonstrating the extent of invasive and non-invasive breast disease. It has proven to be a valuable screening tool for selected high-risk women between the ages of 35 and 50 and is the optimum method for imaging the breast in patients with implants. It can also be of value in assessing early treatment response ^[106].

1.2.3 Breast Biopsy

If clinical examination and/or imaging raise suspicion of a malignancy, a breast biopsy is required. In cases where there is difficulty due to small or impalpable lesions, image guidance using mammography or ultrasound scan may increase the accuracy of the biopsy.

Core Needle Biopsy

Core biopsy is generally preferable to FNAC because it provides an architecturally intact specimen for histological diagnosis and can differentiate between invasive and non-invasive cancers. Whilst core biopsy samples may take a number of days to process in the laboratory, a portion of the sample can also be rolled onto microscope slides for immediate cytology to distinguish benign and malignant lesions. Providing the core biopsy samples the lesion adequately, this technique is highly sensitive and specific. For this reason core biopsy with touch or roll cytology has superseded FNAC ^[107].

Fine Needle Aspiration Cytology (FNAC)

Needle aspiration can differentiate between solid and cystic lesions and aims to extract sufficient tissue for cytological analysis, which can quickly differentiate between benign cells and malignancy.

1.2.4 Accuracy of Investigations

The sensitivity of clinical examination and mammography varies with age. Only two thirds of cancers in women aged less than 50 years are correctly identified as 'suspicious' or 'definitely malignant' on clinical examination and mammography alone. In addition, breast cancer in women under 40 years may fail to present with an identifiable discrete lump. Since FNAC suffers from a false negative rate of 4-5%, a histological diagnosis, based on core biopsy, which allows classification of individual cancers is therefore now considered an essential prerequisite to proceeding with treatment ^[107, 108]. All methods of assessment, whether clinical, radiological, or pathological, are subject to human errors in sampling, processing and interpretation. High standards of training, reporting and audit are therefore essential.

1.3 Classification

1.3.1 An Overview

Breast cancer classifications subdivide the disease in different ways, with different criteria and for different purposes. Each of these can influence treatment, predict response to treatment and help define an individual's prognosis as well as providing a standard nomenclature for use in audit and research.

Classical Descriptions

Current standard classification systems for breast cancers incorporate all of the following:

- (i) Histopathological type and grade.
- (ii) Oestrogen and progesterone receptors and HER2 status.
- (iii) The clinical stage of the cancer.

Molecular Descriptions

The introduction of new molecular tests and gene expression profiling are changing current understanding of breast cancer biology and as they do so, new classification systems are likely to evolve which will improve the accuracy of diagnosis as well as improving prognostication and prediction of the efficacy of treatment. As these systems are not yet in routine clinical use, they will be considered further in a later section on the molecular biology of breast cancer.

1.3.2 Histopathological Classification

Type

Breast cancers whose cells remain within the basement membrane of the terminal duct lobular units and their draining ducts are described as 'in-situ' or 'non-invasive'. An invasive cancer on the other hand, is one in which the basement membrane has been breached and there may be dissemination of cancer cells into the surrounding breast tissue and beyond. Both in situ and invasive cancers have characteristic histopathological appearances by which they can be recognised and classified on light microscopy ^[109].

In situ carcinoma was until recently, classified as ductal or lobular (DCIS and LCIS), each having differing characteristics.

DCIS is the most common type of non-invasive breast cancer, accounting for approximately 1 in 5 new breast cancers and, whilst not life threatening in itself, does confer an increased probability of progression towards invasive carcinoma. DCIS can be subdivided histologically into solid, cribriform, papillary, and micropapillary types of high, intermediate or low grade, and with the presence or absence of comedo histology. The latter subtype being characterised by ducts plugged with necrotic cellular debris and exhibiting a tendency towards higher grade and aggressive behaviour.

DCIS is almost always detected by microcalcifications on screening mammography as it rarely presents symptomatically as a discrete mass. The natural history of DCIS varies by grade such that approximately 40-60% of low-grade DCIS lesions will become invasive at 40 years of follow-up. On the other hand, inadequately treated high-grade DCIS has a 50% chance of becoming invasive within seven years ^[27, 110].

LCIS is considered to be a marker for increased risk of invasive cancer development. Estimates of this increase suggest a 3-4 times relative risk compared to average women or a 21% risk of developing invasive disease within the 15 years following diagnosis. A lack of clinical and radiological signs however, means that LCIS is almost always diagnosed as an incidental finding on biopsy specimens ^[111].

Invasive breast cancer is also subdivided into ductal and lobular types, based on the now somewhat outdated belief that these cancers arise from the ducts and lobules respectively. It is now clear that both invasive ductal (IDC) and lobular cancers (ILC) arise from the terminal duct lobular unit. Nevertheless this terminology remains in common use and is incorporated into the 2003 World Health Organisation (WHO) histopathological classification system for both benign and malignant breast tumours ^[112].

Together, the invasive ductal and lobular carcinomas, along with ductal carcinoma in situ (DCIS) account for approximately three quarters of all breast cancers. The ILC's can be difficult to diagnose because their pattern of single-file cell infiltration, like that of lobular carcinoma in situ, does not form a well defined mass lesion that can be readily detected clinically nor microcalcifications that can be seen radiologically ^[109].

Inflammatory breast cancer (IBC), a form of invasive ductal carcinoma, is distinguished clinically from other carcinomas by the presence of oedema and erythema in the overlying skin secondary to dermal lymphatic involvement and is associated with increased cancer aggression and a particularly poor prognosis ^[113].

Some tumours show atypical, distinct patterns of growth and cellular morphology and are sub-classified as 'invasive carcinomas of special type'. This classification has clinical relevance in that certain special type tumours have a much better prognosis or different clinical characteristics and behaviour compared with those of no special type ^[27].

Grade

Grading focuses on the appearance of breast cancer cells with respect to that of normal breast tissues. Normal cells within the breast are 'differentiated', meaning that they have specific morphological characteristics that reflect their function as part of the breast as a whole. Cancers lose their differentiation and cells become disorganised. Control of cellular division is lost and the cell nuclei become larger and less uniform (nuclear pleomorphism).

Pathologists grade the differentiation of breast tumours by considering glandular formation, nuclear pleomorphism and the frequency of mitoses. Each of these are scored using the Nottingham Modification of the Scarff-Bloom-Richardson grading system from 1-3 and their values are added to produce three overall grades ^[114, 115] :

Grade 1 cancers (scoring 3-5) are low-grade, well differentiated tumours. These confer the most favourable prognosis, can be treated less aggressively and have the best survival rate.

Grade 2 cancers (scoring 6-7) are of intermediate grade and moderately differentiated.

Grade 3 cancers (scoring 8-9) are of high grade. These are poorly differentiated and have lost many of the features seen in normal breast cells. They appear large and immature, will divide more rapidly and have a greater tendency to spread, thus conferring the poorest prognosis.

Among tumours of no special type, the derived grade yields important prognostic information, guides treatment and is an important predictor of both disease free and overall survival ^[109].

1.3.3 Molecular Classification

Cells express proteins on their surface and in their cytoplasm and nuclei that act as receptors for chemical messengers, such as hormones. The interaction between the receptors and their ligands cause responses in cellular behaviour. Breast cancer cells variably exhibit many different types of receptor but presently, three receptors are known to influence both prognosis and treatment and are therefore included in standard breast cancer classification systems. These receptors are the oestrogen receptor (ER), the progesterone receptor (PR) and the Human Epidermal Growth Factor receptor (HER2/neu). Cells with or without these receptors are labelled with the + or – suffix respectively. For example ER+ or ER-. Cells with none of these receptors are known as basal-like or triple negative breast cancers (TNBC). The latter include metaplastic and apocrine cancers.

The receptor status of breast cancers has traditionally been measured by immunohistochemistry (IHC), for inclusion in the pathologist's report. This laboratory technique identifies receptors on the cancer cells using labelled antibodies against the ER, PR and HER2 receptors. A similar technique known as fluorescent in-situ hybridisation (FISH) is also used to assess HER2 status. Receptor status is critical for all breast cancers as it determines the appropriate and targeted use of some of the most effective treatments.

Oestrogen Receptor

Approximately 75% of breast cancers express the oestrogen receptor (ER) ^[116]. ER can be quantified by using the Allred histochemical scoring system (detailed later). Most ER positive cancers tend to be ER rich, with an Allred score of 7-8. Most ER positive cancer cells depend on oestrogen for their growth and are therefore susceptible to drugs that either reduce the levels of oestrogen, or modify the activity of the receptor. These cancers consequently have a better prognosis with respect to their ER negative counterparts when appropriately treated.

Progesterone Receptor

Most ER positive cancers also express progesterone receptors (PR), increasing the likelihood of a favourable response to hormone therapy. Very few PR positive, ER negative cancers exist. The progesterone receptor is scored in the same way as the oestrogen receptor through immunohistochemical methods including the Allred score.

Human Epidermal Growth Factor Receptors

Four human epidermal growth factors exist. These are Epidermal growth factor receptor (EGFR / HER1), HER2 and the less well known HER3 and HER4. Among these, HER2 is the only one routinely included in the pathologist's report.

Almost one fifth of all breast cancers are HER2 positive and these respond to agents that target the HER2 receptor. Monoclonal antibodies raised against HER2, prevent HER2 receptor activation and in combination with conventional chemotherapy, improve the prognosis for these patients ^[117]. The majority of ER positive tumours however, are HER2 negative.

Targeted treatment options for patients with triple negative breast cancers, which are relatively common among BRCA1 carriers, are very limited due to the lack of receptor expression. They therefore carry a comparatively poor prognosis although the introduction of PARP inhibitors, which prevent DNA repair in cancer cells, is showing great promise. These tumours respond to cytotoxic chemotherapy in about half of all cases ^[118].

Clinical Staging

Once a breast cancer has been diagnosed and histologically classified, the local and distant extent of the disease should be assessed. This process is known as staging and allows further classification and stratification of disease extent, which will now be discussed in more detail.

1.4 Staging

Staging based upon clinical examination and radiological investigations alone is known as 'clinical' staging as distinct from 'pathological' staging, which can only be completed after biopsy specimens are processed. Since some patients are appropriately treated without surgical intervention, not all will achieve full pathological staging.

1.4.1 TNM and UICC

Two staging systems are commonly applied to breast cancer and are often used together, although neither are perfectly suited to the disease. They are promoted by The American Joint Committee on Cancer (AJCC) and the International Union Against Cancer (UICC) ^[119] (Appendix I).

The Tumour, Node, Metastases (TNM) system depends upon clinical (radiological or pathological) measurements of the tumour size (T) and clinical assessment of the lymph node status (N) and the presence or absence of distant (metastatic) disease (M). Although each is clearly subject to potential errors, the accuracy of the TNM classification can be improved when the tumour size and nodal status, as assessed by a pathologist, are taken into account.

The International Union against Cancer (UICC) system groups various combinations of the TNM factors into 'stages'. Stage 0 equates to non-invasive or in-situ disease; Stages 1-3 represent disease confined to the breast and lymph node basins and stage 4 to metastatic disease.

Patients with small tumours (stage I and II) have a low risk of metastatic disease and require no further investigations, unless prompted by clinical symptoms and signs. For those with larger tumours and in whom there is suspicion of distant spread, CT and CT-PET scans should be considered.

1.4.2 Staging Terminology

Recent Changes

The most recent 2010 staging scheme incorporates a number of new features:

Tumour – As in previous systems, the T values (TX, T0, Tis, T1, T2, T3 or T4) depend on the primary cancer size within the breast. TX describes an inability to assess that site; Tis (in situ) refers to DCIS, LCIS, or Paget's disease, a rare type of cancer involving the nipple areolar complex that is often associated with deeper tumour within the breast. T4d refers to inflammatory breast cancer, a clinical variant with skin involvement over at least a third of the breast.

Node - The lymph node values (NX, N0, N1, N2 or N3) reflect the number, size and location of cancer deposits in regional lymph node basins, such as the axillae, supraclavicular and internal mammary lymph nodes. There is an additional sub-classification for N0 disease, which includes N0 (i+) for nodes containing Isolated Tumour Cell clusters (ITC), which are small clusters of fewer than 200 cells, not greater than 0.2 mm in diameter. N0 (mol-) nodes contain no histologically detectable metastases, but have molecular evidence of tumour deposits, as measured by the polymerase chain reaction (PCR).

Metastases – Classically, M0 and M1 refer respectively to the absence or presence of metastases in locations other than the breast and regional lymph nodes, such as bone, brain and lung. MX formally referred to an incomplete assessment of metastasis but this has been dropped from the present TNM edition, which introduces one additional category, M0 (i+) to denote the detection of ITCs no larger than 0.2mm in circulating blood, bone marrow or non-regional nodal tissue but without clinical or radiographic evidence of metastases. M1 also now includes the detection of any metastasis larger than 0.2 mm. Perhaps surprisingly, M0 (i+) does not affect the cancer staging of the tumour.

1.4.3 Staging the Axilla

Axillary node status is the single most significant prognostic factor in patients with breast cancer. Both the number of involved nodes and their anatomical level have predictive value for survival ^[120]. With up to 40% of early invasive breast cancer patients having involvement of the axillary nodes at the time of diagnosis, the early evaluation of axillary lymph node status is clearly critical for the accurate staging of patients and to provide a basis for planning treatment.

Clinical and radiological assessment of lymph node status are however unreliable, whilst routine US scan followed by FNAC or core biopsy of clinically suspicious nodes, can detect up to 50% of patients with axillary node involvement ^[121]. Pre-operative identification of nodal malignancy allows for prompt and definitive axillary surgery. In many patients however pathological staging is only complete following surgical sampling of the lymph nodes. Standard histopathological examination of excised nodes is therefore essential and when positive, is of greater prognostic significance than metastases only detected using histochemical techniques.

Lymphatic Drainage of the Breast

Approximately 95% of the lymphatic drainage of the breast passes to the axilla. The remaining 5% is via the internal mammary and intercostal systems.

Sentinel Lymph Node Biopsy (SNB)

The principle behind SNB is that lymphatic drainage occurs through a series of nodes in a sequential manner. Therefore, the status of the first node directly draining the tumour bed helps predict the status of the remaining nodes in the axilla. Sentinel nodes can be identified using a combination of radiolabelled colloid that can be visualised on lymphoscintigraphy or detected by means of a hand-held gamma counter together with the use of a blue dye that can be detected intra-operatively. Histological assessment of sentinel nodes predicts overall node status with over 95% accuracy. The false negative rate is less than 5% in experienced hands and can be reduced further by removing any additional palpable or suspicious nodes at the time of SNB.

SNB is standard practice for patients with clinically N0 tumours, particularly those with tumours <2cm, in whom the likelihood of nodal involvement is low.

Axillary Node Sampling (ANS)

Several centres use the blue dye alone to identify and biopsy at least four palpable nodes and claim results equivalent to SNB ^[122].

1.4.4 Clinical Terminology

Whilst it does not form part of a separate classification system, breast cancer is often referred to as either early, locally advanced or advanced depending upon the results of staging investigations.

Early Breast Cancer

Early breast cancer (EBC) is confined to the breast and to the loco-regional lymph nodes (Stages I-III).

Locally Advanced Breast Cancer

Locally advanced breast cancer (LABC) is a heterogeneous and relatively rare form of breast cancer that is located extensively throughout the breast, the ipsilateral regional lymph node basin or both. The tumour is usually large (>5cm), although size is not an absolute criterion, and the characterisation is based upon one or more of the following:

- (i) Skin or chest wall involvement
- (ii) Fixed axillary lymph nodes
- (iii) Involvement of ipsilateral internal mammary, supraclavicular or infraclavicular nodes

Advanced Breast Cancer

Advanced breast cancer (ABC) is metastatic and has therefore spread beyond the breast and the regional lymph nodes (Stage IV).

The prognosis and treatment of breast cancer relate to the stage of the disease at presentation.

1.5 Prognosis in Breast Cancer

A prognostic factor is one that is objectively measurable at the time of diagnosis and can help to predict clinical outcome, independent of therapy ^[123]. Ideally, they should have biological significance, be reliable, easily interpreted and applied. Accordingly, patients are stratified into low or high-risk groups, which aim to provide a relative, but not absolute, prediction of the future behaviour of their disease.

Many of the most powerful prognostic factors in current clinical use are reflected in the TNM staging and other classification systems already discussed. These can be divided into clinical, histological and molecular factors, which often interrelate with one another.

1.5.1 Clinical Factors

Tumour Size

The size of a cancer, as measured by a pathologist following surgical excision and confirmed or amended after histological examination, correlates with survival. Patients with smaller cancers have better survival than those with larger cancers ^[124].

Node Status

Histologically proven axillary node status is the most powerful prognostic factor in breast cancer. Survival is correlated directly with the number and level of axillary nodes involved ^[125].

Recent interest has focused on the assessment of small deposits of tumour within axillary lymph nodes known as micrometastases. The new TNM system uses a pragmatic definition of a micrometastasis as one measuring between 0.2mm and 2 mm in diameter. The importance of these is however unclear and such metastatic foci are currently treated in a similar manner to node negative disease ^[126].

Metastatic Disease

Patients with metastatic disease, particularly the 10% of patients with metastases at the time of presentation, have a poorer prognosis than those with apparently localised disease. Survival differs according to the site of spread such that patients with supraclavicular fossa disease have a better survival than those with disease at other sites ^[127]. Similarly, patients with bony metastases alone, have a better outlook than those with visceral disease ^[128].

Survival outcomes for breast cancer according to stage are shown in the Table 1.5A. These figures are however based on women diagnosed in the early 1990's. The current outlook for equivalent stage cancers has improved thanks to advances in screening and treatment ^[22].

Breast Cancer Stage	5-year Survival	10-year Survival
Stage I	90%	85%
Stage II	70%	60%
Stage III	50%	40%
Stage IV	13%	10%

Table 1.5A: Breast Cancer Survival According to TNM Stage.

1.5.2 Histological Factors

Histological Type

Special types of invasive breast cancer, including tubular, mucinous and invasive cribriform cancer, are associated with a better prognosis than invasive ductal cancer of no special type. Additionally the histological type may also provide information about the biological behaviour of that cancer and its particular pattern of spread ^[129].

Histological Grade

The histological grade (1-3) as assessed by tubule formation, nuclear pleomorphism and mitotic frequency correlates with survival. Low grade cancers (grade 1) have a better prognosis than intermediate (grade 2) or high grade (grade 3) tumours ^[130].

Lymphatic or Vascular Invasion (LVI)

The presence of cancer cells in the blood or lymphatic vessels and into perineural tissue within a breast biopsy occurs in approximately 25% of cases and is a marker of more aggressive disease that is associated with both an increase in the rate of local disease recurrence and short term systemic relapse ^[131].

Extensive In-Situ Component

If more than 25% of the main tumour mass contains non-invasive disease and there is in-situ cancer in the surrounding breast tissue, the cancer is classified as having an extensive in-situ component. Patients with such tumours are considered to be at increased risk of local recurrence after breast conserving surgery if the margins of excision are not clear ^[132]. If the surgical margins of excision are clear however, then there is no greater risk of recurrence ^[132]. LIN and atypical ductal hyperplasia are not markers for increased recurrence ^[109].

Other Histological Factors

The prognostic significance of peri-tumoural vascular and lymphatic angiogenesis and lymph node micrometastases is not clear at the present time ^[133].

1.5.3 Molecular Biological Factors

A variety of biological factors have also been used in the prognostication of breast cancers (Table 1.5B).

Biological Markers of Prognosis	
Clinically Useful	Experimental
Oestrogen receptor (ER) HER2	DNA alterations RNA expression Intracellular proteins Extracellular proteins

Table 1.5B: Biological Factors Used For Prognostication ^[134].

ER

ER positivity is associated with a good prognosis ^[135], whilst epidermal growth factor receptor (HER 1) correlates inversely with ER status and is associated with poorer survival ^[136].

HER2

HER2 has similarly been associated with a worse prognosis in node negative patients^[137]. The advent of anti-HER2 therapies has, however, produced marked improvements in survival ^[138].

An increased expression of cellular proliferation markers, indicative of uncontrolled growth, is associated with poor clinical outcomes in breast cancer ^[139]. Immunohistochemical (IHC) staining of Ki67, a nuclear antigen that is present only in proliferating cells ^[140], has been shown to be a reliable marker to enumerate the growth fraction of normal and neoplastic cell populations ^[141]. The Ki67 labelling index, based on the percentage of cells with positive Ki67 staining, correlates with the mitotic index. Ki67 is variably expressed in cells throughout the cell cycle but not during the resting G0 phase and is thus most frequently expressed in poorly differentiated tumours with high rates of mitotic activity ^[142]. It is expressed at low levels in normal breast tissue (<3%) and only in ER negative cells. In cancers however Ki67 is expressed in ER positive and negative cells and the level of expression increases with the degree of malignancy from in situ to invasive ^[143].

Multiple studies have shown that baseline tumour Ki67 is a prognostic factor for breast cancer ^[144, 145] with a positive correlation to histological tumour grade and lymph node status ^[146, 147]. In a meta-analysis of 29 studies and over 12,000 patients with early stage breast cancer, Ki67 positivity was associated with a higher probability of relapse and a worse survival ^[148]. Similarly, Stuart-Harris et al. conducted a meta-analysis of 43 studies and over 15000 patients and found a clear correlation between Ki67 and overall survival ^[149]. These studies however were all retrospective and lacked internationally standardised methods for antigen retrieval, staining procedures and scoring. Consequently, there is a lack of confidence in the clinical use of baseline Ki67 for prognostic assessment ^[150]. The American Society of Clinical Oncology does not currently recommend routine use of Ki67 in clinical practice ^[151]. Many other biological factors are of uncertain clinical significance (Table 1.5C).

Biological Markers of Uncertain Clinical Significance	
Proliferation markers	Ki67 MIB-1 (monoclonal antibody to Ki67) Thymidine labelling % S phase Mitotic Activity Index
Apoptosis regulating genes	Bcl-2 Bcl-x Bax Bak Survivin
Cell cycle regulatory genes	Cyclin A, B, D, E Overexpression of p21, p27, p53
Cell adhesion molecules	E cadherin Integrins Fibronectin MSF
Proteases	Cathepsin D Matrix metalloproteinase Tissue inhibitor of matrix metalloproteinase
Oncogenes	HER3 HER4
Oestrogen receptor related	Progesterone receptor (PR) pS2
Signal transduction pathways	Extracellular signal regulated kinase ½ J N-terminal kinase P38
Allelic imbalance	-1p,7q,8p,10q,11q,15q,16q,17p,17q

Table 1.5C: Biological Markers of Uncertain Clinical Significance^[134].

1.5.4 Prognostic Indices

Many clinical, histological and biological factors that determine prognosis are interrelated. Some are difficult to determine and many do not have confirmed independent prognostic value. A number of tools in current clinical use attempt to combine these factors in order to estimate prognosis and plan treatment effectively.

The Nottingham Prognostic Index (NPI)

The Nottingham Prognostic Index (NPI) incorporates invasive tumour size, lymph node status and, histological grade and is calculated using the following equation ^[152, 153].

$$\text{NPI} = 0.2 \times \text{invasive size (cm)} + \text{lymph node stage (score 1 for no nodes, 2 for 1-3 nodes, 3 for > 4 nodes)} + \text{grade (score 1 for grade 1, 2 for grade 2, 3 for grade 3)}$$

Originally the NPI was used to divide patients into good, intermediate, or poor prognostic groups. Subsequent studies have led to a refined NPI with six categories (Table 1.5D) ^[124]. Nonetheless, its use in clinical practice has declined with the development of more sophisticated online tools.

Nottingham Prognostic Index (NPI)		
Group	Index Value	Survival at 10 years (%)
Excellent (EPG)	2.0-2.4	96
Good (GPG)	2.41-3.4	93
Moderate 1 (MPG1)	3.41-4.4	82
Moderate 2 (MPG2)	4.41-5.4	75
Poor (PPG)	5.41-6.4	53
Very Poor (VPPG)	>6.41	39

Table 1.5D: Breast Cancer Specific Survival According to NPI ^[124].

1.5.5 Predictive Indices

In contrast to a prognostic factor, a predictive factor is one that provides information on the likelihood of a favourable response to a particular treatment. Thus, it can be used to identify subpopulations of patients likely to benefit from any given therapy.

Some prognostic factors are predictive and vice versa. For example in addition to providing prognostic information, ER status predicts response to hormone treatment ^[154]. Similarly HER2 positivity predicts response to immunological treatment with Trastuzumab (Herceptin), Lapatanib and Pertuzumab and HER1 (EGFR) positivity to treatment with Lapatanib, Pertuzumab and Gefitinib ^[134].

With this in mind, a number of attempts have been made to incorporate prognostic and predictive information into clinically useful tools, which aim to provide prognostic and predictive information.

Adjuvant! Online

Adjuvant! online is a frequently updated site that provides information on the probability of relapse and survival following the input of patient details including age, general health, tumour size, nodal status, oestrogen receptor status and also HER2. This information is processed to provide details of recurrence rates and survival with and without adjuvant therapy and the likely benefits from alternative therapies (www.adjuvantonline.com).

PEPI Score

The Preoperative Endocrine Prognostic Index, though not widely used clinically, is designed for use in patients who have undergone 3-4 months of treatment with endocrine therapy prior to surgery. It incorporates information regarding pathological tumour size, lymph node status and tumour expression of Ki67 in order to predict the long-term outcome for ER positive breast cancer patients and in particular the likelihood of disease recurrence. Patients with a score above 4 are thought to benefit from additional treatments following endocrine therapy and surgery ^[155].

ICH-4

Recently, a test combining measurement of ER, PR, HER2 and Ki67 has been described that can predict low risk of tumour recurrence and may obviate the need for post-operative chemotherapy in these patients ^[156].

1.5.6 Future Developments

Microarrays

Tissue microarrays, which will be discussed in detail later, allow the rapid evaluation of the expression of an enormous number of genes on a single chip. Presently such technology is being used to investigate the molecular mechanisms of breast cancer development, response and resistance to treatment.

These microarray studies have identified clusters of genes whose expression is significantly up- or down-regulated during treatment. A few of these studies have led to the development of commercially available molecular testing kits for clinical use, that reflect the pattern or signatures of gene expression that are thought to be prognostic or predictive. Examples include MammaPrint® (Agendia, Amsterdam) ^[157] and Oncotype Dx® ^[158], which will be described later. Ongoing investigations aim to determine whether these will prove to be clinically useful prediction tools.

Proteomic Methods

Proteomic arrays and mass-spectrography, which examine the expression of known and novel proteins in tissues from cancer patients, present possible alternative means to detect 'markers' of response to therapy and may be related to prognosis. These are not yet in routine clinical use however and fall outwith the scope of this thesis.

1.5.7 Outcome Measures

Mortality and Survival

Clinical outcomes are generally defined by mortality and survival rates. These are easily measured, objective endpoints that can reflect both the beneficial and harmful effects of an intervention. They are considered to be the gold standard for measuring efficacy in clinical trials evaluating cancer treatments.

Mortality rates are expressed as the number of deaths in a breast cancer population of a defined size over a defined period of time. They may be breast cancer specific or include death from all causes. Survival rates, on the other hand, indicate the proportion of breast cancer patients still living at a defined time, usually 1, 5 or 10 years following diagnosis.

Secondary Endpoints

Secondary clinical outcomes, also known as 'Interim end points', may help with the prediction of survival and mortality. Such measures include:

- (i) Disease (or Event) free survival (DFS) - the length of time between an intervention and the moment a patient experiences a recurrence, a new primary cancer or death.

- (ii) Progression-free survival (PFS) - the time from the beginning of an intervention until a patient shows sign of disease progression. This is also sometimes known as Time to progression (TTP) - This is typically measured in treatment trials involving advanced breast cancer or metastatic disease.
- (iii) Tumour response rate - a radiographic 'calliper' measurement of the reduction in tumour size used in the setting of metastatic disease or neoadjuvant therapy.
- (iv) Recurrence rate, which may refer to local or regional recurrence of a breast cancer during or after a treatment at a selected follow up time point.

Clinical trials using secondary endpoints usually require a smaller number of patients and shorter follow-up than those measuring mortality or survival directly, but do not provide information on possible long-term effects of an intervention.

1.6 Management Pathways – Local Treatments

Treatment algorithms depend upon breast cancer classification and predictive indicators to define how specific subgroups should be treated according to the best available evidence. Most patients benefit from a combination of local treatments (surgery and radiotherapy) to control disease in the breast and axilla, and systemic treatment to treat potential metastatic disease. The aim is to achieve long-term local disease control with minimal morbidity.

1.6.1 Surgical Management of the Breast

Surgery is described as breast conserving (BCS) when excision of the tumour with a margin of surrounding tissue preserves some of the breast volume in distinction to mastectomy. BCS is always followed with radiotherapy.

At least 12 randomised trials have compared morbidity and mortality in mastectomy versus BCS and shown a non-significant reduction in mortality with BCS. Local recurrence rates are similar albeit with a non-significant but reduced incidence after mastectomy. Two large randomised trials have shown no significant differences in survival after 20 years of follow up [159, 160]

Selection for BCS or Mastectomy

Patients deemed suitable for BCS are those with:

- (i) Tumours staged as T1 or T2 < 4cm, without nodal disease or metastatic spread.
- (ii) Tumours > 4cm in a large breast.

Patients not suitable for BCS will be offered mastectomy. Others may choose to have a full mastectomy when BCS could be offered, because of personal preference or in some cases where BCS would produce an unacceptable cosmetic result. Other indications for mastectomy include those with multifocal or multi-centric operable breast cancer, patients unsuitable for radiotherapy and those with a strong family history of breast cancer including BRCA1 or 2 carriers.

Finally, some consideration of various clinico-pathological features of the patient or their tumour may further influence the decision on breast conservation.

Breast Conserving Surgery

Breast conserving surgery aims to excise the entire tumour and any surrounding DCIS with a 1 cm margin of macroscopically normal tissue (wide local excision or WLE). This is combined with removal of the sentinel node or all of the axillary nodes. Impalpable lesions can be localised pre-operatively by several techniques such as the radiologically guided placement of a hook wire.

Up to 10% of the breast volume can be removed without significant deformity whilst adequate excision of lesions over 4 cm is generally associated with poor cosmesis. Extensive excisions of larger tumours by quadrantectomy were performed in the past, but these are no longer favoured as they do not significantly lower local recurrence rates and are associated with poor cosmetic outcomes^[161]. For this reason, most breast units offering BCS tend to limit it to lesions less than 4 cm. An exception may be made in large breasted women in whom larger tumours may be excised with acceptable cosmetic results^[120, 132].

Where tumours are too large for routine BCS, consideration should be given to either neoadjuvant medical treatment designed to shrink the tumour pre-operatively or to oncoplastic surgery. The latter, also known as 'therapeutic mammoplasty' incorporates tumour excision with breast reduction and re-shaping, usually necessitating a symmetrising reduction mammoplasty on the contralateral breast. Alternatively, breast volume after tumour excision can be replaced using a pedicled Latissimus dorsi flap reconstruction. These procedures are associated with less psychological morbidity than that following mastectomy [162].

Complete excision of all invasive and in-situ disease is of course essential and intra-operative imaging of the specimen helps to confirm this. A negative margin of 1-2mm is considered adequate [132]. Wider margins do not reduce local recurrence further but may adversely affect cosmesis. Approximately 20-25% of cancers are incompletely excised at initial surgery. Whilst re-excision of a single positive margin usually results in complete excision, the greater the number of margins involved the less the likelihood of clearance and the greater the chance of recurrence. Other risk factors for local recurrence include younger age, high-grade tumours and the presence of lymphovascular invasion. Such patients may be advised to proceed to mastectomy. On the other hand, the presence of atypical ductal hyperplasia or LIN at the margins does not necessitate re-excision [132, 163-165].

BCS is possible in multifocal cancers provided that all disease is excised.

BCS Following Neoadjuvant Therapy

Large breast cancers, initially unsuitable for BCS can be downsized with neoadjuvant chemotherapy or endocrine therapy to enable breast conservation. Rates of subsequent complete excision are higher with endocrine therapy, which reduces tumour volume in general more effectively than chemotherapy, which has a tendency to reduce cellularity rather than volume [166-169].

Mastectomy

About one third of patients with localised breast cancers are unsuitable for BCS and require or elect for a mastectomy, often including some overlying skin and the nipple-areolar complex. Increasingly however, both skin and nipple sparing mastectomies are performed in conjunction with immediate breast reconstruction to achieve optimal aesthetic results [120, 162, 170, 171].

Down-Staging of Locally Advanced Cancers by Neoadjuvant Therapy

Locally advanced breast cancer was traditionally treated with mastectomy, where possible, but this was associated with a high rate of both local and distant recurrence and was almost never curative. The current approach is to attempt to downstage the primary tumour using chemotherapy or endocrine therapy and thereafter to perform surgery to gain local control of the disease ^[172].

1.6.2 Surgical Management of The Axilla

Breast conserving surgery and mastectomy for early stage invasive breast cancer, need to be combined with an axillary procedure either for staging or treatment purposes.

SNB or Axillary Dissection

Before the introduction of SNB, axillary node dissection (AND, levels 1-2) or clearance (ANC, levels 1-3), was standard practice. The morbidity of such procedures is unacceptably high for node negative disease, which is better assessed with SNB. Where the sentinel nodes are not involved, no further treatment is required to the axilla. Extensive axillary dissection should now be limited to patients with either cytological or histological evidence of axillary node involvement or to those who are highly likely to have involved nodes. Trials comparing SNB with AND have shown that the former is associated with less morbidity and shortens hospital stay ^[173].

Controversy currently surrounds the optimal management of the axilla in patients with only one or two positive axillary nodes at SNB, who are also receiving systemic treatment and whole breast radiotherapy. To date, the evidence from randomised trials suggests that local recurrence and survival rates are similar, regardless of whether or not completion lymphadenectomy or axillary radiotherapy were undertaken ^[174]. These results could be set to change current management for this group of patients.

Studies are ongoing to ascertain the need for axillary therapy where only isolated micro metastases are present in the SNB ^[175, 176]. A conservative approach has been advocated for some patients with low risk invasive cancers but, until the results from studies are available, some form of axillary surgery should be considered.

Treatment of Axillary Disease

In biopsy proven axillary disease, then either completion axillary node dissection or radiotherapy is recommended. Both are effective and associated with similar survival rates. Surgical dissection offers a lower rate of axillary recurrence but the morbidity of the procedure is high. Radiotherapy as a treatment modality can only be used once ^[177].

1.6.3 Radiotherapy

Radiotherapy uses ionizing radiation to control or kill malignant cells. In breast cancer patients, it may be prescribed:

- (i) Before surgery to reduce tumour size.
- (ii) During or after surgery to reduce the chance of local disease recurrence and improve survival.
- (iii) As an alternative to surgery for palliative care and pain control.

Targeted radiation damages the DNA of exposed tissue leading to cellular death. Cancer cells are more susceptible than normal surrounding tissues due to their higher mitotic rate. The radiation field can be shaped and applied from several different directions, all intersecting at the tumour site, to increase the exposure of the tumour whilst minimising that of surrounding healthy tissue.

The draining lymph node basins may also be targeted if they are clinically or radiologically involved with tumour, or if there is thought to be a risk of sub-clinical malignant spread.

The aims of radiotherapy in an individual will vary depending upon the tumour type, location, and stage, as well as the general health of the patient.

Radiotherapy is most commonly delivered as external beam radiotherapy but brachytherapy, in which a radiation source is placed inside the breast or next to the tumour bed, is sometimes used and minimises exposure to healthy tissue. Modern techniques provide more scope for partial breast irradiation, with greater field and marginal dose control.

Radiotherapy Following BCS

All patients should receive radiotherapy following BCS to reduce the rate of local recurrence and improve survival ^[178] ^[179-182]. A novel approach known as accelerated partial breast irradiation (APBI) is used to deliver radiation as part of breast conservation therapy. It limits treatment to the tumour bed and adjacent tissue, whilst reducing the duration of treatment to just five days, compared to the typical six or seven weeks for whole breast irradiation. Its use is usually limited to women with well-defined tumours that have not spread. A study by the National Surgical Breast and Bowel Project (NSABP) is underway to determine whether limiting the radiation field in this way is as effective as irradiating the whole breast ^[183].

Radiotherapy Following Mastectomy

Post-mastectomy radiotherapy (PMRT) is more common than was the case 10 years ago ^[184]. Patients with a high risk of local recurrence benefit from treatment of the chest wall, as do women with pectoral muscle involvement. PMRT results in a three fold reduction in local recurrence at 15 years, with the most significant difference in the first five years ^[179]. The effect of PMRT on mortality is variable, with 1% improvement in mortality in EBC and a 2-4 % reduction in higher risk groups. ^[179]. At present there is no consensus opinion on which lower-risk patients might benefit from PMRT. The ongoing SUPREMO trial aims to address this (www.supremo-trial.com).

Radiotherapy for LABC

Radiotherapy plays an important role in the treatment of LABC with the aim of reducing the risk of local recurrence following surgery, or to achieve local control where the primary tumour is inoperable despite attempts to downstage it using medical treatments ^[172] ^[185].

Radiotherapy Following Axillary Surgery

The American Society of Clinical Oncology suggests that routine post-operative axillary radiotherapy is not necessary after complete level I/II axillary dissection and may add to morbidity ^[186]. However, the supraclavicular field should be irradiated in all patients with four or more positive axillary nodes as this confers a survival benefit. The evidence regarding the treatment of the internal mammary node chain (IMN) with radiotherapy is unclear but on balance, the risks probably outweigh the benefits ^[178].

Timing of Radiotherapy

The optimal timing of post-operative radiotherapy has not been established. It is usual for trial eligibility criteria that radiotherapy is received within 12 weeks of surgery unless receiving chemotherapy. In Scotland, the stated target is that 95% of patients receive radiotherapy within four weeks of BCS or final dose of chemotherapy^[187, 188]. There is insufficient evidence to recommend the ideal sequencing of PMRT and systemic therapy.

1.7 Management Pathways - Systemic Treatments

More than half of patients with operable breast cancer, who undergo loco-regional treatment alone, will die from metastatic disease. Systemic therapy improves survival of these patients, suggesting that micro-metastases are present at diagnosis. Systemic treatment can be classified as chemotherapy, endocrine or targeted therapy. It may be delivered before or after loco-regional treatment. Both chemotherapy and targeted therapy will be discussed in brief in this chapter. Endocrine therapy will form the focus of discussion in subsequent chapters.

1.7.1 Chemotherapy

Traditional chemotherapeutic agents specifically target cells undergoing mitosis, thereby preferentially affecting rapidly dividing cancer cells. Some drugs cause cells to undergo apoptosis (programmed cell death). By its very nature however, chemotherapy also harms normal cells undergoing mitosis – with the bone marrow, digestive tract, and hair follicles being especially susceptible, resulting in the most common side-effects of chemotherapy: myelosuppression, mucositis and alopecia.

The benefits of chemotherapy depend largely on the biological subtype of the tumour and are greatest in women with ER negative, HER2 positive disease and in younger women. Adjuvant chemotherapy is widely used in the over 50's, but its efficacy appears to reduce with increasing age and as yet there is no clear evidence of benefit in women over 70 years^[178, 189]. All women under the age of 70 years with early breast cancer should be considered for adjuvant chemotherapy^[178].

In contrast, there is little evidence to suggest chemotherapy is of additional benefit to many women with ER positive, HER2 negative disease, already taking endocrine therapy. This is particularly true for grade I and II tumours.

Some of the newer therapeutic agents, for example monoclonal antibodies, are not indiscriminately cytotoxic, but rather target proteins that are abnormally expressed in cancer cells and that are essential for their growth. Such treatments are often referred to as targeted therapy, as distinct from classic chemotherapy, and are often used alongside traditional chemotherapeutic agents in treatment regimens.

Adjuvant Chemotherapy in EBC

The ability of adjuvant chemotherapy to reduce recurrence and improve survival from breast cancer has been established by a series of meta-analyses of many clinical trials^[189]. Whilst it was once believed that the benefits of chemotherapy were age related, and greatest in young women, it is now clear that other biological factors have a strong influence. For example, chemotherapy is known to be most effective against ER negative and/or HER2 positive disease. Whilst these subtypes are more prevalent in young women, it does not however preclude older women with such disease receiving appropriate treatment.

The decision to offer chemotherapy is based upon a risk benefit analysis and on the basis of the patient's age and tumour characteristics as well as the type of therapy being considered. The absolute benefit of chemotherapy appears to increase with an higher number of adverse prognostic factors, including: Tumour size >2cm; Nodal involvement; Oestrogen and progesterone receptor negativity; HER2 positivity; High grade (grade III); the presence of LVI and younger age (<35 years). Risk/benefit considerations are important because of the potential toxicity of these agents and so, consensus criteria have been defined to aid in patient selection^[190].

Chemotherapy does not appear to offer an advantage over endocrine therapy alone in postmenopausal women with grade I or II, ER rich, HER2 negative breast cancer. However, a key challenge is to identify a subset amongst this large subgroup in which it may confer additional benefit. Gene expression assays may have an important role in defining this subset of patients.

Adjuvant Regimens

Whilst there is no single gold standard chemotherapy regimen in early breast cancer, an increasing body of evidence suggests that combinations of Anthracycline with Doxorubicin or Epirubicin achieve a significant survival advantage of 4-5% over traditional non-anthracycline regimens containing Cyclophosphamide, Methotrexate and Fluorouracil (CMF) ^[178, 189, 191]. In the UK, a sequential combination of Anthracyclines followed by CMF is widely used and, is superior to CMF alone ^[192]. The Anthracyclines are however more toxic, with higher rates of myelodysplasia and neutropenic sepsis. They are also associated with a modest risk of cardiac impairment ^[191]. Short duration regimes combining 5FU, Epirubicin and Cyclophosphamide (FEC) provide similar benefits to the longer sequential regimen ^[193].

The addition of Taxanes (Docetaxel or Taxotere), to the Anthracyclines has further improved survival in node positive disease, especially in women with ER negative and or HER2 positive tumours ^[194]. Again, these can be used sequentially or in combination ^[195]. With new treatment regimens, five-year survival in node positive women has risen from 65% to over 85%. Currently, there is no evidence to demonstrate a benefit of the Taxanes in node negative disease.

A frequently used regimen in the UK for high-risk node positive patients is the FEC-T combination, which involves 3 cycles of FEC (or Anthracycline) followed by 3 of Docetaxel. This has shown a 27% reduction in the relative risk of death in patients with a high tumour Ki67 ^[196].

Dose Density

There is no clear consensus on how individual chemotherapy drugs should be sequenced. Drugs are often combined and there is some evidence to suggest that block sequential therapy may be better ^[197].

Dose density therapy intensifies the chemotherapy regimen by reducing the inter-treatment interval to two instead of three weeks and this may improve disease free survival ^[198]. Finding the minimum duration for effective treatment is an important challenge.

To counter the negative effect of chemotherapy on white blood cell production, granulocyte colony-stimulating factor (G-CSF) is administered to stimulate the production of new white blood cells. This has been shown to reduce, though not completely prevent, the rate of infection and low white cell count. Most adjuvant breast cancer chemotherapy regimens do not routinely require growth factor support except for those associated with a high incidence of bone marrow suppression and infection.

Neoadjuvant Chemotherapy in EBC

Neoadjuvant treatment was initially given to patients with locally advanced, inoperable breast cancers but its use has been extended to patients with early, operable breast cancer to reduce the size of the tumour and the need for mastectomy and to allow for breast conserving surgery^[199]. Neoadjuvant therapy may also potentially treat occult micro-metastatic disease. The clinical radiological and pathological response of the tumour can be assessed over time – a quality that is useful in a research setting as it may lead to the identification of short-term surrogate markers indicative of long-term disease outcomes.

Neoadjuvant chemotherapy achieves clinical regression of tumours in 70-80% of patients, suggesting that early breast cancers are more chemo-sensitive than metastatic disease. Complete pathological response, which equates to disappearance of the tumour from the breast, is seen in up to 20% of patients, particularly in oestrogen receptor negative patients in whom it is a predictor of better long-term outcome^[200]. Nodal disease is cleared in approximately 35% of patients overall, but may be as high as 50% in those with TNBC and less than 10% in those with ER positive disease. The addition of Trastuzumab may lead to clearance in up to 70% of HER2-positive patients^[201].

Trials have shown equivalence in survival between neoadjuvant and adjuvant regimens using similar agents^[202]. There may however be an improvement in long-term outcome with neoadjuvant therapy^[203], though this may not be the case for patients also treated with radiotherapy but not surgery^[202].

Neoadjuvant Regimens

Neoadjuvant chemotherapy regimens are generally the same as those used in the adjuvant setting. There is some evidence to suggest that the type of chemotherapy given may affect the number of complete pathological responses seen, although the difference between regimens is not always apparent^[204, 205].

Progressive disease, in which the tumour continues to enlarge despite treatment, is a rare phenomenon during neoadjuvant chemotherapy and is an indication for second line chemotherapy or surgery. Around 50% of patients will have adequate tumour regression to avoid mastectomy^[206].

Chemotherapy in LABC

Chemotherapy prior to loco-regional treatment is standard management in LABC, with the exception of older patients with ER+ disease, where endocrine therapy may be more appropriate. Most trials in this disease setting have however been of relatively poor quality.

Anthracycline based schedules are the mainstay of treatment but Taxanes are increasingly used subsequently^[172] and should be considered in patients with advanced disease including intractable residual chest wall disease^[207] as should Capecitabine or Vinorelbine^[208, 209].

High rates of clinical response and complete pathological response have been noted with the addition of cyclophosphamide to Adriamycin and Docetaxel in some studies^[210].

Chemotherapy for Metastatic Disease

In the presence of metastases, a balance must be achieved between the beneficial response and the side effects. Both survival and palliation are improved by active regimens. Response rates of up to 60% have been achieved, with a median time to relapse of six to ten months. The regimens used are similar to those in both the neoadjuvant and adjuvant settings. Epirubicin is often preferred to other Anthracyclines because it has less cardiac toxicity^[211].

1.7.2 Biological Therapies

Adjuvant Trastuzumab (Herceptin)

The transmembrane growth factor receptor HER2/neu (also known as c-erbB-2) is expressed in around 20% of breast cancers and is associated with more aggressive disease and a poor prognosis. Trastuzumab (Herceptin) is a monoclonal antibody against the external domain of the receptor, with clinical effect against tumours expressing HER2^[212]. Trials have shown that it reduces recurrence and improves disease free survival by 50% in patients with early breast cancer when given after or during chemotherapy^[213-216]. There is possible survival benefit to concurrent treatment^[217].

The only significant side effect appears to be the risk of cardiac dysfunction in a small number of cases (<2%). This may be related to concomitant treatment with the Anthracyclines and so, the two are never given together. Cardiac function is monitored through treatment.

Importantly, Trastuzumab has also been shown to improve survival in patients with metastatic disease when given in combination with the Taxanes and achieves higher rates of tumour regression with other agents such as Vinorelbine^[218].

Duration of Treatment

The optimal duration of adjuvant treatment with Trastuzumab is uncertain. Most major trials have incorporated one year, but this is on an empirical basis and by analogy with endocrine therapy. A number of trials currently aim to compare this regimen with shorter treatment periods and recent data suggest that two years treatment is no better than one ^[219].

Neoadjuvant Trastuzumab (Herceptin)

A combination of Trastuzumab and chemotherapy given neoadjuvantly, achieves superior response rates compared to chemotherapy alone, with improvements in pathological complete response of 40-50% and should now be considered standard care for HER2-positive patients. Again, there are no clear indications to date as to the optimal duration of therapy ^[218].

Other Anti-HER2 Therapies

Lapatinib, given orally, is an anti-HER2 drug that acts as a tyrosine kinase inhibitor, stopping phosphorylation and thus activation of HER2. Its role in EBC instead of, or in combination with Trastuzumab, is currently being investigated (www.allttrials.com). It is occasionally used in patients with metastatic disease who have relapsed on Trastuzumab. Its main toxic side effect is diarrhoea.

Pertuzumab is a human monoclonal antibody with affinity for the HER 1, 2 and 3 receptors. It has been shown to work best when given in combination with Trastuzumab without increasing toxicity ^[220]. The use of Lapatinib and Pertuzumab in the neoadjuvant setting in HER2 positive disease, may significantly improve pathological tumour response ^[220]. The long-term survival outcomes from adjuvant trials are still awaited.

Breast Cancer Endocrinology

1.8 Endocrine Dependence and the Breast

This thesis specifically relates to endocrine therapy in post-menopausal women, which will be the subject of the remainder of the introduction.

The relationship between oestrogen and the breast is fundamental to understanding endocrine therapy in breast cancer. Oestrogens have been implicated in the development of breast cancer since their discovery in the 1920s, though Beatson employed oophorectomy to induce regression of breast cancers in the 19th century ^[221]. A considerable body of knowledge and evidence now exists regarding the role of oestrogen in human breast development and tumourigenesis ^[116, 222].

1.8.1 Breast Development and Tumourigenesis

The breast begins to develop in early puberty when primitive ductal structures enlarge and branch. After menarche, branching of the ductal system becomes more complex and lobular structures form at the ends of the terminal ducts to produce terminal duct lobular units (TDLUs). These become increasingly complex with successive menstrual cycles. During early pregnancy, a further burst of activity takes place during which the ductal trees expand further and the number of ductules within the TDLUs greatly increases. These ductules differentiate to synthesise and secrete milk during late pregnancy and subsequent lactation ^[223].

A continuous layer of luminal epithelial cells lines the breast ducts. A layer of myoepithelial cells and then a basement membrane, in turn, surrounds these. The TDLU system as a whole is then surrounded by delimiting fibroblasts and embedded in a specialised intralobular stroma.

Histological studies have shown that most breast tumours appear to be derived from the TDLUs and, have morphological characteristics in keeping with luminal epithelial cells. Moreover, the majority of human breast cancers retain the biochemical features of luminal cells ^[224] including receptors for oestrogen and progesterone that, in the normal breast, are expressed only in the luminal epithelial compartment ^[222]. Luminal epithelial cells must therefore be regarded as the primary source of malignant transformation and subsequent tumour formation.

The process of breast tumourigenesis is thought to result from the accumulation of multiple genetic changes corresponding clinically to the evolution from normal breast epithelium through benign proliferative lesions, atypical proliferative lesions, carcinoma in situ and finally, invasive tumours ^[225]. Although the exact mechanism by which this occurs remains unclear, both oestrogen and the oestrogen receptor are implicated in this process.

In the absence of normal ovarian function, there is complete failure of breast development, such that oestradiol replacement therapy is necessary to induce breast development ^[226]. Furthermore excess exposure to oestrogen due to early menarche, late menopause or a late first pregnancy, increases breast cancer risk, as does the exogenous use of hormones ^[227]. The obligate role of oestrogen has also been demonstrated in mice studies where the gene coding for the ER protein has been rendered non functional ^[228]. The mammary glands in these ER 'knock-out' mice fail to develop, even with exogenous oestrogen treatment, and are resistant to malignant transformation following transduction with oncogenes (genes that promote cell transformation).

Whilst there is far less evidence for the role of progesterone in human breast development, studies on PR knock-out mice suggest that whilst oestrogen stimulates ductal elongation and PR expression, progesterone may stimulate lobulo-alveolar development ^[229]. Human studies in this area are limited due to practical difficulties of studying breast development at a young age, it is nonetheless assumed that progesterone plays a similar role in human breast development as well as stimulating TDLU formation and expansion during puberty and pregnancy. With respect to tumourigenesis, there are data to suggest that progestagens, taken in the form of HRT increase the risk of post-menopausal breast cancer to a greater extent than the use of oestrogen only HRT ^[39, 230].

1.8.2 Oestrogen Production

The Oestrogens are a family of related steroid hormones that stimulate the development and maintenance of female secondary sexual characteristics and reproduction. They occur as the end products of a sequence of steroid transformations ^[231]. The most prevalent forms of circulating oestrogen are Oestradiol and Oestrone (the highest levels in the circulation are of oestrone sulphate) but their primary site of production differs according to menopausal status. In pre-menopausal women, the ovaries are the primary source of oestrogen production, with smaller amounts produced by the adrenal cortex and other organs. In postmenopausal women however, with declining ovarian function, oestrogen production occurs predominantly by the peripheral conversion of ovarian and adrenal androgens to oestrogen in peripheral fat, muscle and bone as well as within breast tumours ^[232]. This conversion is a function of the aromatase enzyme, a product of the CYP-19 gene on chromosome 15.

1.8.3 The Oestrogen and Progesterone Receptors

Regardless of their source, most of the major effects of the steroid hormones on the breast are mediated by their respective protein receptors, which function as transcription factors within cell nuclei to activate oestrogen dependent genes. Transcription and translation constitute the two crucial steps by which functional proteins are made within cells from their encoding genes. Transcription is the process whereby messenger RNA (mRNA) is produced from DNA in the cell nucleus, by the action of the enzyme RNA polymerase. The mRNA migrates to the cell's cytoplasm where a process of translation by ribosomes, converts it into functional proteins. A transcription factor is a protein that binds to specific DNA sequences by means of one or more DNA-binding domains (DBDs), to promote or inhibit DNA transcription of the genes that they regulate.

Two forms of oestrogen receptor ($ER\alpha$ and $ER\beta$) have been recognised and are implicated in breast development and the pathogenesis of breast cancer. The ESR 1 and ESR 2 genes on chromosome 6 and 14 encode these respectively. Both receptors have a modular structure which typifies the steroid receptor family and which includes domains that mediate binding to ligands (steroids) and to intracellular DNA and are therefore known as ligand dependent nuclear transcription factors ^[233]. There are six functional domains, designated A-F (Figure 1.8A). The Amino-terminal domain, AF1, exhibits hormone independent activation function and shows the highest variability among all steroid receptors ^[234, 235]. The middle or C domain is the DNA binding domain (DBD) ^[236], whilst the adjacent D domain or 'hinge region', is implicated in co-regulatory protein binding. The carboxy-terminal domains E and F contain the ligand (steroid) binding domain (LBD). This region is also implicated in modulating the agonist activity of anti-oestrogens to be discussed later and binding to further co-regulator proteins ^[237].

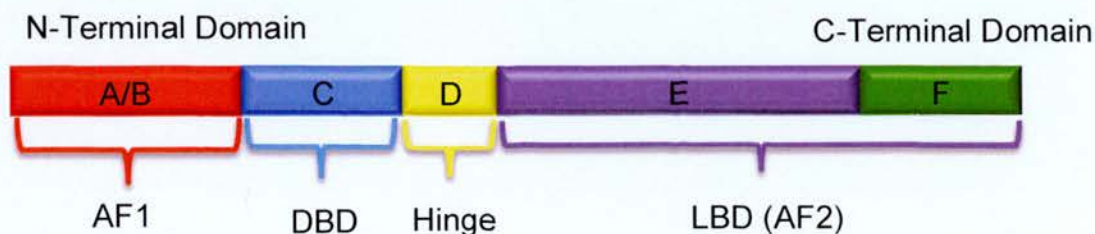


Figure 1.8A: Functional Domains of the Oestrogen Receptor

Whilst both receptor types are homologous in their binding domains, the ER β gene is smaller and encodes a shorter protein ^[234, 238]. It is likely that ER α is the key mediator of oestrogen in the normal mammary gland and that ER β mediates some of the non-classical effects of the oestrogens and may negatively modulate the activity of ER α but its activity has yet to be further defined ^[239].

Progesterone also has two receptors, PRA and PRB. These are again members of the steroid hormone receptor family but both are transcribed from the same gene. Whilst PRB is longer in its amino acid sequence, the two are otherwise identical ^[240]. It appears that PRB is the major activator of gene transcription and that PRA is a repressor of PRB activity, however it is possible that both mediate differing gene transcription ^[222].

1.8.4 The Biology of Oestrogen Receptor Signalling

The steroid hormones are lipophilic and they enter cells primarily by diffusing through the plasma and nuclear membranes. Once in the cell they bind their respective receptors with high affinity and specificity. In the absence of this highly specific binding, the receptors are inactive and exist in a monomeric form bound to heat shock proteins such that they exert no influence on a cell's DNA expression. On binding oestrogen however the oestrogen receptor dissociates from the heat shock protein and undergoes a process of homo-dimerisation and protein phosphorylation on a number of specific residues. This leads to a conformational change in the shape of the AF2 region. Consequently, ER dimers translocate to the nucleus and bind to specific DNA sites, known as oestrogen response elements (EREs). These consist of palindromic nucleotide sequences and they lie upstream of oestrogen responsive genes. The ER dimers simultaneously associate themselves with receptor co-activator proteins and together these associations are responsible for the activation of a large number of oestrogen dependent genes, the transcription of their messenger RNA (mRNA) and the subsequent translation of this mRNA into proteins which invariably impact on the properties of the cell, the cell cycle and that of adjacent cells (Figure 1.8B) ^[241].

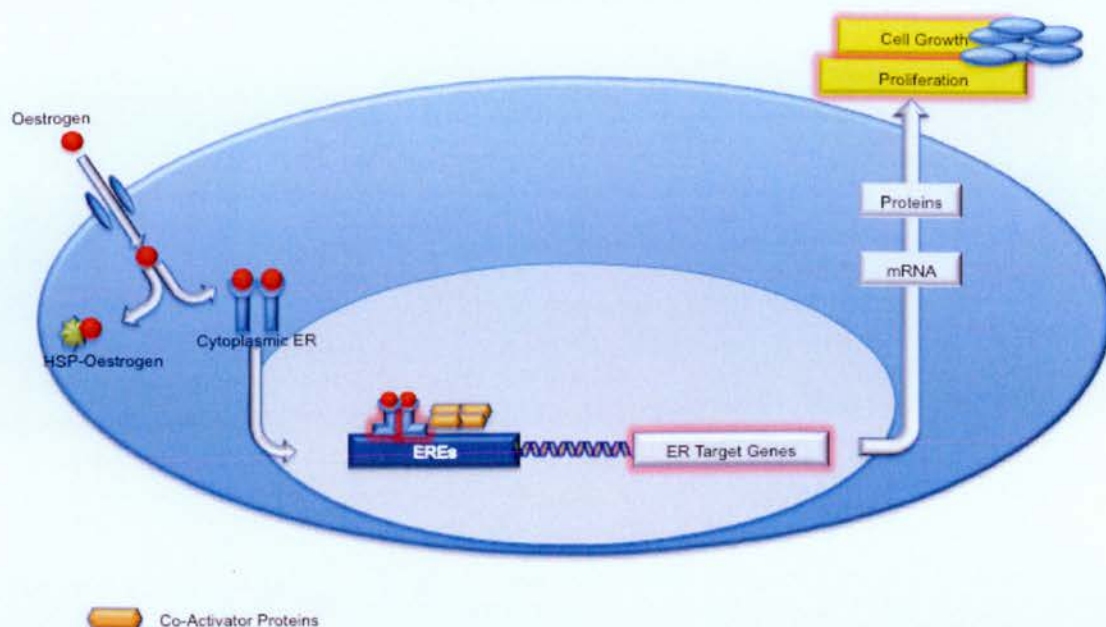


Figure 1.8B: Classical Oestrogen Signalling Pathway.

Since the now widely accepted description of this 'classical oestrogen signalling pathway', additional pathways have been described (Figure 1.8C). Firstly, it has been suggested that some oestrogen receptors may be located on the cell membrane; upon binding with oestrogen these can activate an intracellular signal transduction chain ending at two further transcription factors, *jun* and *fos*, which are capable of recognising alternative receptor binding sites and which may perpetuate cell proliferation signalling. Secondly and most recently it has been suggested that oestrogen receptors within the plasma membrane may interact with growth factor receptor signalling, leading to activation of the *ERK1/2* gene and thus influencing the nearby activity of dividing cells via protein kinase signalling pathways. Interestingly in cell lines it has been noted that the growth factors *TGF α* and *IGF1* are incapable of promoting proliferation in the absence of oestrogen, so clearly these mechanisms are complex and cross talk between various signalling pathways may occur ^[241].

Whilst not experimentally proven, the consensus supports the idea that oestrogens act via one or more of these pathways to trigger uncontrolled cell growth and proliferation by regulating the expression of key oestrogen dependent cell cycle genes. In doing so they either directly or indirectly promote the replication of defective cells or, increase the likelihood of spontaneous mutations and their replication, with resultant tumourigenesis. They may even promote invasion and metastases through the production of proteases that degrade the extracellular matrix. These along with other issues have yet to be elucidated and will be discussed further in a later section.

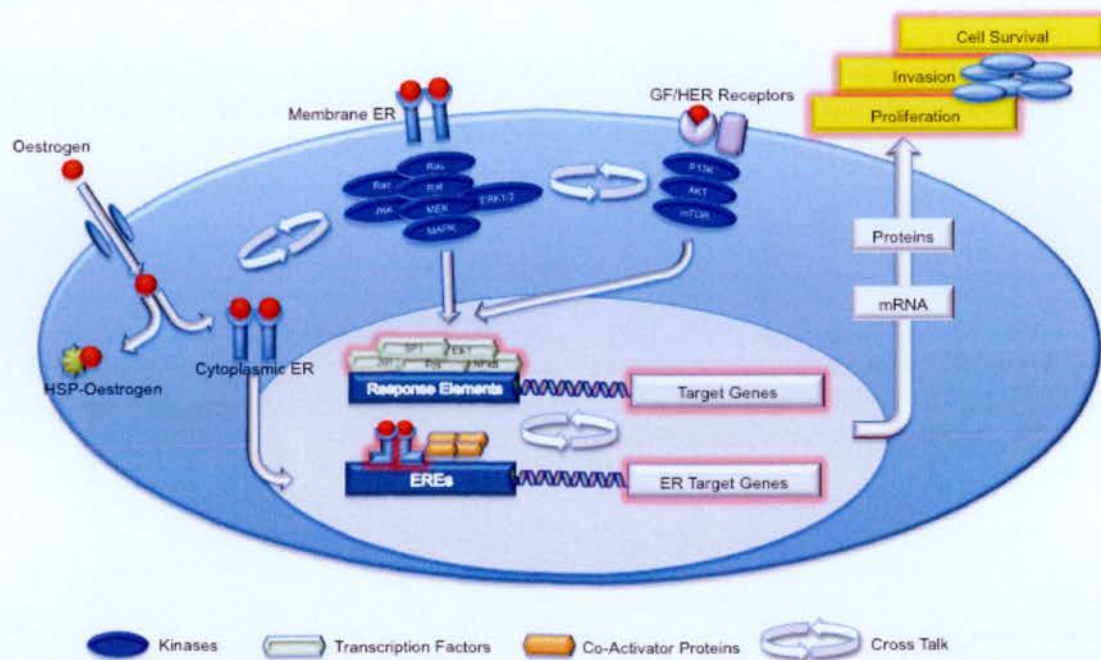


Figure 1.8C: Alternative Oestrogen Signalling Pathways.

1.8.5 Quantification of ER

Measurement of ER status is generally made by immunohistochemistry (IHC) using one of a number of monoclonal antibodies. Its interpretation and scoring however remains contentious, with a variety of scoring systems and end-points used. This must be taken into account when considering the evidence base for treatment protocols. There are data to indicate that ER expression in as few as 1% of cells of a particular tumour may have consequences for clinical outcome and that this should therefore be the threshold set for positivity ^[154]. Modern ER analysis however suggests that there is generally a bimodal distribution of ER values in which almost all tumours are found to be either strongly ER positive or negative, with only a small proportion showing intermediate values ^[242].

The most commonly applied system in UK practice is the Allred scoring system. This reflects the percentage of tumour cell nuclei staining positively and the intensity of staining:

- | | |
|------------------------------|----------------------|
| 0 = No staining | 0 = No staining |
| 1 = Staining in <1% of cells | 1 = Weak |
| 2 = 1 to 10% | 2 = Intermediate |
| 3 = 10 to 33% | 3 = Strong |
| 4 = 33 to 67% | |
| 5 = 67% to 100% | |
| | Total Score = 0, 2-8 |

Alternatively, the Histoscore multiplies the intensity of staining (absent = 0, weak = 1, moderate = 2, strong = 3) by the percentage of positively staining cells to give a total score of up to 300. Tumours with an Allred score of 6-8 are considered to show strong ER positivity and those with a score of 2-5 are weakly positive.

1.9 Endocrine Therapy

The dependence of many breast cancers on oestrogen for their development and continued growth, has naturally led to a range of therapeutic interventions that aim to reduce tumour growth either by reducing oestrogen production or by antagonism of its action at the oestrogen receptor ^[243]. Virtually all breast cancers that respond to endocrine treatment express significant levels of ER α . However, the expression of ER is age dependent: whilst only 60% of patients under 35 years of age will show ER positivity, 80% of women diagnosed over 65 years have ER positive disease ^[241]. Indeed, the mechanistic approach by which oestrogen deprivation therapies are applied depends largely on the primary site of oestrogen production and thus on the age and menopausal status of the patient.

1.9.1 Limiting Oestrogen Production

Ovarian Suppression

In premenopausal women, the ovaries are the primary source of oestrogen production. Bilateral oophorectomy was first performed by Glasgow Surgeon George Beatson over one hundred years ago and is still an option for reducing oestrogen production in premenopausal women. Whilst effective and technically simple, this procedure is however irreversible and is therefore reserved for a select few ^[244]. The same is true of the less effective and more risky alternative of ovarian ablation by radiotherapy ^[245, 246]. These have now largely been superseded by gonadotropin releasing hormone agonists (GnRH agonists), which produce an effective and reversible medical oophorectomy and currently provide the only available method of reversible ovarian suppression ^[247].

Several GnRH agonists have been developed including Goserelin, Buserelin, Leuprorelin and Triptorelin. Whilst no direct comparative studies have been published, all show efficacy as mono-therapy in advanced breast cancer ^[247, 248]. Goserelin (Zoladex) is the drug of choice currently recommended by the FDA for ovarian suppression in premenopausal women with ER positive breast cancer.

It is noteworthy that whilst premenopausal women often experience ovarian suppression secondary to cytotoxic chemotherapy, this does not constitute adequate endocrine therapy in many, since menses will resume following chemotherapy in most young women and in 10-20% of women over the age of 40 years [249].

Inhibition of Aromatase

In post-menopausal women, where ovarian steroid biosynthesis is reduced markedly and oestrogen levels in the circulation fall dramatically, inhibition of the aromatase enzyme, which catalyses the last step in peripheral oestrogen biosynthesis, has long been a target for anti-oestrogen therapy (Figure 1.9A). Aromatase converts androgens to oestrogens by creating an aromatic ring in the steroid molecule and is present in many tissues including adipose, muscle, liver, brain, bone and most importantly, within tumours themselves [250-254].

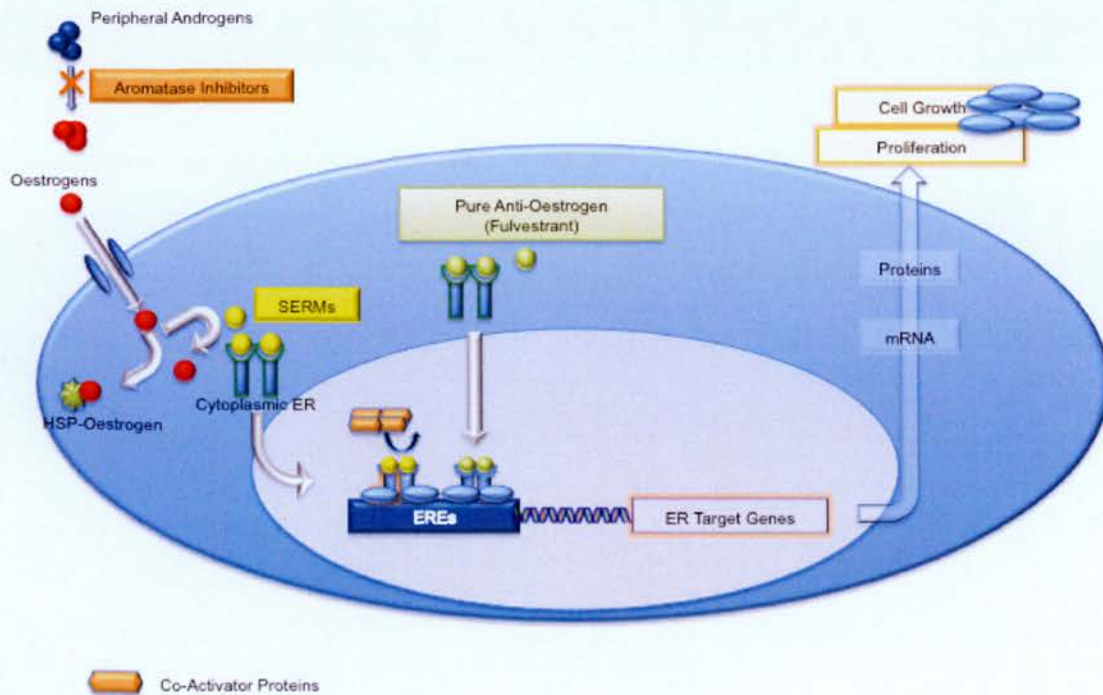


Figure 1.9A: Mechanisms of Action of the Endocrine Therapies.

Three generations of aromatase inhibitors (AIs) have been developed. Whilst the first generation Aminoglutethimide induced inhibition of the enzyme, it was not particularly potent and lacked specificity, often producing side effects unrelated to oestrogen deprivation including cortisol suppression ^[255]. A second generation was developed that showed greater selectivity but their use was limited due to poor bioavailability ^[256]. However, the current third generation agents Letrozole, Anastrozole and Exemestane possess remarkable specificity and potency. They may be subdivided into steroidal (Type I) and non-steroidal (Type II) inhibitors, each exhibiting different interactions with the aromatase enzyme. Steroidal AIs (Exemestane) bind non-covalently and reversibly to aromatase, whereas the non-steroidal AIs (Letrozole and Anastrozole) bind covalently and irreversibly ^[231].

All three of the current available AIs switch off peripheral oestrogen biosynthesis almost completely and are highly effective in the treatment of postmenopausal breast cancer ^[257, 258]. Letrozole is the most potent amongst these, inhibiting peripheral aromatisation by almost 99% ^[259-262]. Exemestane and Anastrozole inhibit aromatase by 97% and 98% respectively, although the clinical significance of this difference is not clear ^[263]. The steroidal AI Exemestane may exert androgenic effects, but the clinical relevance of this has yet to be determined. Whilst in pre-menopausal women the AI's can be used in combination with oophorectomy or the GnRH agonists, they cannot be used in isolation since they do not effectively inhibit *ovarian* oestrogen production.

The AIs are well tolerated though not without side effects, and are rapidly becoming the endocrine therapy of choice in postmenopausal women with ER positive disease.

1.9.2 Blocking Oestrogen Function

The 'Anti-Oestrogens'

The 'anti-oestrogens' as their name suggests do not block oestrogen biosynthesis but rather antagonise its action by competing for binding sites on the oestrogen receptors ^[264]. These agents are considered in two groups depending on their interaction with the receptor.

The Selective Oestrogen Receptor Modulators (SERMs)

The SERMs, which include Tamoxifen, Raloxifene and Toremifene, compete with oestrogen for binding at the ER ^[264]. Indeed, when given at the usual daily dose of 20mg, Tamoxifen occupies at least 99% of ER binding sites ^[241]. The effects of Tamoxifen binding are similar to that of oestrogen, with dimerisation causing a conformational change in the shape of the ER. However rather than activating the AF2 module, the Tamoxifen bound receptor reduces or eliminates the activity of AF2. Consequently, the association of the ER with co-activator proteins does not take place; co-repressor proteins are not displaced and the transcription of AF2-oestrogen-dependent genes is greatly reduced. They can be used in both pre and post-menopausal breast cancer patients (Figure 1.9A).

Whilst Tamoxifen is an effective antagonist in the presence of high levels of oestrogen, it can also exhibit some agonist effects on the breast and on other oestrogen-dependent target tissues in the absence of oestrogen ^[265]. This is probably mediated by the activation of the AF1 ER module. The activity of further molecules such as the steroid receptor co-activator A1B1 and the type 1 growth factor receptor HER2 may also influence the degree of agonism. The effects of enhanced signalling through the ER pathway are evident in the side effect profile of Tamoxifen. The anti-oestrogenic effects have a negative impact on the bone stock of premenopausal women in particular who incur a risk of osteoporosis following treatment ^[266]. There is however a positive effect on bone in postmenopausal women through agonist effects. Tamoxifen also exerts positive effects on blood lipid profiles and is associated with a small increased risk of venous thromboembolism in post-menopausal women. The proliferative effect of Tamoxifen on endometrium accounts for a small increase in the risk of developing endometrial carcinoma also only in postmenopausal women ^[267-269]. Raloxifene on the other hand, has no stimulating effect on the endometrium ^[264, 268].

Tamoxifen was established as the primary endocrine treatment for both early and advanced breast cancer and as an adjunct to surgery in both pre and post menopausal ER positive groups until fairly recently. The AIs with their superior efficacy and tolerability have now superseded Tamoxifen for postmenopausal women in many centres.

‘ The Pure Anti-Oestrogen ’

The second group of anti-oestrogens includes only one drug, Fulvestrant, also known as a selective oestrogen receptor down-regulator (SERD). This also competes with oestrogen for binding to its receptor but unlike the SERMs does not demonstrate any agonist activity, thereby avoiding some of their unwanted side effects ^[270] (Figure 1.9A). Furthermore, by destabilising the receptor it may actually reduce levels of ER. It has been tested primarily on metastatic or locally advanced breast cancers and its efficacy in this group appears to be similar to that of the AIs ^[271, 272].

1.9.3 Adjuvant Endocrine Therapy

It has become clear that adjuvant endocrine therapy is generally only effective in patients with ER positive disease with a clear correlation between the degree of ER expression and the efficacy of therapy ^[154].

Until recently Tamoxifen was the agent of choice in the adjuvant setting for both pre and post-menopausal women and additionally has been shown to reduce the risk of contralateral breast cancer by 40-50% when given for five years, although it may be less effective against HER2 positive tumours ^[273, 274].

Postmenopausal Adjuvant Endocrine Therapy

The advent of the third generation selective AIs has provoked major changes in adjuvant therapy for post-menopausal women with early breast cancer. These AIs reduce the risk of contralateral breast cancer by a further 40-50% when compared to Tamoxifen and have been shown to improve disease free and metastatic free survival, to a greater degree than Tamoxifen ^[275, 276].

Data from the first trial to compare Tamoxifen with Anastrozole alone or in combination (The ATAC trial), demonstrated a significant disease free survival benefit with Anastrozole alone ^[277]. This has remained the case at the most recent 10-year analysis of the data. Subsequent trials have also shown beneficial effects on disease free survival and overall survival when Letrozole is used as first line treatment over Tamoxifen ^[278]. The on-going FACE trial is comparing the efficacy of Letrozole and Anastrozole in over 4000 women with node positive disease and should demonstrate which is most clinically effective ^[279]. Exemestane is not currently licenced for use as first line therapy in the adjuvant setting and results have shown equivalence with Anastrozole ^[280].

Data from a series of early trials led to the trend of 'sequencing' endocrine therapy, whereby treatment was initiated with Tamoxifen and later 'switched' to an AI after two to three years. It was hoped that this would improve the outcome compared with ongoing Tamoxifen or AI treatment [276, 281]. However, two large randomised trials have since shown no benefit to this approach in women eligible for AI treatment [278, 282]. The current standard of care in this subset of patients is therefore to commence an AI and to continue this for five years, unless contraindicated due to the side effect profiles of AIs. Many women however will still be commenced on Tamoxifen and substituted at 5 years to extended treatment.

Extended Adjuvant Therapy

Further evidence in favour of the AI's comes from the Canadian led MA17 trial. This revealed that Letrozole, when given following five years of Tamoxifen as extended adjuvant therapy, reduces risks of local recurrence in ER positive, node negative and positive patients and, produces a significant survival benefit in patients with node positive disease [283, 284]. Two further trials have shown similar benefits with the other AI's [285, 286]. Further follow up data from MA17 suggest that the benefits of Letrozole increase with time, certainly up to four years of extended therapy. A further trial is currently investigating whether treatment for up to ten years is of continuing benefit. For some women, indefinite treatment may be appropriate given the long natural history of the disease. Extended adjuvant therapy with Letrozole in ER positive postmenopausal cancer has therefore become the standard of care for all but low risk women.

Premenopausal Adjuvant Endocrine Therapy

In premenopausal women with hormone sensitive disease, the options include Tamoxifen alone or in combination with ovarian suppression or ablation, most commonly using a GnRH analogue such as Goserelin. The combination of Tamoxifen and Goserelin may improve survival in women with ER positive disease [273, 287] and, is more effective than CMF chemotherapy [288-290]. However, endocrine therapy alone has never been compared with Anthracycline or Taxane based regimens that are now seen as standard [178]. In advanced breast cancer the addition of Tamoxifen to an LHRH agonist improves response rate and overall survival [290]

It is not yet clear whether there is a substantial advantage or indeed a disadvantage in using an AI instead of Tamoxifen in combination with Goserelin.

Combinations of Adjuvant Chemotherapy and Hormonal Therapy

The use of chemotherapy and Tamoxifen in combination is more effective than either alone, for women with high risk ER positive cancer ^[291]. Efficacy is greater when Tamoxifen is given after chemotherapy, rather than concurrently. No data exist on whether the same is true of ovarian ablation or AI's, but it would seem likely.

A recent study suggested that for those patients under the age of 40 years, who remain amenorrhoeic following chemotherapy, prognosis is better ^[292]. This confirms that ovarian suppression is beneficial after chemotherapy in patients with ER positive cancer whose menses persist.

1.9.4 Neoadjuvant Endocrine Therapy

The use of endocrine therapy has become increasingly popular in the neoadjuvant setting. Such treatment was initially reserved for postmenopausal patients with large ER positive tumours who had either locally advanced operable or inoperable breast cancers. Its use has now been extended to patients with earlier operable breast cancer. The aim in these groups is to downstage disease to facilitate a surgical approach in otherwise inoperable tumours or to minimise the extent of surgery in operable tumours ^[293].

Premenopausal Neoadjuvant Endocrine Therapy

Very few neoadjuvant endocrine studies have been conducted in pre-menopausal women. Those reported have examined the use of Tamoxifen versus an AI in women taking a GnRH agonist for 3-4 months, in small series. Whilst they have shown some clinical advantage in patients taking Letrozole, no differences were reported in survival between responders and non-responders ^[294, 295].

Postmenopausal Neoadjuvant Endocrine Therapy

The majority of neoadjuvant endocrine studies have been carried out in the postmenopausal population. Indeed, a number of relatively small but robust randomised neoadjuvant trials have proven the benefits of AI's, particularly Letrozole over Tamoxifen as first line treatment for post-menopausal women with large operable ER positive tumours. Letrozole, given for three to four months pre-operatively, was shown to be superior to Tamoxifen in terms of clinical response (55% versus 36%) and rates of BCS (45% versus 36%), in women who would otherwise have required mastectomy ^[296].

The results of the IMPACT trial, which compared 3 months of neoadjuvant Anastrozole or Tamoxifen alone or in combination, showed no difference in cancer response rates. However, in the subpopulation of patients with large cancers initially warranting mastectomy, Anastrozole, like Letrozole, was significantly more effective than Tamoxifen alone or in combination in achieving BCS (46% vs. 22% vs. 26%)^[297].

A second study (PROACT) compared three months of Anastrozole with Tamoxifen and again showed similar response rates with the two drugs but a higher rate of BCS with Anastrozole^[298]. When the results of PROACT and IMPACT were combined, a significantly greater response rate was seen with the AI's in tumours that were locally advanced or that required mastectomy at presentation.

These results have since been backed up in a head to head trial and further in a meta analysis of neoadjuvant AIs versus Tamoxifen, in which better clinical response rates, objective tumour shrinkage on ultrasound scanning and BCS rates were evident with the AIs^[299]^[300]. Most recently the American College of Surgeons Oncology Group (ACOSOG) carried out a multicentre trial in which 377 patients were randomised to receive Letrozole, Anastrozole or Exemestane for 16 weeks^[301]. Clinical response rates were high in all groups (74.8% with Letrozole, 69.1% with Anastrozole and 62.95% with Exemestane). Disease progression rates were low in all groups but lowest with Letrozole (4.7%) and the overall rate of BCS between groups was 82%. These results are similar to those found in previous studies and suggest that all three agents are suitable for use in the neoadjuvant setting. In the UK, Letrozole is most commonly used in clinical practice, as it is currently the only agent licenced for this indication.

It was initially felt that neoadjuvant endocrine therapy might not be appropriate for all types of ER positive cancer, in particular ILC, which may be difficult to monitor radiologically due to its often occult nature. However, data now show that these also respond well^[302]. Early data from an Edinburgh series, which enrolled 61 patients with 63 ILCs as part of a prospective neoadjuvant Letrozole audit, have shown encouraging results. The mean reduction in tumour volume measured clinically at 3 months was 66% (61% ultrasonically and 54% mammographically). Of the enrolled patients, 81% underwent BCS. With the exception of two patients, the rest who have continued on Letrozole mono-therapy have remained controlled at a median of 2.8 years.

Duration of Neoadjuvant Treatment

The standard duration of neoadjuvant treatment to date has been three to four months. This is based largely on experience with neoadjuvant chemotherapy, without due consideration of how long endocrine therapy might be effective for and what the optimal duration should be. Recent studies with longer durations of chemotherapy have suggested that prolonging treatment increases response rates^[303].

Initial endocrine studies examining patients treated with neoadjuvant Tamoxifen suggested that a lack of tumour response by 3 months, indicates the need to consider alternative treatment^[304]. This idea has been recently challenged. One prospective study published from Edinburgh included 184 women with large operable or locally advanced ER positive breast cancers who were treated with neoadjuvant Letrozole. Of those treated 119 patients underwent surgery at 3 months and, 63 patients continued on Letrozole beyond 3 months. Amongst the latter group, a sustained response to Letrozole with reductions in tumour volume was notable. In fact, prolonged treatment resulted in an increase in the number of patients treatable with BCS from 60% at 3 months to 72 % at 3 years. Three quarters of patients maintained disease control at 3 years on Letrozole^[305].

A further multicentre study similarly found that BCS rates improved with time of neoadjuvant treatment (25% at 5.8 months, 50% at 8.2 months and 75% at 12 months)^[306]. Whilst an initial, rapid response was seen in the first 4 months of treatment, a slower and sustained response was seen thereafter, with a median time to BCS of 8.2 months. The results of both of these studies suggest that the conventional treatment period of 3 to 4 months of neoadjuvant Letrozole may need to be revised.

Sequencing the AI's

The question of whether neoadjuvant sequencing of different classes of AI is of clinical benefit remains the subject of much debate. Results from a recent small head-to-head trial comparing steroidal and non-steroidal AIs suggest no meaningful clinical differences in efficacy or adverse events between the two classes and therefore no conceivable benefit from sequencing^[307]. Similarly, no benefit has been found to date from the addition of an oestrogen receptor blocker in combination with a non-steroidal AI^[308].

Histological Response To Neoadjuvant Endocrine Therapy

One of the main features seen in response to Letrozole therapy is central scar formation within the tumour mass, associated with a reduction in tumour volume. This discovery was one of the first indications that the breast tissue stroma may be an important factor in the development, proliferation and response of breast tumours ^[309]. This was one of the indications to include the assessment of stromal signatures in the present study. The process of signature selection is discussed in more detail later.

In contrast, scattered and diffuse cellular patterns are seen in response to chemotherapy ^[310].

Neoadjuvant Endocrine Therapy versus Chemotherapy

As with adjuvant chemotherapy, evidence suggests that patients with hormone sensitive disease respond better to neoadjuvant endocrine therapy than they do to neoadjuvant chemotherapy ^[311]. One large retrospective study of 1731 patients receiving primary chemotherapy showed significantly lower clinical response rates in its ER positive subset ^[312]. Until recently, there was only one randomised trial of neoadjuvant chemotherapy (doxorubicin and paclitaxel) versus the AIs (Anastrozole and Exemestane) ^[169]. Amongst the 121 women with EBC treated for 3 months, response rates between the chemotherapy and endocrine treated groups were similar, however there was a trend for increasing BCS rates in favour of endocrine therapy with no significant differences in recurrence rates at 34 months. Moreover the endocrine treatments were better tolerated and with fewer serious side effects. Subsequent studies also suggest that endocrine therapy is a safe alternative to chemotherapy in postmenopausal women with hormone sensitive cancers, with similar response rates but less toxicity ^[313].

Response and HER 2 Status

The evidence for whether response rates to AIs and Tamoxifen therapy in ER positive tumours differs in HER2-positive patients remains equivocal. The P024 study comparing neoadjuvant Letrozole with Tamoxifen found that clinical response rates amongst HER2 positive tumours were significantly higher for groups treated with the AI ^[314, 315]. Other smaller studies have however, failed to find a significant difference and the evidence remains unclear ^[316, 317]. Overall however, HER2 positivity may predispose to resistance to endocrine therapy.

1.10 Assessment of Response to Neoadjuvant Endocrine Therapy

There are no current recommendations regarding the method or frequency of monitoring response to neoadjuvant endocrine therapy ^[318]. Most studies use clinical response as an end point. Tumour shrinkage and time to disease progression (progression free survival) are reasonably assumed to be associated with overall survival ^[319], though the association is not strong.

1.10.1 Clinical Response

Clinical response to therapy may be monitored by a change in tumour size using the following modalities (see materials and methods for details):

- (i) Calliper measurements
- (ii) Ultrasound scanning
- (iii) Mammography

Such assessments can be made at various time points although 0, 6 and 12 weeks have been recommended ^[320].

Clinical Response Criteria

Tumour endpoints, such as objective response and time to disease progression, are useful only if based on widely accepted and readily applied standard criteria. Two such systems have been published over the past two decades, the WHO classification published in 1981 ^[321] and subsequently a modified RECIST classification in 2000 ^[322-324]. The use of these is however aimed at the advanced disease setting. Since no system is specifically designed for use in the neoadjuvant setting, trial centres commonly employ locally defined response criteria or 'modify' the WHO criteria ^[320] (See materials and methods for response assessment in this study).

1.10.2 Pathological Response

In trials of neoadjuvant chemotherapy, pathological response is consistently shown to be a powerful determinant of long-term outcome. Patients with a complete pathological response (pCR) have a significantly better disease free and overall survival compared to those with no or incomplete response ^[325]. Neoadjuvant endocrine therapy also produces changes in tumour morphology and histology, including a decrease in cellularity and an increase in fibrosis, so called 'central scar formation' ^[309, 326, 327]. Such pathological responses have been reported in 60-80% of tumours after 3-4 months of treatment ^[328, 329]. Complete pathological responses to endocrine therapy are however rare ^[330]. Currently the criteria defining partial or incomplete pathological response are subjective and variable. Given the low frequency of complete pathological response, some form of standardised assessment of partial pathological response needs to be devised.

Endocrine treatment may also reduce histological grade; mitotic score being the most commonly affected grading feature ^[328].

Rates of clinical and pathological response are similar and significantly associated, but the concordance is not exact. For example, approximately one fifth of tumours either display a decrease in tumour volume without evidence of pathological changes, or appear unchanged in size but display a decrease in cellularity ^[166]. Because of this discordance, clinical response is most often used as the primary end point for response and pathological response reserved for secondary use or to confirm the clinical response.

1.10.3 Proliferative Response

The predictive value of Ki67 as a marker of proliferation and response to neoadjuvant endocrine therapy has been assessed in a number of studies. In the P024 study comparing Letrozole to Tamoxifen, there was a significantly greater reduction in proliferation with Letrozole (87% vs. 75%)^[331]. As part of the IMPACT trial Ki67 was assessed following 2 weeks of treatment with Anastrozole, Tamoxifen or a combination of the two and again at the time of surgery (12 weeks)^[332]. There was a significantly greater reduction in Ki67 in the Anastrozole only group at both 2 and 12 weeks. Moreover, both the percentage change in Ki67 expression and the absolute level of Ki67 expression at 2 weeks, predicted the superiority of Anastrozole over Tamoxifen in disease free survival in IMPACT's adjuvant equivalent, the ATAC trial^[277]. Whilst no significant relationship between Ki67 changes and clinical response was evident initially, the results of both the IMPACT and P024 trials at long-term follow up have since shown a statistically significant correlation between the level of Ki67 after 2 weeks of treatment and relapse free survival^[155, 332, 333]. This implies that there may be predictive value in the residual rate of proliferation after 2 weeks of treatment and that measuring Ki67 at that time point could be a valuable tool for predicting long-term outcome. This theory is being tested in the randomised POETIC study (Peri-Operative Endocrine Therapy for Individualised care) which is currently underway^[334].

1.10.4 Molecular Response

Clear evidence of molecular responses to aromatase inhibitors has also been elicited following treatment.

ER loss after neoadjuvant endocrine therapy was identified as an independent prognostic marker for relapse in both the IMPACT trial and the P024 trial. In the former, the levels of ER following two weeks of neoadjuvant therapy were significantly associated with recurrence free survival (66, 68). In the P024 trial, patients with post treatment ER negative tumours had poorer relapse free and overall survival than patients with tumours that retained their ER positive status after treatment (72).

The progesterone receptor (PR), which is regarded as a marker of oestrogenic activity, is also reduced in 70-80% of tumours during neoadjuvant endocrine therapy, with staining completely disappearing in about half of cases^[166, 330]. These changes are also encountered as early as 14 days following initiation of treatment and are consistent with the anti-oestrogenic effects of the AIs. However, loss of PR expression may occur independently of pathological and clinical response^[166, 315].

1.11 Molecular Mechanisms of Endocrine Resistance

Although there have been significant improvements in the efficacy of endocrine therapy since the introduction of the AI's, drug resistance remains a major obstacle to the effective clinical management of breast cancer. Approximately a third of patients with ER positive disease fail to respond to endocrine therapy in the neoadjuvant setting ^[335-337]. Primary resistance occurs when the tumour fails to respond to initial treatment. On the other hand, acquired resistance occurs after an initial response. The latter is of particular concern, as this subset of tumours appears to be more aggressive, difficult to treat and is therefore associated with poorer prognosis and survival ^[336, 338].

The oestrogen receptor and its signalling pathway play a central role in promoting cellular proliferation, survival and invasion of ER positive cancers, through both oestrogen-dependent and independent mechanisms ^[336, 339, 340]. A number of potential mechanisms of tumour resistance have been proposed and can be broadly considered as occurring within or outwith the tumour.

The extra-tumoural mechanisms include:

- (i) Patient factors such as poor compliance with treatment ^[341].
- (ii) Adverse drug metabolism, reducing the effective dose ^[342].
- (iii) Inaccurate assessment of the ER status ^[343].
- (iv) Stimulation of ER by dietary or other exogenous oestrogen sources ^[344, 345].
- (v) Reduced efficacy of AIs, due for example, to polymorphisms in the aromatase gene ^[346].

Intra-tumoural mechanisms are perhaps more complicated and are subject to greater research efforts, but can be considered to occur at any point in the pathway of oestrogen stimulated tumour growth (Figure 1.11A).

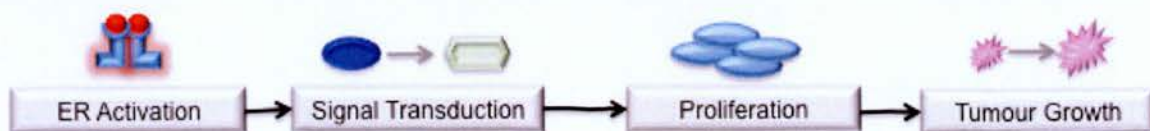


Figure 1.11A: Points in the ER Stimulated Pathway during which Resistance may Develop Within a Breast Tumour.

More specifically, resistance to treatment may develop because of:

- (i) Deregulation of various components of the ER pathway and ligand-independent activation of ER.
- (ii) Alterations in cell cycle signalling proteins ^[347].
- (iii) Activation of escape pathways that provide tumours with alternative cell proliferation pathways ^[348].
- (iv) Alteration of cell survival proteins and signalling pathways.

Each of these will be discussed below.

1.11.1 Deregulation of Components of the Classical ER Pathway

The oestrogen bound ER complex exerts its classical effects by translocating to the nucleus where it binds to oestrogen-response elements (ERE), altering the expression of numerous genes involved in tumour growth, as well as normal cell function ^[349]. The presence of various co-regulatory proteins such as NCoA1-4 can augment or inhibit this activity ^[350-352]. Similarly, ER can itself modulate other transcription factors, such as AP-1, NF-Kb and MYC to effect transcription pathways that are oestrogen independent (Figure 1.11B(i)) ^[340]. Furthermore, other growth factors receptor pathways can activate ER and its co-regulators, through mechanisms collectively known as 'cross talk' (Figure 1.11B(ii)) ^[353, 354].

In this way, tumour cells may lose their dependence on the oestrogen-ER complex and develop resistance to endocrine therapy. Alternatively, overexpression of the co-regulator proteins may increase sensitivity to low levels of oestrogen (Figure 1.11B(iii)) ^[350].

1.11.2 Alterations in the Cyclins and Other Cell Cycle Regulators

A number of cell cycle regulators acting downstream of the ER have been linked to endocrine resistance ^[347]. The Cyclin D gene (CCND1) is amplified in about one fifth of breast cancers ^[355] and its protein is overexpressed in approximately half of breast tumours (Figure 1.11B(iv)) ^[356, 357]. Since Cyclin D plays an important role in the G0-G1-S transition of the cell cycle, it has been hypothesised that this increased activity may be associated with high proliferation rates and poor outcome ^[358]. Conversely however, some groups have associated overexpression of the Cyclin D protein with ER positivity and a favourable prognosis ^[356].

This discrepancy may be explained by the fact that the CCND1 gene is a known ER target [359]. Therefore, high expression of the gene in ER+ disease, in the absence of gene amplification, may be indicative of oestrogen-dependent growth, which is likely to respond to endocrine treatment and to have a good prognosis. In contrast, CCND1 gene amplification is associated with activity of the autonomous ER-independent Cyclin D1, which may contribute to endocrine resistance [360] and is linked to a poor response to endocrine treatment and poor prognosis [361, 362].

Other examples include the overexpression of Cyclin E (CCNE1), downstream of CCND1, which can cause endocrine resistance by bypassing the cell cycle arrest induced by endocrine treatment [363-366].

1.11.3 Activation of Escape Pathways

As well as the classical ER pathways, the receptor can also function outwith the nucleus and is capable of indirectly regulating the activity of genes through interaction or 'cross talk' with other growth factor pathways [353, 354, 367-371]. These pathways can stimulate each other through phosphorylation by cell cycle kinases (Figure 1.11B(v)).

A number of growth factor pathways have been implicated. Prominent among these are the tyrosine kinase growth factors HER1/HER2, IGF-1R, FGFR and their associated downstream pathways. The HER2/EGRF pathway is among the most avidly studied. HER2 is a member of the tyrosine kinase family, regulating many processes that can promote tumour cell proliferation and survival [372]. HER2 does not directly bind with oestrogen but is a common dimerisation partner for the other HER family members [373]. Upon binding to their respective ligands, dimerisation of the HER receptor occurs, followed by phosphorylation of its intracellular tyrosine kinase domain leading in turn, to the activation of downstream signalling pathways such as PI3K-AKT-mTOR and RAS-ERK (MAPK) [374]. The former is a key mediator of growth factor receptor signalling and is important in the regulation of cell metabolism, growth, survival and angiogenesis. The latter is important in mediating the transcription of genes related to cell proliferation and differentiation.

HER2 amplification in ER positive tumours is associated with resistance to endocrine treatment [375, 376]. This can be explained by:

- (i) The direct effect of HER2 on proliferation via the downstream signalling pathways.
- (ii) The cross-talk stimulation between ER and HER2 [339, 377], for example through ligand-independent phosphorylation of ER.

1.11.4 Alteration of Cell Survival Proteins and Signalling Pathways

A number of other molecular mechanisms have also been implicated in endocrine resistance. These involve certain proteins associated with apoptosis (e.g. P53, BCL2 and CASP8) [378, 379] and also micro-RNAs (e.g. mir9, 221, and 222) [380, 381]. The latter are short RNA sequences that have been recognised as important regulators of gene expression through their ability to cause degradation of mRNAs and repress translation.

It is increasingly clear that the mechanisms of endocrine resistance, which are complex in themselves, are rendered ever more so by the fact that they closely interact with each other. For example, micro-RNAs exert control over the ER and cell cycle regulators whilst apoptosis is in part regulated by oestrogen signalling pathways. The combination of mechanisms at play may differ between individual tumours.

In order to understand the mechanisms of tumourigenesis, proliferation and resistance, it is necessary to consider them within the context of the hallmarks of cancer development.

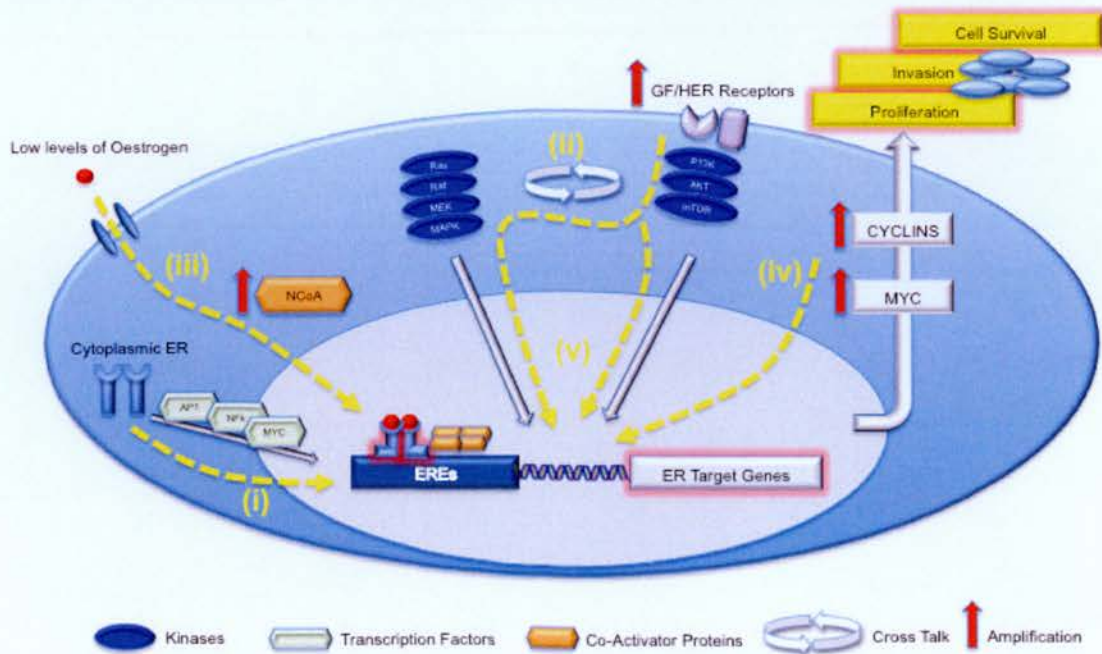


Figure 1.11B: Mechanisms of Resistance to Endocrine Therapy

The Molecular Biology of Breast Cancer

1.12 The Molecular Basis of Tumourigenesis

Multiple events are implicated in tumourigenesis (Figure 1.12A) [382, 383].

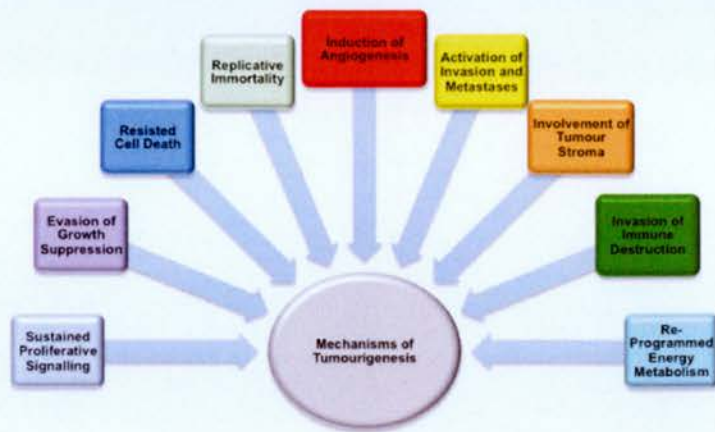


Figure 1.12A: Mechanisms of Tumourigenesis.

These are:

- (i) Sustained proliferative signalling
- (ii) Evasion of growth suppression
- (iii) Resisted cell death
- (iv) Replicative immortality
- (v) Induction of angiogenesis
- (vi) Activation of invasion and metastasis
- (vii) The involvement of the tumour microenvironment or stroma
- (viii) Evasion of immune destruction and,
- (ix) Re-programming of energy metabolism

These provide a framework by which the complexities of tumourigenesis may be considered and each will be discussed briefly.

1.12.1 Sustained Proliferative Signalling

Uncontrolled cell growth and the ability to sustain proliferation are arguably the most fundamental traits of the cancer cell. Entry into the cell cycle is usually highly controlled by the production and release of growth promoting signals, which ensure a homeostasis of cell number and maintenance of normal tissue architecture and function. In order to understand this better, it is necessary to have an overview of the cell cycle and its regulation.

An Overview of the Cell Cycle

The cell cycle comprises two principle stages: Interphase and Mitosis. Cells continually cycle between these two phases, occasionally entering a quiescent or 'resting' phase known as Gap0 (G0). During interphase, the cell grows, replicates its chromosomes and prepares for cell division. Through mitosis, the cell divides its chromosomes and cytoplasm to form two daughter cells. Entry into each phase of the cycle is dependent on the proper progression and completion of the previous one. The majority of the time is spent in interphase, which is subdivided into three components: G1, S and G2 (Figure 1.12B).

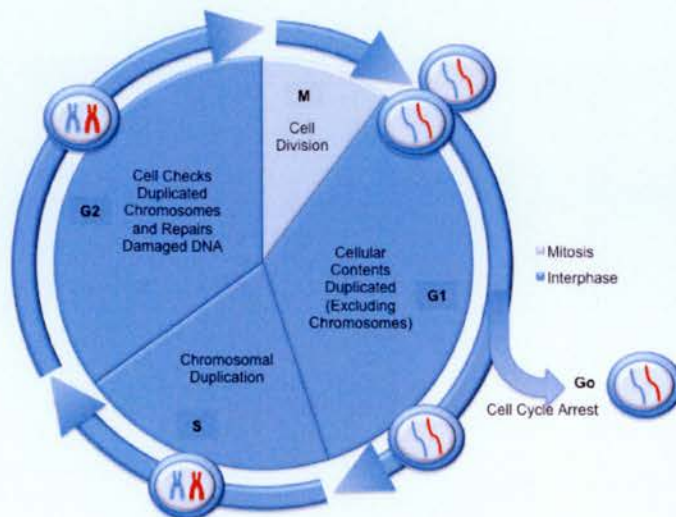


Figure 1.12B: The Cell Cycle.

- (i) G1 is the period between cell division and further chromosomal duplication. During this time the cell functions normally whilst checks take place to ensure its readiness for further division. At this point the cell may either:
- Exit the cell cycle and enter into G0.
 - Continue to remain in this phase for a highly variable period or,

- Continue to grow and to synthesise the proteins and enzymes required for chromosomal duplication during the next stage.
- (ii) The 'S' or Synthesis phase represents the period of chromosomal replication following G1 during which the cell is particularly vulnerable to the influence of mutagens. A doubling of chromosomal number marks the end of this phase.
 - (iii) G2 follows immediately on from the S phase. The cell continues to grow, in preparation for division. Further checks are made to ensure a full complement of DNA and other cellular components prior to mitosis.

Regulation of the Cell Cycle

Regulation of the cell cycle includes processes crucial to cell survival, mainly the prevention of uncontrolled proliferation and the detection and repair of damaged DNA. At least two types of cell cycle control mechanism are recognised. The first involves a cascade of protein phosphorylations that trigger the cell to move from one stage to the next. The second is a series of checkpoints that monitor the completion of critical events and delay progression to the next stage if necessary, for example to allow time for DNA repair. Mutations in many cancer promoting genes known as oncogenes, with resultant altered expression of their protein products, are primarily responsible for these changes [384, 385]. The regulation is governed by positive and negative modulators of the cell cycle Figure 1.12C.

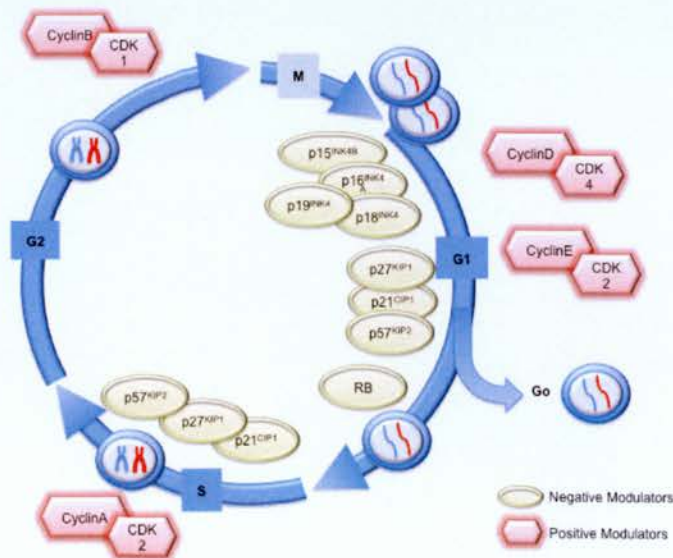


Figure 1.12C: Modulators of the Cell Cycle.

Positive Modulators of the Cell Cycle

The Cyclins

These are a highly regulated family of proteins that are transiently expressed in a cyclical manner, at appropriate times in the cell cycle, in response to growth signals and are subsequently degraded. They play a pivotal role in the activity of another group of proteins known as the Cyclin dependant kinases (CDK's). Several classes of Cyclins are recognised to be active at various stages of the cell cycle. The D and E Cyclins are critically associated with the G1-S phase and the B Cyclins with G2-M transition ^[384, 385].

The Cyclin Dependent Kinases

The Cyclin dependent kinases (CDKs) are an intracellular group of enzymes that form complexes with the Cyclins and are dependent upon them for their ability to phosphorylate other proteins. Active Cyclin-CDK complexes have unique substrate specificity and orchestrate signalling pathways that enable a cell to pass into the next phase of the cycle, whilst also influencing other biological processes (Figure 1.12C). Cyclins and CDKs are therefore positive modulators of the cell cycle. When a cell ceases to divide, the Cyclins are degraded, with consequent deactivation of the CDKs and arrest of the cell cycle.

Negative Modulators of the Cell Cycle

The CDK Inhibitors

In contrast to the Cyclins and CDKs, another group of proteins, the CDK inhibitors, negatively modulate the cell cycle by disabling the phosphorylating activity of the Cyclin-CDK complexes ^[386]. Two classes of CDK inhibitors are known to exist:

- (i) Class 1 specifically bind to CDK 4 and 6 and inhibit their association with the D-type Cyclins.
- (ii) Class 2 are inhibitors of the Cyclin A-CDK, D-CDK and E-CDK complexes. These are otherwise known as kinase inhibitor proteins (KIPs) and include the proteins products of the p21, p27 and p57 genes.

Whilst many CDK inhibitors are known, the above are prominent amongst them and are known to play a pivotal role in breast cancer development ^[384, 385].

Deregulation of the Cell Cycle

Since the production of the Cyclins, which promote progression through the cell cycle, is dependent upon stimulation by growth factors and in the case of breast cancer, upon steroid hormones, it has been hypothesised that cancerous cells may induce and sustain their proliferative signalling by:

- (i) Up-regulating / changing the configuration of their receptors rendering them hyper-responsive to low levels of stimulation ^[374, 387].
- (ii) Producing and responding to their own growth factors in an autocrine fashion ^[374, 388, 389].
- (iii) Activating neighbouring stromal cells to produce growth factors, which act in a paracrine manner ^[390, 391].
- (iv) Loss of the intracellular negative feedback mechanisms that ordinarily act to attenuate proliferative signalling ^[392].
- (v) Achieving growth factor independence via the activation of downstream pathways directly.

The latter may occur as a result of 'cross talk', whereby activation of the intracellular kinases or other proteins, not only promote entry into the cell cycle but also result in the activation of multiple other intra and inter cellular signalling pathways, which intercommunicate resulting in complex biological consequences ^[383].

1.12.2 Evasion of Growth Suppression

Cancers are capable of circumventing the signals that negatively regulate cell proliferation. Many of these signals are dependent on the actions of tumour suppressor genes, whose protein products such as the CDK inhibitors have an inhibitory effect on the cell cycle. Dozens of tumour suppressors operate in this way and have been discovered through their characteristic inactivity in various types of cancer. Two prototypical tumour suppressor genes encode the RB (retinoblastoma associated) and TP53 proteins; they operate as central controls within two key complementary circuits that govern the decisions of cells to proliferate, rest or undergo apoptosis ^[383, 393].

Malignant cells do not exhibit 'contact inhibition', whereby cell-to-cell contact inhibits further cell replication. The mechanisms underlying this probably relate to the defective actions of intracellular proteins that ordinarily orchestrate growth inhibition through the coupling of cell surface adhesion molecules (e.g. E-cadherin) to trans-membrane tyrosine kinases ^[394].

1.12.3 Resisted Cell Death

Programmed cell death or apoptosis, serves as a natural barrier to tumour development ^[395, 396]. Apoptosis is triggered in response to various physiological stimuli including:

- (i) Signalling imbalances resulting from elevated levels of oncogene expression.
- (ii) DNA damage associated with hyper-proliferation or otherwise.

Yet, apoptosis is attenuated in high-grade malignancy and in cancers that are resistant to therapy ^[395, 396].

The apoptotic machinery is composed of pathways that are either extrinsic or intrinsic to the cells, each ultimately culminating in the activation of a normally latent protease. Proteolysis, disassembly and consumption of the cell by phagocytosis then follows ^[395]. Autophagy, an intracellular catabolic process by which redundant or dysfunctional cellular components are degraded by lysosomes may mediate both tumour cell survival and death ^[397].

Tumour cells evolve a variety of strategies to limit or circumvent apoptosis. Most common is the loss of TP53 tumour suppressor function that eliminates this critical element from the apoptotic circuitry. Resistance to apoptosis may also be achieved by:

- (i) Increased tumour expression of anti-apoptotic proteins.
- (ii) Increased expression of survival signals.
- (iii) Decreased expression of pro-apoptotic proteins.
- (iv) Short-circuiting the extrinsic pathway.

These strategies reflect the diversity of apoptotic signals within each tumour ^[383].

1.12.4 Replicative Immortality

Cancer cells require unlimited replicative potential in order to form macroscopic tumour whilst normal cells are limited to a defined number of successive cell cycles. With each cycle the protective terminal sequences of the chromosomes (telomeres) shorten, eventually allowing fusion to their neighbouring chromosomes. The resultant DNA is unstable and the cell becomes senescent or undergoes apoptosis. Cancer cells employ the enzyme telomerase to lengthen their telomeres and thus become immortalised ^[398].

1.12.5 Induction of Angiogenesis

Tumours, like any tissue, are dependent on a blood supply to survive and grow. During tumour progression, an imbalance in certain regulatory proteins, which bind to endothelial cell surface receptors, promotes continuous angiogenesis at an early stage ^[399]. Vascular endothelial growth factor (VEGF) and Thrombospondin (TSP-1) are central regulators in this process ^[400, 401]. A variety of bone marrow derived cells also contribute to tumour angiogenesis ^[402].

1.12.6 Activation of Invasion and Metastasis

The process of tumour propagation involves several stages beginning with local invasion and followed by intravasation of cancer cells into blood and lymphatic vessels from where they are dispersed and subsequently extravasate into the parenchyma of distant tissues. Small micrometastases form and eventually become macroscopic tumours ^[403].

Recently a normal developmental pathway known as the 'epithelial-mesenchymal transition' (EMT) has been implicated in tumour invasion and spread ^[404]. Epithelial cells lose their polarity and cell-cell adhesion and instead acquire the ability to migrate and invade other tissues. This process is essential for numerous normal developmental and wound healing pathways but appears to be exploited by tumours. EMT is dependant on a number of key transcriptional proteins that evoke changes in cellular biology including:

- (i) Conversion from an epithelial to a fibroblastic morphology
- (ii) Loss of cell-cell and cell-ECM adhesion
- (iii) Expression of ECM degrading enzymes
- (iv) Increased motility
- (v) Resistance to apoptosis

Current evidence suggests that these transcription factors regulate each other and share regulation of other genes. Little is known about these interactions although cross talk between the cancer cells and the surrounding stroma may induce and promote these changes ^[405].

1.12.7 Involvement of the Tumour Microenvironment

In recent years, the tumour stroma has been recognised to play a critical role in tumour development ^[406, 407]. There is now a growing appreciation that tumours are not simply an independent mass of proliferating cells but rather a complex collection of multiple different cell types including stromal cells, which participate in complex interactions with one another.

In the breast, the stromal system is composed of the connective tissue matrix and fibres and supporting cells that surround the parenchyma including fibroblasts, immune cells, adipocytes and endothelial cells ^[406, 408]. Each of these may contribute individually or collectively to tumourigenesis via mechanical and biochemical regulatory signalling pathways ^[409]. In this way, the stroma may be activated by the tumour epithelium to encourage and support tumour growth through secretion of growth factors, stimulation of angiogenesis, or by facilitating cell migration and metastasis ^[409].

Cancer Associated Fibroblasts (CAFs) and Collagen Re-Alignment

One of the earliest indicators of a link between stromal biology and tumour progression was made by Boyd *et al*, who found that women who had mammographically dense breasts also had a two to fourfold increase in their susceptibility to breast cancer ^[410]. This appears to relate to an increase in collagen density ^[411]. Furthermore in breasts of variable density, tumours most commonly arise within the densest areas ^[412, 413].

Three-dimensional optical imaging of breast tumours in mice suggested that collagen changes at the tumour-stromal boundary might be important. Certainly, the invasion of the collagen-based basement membrane in humans appears to correlate with exponential tumour growth and invasion ^[414, 415]. Three patterns or 'tumour associated collagen signatures' (TACS) were identified which reflect tumour progression as follows:

- (i) TAC1 – Tumours exhibit a localised increase in the deposition of surrounding stromal collagen.
- (ii) TAC2 – With an increase in tumour size, the collagen fibres straightened and aligned themselves parallel to the tumour boundary.

- (iii) TAC3 – Remodelling of the stroma occurs with re-orientation of the collagen fibres perpendicular to the tumour boundary, the preferred orientation for tumour cell invasion ^[415].

TAC-3 has subsequently been associated with poor disease free survival in human breast cancer ^[416].

The mechanisms contributing to re-alignment of the collagen matrix remain unclear, however recent evidence points to the role of cancer associated fibroblasts (CAFs) and syndecan signalling ^[417, 418]. CAFs comprise at least two cell types; those like fibroblasts and those like myofibroblasts. It is unclear whether CAFs are normally resident within the stroma or arise from mesenchymal stem cells in response to tumour-secreted growth factors and thence recruited into the stroma or by the conversion of adipocytes in response to other tumour derived factors ^[419].

The syndecans comprise a family of transmembrane cell surface proteins, expressed on fibroblasts and epithelial cells. Syndecan-1 is normally expressed on mammary epithelial cells but not in the stroma. In three quarters of breast cancers however, this pattern is reversed ^[420]. It has been hypothesised that growth factors secreted by malignant cells activate stromal syndecan associated with CAFs to promote collagen re-alignment via unknown mechanisms ^[421]. In addition, Syndecan-1 appears to stimulate the proliferation of tumour epithelium and angiogenesis, and may be closely correlated with patient outcome ^[422-424]. The full spectrum of activity of the CAFs has yet to be elucidated.

Stromal Re-Modelling, Inflammation and the Immune Cells

The normal breast undergoes extensive cell death and remodelling on cessation of breast-feeding in an attempt to restore normal architecture ^[425]. The associated inflammatory changes of the stroma are similar to that seen in aggressive breast cancers ^[426]. Indeed, pregnancy associated breast cancer (PABC) has been associated with a particularly poor survival ^[427, 428], suggesting that inflammatory stroma may be in part responsible for invasion and metastases.

The involution of breast stroma is linked to the activity of cyclooxygenase-2 (COX-2), an enzyme that catalyses the synthesis of the prostaglandin mediators of inflammation. Moreover, a positive feedback loop exists as denatured collagen binds to the receptors of epithelial cells stimulating the production of COX-2 and prostaglandin signalling which in turn, increases the inflammatory response, resulting in further collagen deposition ^[426]. Inhibition of COX-2 in animal models results in stroma with severely diminished tumour-sustaining ability ^[429]. Given the link between breast density and survival, this signalling cascade may explain the decreased survival in PABC. The role that COX-2 plays in breast tumour progression appears to be significant ^[430]. In human breast cancer patients, moderate to strong expression of the COX-2 protein occurs in 40 % of patients and is associated with a poor disease free survival ^[431]. Agents like Aspirin, that inhibit COX-1 and COX-2, may inhibit breast cancer development.

The process of breast involution is also known to release pro-inflammatory colony stimulating factor (CSF-1) into the tumour micro-environment, which attracts inflammatory cells and in particular stromal macrophages ^[432]. The presence of large numbers of macrophages is associated with angiogenesis, tumour invasion and the stimulation of matrix degradation enzymes known as matrix metalloproteinases ^[407, 433, 434] whose presence is also correlated with poor outcome ^[435].

The matrix metalloproteinases (MMPs) are a family of nine or more highly homologous Zinc-endopeptidases that collectively cleave most if not all of the constituents of the extracellular matrix. Their specific role in tumour progression is still highly debated yet clinical studies have firmly established a correlation between survival and MMP expression ^[436-439]. The hypothesis however is that these enzymes contribute to the breakdown of the basement membrane and degrade stromal collagen to facilitate migration of tumour epithelial cells.

This strong indication of stromal involvement in breast cancer led to the decision to include multi-gene stromal signatures in this study.

1.12.8 Immune Destruction

The ability of cancer cells to evade the immune system highlights the dichotomous role of this system in tumour formation, at once trying to eliminate the tumour through various mechanisms whilst enhancing tumour development and progression through others ^[382].

Both B cell and T cell lymphocytes are likely to influence tumour progression and have been linked to overall survival in some studies ^[440, 441]. As yet however, their role is not clearly defined.

1.12.9 Re-Programming of Energy Metabolism

Another attribute that is currently under investigation is the ability of tumours to instigate a major re-programming of cellular energy metabolism in order to support continuous cell growth and proliferation^[382]. These are beyond the scope of this thesis.

The very many and complex biological pathways and signalling mechanisms that occur within and between the tumour parenchyma and associated stroma, are yet to be fully elucidated. However, it appears that the acquisition of each hallmark of malignancy is dependant on a combination of many processes and most importantly, the development of genomic change and instability. Molecular features have therefore been extensively applied to classify breast cancer and to design new biomarkers. In the breast cancer field the main focus of translational molecular profiling has been directed toward the development of multi-gene transcriptional signatures. This will be explored in the following section.

1.13 The Molecular Classification of Breast Cancer

1.13.1 The Limitations of Current Classifications

The ultimate goal of breast cancer management is to provide a perfect classification, to give an exact prognosis and to predict the best treatment regimen for each individual with precision. However, our current inability to consistently and accurately predict tumour behaviour and response to therapy indicates that the classification and prognostication systems currently in use are far from perfect.

The standard prognostic variables such as tumour size, grade and lymph node status combined with the predictive indicators of receptor status and Ki67 may direct treatment, but the response to that treatment can be remarkably variable.

There are several likely reasons for these inconsistencies:

- (i) Breast cancers are comprised of a mosaic of heterogeneous cell lines, such that the behaviour and response to treatment of the tumour as a whole is a reflection of all the cell lines within it.
- (ii) Inter-observer variation in the sampling and interpretation of clinical and pathological findings can result in misclassification.
- (iii) The mechanisms underlying cancer development, progression and response to treatment are more complex than the current classification systems allow for.

The main reason however for the shortcomings of the current classification systems, is that none of them fully incorporate our understanding of breast cancer at the molecular or genetic level.

1.13.2 The Genetic Basis of Cancer

Each of the hallmarks of cancer development discussed in the previous chapter, are ultimately regulated through complex mechanisms, under genetic control. When mutations occur in the genes responsible for this regulation or when their products are over or under expressed, tumourigenesis can occur.

There are several means by which cancer-enabling genes can malfunction. An abnormal variation in the number of whole or parts of chromosomes within the cell nucleus, known as aneuploidy, may arise through errors in mitosis. This results in an increase or decrease in the number of copies of specific genes within the nucleus (copy number alterations or CNA), increasing the potential for the over expression of their gene products.

Gene expression may also be altered by mutations within the genes themselves. These include point mutations, insertions, deletions and translocations of DNA sequences, as well as epigenetic events. The latter are defined as heritable changes in gene expression or cellular phenotype caused by mechanisms other than changes in the underlying DNA sequence - for example, changes in DNA methylation and histone modification, both of which are involved in the regulation of gene expression, without altering the underlying DNA sequence.

Whilst DNA repair mechanisms exist and may in some cases correct these mutations before they become phenotypically apparent, this is not always achieved successfully. In addition, genetic mutations may similarly arise in the very genes responsible for DNA repair.

In the majority of cases, the genetic changes that lead to malignant transformation are acquired over a lifetime and are present only in certain cells. These 'somatic' mutations are not inherited. Less commonly, 'germ line' gene mutations are inherited resulting in susceptibility to cancer from birth. Most often however, cancers only develop following the acquisition of additional somatic mutations, as part of the multistep process of tumourigenesis.

Our understanding of the molecular biology of breast cancer and of the alterations in chromosomal and genetic structure underlying it, is rapidly improving. In time, the accuracy of classification systems, prognostication and prediction of individual treatment response will become more sophisticated and accurate. An enormous amount of research is currently focusing on this objective, employing new techniques of gene expression analysis.

1.13.3 Gene Expression Profiling and the Intrinsic Subtypes

The advent of microarray technology has for the first time, allowed a rapid assessment of the activity of many thousands of genes from small tissue samples.

The concept behind this technology is that thousands of different short segments of nucleic acids or 'probes' are bonded in an organised pattern to a solid surface (microarray). These hybridise with mRNA-derived genetic material extracted from a tumour / tissue sample and this is then detected and quantified by means of chromogenic or luminescent markers. The process is known as gene expression profiling and the patterns of selected gene expressions thus derived are known as 'signatures'.

Microarrays can be used to assess gene expression in any type of breast cancer and changes in the signatures obtained before, during or after treatment can help us to understand the mechanism and efficacy of any particular therapy as well as guiding the development of new therapeutic approaches. This technology has already promoted the concept of breast cancer not as a single disease with variations in behaviour and treatment response, but rather a collection of different cancers affecting the breast, each with distinct patterns of gene expression, behaviour and histopathological features, whose response to treatment is determined by molecular characteristics as opposed to anatomical ones^[442-445].

One approach to microarray analysis in breast cancer has been to look empirically at a large range of genes to determine which are expressed in breast cancers and how their expression changes with tumour progression and treatment, in order to identify prognostic and predictive markers. Some of the earliest studies by Perou *et al.* demonstrated marked differences in the expression of many hundreds of genes between ER positive and ER negative tumours^[446, 447]. Surprisingly, it appears that a relatively small proportion of these are directly oestrogen regulated. In the majority of cases the co-expression of ER exists by virtue of common cell lineages during the evolution of an ER positive or negative tumour^[446-448].

Analysis of the genes whose expression varied more between tumours as opposed to between repeated samples from the same tumour, suggested at least five subtypes of tumours: Luminal; HER2 enriched; Basal-like; Normal breast like and Claudin-low. The luminal subtypes (A and B) are characterised by high ER gene expression, the remaining subtypes are ER negative ^[446, 447]. All of these subgroups exhibit marked differences in biology, survival and recurrence rates ^[446, 447, 449]. More recent studies have varied in their number and definition of molecular subtypes as well as their prognostic and predictive value ^[450-455] but the patterns of molecular classification are remarkably similar and confirm that ER status is a dominant molecular feature ^[446, 447, 449, 456, 457]. In addition, the normal breast like subtype may be due to a mix of normal breast and cancer within the same sample.

One empirical study identified a 70-gene prognostic signature in systemically untreated, node-negative breast cancer patients ^[158], which was subsequently validated in treated and node positive breast cancers ^[157]. Another 76-gene signature was later identified and validated, with prognostic implication for metastatic spread ^[458]. Only three genes are common to both signatures ^[459, 460]. Impressively, both signatures outperformed the St Gallen and NIH clinical risk classification systems, particularly in low risk patients ^[150, 461]. Nevertheless, the fact that half of the patients classified as high risk did not develop disease recurrence demonstrates that these signatures are still imperfect predictors.

An alternative approach to microarray analysis has been to consider which genes are involved with a particular cellular pathway and to specifically study expression of these genes ^[462]. This 'hypothesis led' approach was used to demonstrate that a 'wound response' gene expression signature is associated with poorer clinical outcome in breast cancer patients and is able to supplement the NIH and St Gallen classifications by identifying relatively low risk patients within the clinically high risk groups ^[463, 464].

Similarly, Glinsky *et al.* used microarray analysis to demonstrate the activation of 'self-renewal pathways' in malignant cells. These pathways are associated with normal stem cells and might contribute to the survival of cancer cells as well as promoting tumour progression. This investigation was based on the hypothesis that the presence of rare stem cell like populations of cancer cells amongst the heterogeneous mix of cells within a tumour is essential for tumour progression and metastasis of epithelial malignancies. An 11-gene signature was identified whose presence in primary tumours is a consistently significant predictor of rapid disease recurrence, distant metastasis, and death after treatment in a variety of cancers including those of the breast ^[465]. The genes associated with this signature have been shown to be involved in cell cycle control ^[466].

Desmedt *et al.* studied the expression signatures associated with breast cancer grading to determine if these could be used to refine the histopathological grading system ^[467]. This concept followed from the increasingly accepted views that high and low grade tumours are distinct diseases with differing clinical outcomes; that tumour progression is independent of grade ^[468, 469] and that intermediate grade tumours (30-60% of all breast cancers), display intermediate behaviour for which prognostication and prediction are difficult. A 97-gene signature, known as the Genomic Grade Index, was identified consisting of genes that are expressed to different extents in high and low grade tumours and which are associated with cell cycle progression and proliferation ^[467]. It was also confirmed that intermediate grade tumours are composed of a mixture of high and low-grade tumours. Not surprisingly, a significant overlap is evident between these genes and those of the separate 70 and 76 gene signatures ^[470].

Dai *et al.* demonstrated that the occurrence of metastases within a subset of patients with high ER is strongly predicted by a signature largely consisting of cell cycle genes, the overexpression of which is associated with an extremely poor outcome ^[471]. The Dai signature was selected for assessment in this study, but was found to be significantly less well correlated with the outcome of neo-adjuvant treatment in other patient subpopulations (see results section).

Many other studies have identified different gene signatures that are of prognostic value in breast cancer. These share several characteristics:

- (i) They mainly identify the same group of patients with poor prognosis, [472-474] corresponding to high expression of genes involved with proliferation [442, 473, 475], which have been shown in meta-analysis to have the highest predictive value [473].
- (ii) The level of expression of proliferation genes is the strongest prognostic indicator in ER positive tumours [442, 473, 475, 476].
- (iii) They mostly identify ER negative tumours as having poor prognosis due to high levels of proliferative gene expression [442, 473, 475, 476].
- (iv) Most correlate with response to traditional therapeutic agents, which is not surprising as these target proliferation [442, 445, 476].
- (v) Meta-analysis suggests that clinical and pathological criteria provide prognostic information distinct from these signatures [443, 444, 473, 477].
- (vi) Prognostic accuracy is time dependent with predictions at 5 years being more accurate than at 10 years [459, 478].
- (vii) Prognostic information above and beyond that provided by ER, PR, HER2 and Ki67 is currently limited [156, 455].

It seems likely that the prognostic and predictive power of various signatures would be increased if they were carefully combined with one another or indeed with other non-genetic factors. Recently, some benefit has been demonstrated by combining both empirical and hypothesis driven signatures, for example in the Sensitivity to Endocrine Therapy (SET) index [479]. In this way, a 165-gene signature was identified and validated on ER positive patients who received either adjuvant Tamoxifen for 5 years or neoadjuvant chemotherapy followed by Tamoxifen or an AI as well as patients receiving no adjuvant therapy. The signature was predictive for patients receiving either of the endocrine treatments but not in untreated patients. The signature may therefore help to determine which patients should receive endocrine treatment.

1.13.4 Clinical Application of Microarray Technology

Despite all these findings, molecular classification is still in its infancy and processing specimens is both expensive and time consuming. Classification is further hindered by the fact that most breast cancers contain multiple cell lines with different pheno- and genotypes which, in part, explains the variability in response to therapy in these cancers ^[480]. Therefore, the clinical use of molecular classification is currently limited despite the promise it holds. There are however, several clinically available kits for identifying molecular signatures in breast cancers including MammaPrint® ^[157, 158], Oncotype DX® ^[481-485], the Veridex 76-gene signature ^[458], Mapquant DX™ ^[467] and the Breast Cancer Index™ ^[486]. Only the first two are currently supported by level II and level I evidence respectively and approved by the FDA ^[459, 460]. Many more are sure to be developed.

1.13.5 Prognostic Signatures

MammaPrint® (Agendia, Amsterdam) is a microarray based gene expression profiling assay and is the same 70-gene signature described above ^[157, 158]. It relies on RNA from fresh tissue samples and can be used for patients with stage 1 or 2, node negative breast cancer of less than 5cm, irrespective of oestrogen receptor (ER) status ^[157, 459, 487]. MammaPrint® categorises tumours as either high or low risk and provides independent prognostic information beyond standard clinico-pathological variables ^[459].

Oncotype DX® (Genomic Health, California) is a 21-gene signature (16 cancer-related genes and 5 normal comparator reference genes) based on quantitative reverse transcription PCR analysis of formalin fixed, paraffin embedded tumour samples ^[481, 482, 485]. The result is reported as a Recurrence Score (RS) score that classifies patients as high, intermediate or low risk of recurrence at 10 years in ER positive, node negative disease as well as having predictive relevance when these tumours are treated with Tamoxifen. Oncotype DX® may also be applicable to ER positive tumours treated with aromatase inhibitors ^[488] and those with positive nodes ^[489]. The ability to retrospectively apply this test to stored fixed tumour samples has enabled further validation and resulted in level I evidence leading to the incorporation of Oncotype DX® into several national guidelines in concert with standard anatomical and pathological staging ^[485, 490-493].

Sotiriou described the currently available signatures as adding "modest prognostic information for patients with HER2-positive and triple-negative tumours", but added that "when measures of clinical risk are equivocal (e.g., intermediate expression of ER and intermediate histologic grade), these assays "could guide clinical decisions" ^[443].

The promise of the future is to develop more specific, clinically applicable signatures that can identify the most significant molecular prognostic and predictive markers in breast cancer and guide the most appropriate therapy. However, it is likely that histopathological classification will remain the baseline standard for some time, with the gradual introduction of molecular tests which may be useful only in specific histopathological subtypes.

1.13.6 Second-Generation Prognostic Signatures

The first generation prognostic signatures are useful in patients with ER positive disease and have pointed to genes largely associated with proliferation. Therefore, proliferation signatures were selected for assessment in this study. The hope is that the next generation of signatures will provide both prognostic and predictive information for ER positive and negative disease. The principal difficulty with ER negative disease is in identifying a population of patients with significantly good prognosis to justify withholding chemotherapy, since the incidence of disease relapse within 5 years exceeds 20%. Whilst many genes related to the cell cycle and proliferation have been shown to predict outcome in ER positive disease, significantly fewer such genes have been identified in ER negative disease. The latter include genes relating to immune response ^[445, 475, 477, 494].

Stromal gene signatures are also of increasing interest and several of these have been identified ^[495-498], although as yet, their usefulness in clinical practice has not been determined. An appreciation of the importance of stromal rearrangement in the biology of breast cancers treated with neo-adjuvant endocrine therapy influenced the selection of these signatures for evaluation in the present study.

1.13.7 Predictive Signatures

Clinically useful gene signatures should also be able to predict response to treatments including endocrine and chemotherapy. Indeed, the first generation signatures, by identifying poor prognosis tumours with high expression of proliferation genes have proved useful in predicting response to combination therapy that targets proliferation ^[442-445, 476]. Prediction of response to specific agents has been less successful ^[499-504] and there are currently no commercially available signatures for this purpose. The search for such signatures may be confounded by the fact that the mosaic of cell lines within a particular tumour confounds microarray analysis. For example, a cell line within a tumour may express genes that confer resistance whilst the remaining cell lines could be relatively sensitive. Microarrays are not capable of distinguishing between the different cell lines within a tumour and instead present the 'average' pattern of expression for the tumour as a whole.

Many of the molecular sub-classifications of breast cancer whether based on the tumour epithelium or stroma or both, are essentially descriptive of the cell type rather than the underlying biological changes associated with tumour progression or response to treatment. As a result, despite being predictors of survival, they are not yet clinically useful predictors of treatment.

Efforts are now underway, to categorise gene expression signatures based upon their responsiveness to adjuvant or neoadjuvant therapy ^[505]. In this way, the signatures can be used to determine the optimally effective treatment plan for individual patients.

1.13.8 Predictive Signatures in ER Positive Tumours

To date, no multi-gene test predicting response to neo-adjuvant endocrine therapy in advanced ER positive tumours has been validated for clinical use. Recently however a number of exploratory studies have applied pre-operative and advanced disease protocols to examine the relationship between endocrine response and gene expression profiles ^[506-510].

Jansen *et al.* evaluated clinical response to Tamoxifen in advanced breast cancers and found that genes associated with clinical response included classical oestrogen related genes and elements of 'stromal signature' including genes associated with extracellular matrix production and immune response ^[506].

A series of studies undertaken in Edinburgh have looked at molecular response to aromatase inhibitors. Miller *et al.* studied sequential core biopsies taken before and after two weeks of neoadjuvant treatment with Letrozole and related the early changes in gene expression to the clinical response observed at the time of tumour excision ^[508]. At present this is the largest study to report a predictive signature for neo-adjuvant endocrine treatment. However as yet this signature lacks independent validation. Assessment of the performance of this signature was therefore a major aim of this study (see Aims and Objectives).

Similarly, Mackay *et al.* studied gene expression profiles in sequential biopsies taken before and after two weeks of neo-adjuvant treatment with either Anastrozole or Letrozole and followed immediately with surgery ^[507]. However, they developed a signature that used Ki67 as a response surrogate to evaluate changes in proliferation.

The molecular responses observed in both Miller's and Mackay's studies were generally consistent. Aromatase Inhibitors suppressed classical oestrogen dependent and proliferation associated genes. On the other hand, the most consistently up-regulated genes were enriched by the stromal signature, including genes encoding specific types of collagens and genes associated with cell adhesion, immune response and turnover of the intercellular matrix. Significantly, these molecular changes were observed as early as two weeks post treatment, before morphological changes were evident.

1.13.9 Stromal Gene Signatures

Breast stroma is increasingly recognised to play an important role in tumour progression and invasion, and possibly initiation. The mechanisms by which this may occur are unclear but are certainly complex and are the subject of a comprehensive review by Muschler and Streuli^[511]. These mechanisms are again influenced at the genetic level and significant effort is being focused in this area.

Early gene expression studies used tissue derived from whole tumours composed of both the tumour epithelium and stromal cells, making it impossible to distinguish the gene expression profiles of each^[157, 158, 446, 447]. Given that the tumour epithelium often only comprises about half of the total tumour volume, this was a significant shortcoming^[512]. The relevance of this was reflected in the observation that the gene signature profile from the same tumour differs when the biopsy is performed using FNA (stroma poor) or core biopsy (stroma rich)^[495]. Furthermore, comparison of the gene expression profiles of stroma and epithelium from normal breast and breast cancers, demonstrated that the stroma plays a role in tumour progression even before invasion with 90% of the alterations occurring during the transition from normal breast tissue to in situ carcinoma^[513].

The importance of the stroma is confirmed by a number of other studies^[495, 514, 515]. One of these looked at the expression of stromal genes in aggressive breast cancer and found that genes encoding Syndecan (Sdc-1), Fibronectin and some Collagens were significantly up-regulated compared with less aggressive tumours^[516].

In fact, stromal gene expression signatures may prove to be better predictors of patient outcome because the stroma and not the epithelium may be the primary target of carcinogens. When isolated normal mammary stroma was exposed to a carcinogen and recombined with normal epithelium, a tumour developed. The same did not occur when isolated carcinogen-exposed epithelium was recombined with normal stroma^[517].

Recently, microarray analysis was used to compare gene expression profiles of normal breast tissue and breast tumours and these were correlated with patient outcomes^[497]. A 26-gene stromal signature (Stroma Derived Prognostic Predictor, SDPP) was identified and subsequently trained to a 23-gene signature that stratified disease outcome independently of standard clinical prognostic factors including lymph node status. The signature was also found to be a better predictor of survival than whole tumour signatures from amongst patients with multiple clinical subtypes of breast cancer, having even greater accuracy than predictors that were derived from whole tumour samples. Five and later six, categories of altered stromal gene expression were identified: matrix remodelling, hypoxia, fibroblast signalling, ER signalling, angiogenesis and the immune response. Each category exhibits distinct molecular and biological phenotypes, carrying prognostic information that is independent of existing therapeutic biomarkers and tumour subtypes. The presence of a defined stromal signature such as this, allows whole tumours to be analysed for predictive purposes, bypassing the need to isolate stromal cells^[497, 518]. Another gene expression study of stroma tissue has identified a wound response signature that predicts poor survival in whole tumours^[463, 497].

In terms of predicting therapeutic responses, Farmer *et al.* developed a stromal gene expression signature that predicts resistance to pre-operative FEC chemotherapy for patients with ER negative tumours. The genes expressed in this signature are associated with reactive stroma and this was taken to indicate that anti-stromal agents could be useful in mitigating tumour resistance to chemotherapy^[496].

Since the complex communication between the stroma and the tumour epithelium underlies tumour progression and metastasis, better understanding of stromal gene expression patterns may lead to novel strategies for improving patient outcome.

2. Aims and Objectives

2.1 Aims

Approximately 75% of all breast cancers are ER positive and the vast majority of these occur in postmenopausal women. Amongst this group of patients, neoadjuvant endocrine therapy is used widely in an attempt to reduce tumour size prior to surgery. For some patients, this will permit breast-conserving surgery to be performed instead of mastectomy, whilst for others it can make an inoperable cancer operable. A further group of patients with advanced inoperable disease may be palliated with endocrine therapy.

It is clear however that not all patients respond to endocrine therapy and for those that do, the response may vary in magnitude and longevity. It is therefore essential to find an accurate and consistent means of predicting which subset of ER positive patients will respond well to endocrine treatment, so that the others can be spared unnecessary delays in receiving more appropriate forms of treatment.

Molecular classifications of breast cancers using microarray technology and other means show great potential to refine the current clinical classification systems. Gene expression signatures have been largely derived from whole tumour specimens and have focused primarily on tumour epithelium. More recently, the importance of peri-tumoural stroma has been demonstrated and several stromal signatures derived. Not only do these signatures help with the sub-classification of breast cancers into clinically useful groups, but they may also be used to provide a prognosis and help predict which patients will respond to specific treatments. Furthermore ontological assessment of the genes comprising these signatures may help elucidate the control mechanisms underlying the development, progression and response or resistance to treatment for breast cancer.

Nonetheless, at the time when the present study was conducted, the clinical utility of only one of the published gene expression signatures had been tested in the neo-adjuvant endocrine setting ^[510] and this had yet to be validated in an independent dataset. Since then however, several published studies have investigated response prediction to AIs in this setting ^[519-521].

The aim of this study therefore was to validate the efficacy of the published neo-adjuvant endocrine response predictive signature and to compare it to other potentially informative signatures.

Gene expression signatures reflective of proliferation and stromal rearrangement were selected for comparison. Proliferation signatures were chosen because they reflect the principle component of the majority of previously described adjuvant signatures^[522]. Additionally, they represent the main gene cluster change following two weeks of neo-adjuvant treatment with Letrozole^[508, 523]. Stromal rearrangement signatures on the other hand, were selected for their important biological role in response to endocrine treatment^[168, 309, 326] and for their potential to predict neoadjuvant response to cytotoxic chemotherapy. To ensure that the results were pathway related, rather than being related to a specific signature, two signatures were reported for each pathway.

In light of the strong preliminary evidence supporting the clinical relevance of proliferation, the following were investigated further:

- (i) Shifts in mRNA expression of selected Cyclins and CDKs during neo-adjuvant treatment with Letrozole using real-time quantitative PCR.
- (ii) The association of proliferation with long-term outcomes in a sub-group of AI treated patients, using immunohistochemical measurements of Ki67 as a surrogate marker of tumour proliferation.

2.2 Objectives

The objectives of the study were therefore as follows:

- (i) Identification and enrolment of patients.
- (ii) Collection of response and survival data.
- (iii) Assay of tumour samples by microarray analysis.
- (iv) Validation of the 205-covariables signature published by Miller *et al.* ^[510] against a new and independent dataset.
- (v) Identification of four relevant published signatures for comparison: Two proliferative and two stromal ^[471, 496, 508, 524, 525].
- (vi) Comparison of the predictive capacity of each signature using all available datasets derived from this and previous studies based on the Edinburgh Letrozole Audit series.
- (vii) Measurement of changes in the expression of mRNA of Cyclins B1, A2, D1, CDKs 1,2 and 4 and the NUSAP1 gene (a mitotic spindle associated protein) by real-time PCR in a sub-group of cases. Then to relate these changes in expression of the Cyclins/CDKs to changes in Ki67 (measured by ICH) and clinical response to neoadjuvant therapy.
- (viii) Relate neoadjuvant changes in Ki67 within a subgroup of patients, to breast cancer specific survival and to compare Ki67's prognostic capacity with classical clinical parameters, such as tumour size and lymph node status.

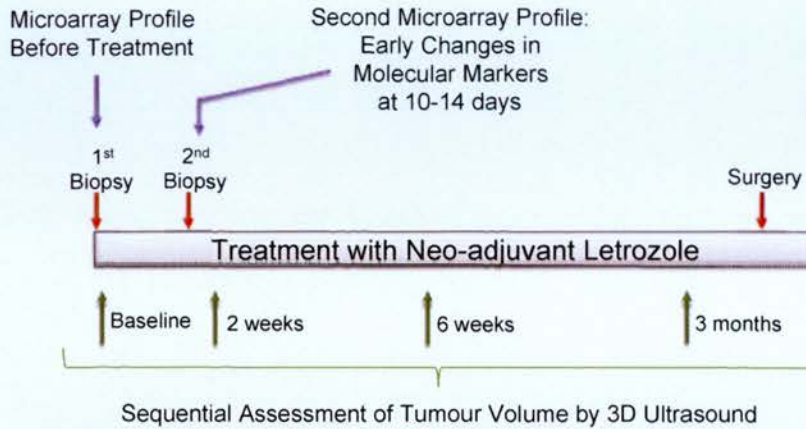
3. Methods and Results

3.1 Study Design

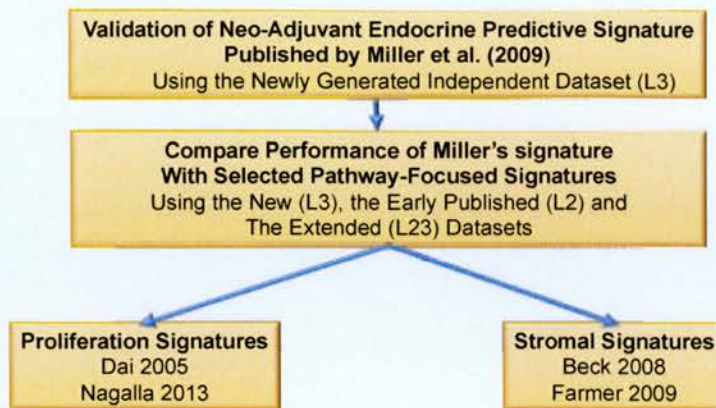
The study design is summarised in Figure 3.1A-C below. In brief, enrolled patients were treated with neoadjuvant Letrozole whilst sequential tumour biopsies were taken before and early in the course of treatment. Tumour response was assessed during treatment using 3D ultrasound and whole genome mRNA expression profiles were obtained using high-density microarrays. The performance of the published gene expression signatures was assessed using the microarray and clinical response data. Two focused sub-analyses were performed using real-time qPCR and IHC data. Each of the study elements will be described in more detail in the subsequent sections.

Figure 3.1A-C: Study Design Overview

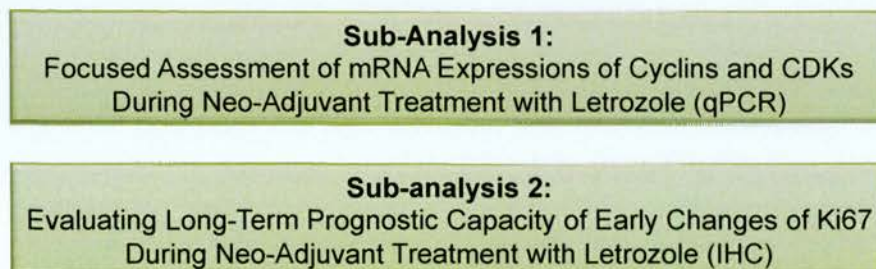
A: Data Collection



B: Main Analysis



C: Focused Sub-Analyses



3.2 Patient Selection and Enrolment

This study involved the prospective enrolment of postmenopausal patients presenting with locally advanced or large yet operable, ER positive breast cancer for whom the current clinical guidelines indicated neoadjuvant treatment with the aromatase inhibitor Letrozole.

All patients included in this study were diagnosed and treated at the Edinburgh Breast Unit of the Western General Hospital. Those who met the eligibility criteria were discussed at a weekly multi disciplinary meeting (MDM) and referred to the author or to one of the other medical and research staff involved.

The inclusion criteria required that patients had:

- (i) Previously untreated, biopsy proven invasive breast cancer.
- (ii) Oestrogen receptor (ER) rich tumours (Allred score 5-8).
- (iii) No evidence of metastatic spread.
- (iv) Tumours that were amenable to clinical examination, mammography and ultrasound scan.
- (v) Ceased hormone therapy, such as HRT, at least four weeks prior to commencement of Letrozole.
- (vi) Capability to give written, informed consent.

Patients excluded from this audit included those:

- (i) With premenopausal disease.
- (ii) With Oestrogen receptor negative (or ER poor) tumours.
- (iii) Having had prior treatment with endocrine therapy or treatment with another investigational drug within 30 days of commencing Letrozole.
- (iv) With transmissible disease (e.g. HIV, Hepatitis C).
- (v) With severe renal or hepatic impairment or other co-morbidities.
- (vi) Unable to give informed consent.
- (vii) Deemed unlikely or unable to maintain compliance with the treatment or follow-up regimen.

Following the MDT, the author met with the patients to explain the purpose and aims of the study, to provide details of the nature and timings of additional investigations that they would be subject to, should they agree to participate and to determine their willingness to participate. Written documentation was provided with the same information (Appendix III) and follow-up appointments were arranged for further discussion and to obtain consent. A detailed consent form was signed for each participant (Appendix II). Information about the study and patient enrolment was sent to each patient's GP for their records.

Letrozole has been established as a front-line neoadjuvant treatment within this population of patients since 2001. In the present study, a standard daily dose of 2.5mg oral Letrozole was prescribed for a period of three months. The duration of treatment was based upon previous clinical experience and allows sufficient time to effect meaningful tumour response in most cases, whilst avoiding prolonged, ineffective treatment and tumour progression in patients that do not respond.

The use of the neoadjuvant setting allows for the *in vivo* and *ex vivo* assessment of the clinical, pathological and molecular response to Letrozole at multiple time points, which are independent of surgery or other treatments. Clinical and radiological assessments of the tumours were made using callipers, USS and mammography and serial core biopsies were taken at the time of diagnosis and repeated after two weeks and after three months of treatment, the latter generally corresponding to the time of surgery (Figure 3.2A). The core biopsies were fresh frozen or formalin-fixed for further analysis.

Long-term clinical outcome measures were recorded as part of an ongoing audit – the Letrozole Audit series.

A sub-group of patients treated with various neo-adjuvant AIs which included Letrozole, Exemestane and Anastrozole for whom long term follow-up was also available, was used in a sub-analysis which related neo-adjuvant Ki67 changes to breast cancer specific survival.

It is important to note that several studies were being undertaken simultaneously to maximise the research potential of the information and tissue samples collected from these patients. Whilst the author was involved in many of these, their purpose, conduct and findings are outwith the scope of this thesis and are therefore not described.

These Studies were undertaken in accordance with the ethical principles of the declaration of Helsinki and approved, in advance, by the Lothian Research and Ethics committee (LREC 06/S1103/65, 2001/W/BU/09, 07/S1103/26, 07/S1104/23 and 2001/W/BU/10) and the NHS Trust Research and Development Department. Patients gave their written informed consent to participate before study enrolment.

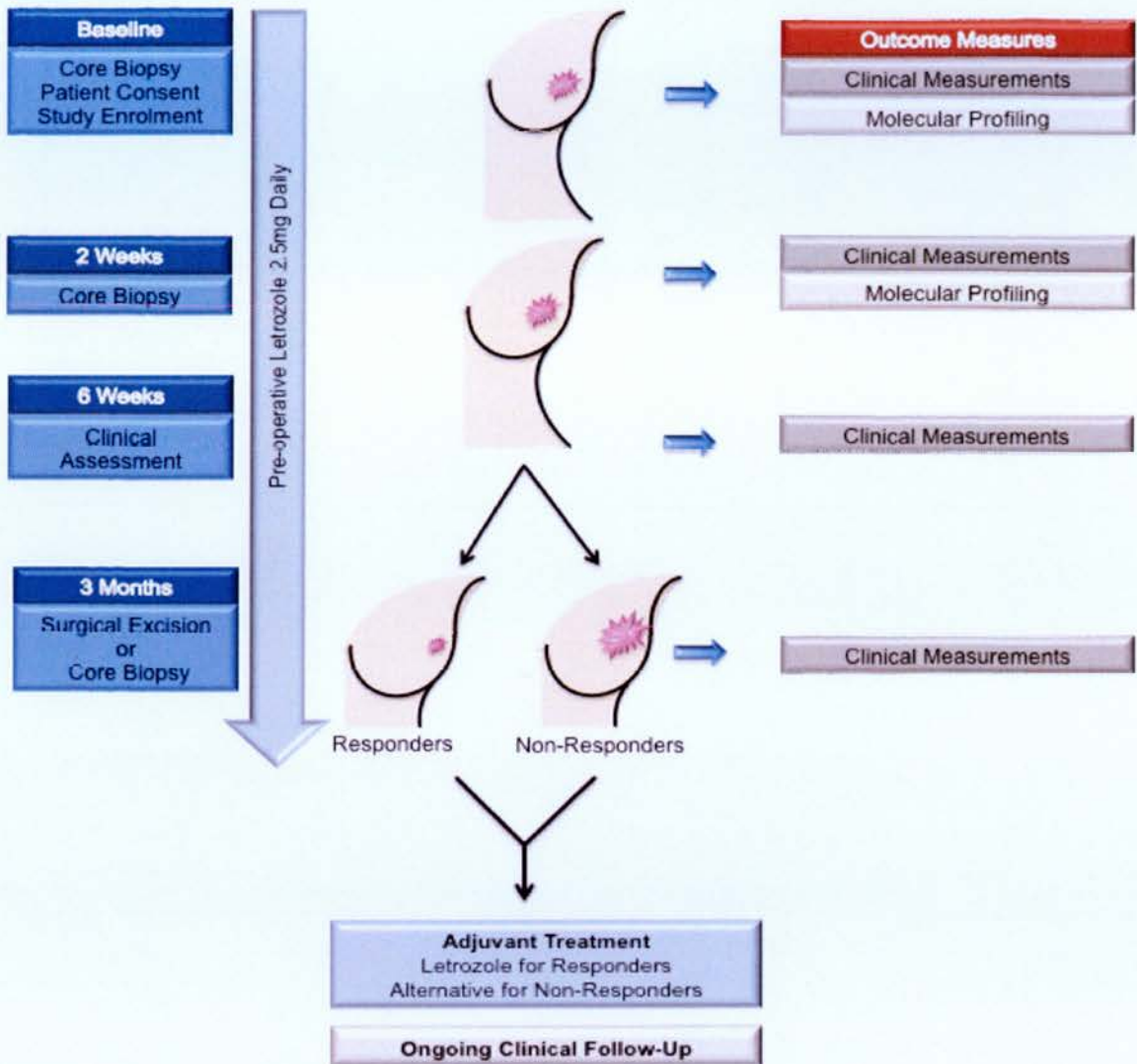


Figure 3.2A: Summary of Study Chronology

Note: Additional ultrasound measurements were taken when possible

3.3 Assessing Tumour Response and Clinical Follow-up

A variety of complimentary methods were utilised to evaluate initial tumour characteristics and subsequent response to Letrozole treatment. These included clinical, radiological and pathological measures. In addition, patients were asked to confirm their compliance with the treatment and report any side effects. Outwith the follow-up appointments, all patients in the study were given contact details for the investigators with whom they could discuss side effects or any other complications or concerns.

3.3.1 Clinical Examination

The maximum diameter of the tumour was measured in two perpendicular planes using callipers. Based on the assumption that a tumour is approximately spherical, its volume (V) can be related to its diameter (D) by the equation $V=(\pi D^3)/6$. Tumour volumes were recorded at each time period. Any change in the overlying skin was also documented.

Measurements taken in this way are subject to intra and inter-observer variability and are notoriously inaccurate, nonetheless they remain in common clinical use and it was therefore considered to be an important variable to document in the present study.

3.3.2 Ultrasound

Ultrasound offers the potential to more accurately measure tumour volume by utilising the third dimension of tumour thickness as measured from superficial to deep. In addition, the tumour dimensions taken by USS do not include those of the skin and subcutaneous fat and are therefore subject to less intra and inter-observer variability than calliper measurements. The USS measurements for this study were undertaken using a Honda Convex scanner (HS-2000) with a linear array probe. Tumour volume was calculated according to the formula $V=(d\pi D^2)/6$, where d is the tumour thickness and D is the mean of the tumour width and height. The results were recorded as above.

3.3.3 Mammography

Oblique and cranio-caudal mammographs were taken at the time of diagnosis and prior to surgery. The diameter of the tumour was measured and used to calculate the tumour volume in the same way as that described for calliper measurements above. This technique suffers from some distortion associated with compression of the breast and tumour between the mammographic plates.

All three methods of volumetric assessment were applied at the time of diagnosis and surgery. Additional USS and Calliper measurements were taken whenever possible, and at least at 6 weeks to assess tumour response to Letrozole. Where these measurements suggested a lack of response or indeed expansion of the tumour, consideration was made to changing treatment as appropriate. Responsive patients however, were scheduled for surgery at around three months. Dynamic changes in tumour volume measured by 3D ultrasound were considered as the primary measure for response assessment as described later.

3.3.4 Pathological Assessment

At the time of tumour extirpation, the examining pathologist determined the exact invasive tumour dimension both macro and microscopically, for comparison with those taken by the pre-operative clinical methods (callipers, USS and mammography).

In addition, the tumour was assessed histologically by Haematoxylin and Eosin staining. Changes in cellularity and fibrosis, known as central scar formation were assessed by comparison with the biopsies taken at the time of diagnosis and after two weeks of treatment. Such pathological responses have been reported in 60-80% of tumours after 3-4 months of treatment ^[328, 329]. Response may also be indicated by a reduction of histological grade; mitotic score being the most commonly affected grading feature ^[328].

3.3.5 Classification of Tumour Response

Tumour endpoints, such as objective response and time to disease progression, are useful only if based on widely accepted and readily applied standard criteria, which are in turn, based upon assessment of anatomical tumour burden. Two such systems have been published over the past two decades: the WHO classification, published in 1981 ^[321] and subsequently a modified version known as the RECIST classification in 2000 ^[322-324]. These are however designed for use in the advanced disease setting.

Since no defined classification systems currently exist for measuring response in the neoadjuvant setting, trial centres commonly employ locally defined criteria. The following modified WHO classification is routinely used in Edinburgh for clinical purposes ^[320].

- Complete Response (CR): No measurable tumour.
- Partial Response (PR): Reduction in tumour size of 50% or more.
- Minor Response (MR): Reduction in tumour size of 25% to 50%.
- No Change (NC): Less than 25% reduction or increase in size.
- Progressive Disease (PD): Increase in size over 25%.

Minor response and no change are also regarded as being indicative of stable disease.

The present experimental analysis used a highly selective approach based on several sequential ultrasound measurements. Essentially, only tumours exhibiting extreme responses (i.e.- rapid and stable responders or long-term non-responders) were selected. Tumours with intermediate response were excluded from further analysis, to ensure true biological comparison between responders and non-responders.

3.3.6 Outcome Monitoring

Patients were, and continue to be, monitored for disease recurrence (local and/ or distant) and survival (disease-free, distant disease free and overall). Local recurrence was classified as that occurring in the ipsilateral breast, adjacent chest wall or ipsilateral axillary nodal basin.

A sub-group of patients for whom long-term outcome data was available at the time of the present analysis was used to relate breast-cancer specific survival to changes in Ki67 observed during neo-adjuvant treatment.

3.4 Tissue Collection and Storage

Core biopsies were taken under ultrasound guidance at the time of diagnosis and again approximately two weeks later, following initiation of Letrozole treatment. This second biopsy was optional although the majority of patients agreed to it. The final tumour specimen was taken at the time of surgical excision, usually at three months following initiation of treatment.

When taking a core biopsy, local anaesthetic solution (10-20mls of Lignocaine 1% with 1:200,000 adrenaline) was infiltrated around the tumour to minimise discomfort and bleeding. Next, the skin overlying the tumour was prepared with Betadine antiseptic solution and a fenestrated disposable drape applied to the chest to minimise the risk of infection. A small incision was made over the palpable tumour through which a single use 14-gauge Bard Max core biopsy cannula was inserted into the tumour substance. Several passes (approximately 8) were made to obtain sufficient and representative samples of the tumour. Pressure was then applied to the biopsy site to minimise bruising and bleeding and a dressing applied.

On several occasions, tumour biopsies were taken by the lead clinician, using the ultrasound guided Mammotome™ biopsy device. This device employed a larger 8 or 11-gauge biopsy cannula and suction to permit harvest of larger specimens than those from conventional core biopsy.

The final (three month) tissue specimens were collected from the operating theatre at the time of wide local excision or mastectomy, either by core biopsy on the excised specimen or direct excision from the specimen performed by the surgeon or pathologist.

Each specimen was labelled with a unique, anonymised identifier. A part of the sample was immediately fixed in Formalin and sent to the pathology department at the Western General Hospital to be embedded in paraffin and undergo further histological analysis. The remaining cores were immediately snap frozen in liquid nitrogen to preserve nucleic acids and maintained in liquid nitrogen in the on-site tissue bank for later genetic analysis.

A spreadsheet was kept of patient and sample details to allow identification for later analysis.

3.5 Tumour Processing and RNA Extraction

Prior to molecular analysis, unit pathologists assessed frozen sections for tumour content. Samples in which the malignant component exceeded 20% of the cross-sectional area were chosen for RNA analysis. Tissue samples were not enriched for their malignant component only, because the intention was to include stromal tissue for analysis in the belief that the latter may also produce relevant genetic information.

3.5.1 Extraction of Specimens from OCT and Stabilising RNA

The frozen-section specimens were returned from the pathology department embedded in frozen OCT (Optimal Cutting Temperature buffer) and transferred on dry ice (Figure 3.5A). The volume of OCT significantly exceeded the volume of tissue embedded within it by a factor of 10 and therefore, the OCT had to be removed prior to RNA extraction so that the appropriate ratio of reagents to tissue could be maintained throughout, to optimise the extraction process.

This was achieved by placing the OCT-tissue blocks in pre-chilled (-20°C) RNA^{later}®ICE solution and storing them overnight at -20°C. The now stabilised tissues were defrosted to room temperature immediately prior to use. Once the OCT had melted, the tissue biopsies were collected using forceps.



Figure 3.5A: OCT Embedded Tissue.

RNA^{later}®-ICE (Ambion, Texas, US) is a reagent used in the transitioning of frozen tissue to room temperature in order to preserve high quality RNA for further extraction.

3.5.2 Procedure for RNA Extraction using Qiagen RNeasy Mini-kits

RNA extraction was performed by the author for a subset of tumour samples collected between 2005-2008 and used in this study. The remaining samples were extracted by trained laboratory personnel employed in the unit at the time.

Weighing and Homogenising the Tissue

1. The defrosted tissue was weighed and 100mg used for RNA extraction, the remainder was stored in RNA/ater®-ICE according to the manufacturer's recommendations.
2. The tissue was then disrupted by finely cutting into small pieces and
3. Transferred to a tube containing 2ml RLT buffer (Qiagen) and 2-Mercaptoethanol.
4. This was then homogenised using a hand-held rotor-stator homogeniser (Qiagen TissueRuptor or Silverson) for 45 seconds.
5. The homogenate was cleared by centrifugation for 3 min at room temperature.
6. 2 ml of the supernatant was then dispensed into pre-labelled tubes for RNA extraction on Qiagen RNeasy mini columns according to the manufacturer's recommendations as described below.

RNA Extraction and Purification with RNeasy Mini Columns (Qiagen)

7. The lysates were mixed with 2ml of 70% ethanol by vortex for 2 minutes.
8. 700µl aliquots were transferred into the pre-labelled RNeasy-mini columns, which were centrifuged for 15 seconds and the flow-through discarded. This process was repeated until the entire sample had been applied to the column.
9. 700µl of Buffer RW1 (Qiagen) was applied to the column, which was centrifuged for 15 seconds, and the flow-through discharged.
10. 500µl of Buffer RPE (Qiagen) was applied to the column, which was centrifuged for 15 seconds, and the flow-through discharged.
11. Another 500µl of RPE was applied to the column, which was centrifuged for 2 minutes, and the flow-through discharged.
12. The mini columns were left open and centrifuged for a further minute to dry them.

Elution of RNA

13. 30µl of RNase-free water was added to the mini columns and centrifuged for 1 minute to elute the RNA, which was then placed on dry ice for transfer and storage in a freezer at -80°C.

After purification, the quantity and quality of the RNA was evaluated on an Agilent Bioanalyser (Agilent Technologies UK Ltd, Berkshire) and on a Nanodrop-1000 spectrophotometer.

3.6 Microarray Analysis

Following RNA extraction, the tissue samples were subjected to whole-genome gene expression microarray analysis in order to identify which genes, of many thousands, were expressed and either up or down-regulated by Letrozole treatment. Qualified personnel at the local Genetic Core Facility carried out both the labelling and amplification of the RNA samples and microarray hybridisation.

3.6.1 Microarray Platforms

In principal, microarray analysis depends upon the hybridisation of thousands of specific DNA probes on a microarray chip to fragments of DNA or RNA within the sample being analysed. Prior to hybridisation, the RNA is amplified and labelled as recommended by the array manufacturer.

The probes are short oligonucleotide sequences that correspond to a known gene or other DNA element. These are synthesised and bound to a solid substrate surface such as glass, known as a chip or array (Figure 3.6A). The surface of the array comprises thousands of spots or 'features', typically measuring less than 10 microns in diameter and each containing millions of copies of a single probe sequence in a specific position on the array.

When the tissue sample extract is added to the surface of the array under specific conditions, the probes hybridise to their complementary sequences within the sample. Unbound sequences are washed off the chip and the remaining hybridised sequences are detected by the addition of a fluorescent tag, which adheres to biotin, attached to the fragments during the amplification and labelling process, prior to hybridisation. The fluorescent features, representing expressed genes, are detected using a microarray scanner and their signal strength can be quantified by computer analysis.

Various microarray platforms are commercially available. The Affymetrix® (Santa Clara, California) and Illumina® (San Diego, California) systems are leaders in this field and both were used in this study.

Affymetrix Arrays

The Affymetrix® HG-U133A (22283 features) system was used to analyse tissue samples collected from the first 137 patients enrolled in this study (L2 series, see Section 3.8 for details). The Affymetrix® chips contain a high density of multiple short probes, each of approximately 25 nucleic acids. These are synthesised in-situ directly onto specific pre-defined locations on the chip by a process of photolithography and can hybridise to complimentary aRNA (amplified, labelled RNA) within tissue extracts. Because the probes are short, it is possible for several different probes to hybridise with the same sequence of target RNA. To improve specificity, multiple 'perfect match' probes are used for each gene along with multiple 'single-base mismatch' probes. The chips were scanned using the Affymetrix 3000 laser scanner.



Figure 3.6A: Affymetrix Microarray Chip.

Illumina® Arrays

The newer Illumina® array system (HT-12 chip) became available at the genomics laboratory of the Roslin Institute at Edinburgh University and was therefore used for gene expression analysis on all subsequent tumour samples (L3 series). This system uses pre-synthesised probes that are attached to microbeads of 2-3 microns in diameter to a density of 800,000 probes per bead. The beads are randomly applied to micro-wells etched on a silica array surface, each 5.7 microns apart (Figure 3.6B). These locations are 'decoded' after manufacture and a unique 'bead map' is supplied with each array. The probes are longer than those of Affymetrix at around 50 nucleotides each and approximately 30 beads at different sites on each array contain the same probe to improve the quality of the data. In

general, the Illumina HT-12 system allows for higher density analysis than the Affymetrix HG-U133A system and can therefore produce even more reliable results.

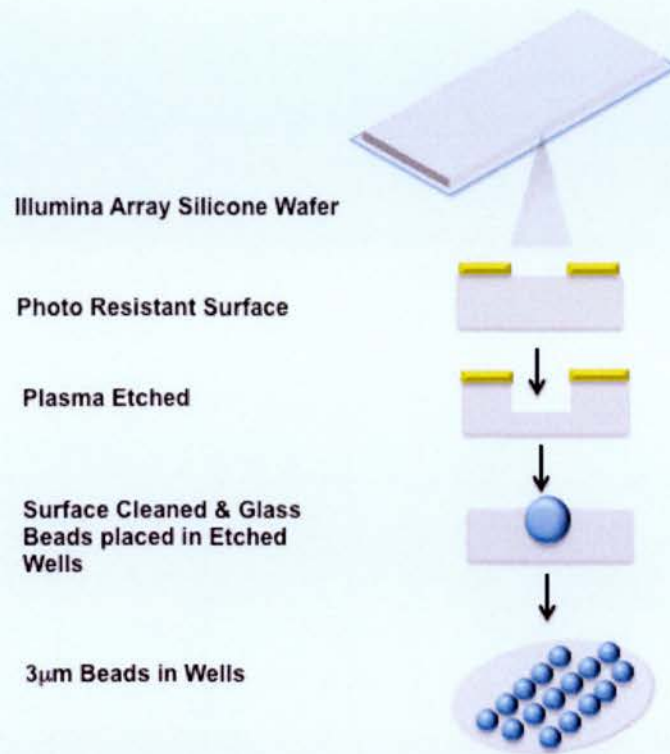


Figure 3.6B: Illumina Microarray System.

The present study used data generated using HG_U133A Affymetrix arrays (L2 dataset ^[508, 510, 523, 525]) and Human HT-12 V3 expression Beadarrays (L3 dataset ^[526]).

3.6.2 Microarray Data Processing

Microarrays scans produce fluorescence intensity measurements for each probe spot, with the intensities corresponding to the levels of hybridisation.

By their very nature this produces vast amounts of raw data, which needs to be processed carefully in order to yield useful information. This process is highly complex and belongs to the realm of bio-informatics and computer analysis but in essence involves a series of stages described below. These are specific to the array technology used. The microarray data processing for this study was carried out with the support of local bioinformaticians.

Pre-Processing

There are several potential causes of variation in the intensity of an individual spot or feature on a microarray. Whilst this may be due to an actual difference in gene expression, it is often influenced by background 'noise', which needs to be eliminated or reduced, prior to final data analysis.

'Noise' may be caused by non-specific hybridisation or other fluorescent artefacts. The initial steps of pre-processing are usually carried out by computer software, which is provided with the commercial microarray scanning equipment. Later steps may be carried out using multiple third-party tools, for example the R-statistical software environment. In general, pre-processing aims to extract background intensity caused by non-specific gene probe hybridisation. It also permits the comparison of data from multiple arrays. Microarrays processed at the same time and in the same environment to minimise the variability in analysis conditions, are referred to as a 'batch'. However, it is often necessary to process samples at different times, in different batches. In order to negate or minimise the effect this can have on data quality, some form of 'batch correction' processing must be applied. This was performed by computer software.

Normalisation

A further source of noise relates to labelling and detection inefficiencies. Normalisation aims to correct for differences that do not represent true biological variation between samples. It also aims to remove systemic or technical variations in the relative intensities of each feature and to correct for differences in intensities between samples, whether on the same or different arrays. For example, with Affymetrix, the intensity of the probes on a given array can be consistently higher or lower than those on other arrays. This is referred to, as array bias and failure to correct for such bias would lead to serious errors in data interpretation.

Normalisation often makes use of genes that are constitutively expressed in all samples at a constant rate (e.g. housekeeping genes) to act as intensity controls. Alternatively, the summed intensity values of each whole array can be compared and used to normalise the difference between them, based on the assumption that most genes are equivalently expressed between different arrays and, the remainder, which are either up or down-regulated, will balance each other out.

The following pre-processing was used for each dataset:

L2 dataset

Raw expression values were read from Affymetrix' CEL files and normalised using the Robust Multichip Average (RMA) algorithm, including background correction, normalizing, log-2 transformation and summarisation, as implemented in the Affymetrix R Bioconductor package (<http://www.bioconductor.org/>). Doctor A. Krause, the Novartis bioinformatician responsible for the Letrozole audit project at that time, performed this pre-processing. Further detail and references can be found in the papers ^[508, 510, 523].

L3 dataset

The background subtracted, non-normalised data were exported from Illumina Genome Studio for downstream processing in Lumi R-package ^[527]. Seven arrays were excluded because of their low global detection rate (<10%) or low mean intensity (<6.5). After applying force-positive background correction, the data were log-transformed and quantile-normalised. Low-expression probes were removed (detected with $p < 0.01$ in less than 25% of cases). Only probes classified as 'Perfect' or 'Good' according to the ReMOAT probe annotation were used for the analysis ^[528]. This reduced the number of probes from approximately 47,000 to 11,000. Batch correction was performed using SVA R-package: ComBat correction accounting for treatment response and sample type (baseline, 2 weeks or 3 months) without interaction ^[529, 530]. This pre-processing was performed by Dr. A. Larionov specifically for this MD project in 2013.

L23 dataset

This dataset was produced by Dr. A. Turnbull through merging of the L2 and L3 datasets at signal level. The detailed data processing is available in Turnbull *et al* (2012) ^[526]. Briefly, prior to the integration, the L2 raw data were pre-processed using RMA from Affymetrix R package and the L3 raw data were quantile normalised using lumi R package. A sophisticated in-house bioinformatics algorithm was used to map compatible Affymetrix and Illumina probes ^[526]. Only 7000 probes were found to be compatible out of 24,000 Affymetrix and 43,000 if Illumina probes. Pre-processed expression values for the compatible probes were merged using the XPN (cross-platform normalisation) algorithm described in Shabalín *et al* 2008 ^[531].

Filtering

Filtering allows the identification and removal of non-informative genes, such as those either not expressed in any studied sample or uniformly expressed in all samples. It also corrects for poor quality or empty spots, outliers and control probes. Data are also processed by three different means to determine which genes are significantly up or down regulated. These are:

- i) Direction of change.
- ii) Magnitude of change.
- iii) Statistical significance of the change.

The function of these genes can later be ascertained by comparison with previously identified biological pathways or gene ontology.

Downstream Analyses

Multiple methods of downstream analysis can be applied to the microarray data depending on the specific study aims, for instance to detect differentially expressed genes or to visualise the results. The present study calculated multi-gene expression scores and evaluated their association with clinical response using Receiver Operating Characteristic Curve Analysis. Details of these study-specific analyses are described later in section 3.11.

3.7 Reverse-Transcription and Quantitative Real-Time PCR for Selected Genes

The microarray analyses indicate the pattern of expression of genes within a tissue sample. Changes in expression can be determined by analysis of the tissue samples collected at different time points during treatment with Letrozole. However, microarray probes can become saturated by target sequences and are not necessarily a reliable means by which to *quantify* gene expression. In order to validate and quantify gene expression based on the microarray results, quantitative polymerase chain reaction (qPCR) also known as real time PCR was performed for selected genes (Section 3.16). The same technique was also employed for the focused sub-analysis of the cell-cycle associated genes (Section 3.17).

Standard PCR can be used to target and amplify any specific nucleic acid sequence within a tissue sample, which can also be quantified by different means. PCR comprises a series of steps (Figure 3.7A):

- 1) The target gene for amplification is identified and used to design primers. These are short nucleic acid sequences that are complementary to sequences flanking the target of interest.

- 2) The primers, buffer, a supply of nucleotides and the enzyme Taq polymerase are mixed in appropriate quantities within a 'thermocycler'.
- 3) The thermocycler controls a cyclical variation in temperature to facilitate the replication of the target gene. Each cycle consists of the following:
 - (i) The solution is raised to a high temperature (95°C), causing complementary DNA strands to separate.
 - (ii) The temperature is then lowered to 50-60°C to promote hybridisation or 'annealing' of the synthetic primers to the separated DNA strands.
 - (iii) Finally, at a temperature of 68-72°C, the polymerase enzyme replicates the DNA strand commencing at the annealed primer sequence.

At the end of each round, the number of target gene DNA strands is theoretically doubled. Typically 25-50 cycles are performed to achieve amplification of the order of 1,000,000 fold.

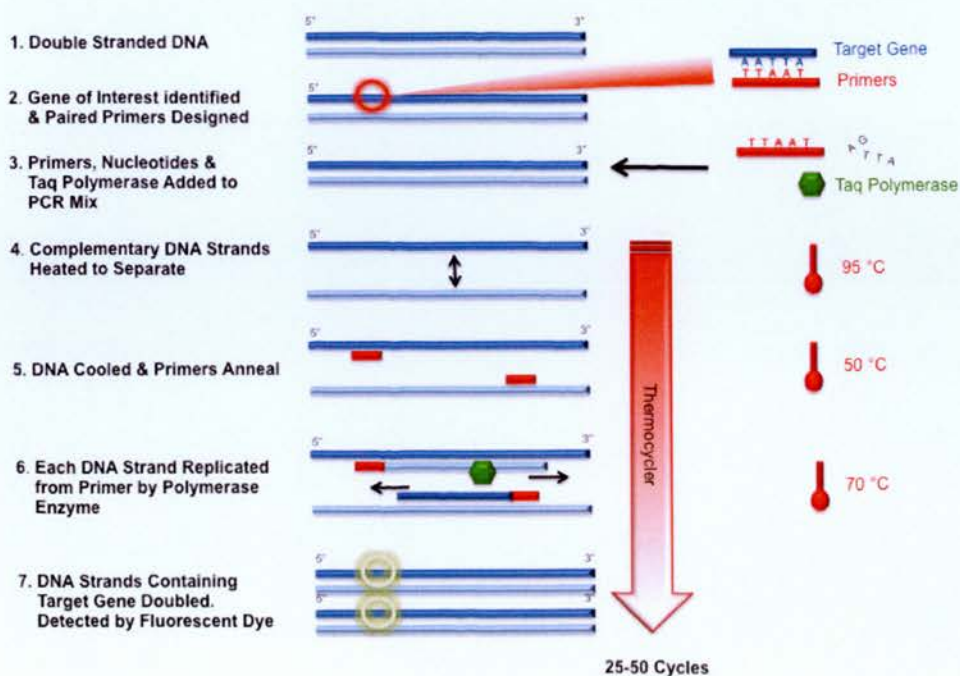


Figure 3.7A: Amplification of Specific DNA Sequences by PCR.

Standard PCR does not however allow for quantification of the amplified DNA. Real Time or quantitative qPCR does permit relative or absolute quantification of the DNA by means of either double-stranded DNA intercalating-dyes such as SYBR green, or by the use of fluorescently labelled oligonucleotide probes which can be quantified during the PCR process. The latter technique is more specific: the bases at either end of the probe are labelled with a fluorescent reporter tag and a 'quencher', which serves to absorb the fluorescent energy from the tag when in close proximity to it. As the polymerase reaction proceeds, the fluorescent probe is cleaved off the DNA strand as it is replicated. The tag and quencher are therefore separated and the resulting fluorescence can be detected with a fluorometer. The SYBR green technique used in the present study does not require the synthesis of specific labelled probes and can still be used for reliable quantification of PCR targets.

The *relative* quantity of target within a sample can be measured by comparison to one of several standard, constitutive genes that are considered to be expressed at consistent rates. In breast cancers undergoing endocrine treatment, the TBP, KIAA0674 and PUM1 genes show low variation in expression and can be used as references. Alternatively, an *absolute* quantification can be obtained by comparison with DNA standards, using an absolute calibration curve.

Whilst DNA is the natural substrate for PCR, the samples obtained in this study were of mRNA. It was therefore necessary to convert them to cDNA using the enzyme reverse transcriptase, prior to the application of a PCR technique similar to that described above.

Reverse transcription and qPCR were performed by the author for a subset of tumour samples collected between 2005-2008 and used in this study. The remaining samples were processed by trained laboratory personnel employed in the unit at the time.

3.7.1 Reverse Transcription Protocol

The following protocol for the superscript® VILO™ cDNA synthesis kit (Invitrogen) was used to convert the RNA specimens to cDNA for use on PCR, according to the manufacturer's recommendations:

- 1) The RNA samples were treated with DNase enzyme to degrade all DNA contaminants within the samples.
- 2) The pre-measured concentration of each RNA sample (Nanodrop 1000) was used to calculate the appropriate dilution with RNase-free water to achieve a concentration of 125ng/μl.

- 3) A mastermix was prepared comprising:
 - 6µl of RNase-free water
 - 4µl of 5X VILO™ reaction mix
 - 2µl of 10X superscript® enzyme mix.
- 4) The mastermix was vortexed, briefly spun and aliquots of 12µl were dispensed into 0.5ml Eppendorf tubes for each reaction.
- 5) 8µl of diluted RNA specimen was added to the tube of mastermix solution to achieve a total of 1µg of RNA in a final reaction volume of 20µl, which was gently mixed.
- 6) A PCR thermal cycler (Thermo Hybaid, Ashford, England) was used to control the following incubation sequence:
 - 25°C for 10 minutes.
 - 42°C for 60 minutes.
 - 85°C at 5 minutes to terminate the reaction.
- 7) The resultant cDNA solutions were diluted ten fold by adding 180µl of RNase-free water, which was then mixed by vortex and divided into four 50µl aliquots which were stored at -20°C for later use.

Control samples were also required for the PCR assays. One type of control was prepared using RNA from one of the most abundant tumour samples in which the Reverse Transcriptase component was substituted with DNase-free water prior to the PCR reactions. The second series of controls (Non-template, NTC) were processed using water instead of cDNA samples.

3.7.2 Selection and Design of Primers

PCR primers were designed using the Primer-3 software (open-source online software), avoiding known SNPs. Since the PCR was run on cDNA samples prepared from RNA, only expressed sequences (Exons) could be used in the PCR primers. At least one primer in each pair spanned exon boundaries or the primers were each located in different exons, preferably co-located on the gene with the sequences detected by the microarray probes. The primer sequences were then validated on the public database RTPrimer DB (www.rtprimerdb.org) and the National Centre for Biotechnology Information's e-PCR tools (www.ncbi.nlm.nih.gov/projects/e-pcr/reverse.cgi). Once validated, they were ordered and supplied by Sigma Genosys (Dorset, England).

Finally, the primers were validated using gradient PCR, which essentially completes PCR cycles at a range of different annealing temperatures (50-64°C) in order to determine the optimal annealing temperature.

The Primers sequences are shown below in Table 3.7A

Gene	Forward Primer	Reverse Primer	Function
CCND1	TGTCCTACTACCGCCTCACA	CCAGGTCCACCTCCTCCT	Proliferation/Cell cycle
CCNA2	ACAAAGCTGGCCTGAATCAT	ACTGTTGTGCATGCTGTGGT	Proliferation/Cell cycle
CCNE1	ACAGCTTGGATTTGCTGGAC	GCTCTGCTTCTTACCGCTCT	Proliferation/Cell cycle
CCNB1	CATGGTGCACCTTCTCCTT	CAGGTGCTGCATAACTGGAA	Proliferation/Cell cycle
CDK2	GTGGTGTGGCCAGGAGTTAC	CAGCATTTGCGATAACAAGC	Proliferation/Cell cycle
CDK4	ACTGGCCTCGAGATGTATCC	AGGCAGAGATTCGCTTGTGT	Proliferation/Cell cycle
CDK1/ CDC2	TGGCCAGAAGTGGAAATCTTT	AAATTCGTTTGGCTGGATCA	Proliferation/Cell cycle
NUSAP1	AGTTTGTCTCGTCCCCTCAA	CGTTTTCTCCGTTGCTCTTC	Proliferation/Cell cycle
DCN	GGGCTGGCAGAGCATAAGTA	CAGAGCGCACGTAGACACAT	Stromal adhesion/rearrangement
COLEC12	TGGAACGATTTCCAATGTGA	TTTGGTAATGACGGATGGT	Stromal adhesion/rearrangement
MMP2	TGCTGGAGACAAATTCTGGA	TGATGCTTCCAACTTCACG	Stromal adhesion/rearrangement
KIAA0674	GGGCAACTCTGAAAGCTCTG	CTTCATCCCAGCCAGTCAAT	Reference
PUM1	CGTGGTCCAGAAGATGATTG	TACGAAGAGTTGCGATGTGG	Reference
TBP	GGGGAGCTGTGATGTGAAGT	CCAGGAAATAACTCTGGCTC	Reference

Table 3.7A: Primer Sequences for all PCR Analyses.

3.7.3 Quantitative PCR Protocol

Quantitative PCR was undertaken using the QuantiTect SYBR Green® qPCR kit (Qiagen) and 96x PCR plates on an Opticon™2 DNA engine thermal cycler (MJ Research, Bio-Rad). This kit utilises SYBR Green fluorescent dye to bind to double stranded DNA and enable quantification of PCR targets without the need to synthesise specific probes. All work was undertaken in a safety cabinet to prevent contamination.

The following protocol was applied:

- 1) The prepared cDNA samples, primers and SYBR Green PCR mastermix were thawed and placed on ice.
- 2) All the reagents were spun briefly to ensure collection of the content in the bottom of each tube.
- 3) Thirty-two 0.6ml tubes were pre-labelled and placed in a tray on ice.
- 4) The PCR mix containing RNase-free water, SYBR green and 2 primers was prepared as follows:
 - 2x QuantiTect SYBR Green PCR mastermix 25 μ l
 - Primer A: Variable volume to achieve a final concentration of 0.3 μ M
 - Primer B: Variable volume to achieve a final concentration of 0.3 μ M
 - RNase-free water to make total volume up to 45 μ l in each pre-labelled vial.
- 5) 45 μ l of the PCR mix was pipetted into each of the 32 pre-labelled tubes.
- 6) 5 μ l of cDNA preparation was then added to each tube.
- 7) The tubes were mixed by vortex and spun briefly to collect the content in the bottom of the tube.
- 8) 15 μ l of each PCR/cDNA mix was pipetted into 3 cells on a white 96x PCR plate to ensure the optimal reflection of the fluorescent signal.
- 9) The plates were centrifuged briefly (3000 rpm, 1 min, room temperature).
- 10) PCR was carried out on the Opticon™2 thermal cycler set with the following parameters:
 - Once only activation of Taq DNA Polymerase, 15 minutes at 95°C
 - Denaturation, 15 seconds at 94°C
 - Annealing, 30 seconds at 56°C
 - Extension, 30 seconds at 72°C
 - Fluorescence data collected at 72°C

- 11) The PCR cycle was repeated 50 times.
- 12) PCR cycling was followed by once only 5' elongation at 72°C.
- 13) A melting curve analysis was performed on the thermal cycler following the DNA amplification cycles to verify the specificity of PCR products, based on the binding of SYBR green to double-stranded DNA. During this analysis, the product of qPCR was gradually heated until the DNA strands were dissociated and released the SYBR green dye with a corresponding drop in measurable fluorescence. This allowed detection of the melting temperature of the formed PCR product(s).
- 14) PCR results were normalised by geometric averaging of three housekeeping genes: KIAA0674 (FKBP15), PUM1 and TBP. The genes and primers used for their detection are presented in Table 3.7A. These housekeeping reference genes were validated in previous studies from our unit and found to be expressed stably in clinical breast cancer biopsies undergoing endocrine treatment ^[532].

3.8 Datasets

The datasets used in the present work have been generated during a Letrozole Audit initiative in the Edinburgh Breast Unit between 1999 and 2010. The relationship between each dataset and their allocation within the Letrozole Audit series as a whole, are illustrated in Figure 3.8A. The initial small set of subjects known as LETMA-1 (LETrozole MicroArray) or 'L1', was used for methodology and logistics optimisation and was not used in the present analysis. The author personally participated in enrolment of patients between 2005 and 2008. Qualified personnel employed in the Unit at the time have enrolled other patients used in the present analysis.

Allocation of Patients into Datasets within Letrozole Audit Series in Edinburgh Breast Unit

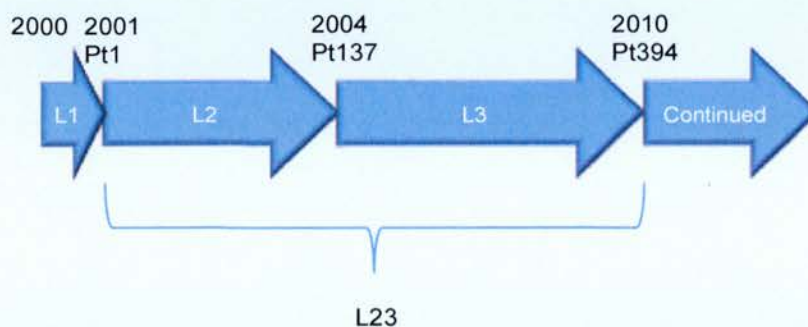


Figure 3.8A: Overview of Datasets.

3.8.1 Letma 2 (L2)

The first 137 patients enrolled into the Letrozole audit between 2001 and 2004 were incorporated into the L2 dataset used in the present study (52 of these were reported by Miller *et al.* earlier^[508, 510, 523]). Tumour sample microarray analysis was performed using the Affymetrix HG-U133A platform. Doctor A. Krause performed the microarray pre-processing using Robust Multichip Average methodology^[533, 534].

Of these 137 patients, only those with a complete set of measurements taken at baseline (pre-treatment), two weeks and three months and high quality RNA samples were included in the final analysis. Sixty-eight paired samples (Baseline and 2-weeks) from 34 subjects were therefore selected for present analysis based on the tumours' response to treatment (see Table1 in Datasets and Signature Translation, Appendix V)

After this period, the Illumina HT-12 assay replaced Affymetrix HG-U133A because it became locally available and additionally includes a greater number of features. Furthermore, the use of an alternative platform permitted platform-independent validation of the L2 series results in the subsequent L3 series of tumour samples.

3.8.2 Letma 3 (L3)

Continued recruitment into the Letrozole Audit after 2004, produced a second dataset, the L3 series. This group was recruited in exactly the same way as the L2 series and treated according to the same protocol, although fewer patients consented to a two-week biopsy or had clearly documented USS measurements at three months. The main difference between L2 and L3 related to the microarray platform used in their analysis. Microarray was performed on Human HT-12 V3 expression BeadArrays (Illumina, Cambridge, UK). Overall, the L3 series recruited a further 257 patients of which 44 had complete, high quality data and tissue samples for inclusion at the time of this study.

Doctor A. Larionov performed the microarray pre-processing for this series. The background subtracted non-normalised data were exported from Illumina Genome Studio for downstream processing in Lumi R-package. Some arrays were excluded because of a low detection rate (<10%) or low mean intensity (<6.5). After applying force positive background correction, the data were log-transformed and quantile-normalised. Low-expressed probes were removed (detected with $p < 0.01$ in less than 25% of cases). Only probes qualified as "Perfect" or "Good" according to ReMoat probe annotation were used for the analysis^[528]. This reduced number of probes from the initial ~47k to ~11k probes. Batch correction was performed using SVA R-package: ComBat correction accounting for treatment response and sample type (baseline, 2 weeks or 3 months) without interaction^[529, 530]. Finally, sixty paired samples (baseline and 2 weeks) from thirty subjects were selected for the analysis basing on their response to treatment (see Table2 in Datasets and Signature Translation, Appendix V).

3.8.3 Letma 2 and 3 (L23)

Whilst the L2 and L3 series were identical in terms of recruitment and inclusion / exclusion criteria, the microarray analysis was performed using different platforms as above. Combining the data from both series therefore has the potential to improve statistical power of subsequent analyses, however direct integration of expression data is complicated by the fact that different platforms were utilised. Even the scanner used in data collection and the day on which samples are analysed can introduce inter-platform bias. In order to combine these datasets, the microarray data from Illumina and Affymetrix analyses was first translated into a 'common language' using the pipeline for Affymetrix and Illumina integration described by Turnbull *et al.*^[526]. This cross platform integration was mediated using gene ID's from the Ensembl genome browser (www.ensembl.org), which was shown to significantly out perform other methods of cross platform integration^[526].

In the present study the combined dataset was used to improve statistical power when comparing the signatures that had not been derived from either L2 or L3. However, this dataset could not be used for assessment of the Edinburgh signature because the signature was partially derived from these samples. A total of 99 patients were included in the L23 dataset.

The method of data integration has been previously described in detail in Turnbull 2012 ^[526]. In summary, the data from different platforms were first pre-processed separately, using tools specific for each platform (i.e. Affymetrix and lumi R packages – see section 3.6.2). Genes with low detection rates were removed. Batch-effect correction was performed using ComBat R script ^[529]. L2 probes were re-annotated to Ensembl Gene IDs using the custom Chip Definition File (CDF^[471]). L3 probes were re-annotated to Ensembl Gene IDs using ReMOAT^[528] and a custom BLAST and BioMart search. This strict methodology ensured that only comparable probes were taken for the downstream data integration; numerically reducing the dataset to only 7000 compatible probes out of 24,000 Affymetrix and 43,000 Illumina probes. Pre-processed, batch-corrected and re-annotated data for the compatible probes were integrated using the XPN (cross-platform normalisation algorithm described in Shabalín *et al.* 2008 ^[531]). Finally, 126 paired samples (baseline and 2 weeks) from 63 subjects were selected for the current analysis based on their response to treatment (see Table3 in Datasets and Signature Translation, Appendix V).

3.9 Tumour Response

Tumour response was assessed by measuring any change in tumour size between the initial assessment at the time of diagnosis, and following two and 12 weeks of Letrozole therapy. Although calliper, mammography and three-dimensional USS measurements were all recorded, the latter is known to be the most accurate and reproducible and was therefore used as the primary determinant of tumour response in the present study ^[535].

In order to classify tumours by response, the change in tumour volume by three-dimensional USS was plotted against time, for each dataset (Figures 3.9A-C). For the purpose of this study, responders were classified as those whose tumour volume reduced by 50% or more at 60 days and 75% or more at three months. Non-responders were classified as those who failed to reach a reduction in tumour volume of at least 50% by the time of tumour excision at 3 months. All remaining patients were classified as intermediate responders and excluded from subsequent analysis in this study with the purpose of separating clear responders from clear non-responders.

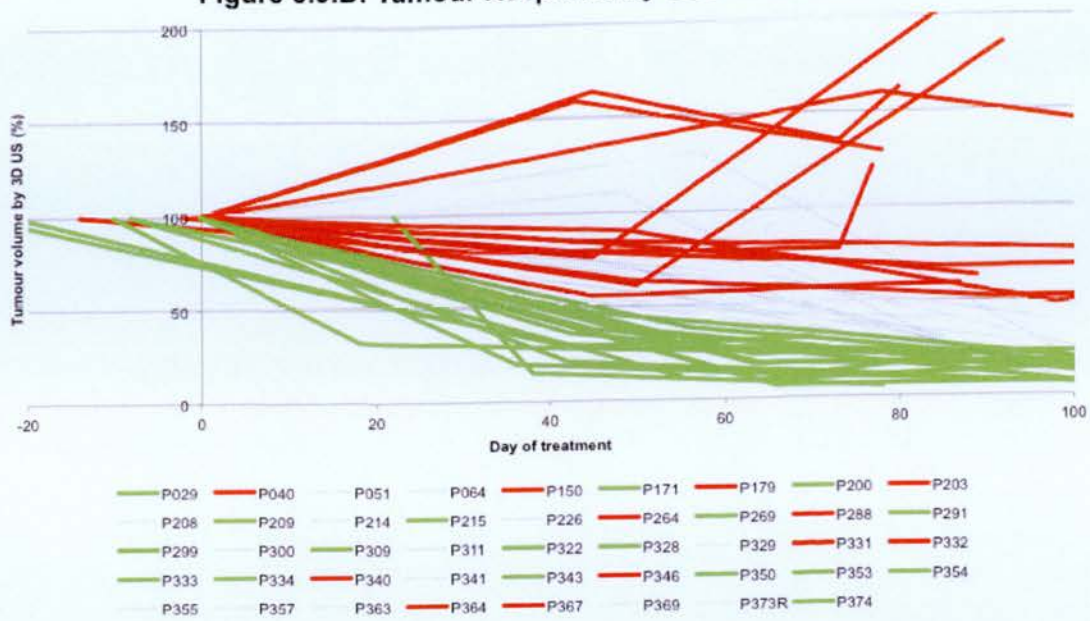
Of the 57 patients included in the L2 dataset, 24 (42%) were classified as responders and 10 (18%) as non-responders. Twenty-three patients (40%) were classified as intermediate responders and excluded from further analysis (Figure 3.9A).

Figure 3.9A: Tumour Response by USS in L2 series



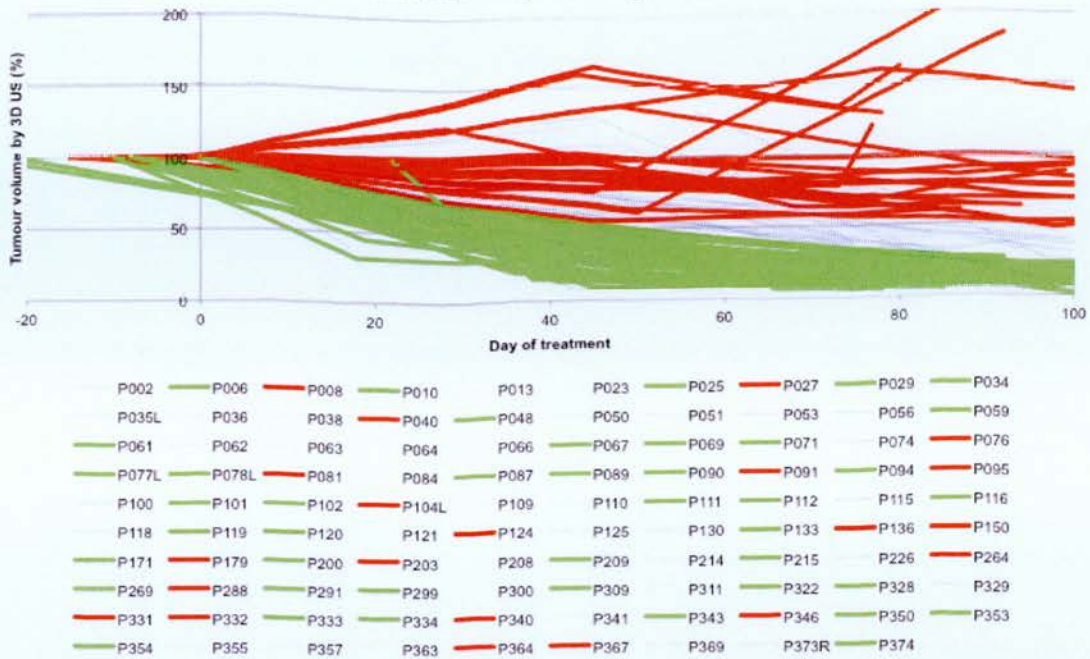
Similarly, in the L3 dataset 18 patients (41%) were clear responders and 12 (27%) non-responders. Fourteen intermediate response patients (32%) were excluded from further analysis (Figure 3.9B).

Figure 3.9.B: Tumour Response by USS in L3 Series



When the combined datasets were considered (L23), there were 42 (42%) responders, 21 (22%) non-responders and 36 intermediates (Figure 3.9C). The discrepancy in the count for intermediate and non-responders (where $L23 \neq L2+L3$), is due to the patients P103 (L2 non-responder) and P117 (L2 intermediate responder) who were included in L2 dataset (Miller 2007, Miller 2009 and Miller 2012) but eliminated from the L23 dataset (Turnbull 2012). This is illustrated in Figures 3.9A and 3.9C.

Figure 3.9C: Tumour Response by USS in L23 Series



3.10 Selection of Gene Signatures

The microarray analyses were used to validate established breast cancer signatures for their predictive power in these datasets. Other groups have published a large number of breast cancer signatures, though none were specifically designed in the neo-adjuvant setting at the time that the present study was undertaken. Therefore, after an extensive literature search, a few of these 'adjuvant' signatures were selected for evaluation in the present study on the basis of their anticipated relevance to the neo-adjuvant setting. Two signatures were selected for proliferation and two for stromal rearrangement, to ensure that the results related to the pathway rather than to any one specific signature.

In order to extend the basis for comparison, the efficacy of these signatures was also compared to that of the Edinburgh signature on an extended series of patients. This included the initial dataset published by Miller *et al.* (2009) and the subsequent L3 dataset.

The overlap of genes within the various signatures is represented in the Venn diagram in Figure 3.10A and listed in Table 3.10A. The full genes lists for each signature are provided in Appendix IV.

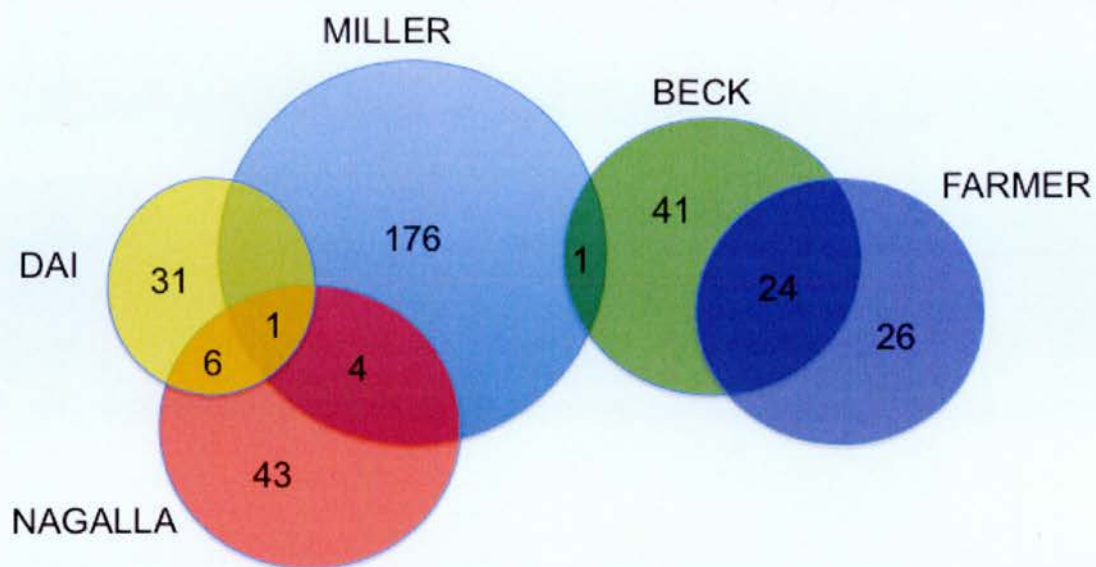


Figure 3.10A Representation of the numbers of overlapping genes within each of the signatures used in the present study. See Appendix IV for gene lists.

Beck Miller	Beck Farmer	Dai Nagalla	Nagalla Miller	Dai Miller
1	24	7	5	1
NBL1	PDGFRL CDH11 COL3A1 SPARC COL6A1 COL1A2 CSPG2 FBN1 ADAM12 AEBP1 ANGPTL2 COL5A2 MMP11 PLAU THBS2 SPON1 CTSK LOXL1 RARRES2 DPYSL3 MFAP2 OLFML2B POSTN ECM2	BIRC5 CCNB2 MAD2L1 CENPA CCNE2 BUB1 PRC1	KIAA0101 MAD2L1 RRM2 NUSAP1 MLF1IP	MAD2L1

Table 3.10A: Table showing number and lists of genes common to paired signatures.

3.10.1 Signature Translation

When assessing the efficacy of any particular gene expression signature within a population of samples, consideration has to be given to the microarray platform used to design the signature and that used to process the samples.

Affymetrix and Illumina microarrays differ in the selection, design and length of their oligonucleotide probes, which therefore may have variable affinity for different regions of any particular gene. In this way, probes intended to hybridise with the same gene may not do so equally effectively on either platform. In addition, the one-base mismatch probes of Affymetrix arrays act as a competitive hybridisation control, whereas the Illumina uses multiple copies of the same probe on each array to provide an internal non-competitive hybridisation control. Other differences such as the probe layout and the ability to process multiple arrays in parallel using Illumina, contribute to the potential difficulty in comparing data derived from either platform. For these reasons, some form of signature translation is necessary. Nevertheless the correlation between these two platforms in particular has been shown to be surprisingly good^[536].

Human genome sequence data are held in a number of open access online databases. The following are part of an international collaboration and exchange data on a daily basis:

- European Molecular Biology Laboratory (EMBL) Ensembl database.
- American National Centre for Biotechnology Information (NCBI) GenBank database.
- Japanese DNA DataBank of Japan (DDBJ).

The HUGO (Human Genome Organisation) Gene Nomenclature Committee is the only global authority assigning standardised nomenclature to human genes including unique gene identifiers (gene IDs) and short codenames or 'symbols' for each.

Database interrogation can be made directly but recent applications such as BioMart (www.biomart.org) can be used as a single online interface to access these multiple sources.

In the present study, the genes within each signature were translated into a "common language" using the databases described above and some further optimisation procedures summarised in Turnbull *et al.* ^[526]. In doing so, some genes from the original signatures chosen for use in this study, were excluded on account of their low cross-platform correlation and reliability. The methods used for the translation of each signature are described below, together with a description of the signature itself.

No translation was needed to apply the Nagalla and Miller signatures to L2 dataset as all were derived with the Affymetrix platform. However, signatures, which reported HGNC Gene Symbols were translated to the required namespaces using the Ensembl Biomart web interface (www.ensembl.org/biomart). In several cases, the initially reported gene symbols/IDs had to be manually verified to account for the nomenclature change since the date of publication. To convert Affymetrix signatures to Illumina and ENSG name-spaces we used the optimised translation table developed by Dr.A.Turnbull ^[526].

Translation always affected the number of features in the signature. Ensembl Biomart translation from HGNC to Affymetrix or Illumina namespaces usually increased the numbers of features, because of the multiple probes associated with the individual genes. Conversion from Affymetrix to Illumina namespace using Dr.Turnbull's translation tables usually reduced the size of signature because it included only probes directly comparable between Affymetrix and Illumina platforms. Of note, the effective number of features might be further reduced if some of the features were not available in the dataset. A summary of the signature translations is shown in Figure 3.10B.

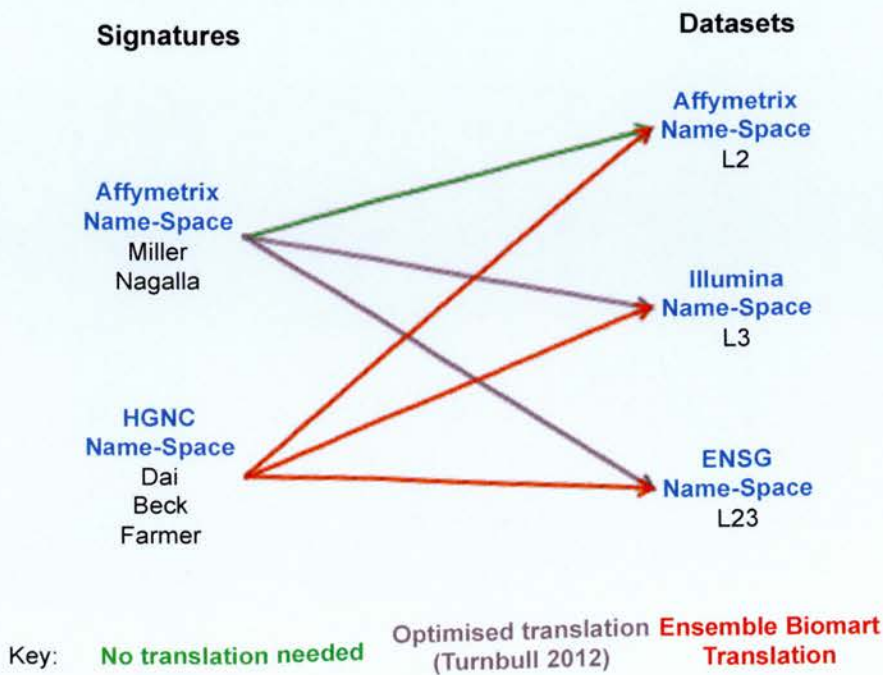


Figure 3.10B: Overview of Signature Translation. Signatures used for this analysis had different feature IDs: Miller's and Nagalla's signatures reported Affymetrix probe IDs, whilst the remaining signatures – Dai, Farmer and Beck - reported HGNC gene symbols. For the downstream analysis the signatures first had to be translated into the datasets' namespaces

3.10.2 The Edinburgh Signature – Miller 2009 ^[510]

The Edinburgh signature, which is the only signature generated from patients treated with endocrine therapy in the neoadjuvant setting, was chosen as a benchmark for validation and for further comparison with pathway focused signatures previously described and validated in the adjuvant setting.

Miller's signature included genes both positively and negatively associated with response. The association with response for the current analysis was obtained from Supplementary Figure 3 of the original publication. The signature was derived using the Affymetrix arrays and translated into Ensembl and Illumina IDs using an optimised procedure described by Turnbull *et al.* ^[526]. The noticeable loss of genes during translation reflects the exclusion of "non-reliable" probes and IDs, which had not passed the translation procedure.

Namespace	Total	Unique Affymetrix IDs	Unique NGNC symbols	Notes
AFFY	205	182	201	Original namespace
ENSG	74	67	64	
LUMI	143	67	70	

3.10.3 Proliferation Signature – Dai 2005 ^[471]

The first proliferative signature chosen for this study was that derived by Dai *et al.* The signature composition is illustrated in Figure 2 and in Supplementary Table 3 of the original publication. It includes a series of 50 genes that were highly correlated with outcome in a population of 311 patients with relatively high ER positivity for their age (all were under 55 years). Almost all of these genes were found to relate to the cell cycle and their over-expression correlated with very poor outcome ^[471]. The 50 genes were presented as HGNC Gene symbols and Gene-bank IDs.

A manual search for each symbol and ID was conducted using Ensembl and NCBI databases online. Only 40 of the 50 genes could be identified. The genes that could not be identified had been described in the initial signature in a non-conventional way (Contig57584_RC, Contig38901_RC, Contig41413_RC, Contig31288_RC, Contig64688, Contig44289_RC, Contig28552_RC, Contig46218_RC, Contig28947_RC and Contig38288_RC). Two of the remaining 40 probes described the same gene and were therefore combined (AURKA: STK15/ STK6 = NM_003600/ NM_003158). The resulting 39 HGNC gene symbols, all of which were associated with cellular proliferation (35 positively, 4 negatively) were translated to Ensembl, Illumina (HT12) and Affymetrix (HG U133A) gene IDs using the Ensembl BioMart online interface.

Namespace	Total	Unique NGNC symbols	Notes
HGNC	39	39	Original namespace
ENSG	40	39	
LUMI	64	39	
AFFY	63	37	

3.10.4 Proliferation Signature – Nagalla 2013 ^[525]

Nagalla *et al.* analysed gene expression profiles from almost 2000 breast tumours to identify three gene signatures (termed metagenes) associated with distant metastasis-free survival, each of which related to different immunological functions. Among this subset, tumours were stratified into high, medium and low proliferation groups by a further proliferative signature. The prognostic strength of each immune metagene was dependent on proliferation and greatest amongst highly proliferative tumours, where increased levels of expression were associated with a markedly reduced risk of metastasis; suggesting that a subset of immune responders exists among highly proliferative tumours, with potential prognostic ramifications ^[525].

The metagene used in this study was obtained from supplementary Table 6 of the original publication, and consists of 61 paired Affymetrix IDs, accompanied by 54 corresponding unique gene symbols. The noticeable loss of genes during translation reflects the exclusion of “non-reliable” probes and IDs, which had not passed the strict translation procedure. All genes were positively associated with proliferation. Translation to Illumina IDs was performed using the procedure of Turnbull *et al.* ^[526].

Namespace	Total	Unique Affymetrix IDs	Unique NGNC symbols	Notes
AFFY	61	61	54	Original namespace
ENSG	26	26	26	
LUMI	39	26	26	

3.10.5 Stromal Signature – Farmer 2009 ^[496]

Farmer *et al.* identified a stromal gene expression signature that predicted resistance to neoadjuvant chemotherapy using FEC. The predictive value of this signature was subsequently validated in two independent cohorts of patients who received chemotherapy, but not in an untreated control group, which was taken to indicate that the signature was predictive rather than prognostic. These 50 genes were all associated with a reactive stroma, which is rich in activated fibroblasts, or CAF's ^[496]. Again, all genes were considered positively associated with the stromal component.

The stromal metagene used in this study was obtained from Supplementary Table 4 in the original publication and includes 50 HGNC gene IDs that were translated to Ensembl, Illumina (HT12) and Affymetrix (HG U133A) gene IDs using the Ensembl online interface.

Namespace	Total	Unique NGNC symbols	Notes
HGNC	50	50	Original namespace
ENSG	51	50	
LUMI	77	48	
AFFY	100	50	

3.10.6 Stromal Signature – Beck 2008 ^[524]

Beck *et al.* developed the second stromal signature selected. This group initially demonstrated that the gene expression patterns in two types of soft tissue tumour, Desmoid Type Fibromatosis (DTF) and solitary fibrous tumours, could be used to identify distinct stromal reaction patterns in a subset of breast tumours ^[537]. They have since identified 66 DTF associated genes, consistently expressed in 25-35% of breast tumours, among a population of 561 tumours, the expression of which is associated with low grade, ER positivity and improved survival. After validation on a separate subset of tumours, this signature has been recognised as a robust predictor of clinical response associated with improved survival outcome ^[524].

The DTF Core Gene Set was obtained from Supplementary Table 1 in the original publication and includes 66 HGNC gene IDs. All were identified on the Ensembl or NCBI web sites and subsequently translated to Ensembl, Illumina (HT12) and Affymetrix (HG U133A) gene IDs using the Ensembl online interface. All genes were considered as positively associated with the stromal phenotype.

Namespace	Total	Unique NGNC symbols	Notes
HGNC	66	66	Original namespace
ENSG	68	66	
LUMI	101	66	
AFFY	124	64	

Gene Ontology

In order to assess the possible existence of proliferation genes within the stromal signatures, Gene Ontology (GO) analysis was performed using the online tool at www.omicslab.genetics.ac.cn/GOEAST ^[538]. The results indicate that neither of the stromal signatures was enriched by any GO term associated with promotion of cell division or proliferation, suggesting that interference between the proliferation and stromal signatures studied here is unlikely.

The full list of detected GO terms is tabulated in Excel and attached in Appendix V. No GO terms associated with proliferation were detected in Beck's stromal signature. On the other hand, the Farmer signature may be enriched by genes associated with the negative regulation of proliferation in the prostate (GO:0060770, GO:0008285).

3.11 Signature Interpretation

3.11.1 Incorporating Dynamic Measurements

The datasets from the present study include pre-treatment tumour profiles as well as those from early treatment samples taken at two weeks. This allows analysis of the signatures' performance in three ways:

- (i) On pre-treatment (baseline) data only.
- (ii) On early-treatment (two week) data only.
- (iii) On the change between these time points.

Such an approach was not necessary for the Miller signature, because it is comprised of separate gene-lists for all three of the above.

3.11.2 Sign of Association

Most of the selected signatures included only genes "positively" associated with the phenotypic feature. For example, the expression of all of the genes included in the Nagalla signature increased in high-proliferation cases. Some other signatures however, included genes both positively and negatively associated with the phenotypic feature. For example, the Dai signature includes some genes whose expression is increased and others decreased in highly proliferative tumours. The sign of association was therefore taken into account when calculating the multigene scores as described in the next section.

3.11.3 Heat Maps

Following pre-processing of the microarray raw data, 'heat maps' were generated as visual representations of gene expression. The intensity of the colours on each spot of the map reflects the quantity of RNA hybridisation and therefore the level of gene expression (Figure 3.11A).

On these heat maps, the x-axis corresponds to the samples within the studied dataset and the y-axis to the genes within the signature being assessed. The colour and intensity of each spot on the matrix represents the sign of association and level of gene expression respectively. Green represents down-regulated genes and red represents up-regulated genes. When assessing the expression of a whole signature for a particular sample, the intensity data, as represented in the heat map, can be converted into a single score for each sample.

3.11.4 Multi-Gene Expression Scores

Signature performance was evaluated using a multi-gene expression score adapted from Bedard ^[539]. This effectively represents the average gene expression within the signature, taking into account the sign of association. In order to equalise the relative contribution of each gene within the signatures and to minimise platform-specific biases, each gene expression was Z-scaled before summation. The following formula summarises the calculation:

$$\text{Gene Expression Score} = \frac{\sum_i W_i X_i}{\sum_i |W_i|}$$

Where W_i is the total expression value of each gene within the signature, X_i is either +1 for up-regulated expression or -1 for down-regulated expression and $|W_i|$ is the absolute value of W_i .

When the resultant scores are plotted in order of increasing value, a graph such as that in Figure 3.11B is produced, with individual tumour samples on the x-axis and the multi-gene expression score on the y-axis. The points on the graph are colour coded into clear responders (green) and long-term non-responders (red) based on their USS measurements (see section 3.9).

Gene expression scores were calculated in this way for each of the signatures against each of the three datasets (L2, L3 and L23). The heat maps and graphs are available in the relevant sections below.

Individual Genes Within the Signat.re

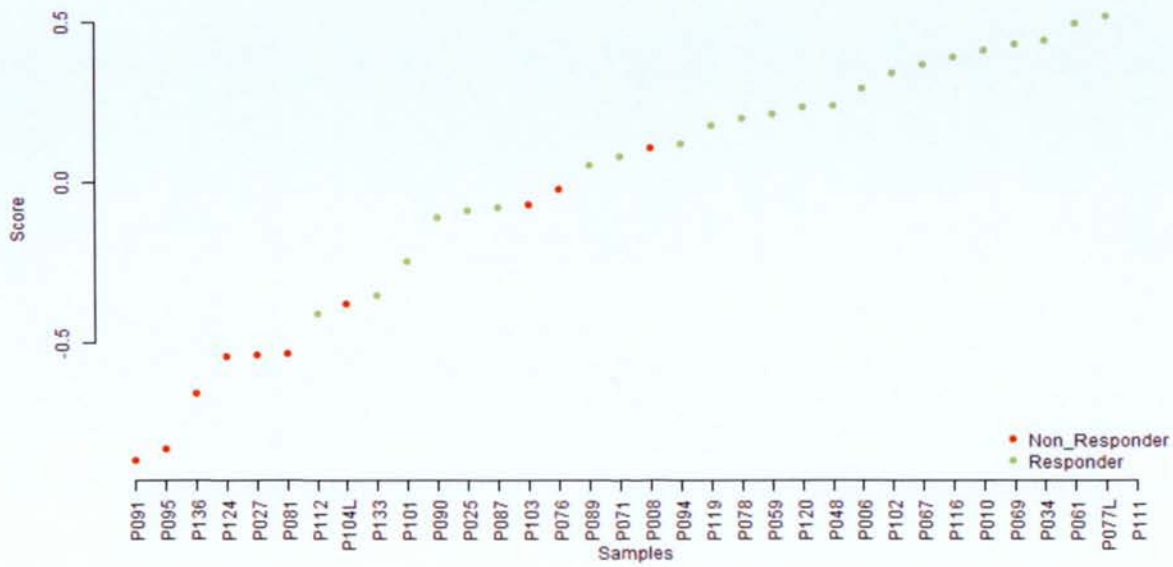
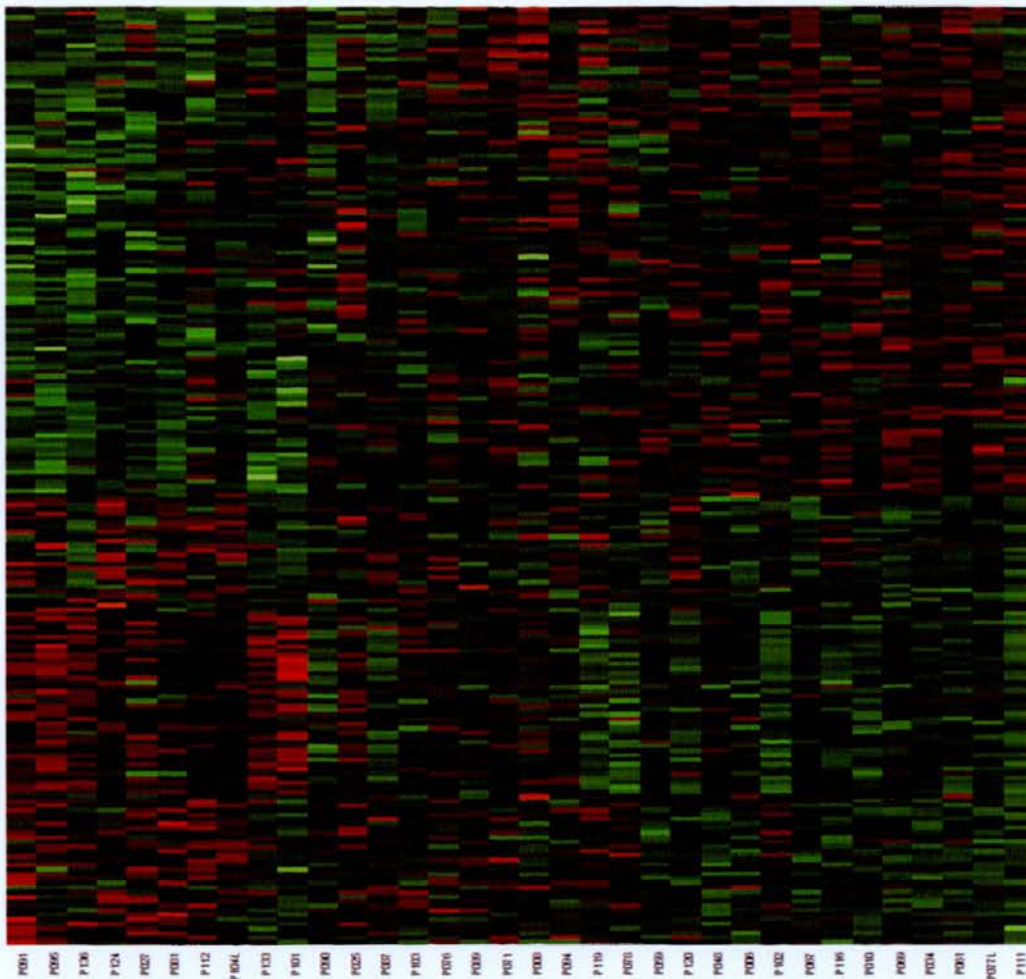


Figure 3.11A (top) & B (below): Corresponding Heat Map and Multi-gene Expression Score Chart for the Edinburgh Signature and Response for Samples from the L2 series.

3.11.5 Receiver Operating Characteristic (ROC) Curve Analysis

In a ROC curve the sensitivity (true positive rate) of a test is plotted as a function of one minus the specificity (false positive rate) for different cut-off points of a parameter. Thus each point on a ROC curve represents a sensitivity/specificity pairing which corresponds to a particular threshold setting. ROC analysis is commonly used as a statistical test of association between the signature and response which alleviates the potential 'over-fitting' of results which can occur when vast amounts of data are generated from relatively few samples.

As illustrated in Figure 3.11C below, the best possible predictive signature would yield a point in the top left corner of the ROC curve, representing perfect sensitivity (no false negatives) and perfect specificity (no false positives). On the other hand, a completely random signature would give a point along a diagonal line from the bottom left to top right corners of the ROC curve, known as the line of no discrimination. This line divides the ROC space into two equal triangles. Points above this line represent a positive correlation of each signature to tumour response (better than random), i.e. a signature with good predictive value. Points below the line represent the opposite association or signatures with negative predictive value. However, it should be noted that consistently negative predictors could be inverted to obtain a positive predictor. Thus only an AUC of 0.5 indicates to a lack of association.

In the present study, calculations were performed using the pROC R package. 95% confidence intervals (CI) of the AUC were computed with 2000 stratified bootstrap replicates^[540]. The association was considered significant if the AUC CI did not include the value 0.5.

The combination of multi-gene scores with ROC analysis has been successfully used in several studies of gene expression signatures in breast cancer^[496, 539]. The advantage of this approach is its simplicity and transparency whilst preserving an equivalent performance to that of many alternative and more complex methods^[541]. Furthermore, this approach circumvents the need to select a specific threshold for signature comparison because the area under the ROC curve represents an integrative estimate that is independent of any one specific threshold.

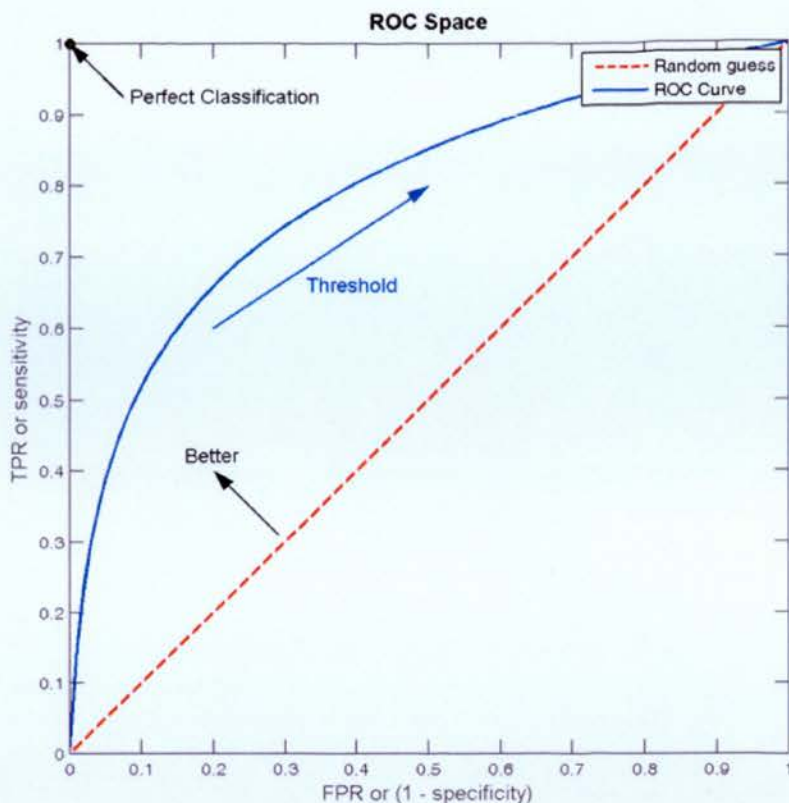


Figure 3.11C: Model of the Receiver Operating Characteristic Curve.

The area under the ROC curve (AUC) reflects how well the parameter, in this case the given gene signature, functions to distinguish between the two groups - responders and non-responders.

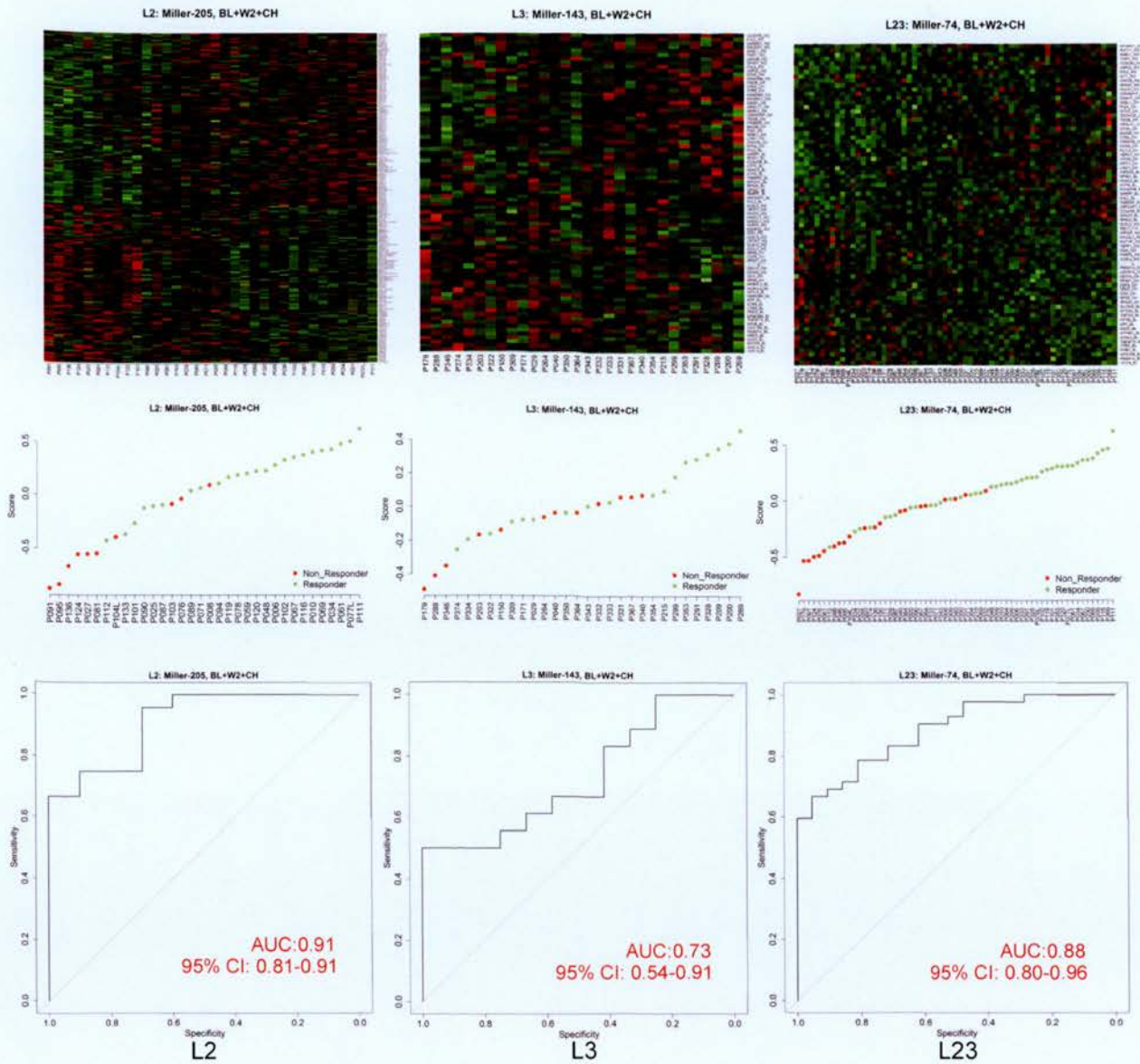
3.12 Validation of the Edinburgh Signature

The predictive capacity of the Edinburgh signature has been evaluated in all three datasets using multi-gene scoring and ROC curves as described above. The results are shown on Figure 3.12A.

Using the L3 series, it was confirmed, for the first time, that this signature has significant association with neo-adjuvant response to Letrozole on an independent dataset. However, the assessment of the Edinburgh signature on the L2 (and L23) datasets is not informative in terms of prediction because the signature was derived from samples within this series and would therefore be expected to perform well. Nevertheless, it is a useful benchmark for comparison of the performance of the other signatures in this regard.

The results show that the ability of the Edinburgh signature to predict response to neoadjuvant Letrozole therapy is statistically significant in all of the datasets. This is the first time that this signature has been validated on independent (L3) and extended (L23) datasets.

Figure 3.12A: Performance of Miller Signature on L2, L3 and L23 Datasets



Key: Dataset: Signature-Number of useable genes for each dataset: Time point (BL=Baseline, W2= Week 2, CH = change between BL and W2). *Note that the Miller signature incorporates all three timepoints whereas the other signatures were assessed at each time point as shown in the figures below.* The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. For example, Miller-143 indicates that 143 genes have been retained in the signature after the translation to the L3 namespace. However some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. An overview of the signature translation is shown on Figure 3.10B and discussed in the section 3.10.2 above. The exact composition of translated signatures is shown in tables 4-18 of the Datasets and Signature Translation file, Appendix V.

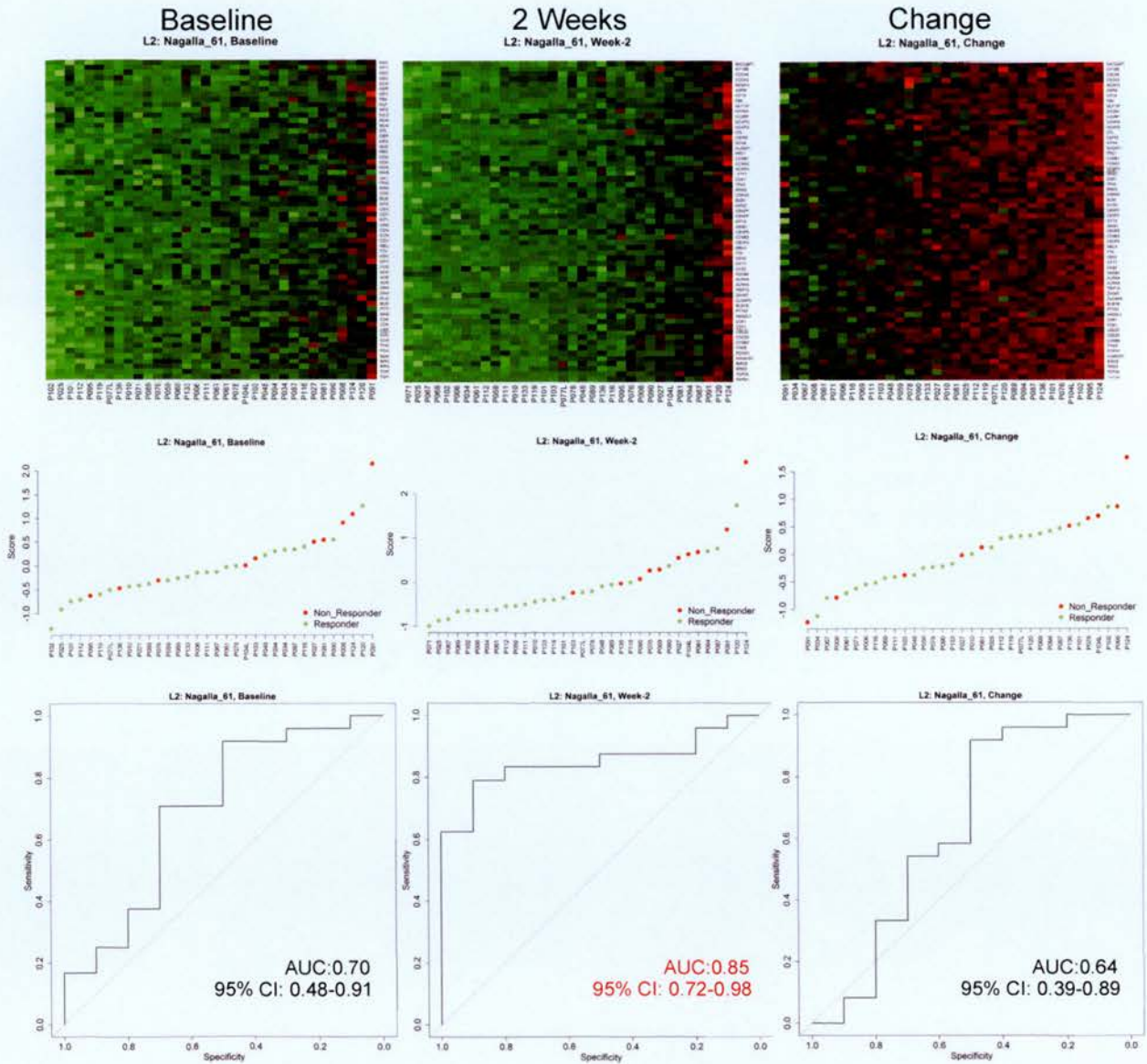
In Figure 3.12A, the heatmaps (top panels), multigene expression scores (middle panels) and ROC curve analyses (lower panels) are shown, indicating the performance of the Edinburgh signature in each of the three datasets. In each case, the AUC CIs do not include 0.5, indicating a significant association. Not surprisingly, the signature performed very well against the L2 dataset from which it was derived and the extended L23 dataset. The reproducibility of the result with the L3 dataset however, represents the first time this signature has been validated against an entirely new dataset. Note that separate analyses for baseline, two-weeks and change were not carried out as for the other signatures below because the Miller signature incorporates these already.

3.13 Performance of Proliferative Signatures

Both of the independent proliferative signatures chosen for this study were designed for use in response prediction to endocrine therapy in the adjuvant setting. The results of this study nevertheless confirm their significant predictive application in the neoadjuvant setting for the first time.

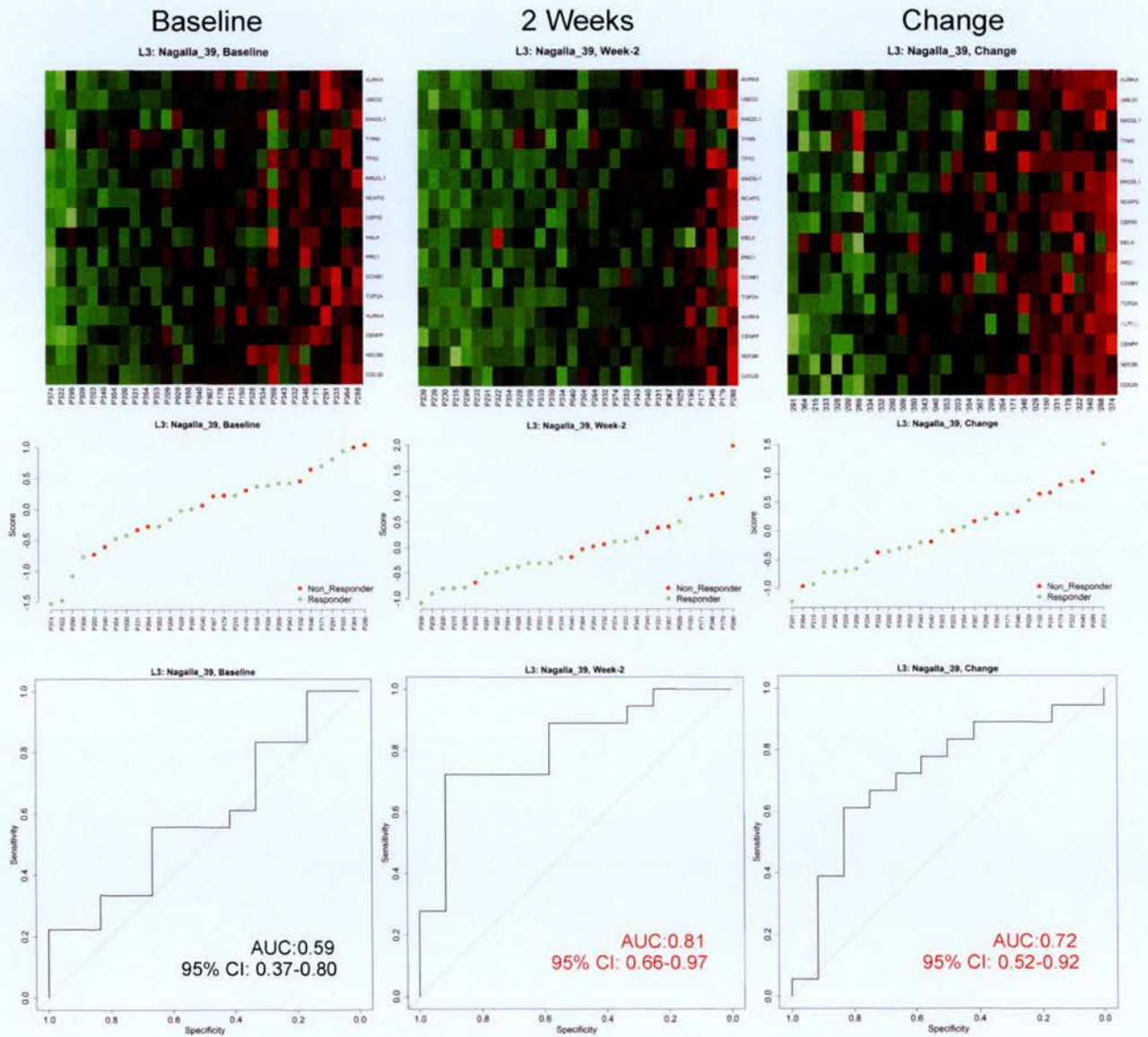
Significance was demonstrated at the two-week time point, but not at baseline for both signatures on all three datasets (Figures 3.13A-F). In addition, the Nagalla signature alone, demonstrated significant predictive power, comparable to that of the Edinburgh signature, when changes in gene expression between baseline and two weeks were considered in the L3 and L23 datasets.

Figure 3.13A: Performance of Nagalla Signature on L2 Dataset



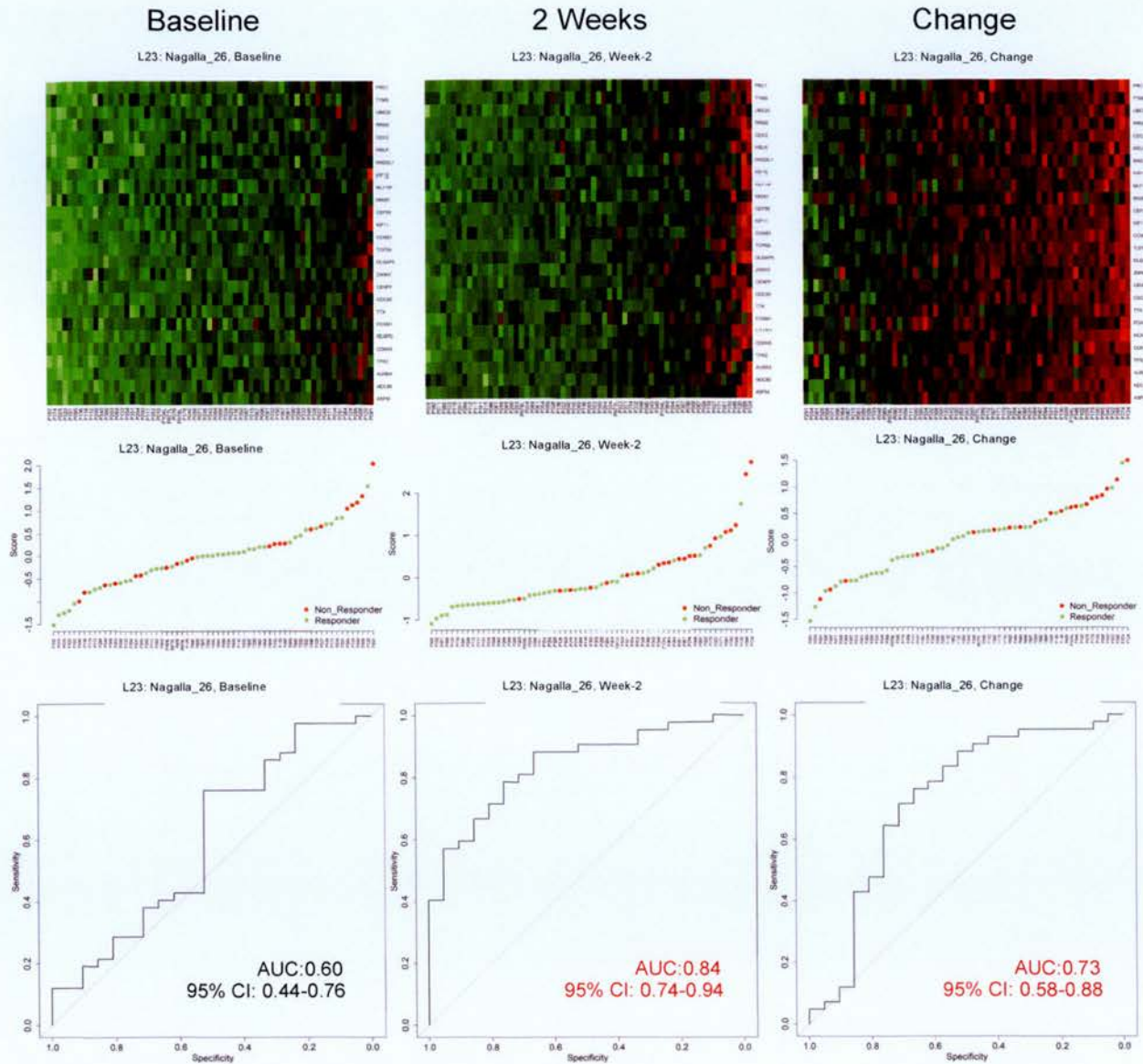
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.4.

Figure 3.13B: Performance of Nagalla Signature on L3 Dataset



Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.4.

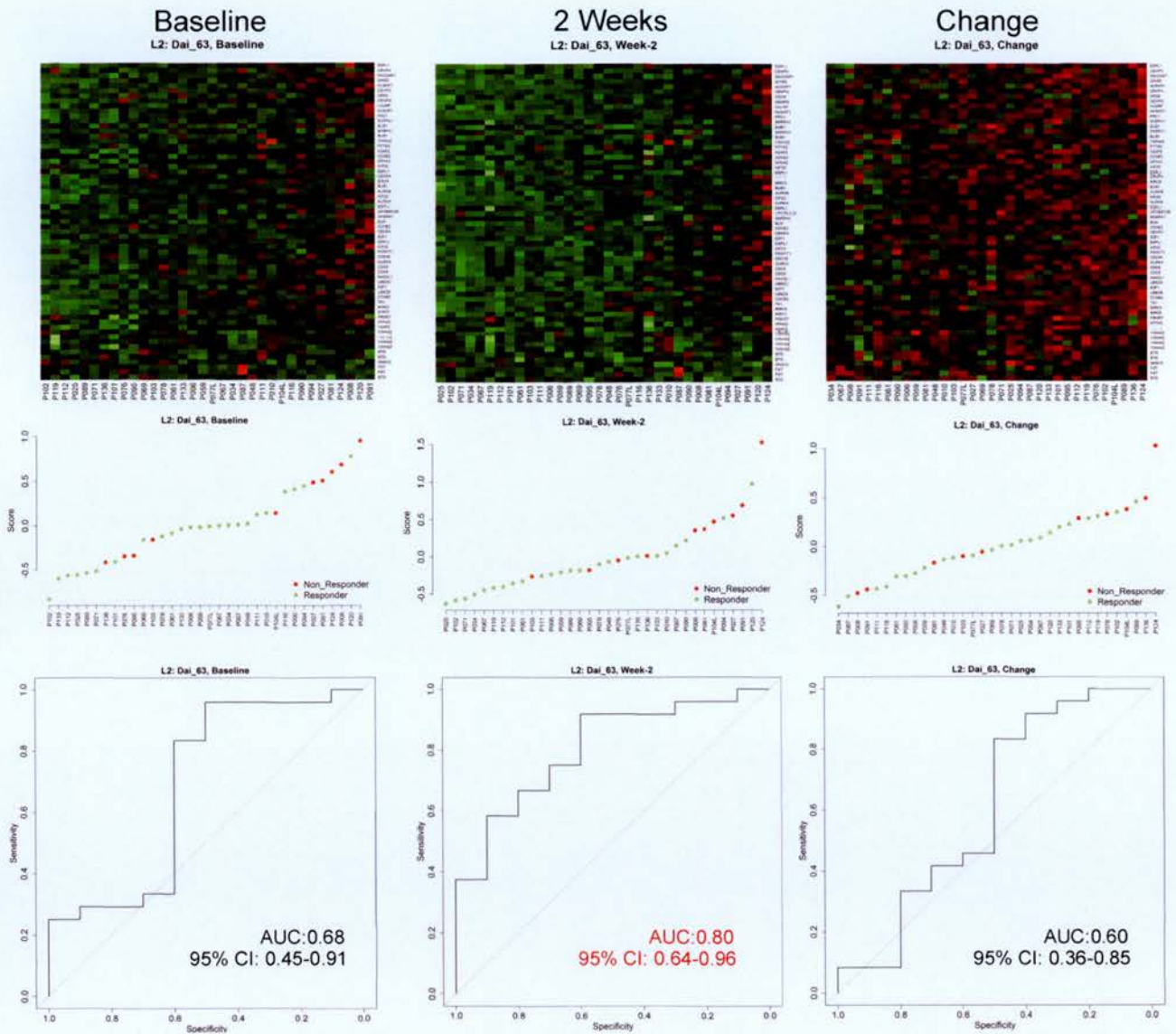
Figure 3.13C: Performance of Nagalla Signature on L23 Dataset



Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.4.

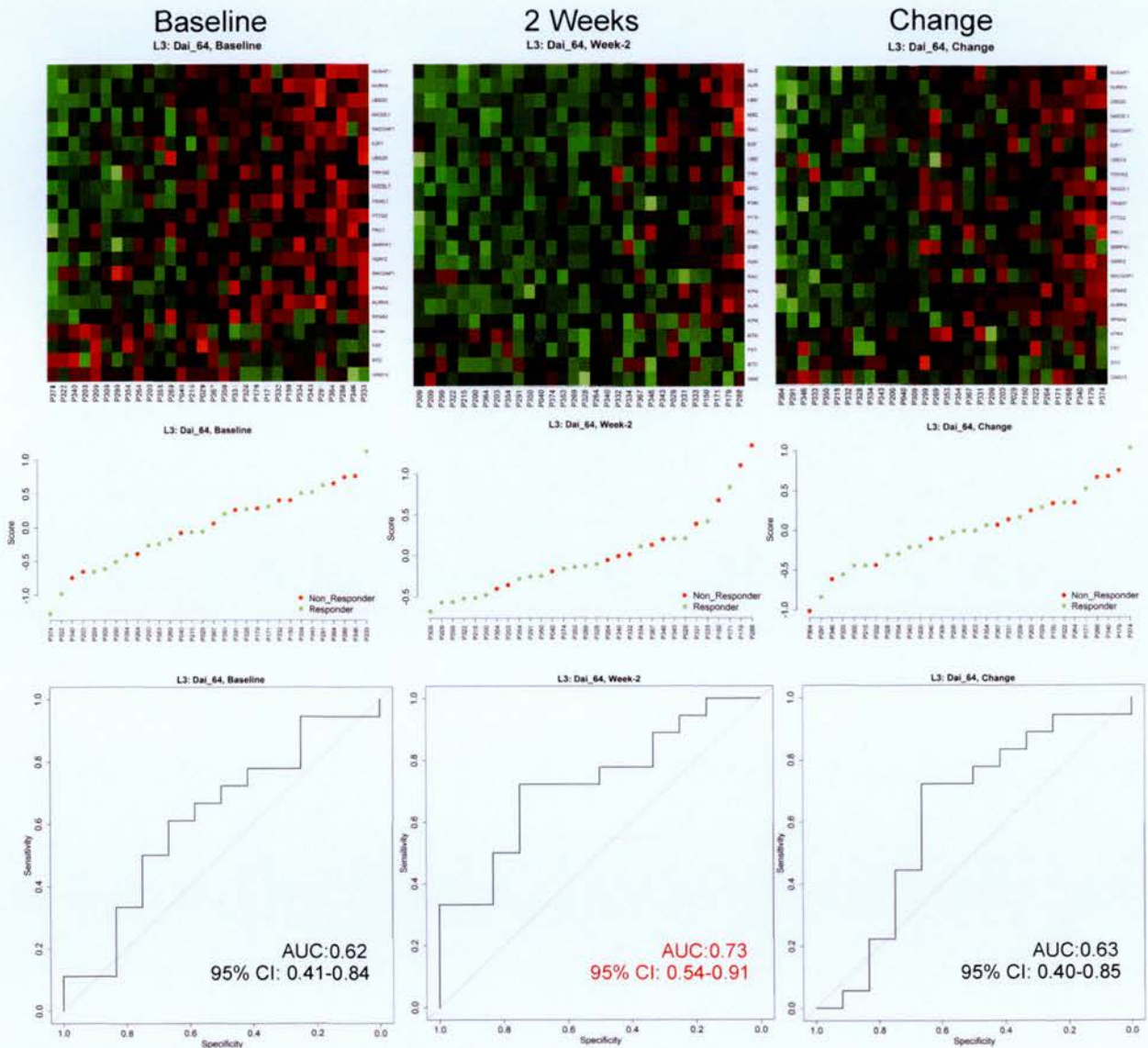
The three figures above (3.13A-C) show the heatmaps, multigene expression scores and ROC curve analyses indicating the performance of the Nagalla (proliferative) signature in each of the three datasets. In each case, the AUC CIs indicate a significant association at the two-week time point and in the case of the L3 and L23 datasets there is an also a significant association in relation to the change between baseline and two-week gene expression.

Figure 3.13D: Performance of Dai Signature on L2 Dataset



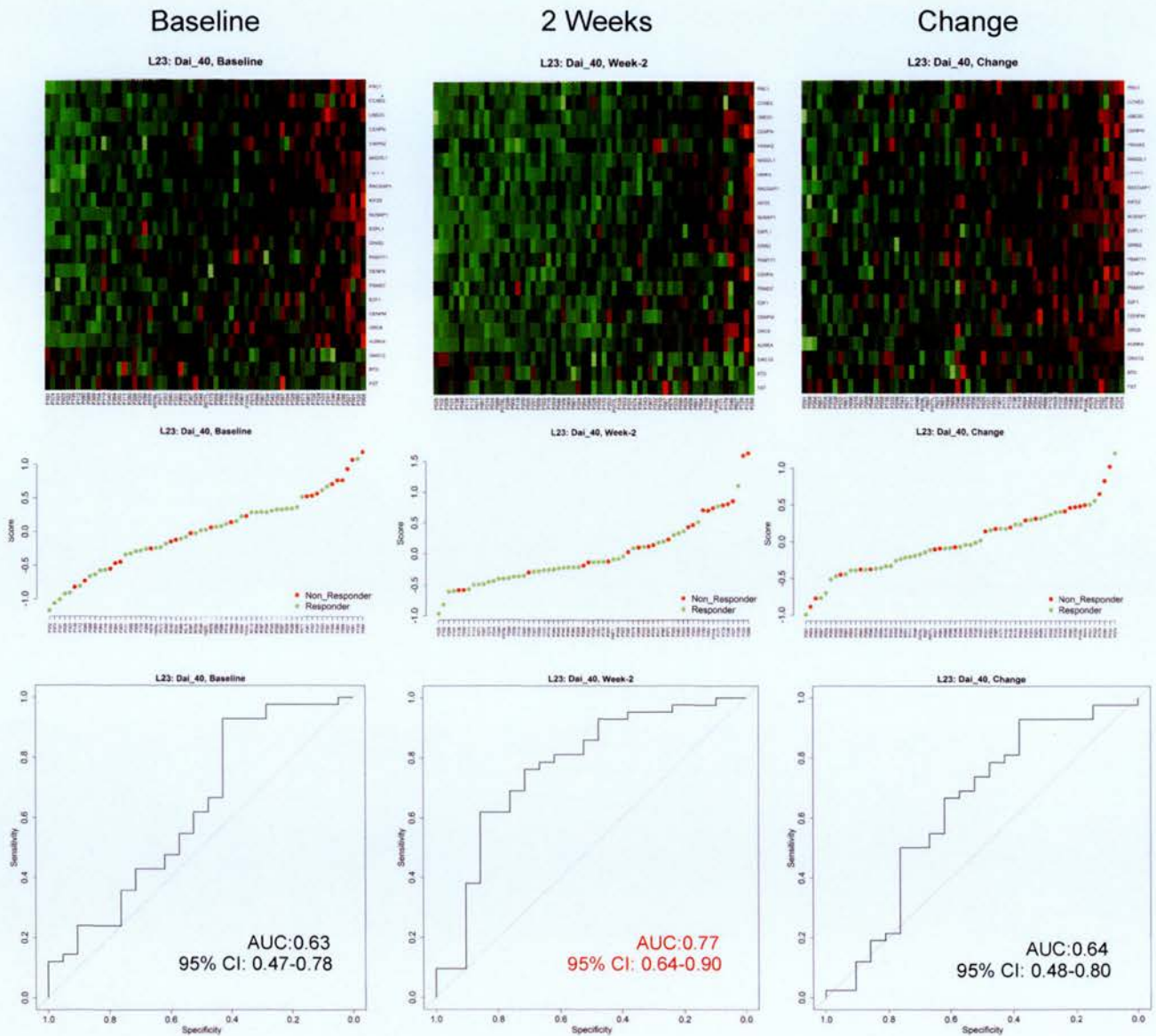
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.3.

Figure 3.13E: Performance of Dai Signature on L3 Dataset



Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.3.

Figure 3.13F: Performance of Dai Signature on L23 Dataset



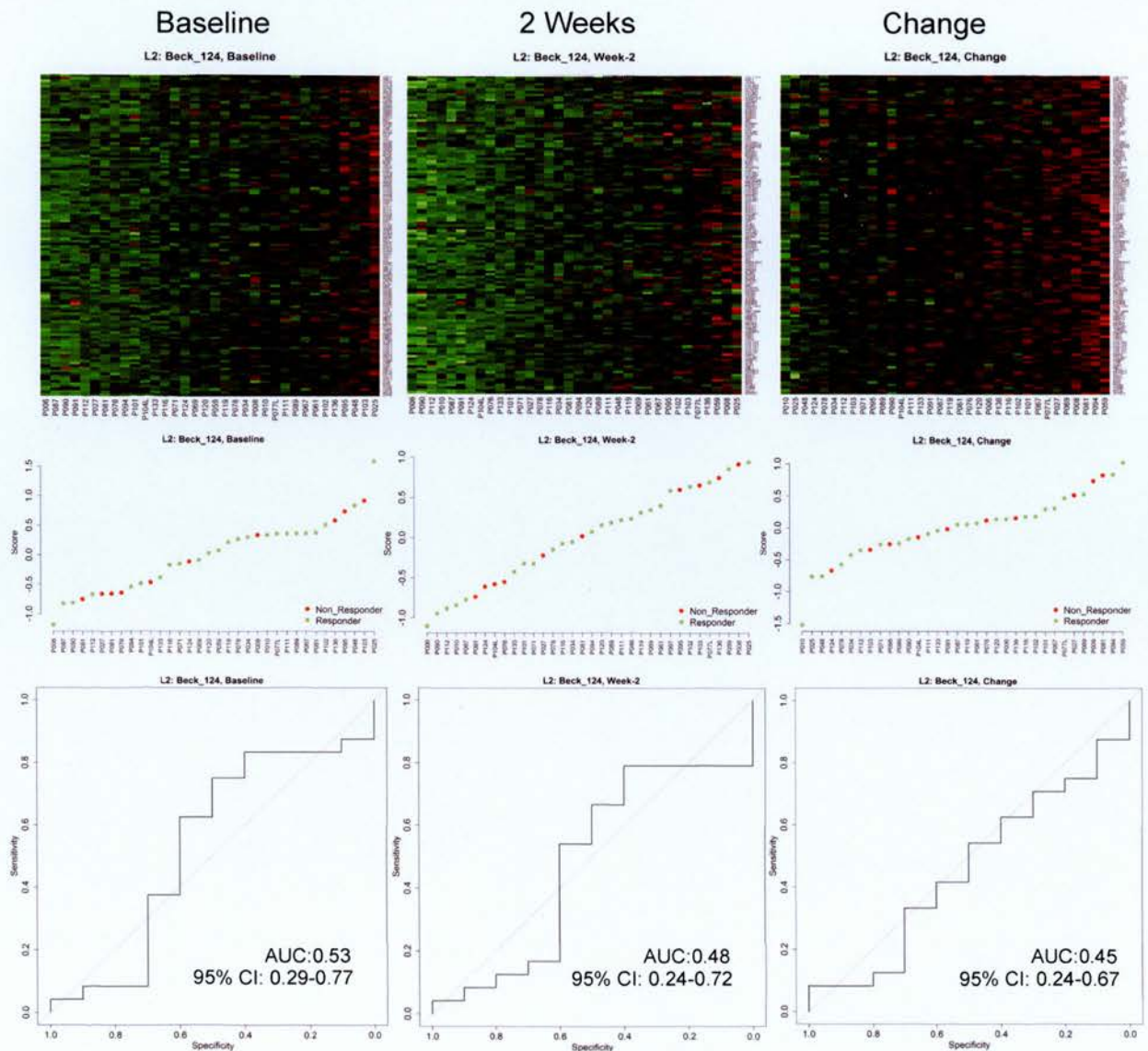
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.3.

Figures 3.13D-F above, show the heatmaps, multigene expression scores and ROC curve analyses, indicating the performance of the second proliferative signature used in this study (Dai) against each of the three datasets. As was seen with the Nagalla signature, the Dai signature AUC analyses also indicate a significant association at the two-week time point. However this association is not seen in relation to the change in gene expression levels between baseline and two weeks, which was not significant.

3.14 Performance of Stromal Signatures

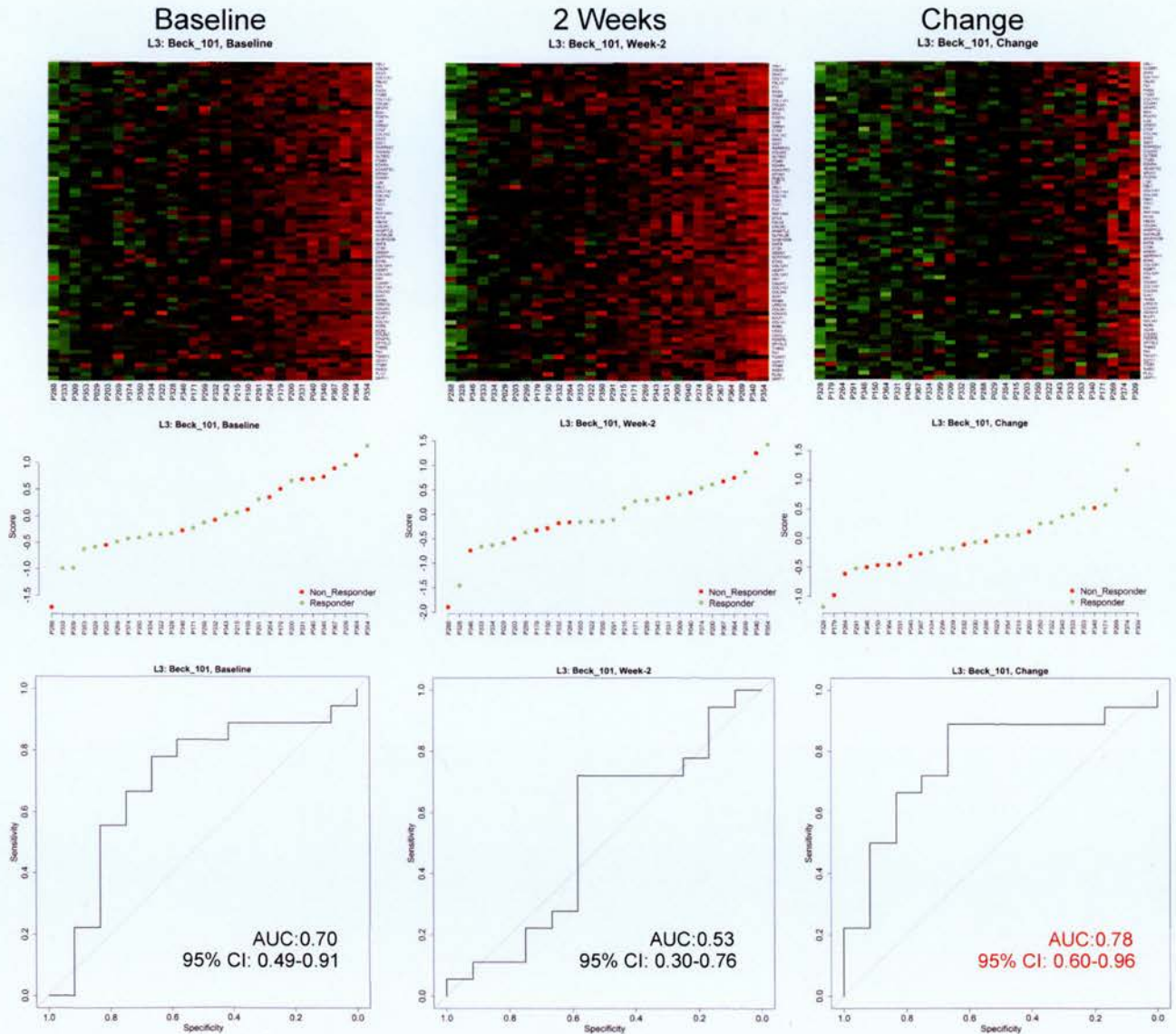
In contrast to the performance of the proliferative signatures, neither stromal signature was significantly informative of response to neoadjuvant Letrozole in the baseline or two-week samples. Both however showed significant predictive power with respect to changes in gene expression between baseline and two weeks in the L3 dataset only. The figures below show the results for the Beck (Figures 3.14A-C) and Farmer (Figures 3.14D-F) signatures.

Figure 3.14A: Performance of Beck Signature on L2 Dataset



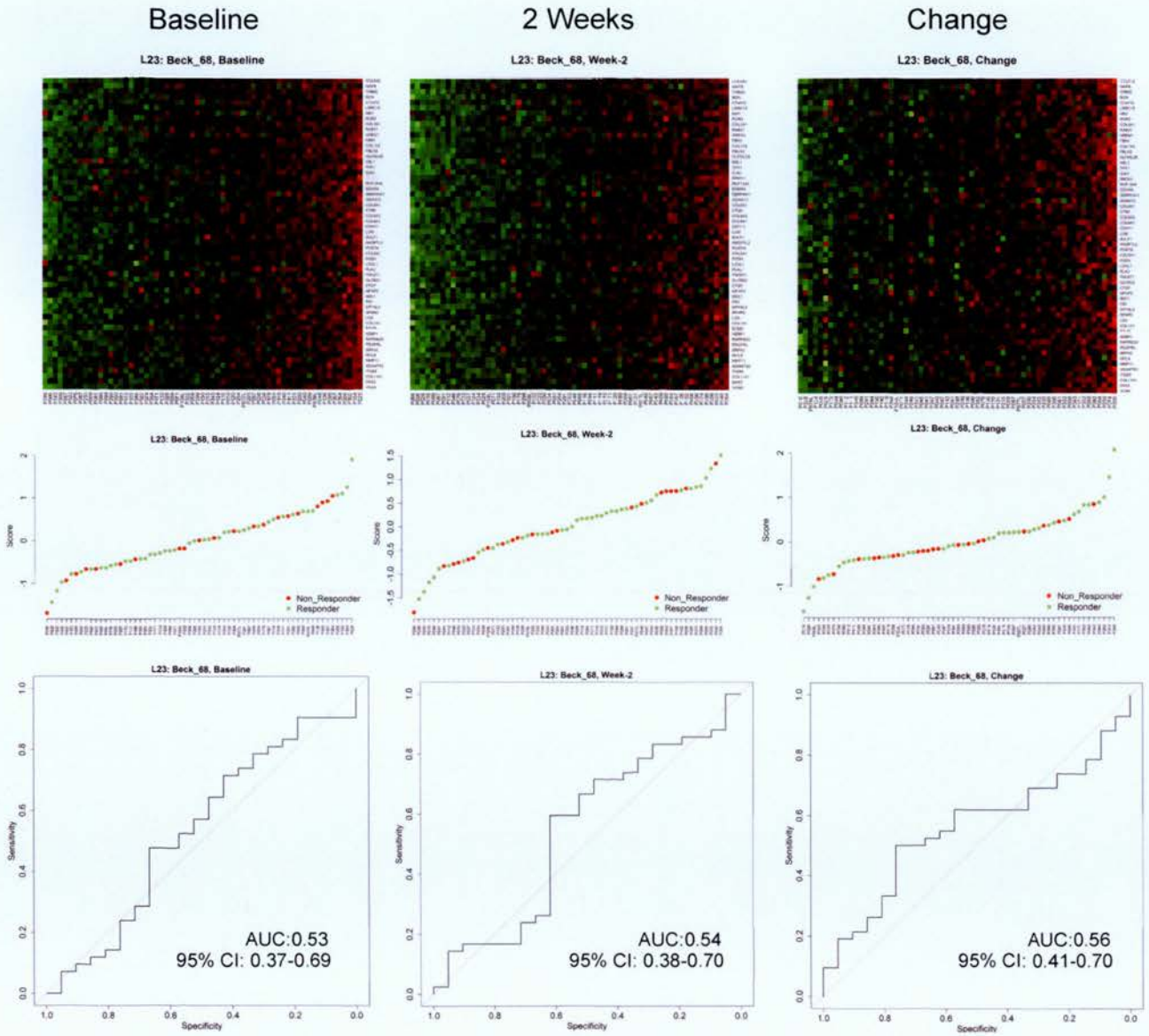
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.6.

Figure 3.14B: Performance of Beck Signature on L3 Dataset



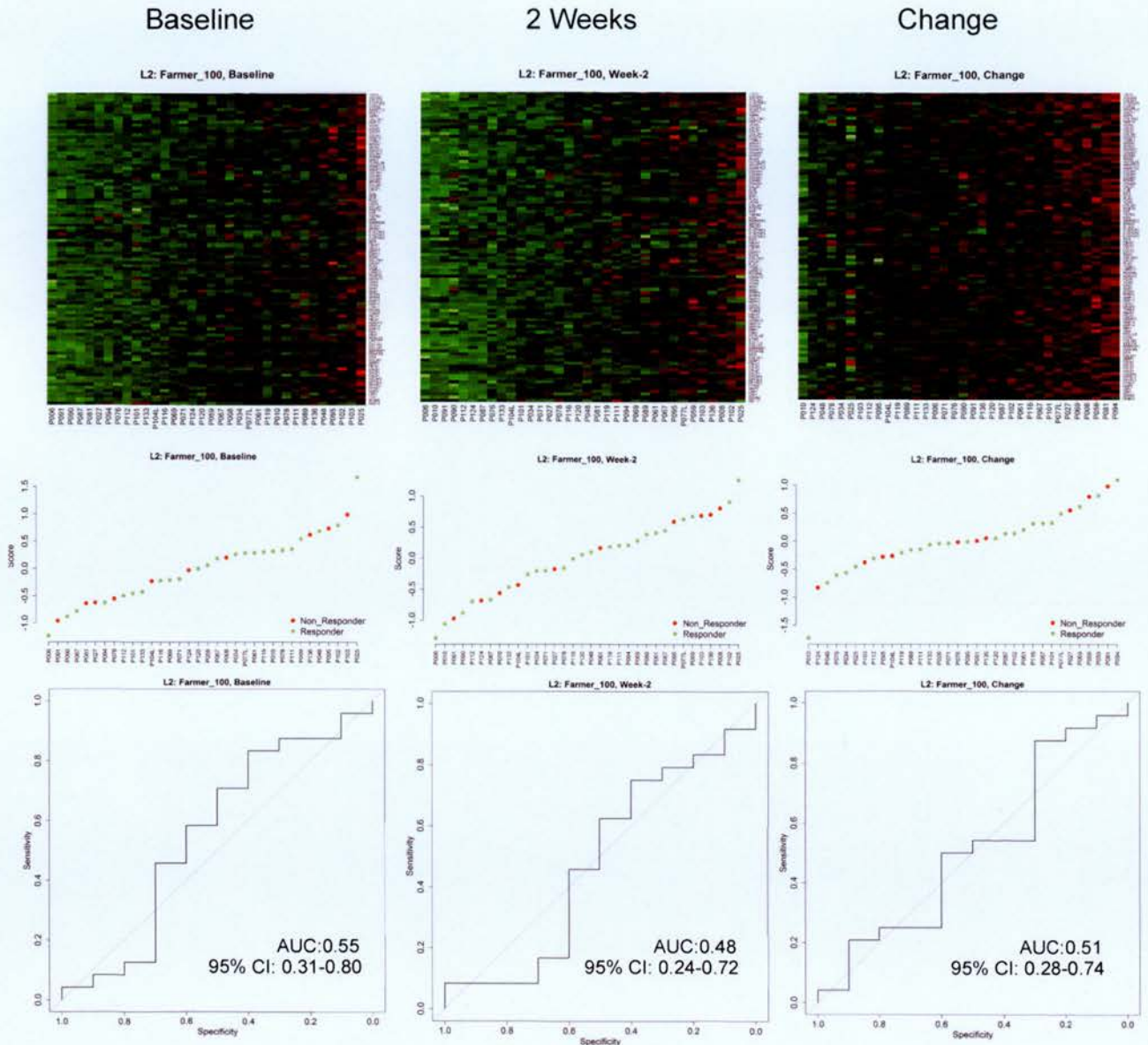
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.6.

Figure 3.14C: Performance of Beck Signature on L23 Dataset



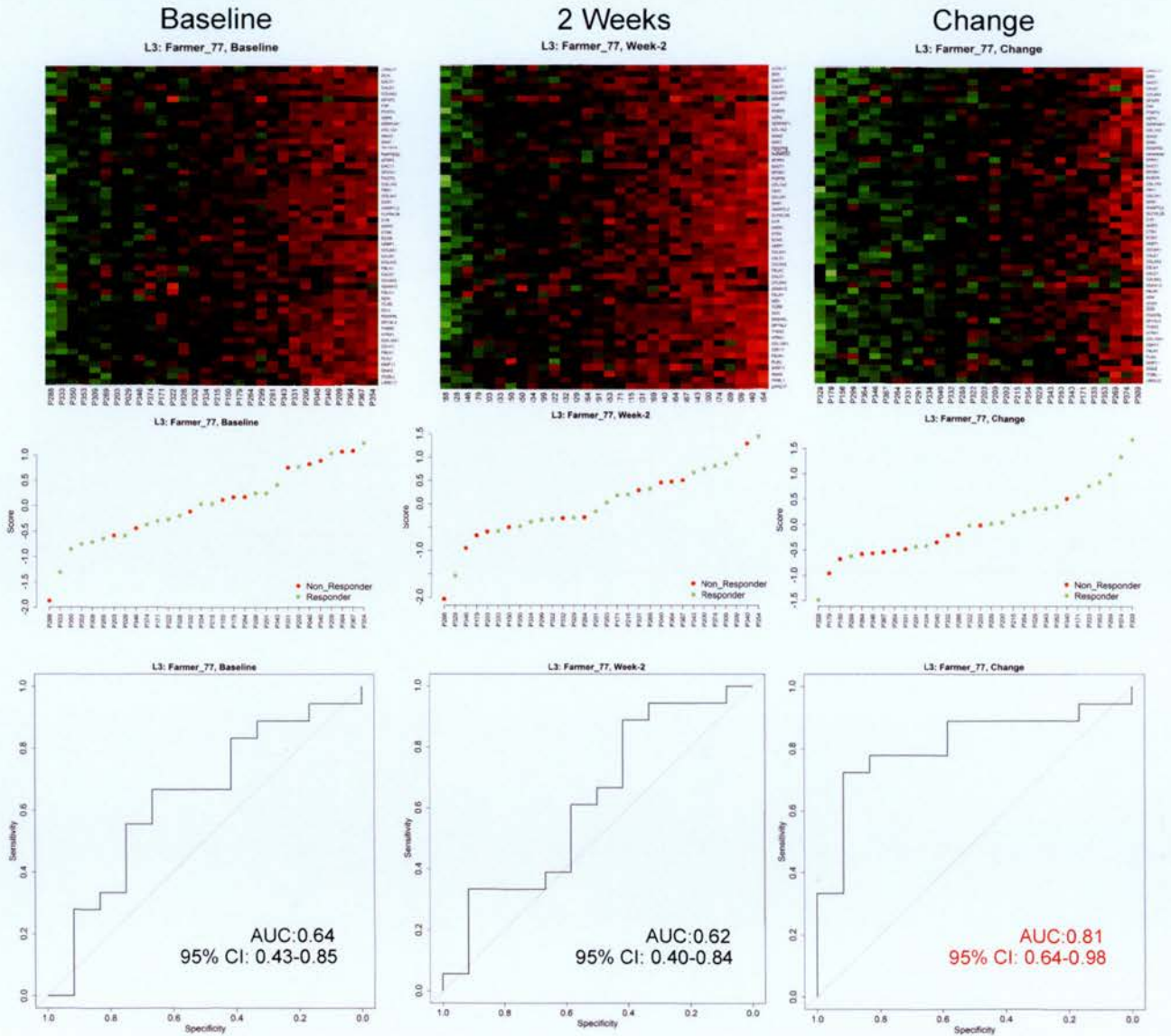
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.6.

Figure 3.14D: Performance of Farmer Signature on L2 Dataset



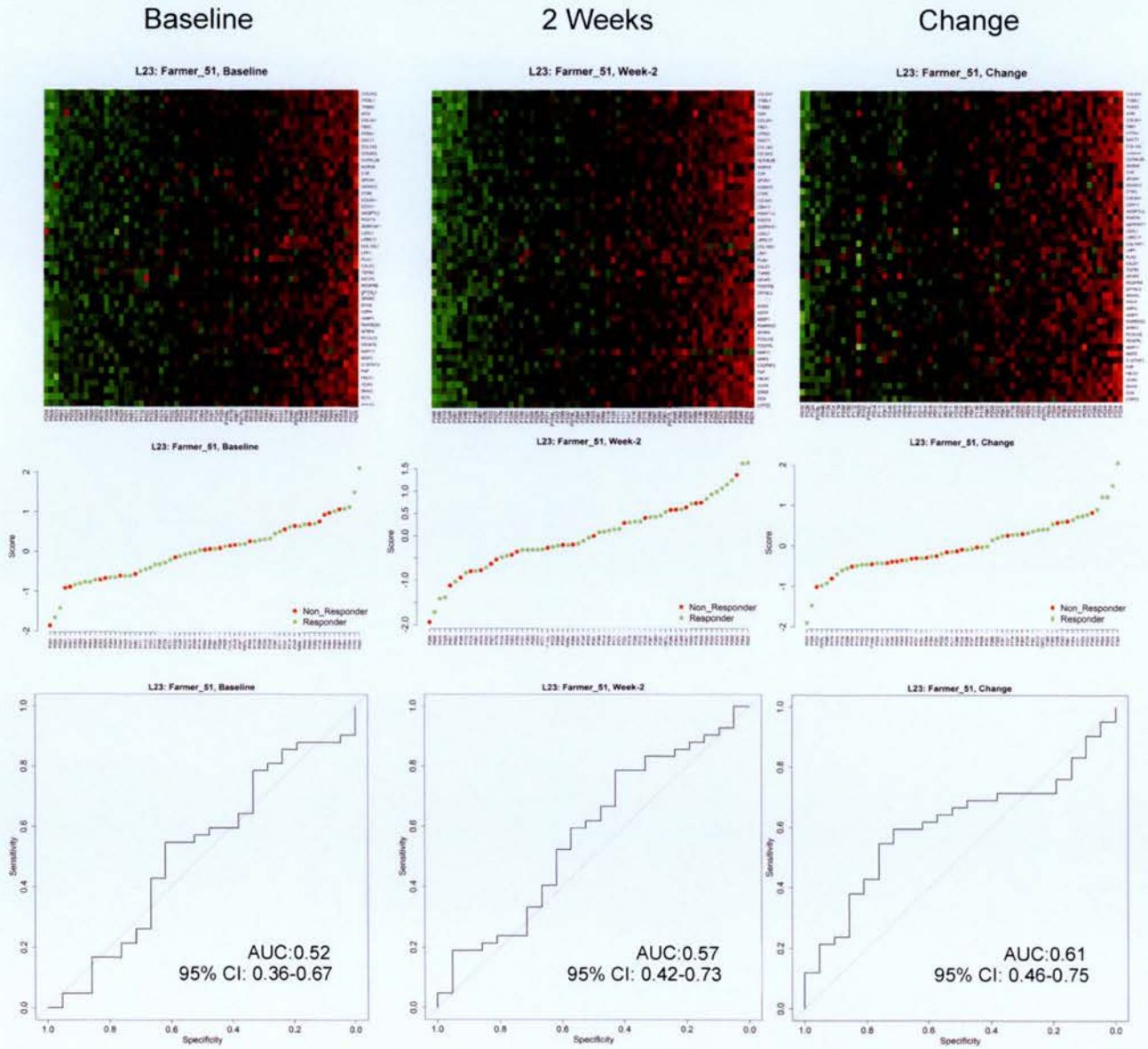
Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.5.

Figure 3.14E: Performance of Farmer Signature on L3 Dataset



Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.5.

Figure 3.14F: Performance of Farmer Signature on L23 Dataset



Key: Dataset (L2 or L3 or L23): Signature Name-Number of genes left in signature after translation: Time point (Baseline, Week 2, Change between Baseline and Week 2). The number shown after the signature name indicates the number of genes remaining in the signature after translation to the appropriate namespace. Some genes were then lost through pre-processing (e.g. poorly performing or annotated probes) and so the number of genes shown on the heatmaps corresponds to those actually used. See section 3.10.5.

3.15 Signature Comparison

The area under the curve was calculated for each ROC curve with 95% confidence intervals and plotted into bar charts. These are shown in Figures 3.15A-C below which correspond to the data derived from the L2, L3 and L23 datasets respectively. These summarise the main results from this study. In each chart, the significant results are highlighted in red and the non significant in blue.

Figure 3.15A represents the signature comparison in the L2 dataset. Not surprisingly the Edinburgh signature shows extremely high predictive power on the L2 dataset with the ROC AUC confidence interval reaching 1. This reflects the fact that this signature was derived from this dataset and this association had to be confirmed within an independent dataset. With respect to the other signatures, only the proliferation signatures showed significant association with response and only at the two-week time point. Their performance at two weeks nonetheless approximated that of the Edinburgh signature.

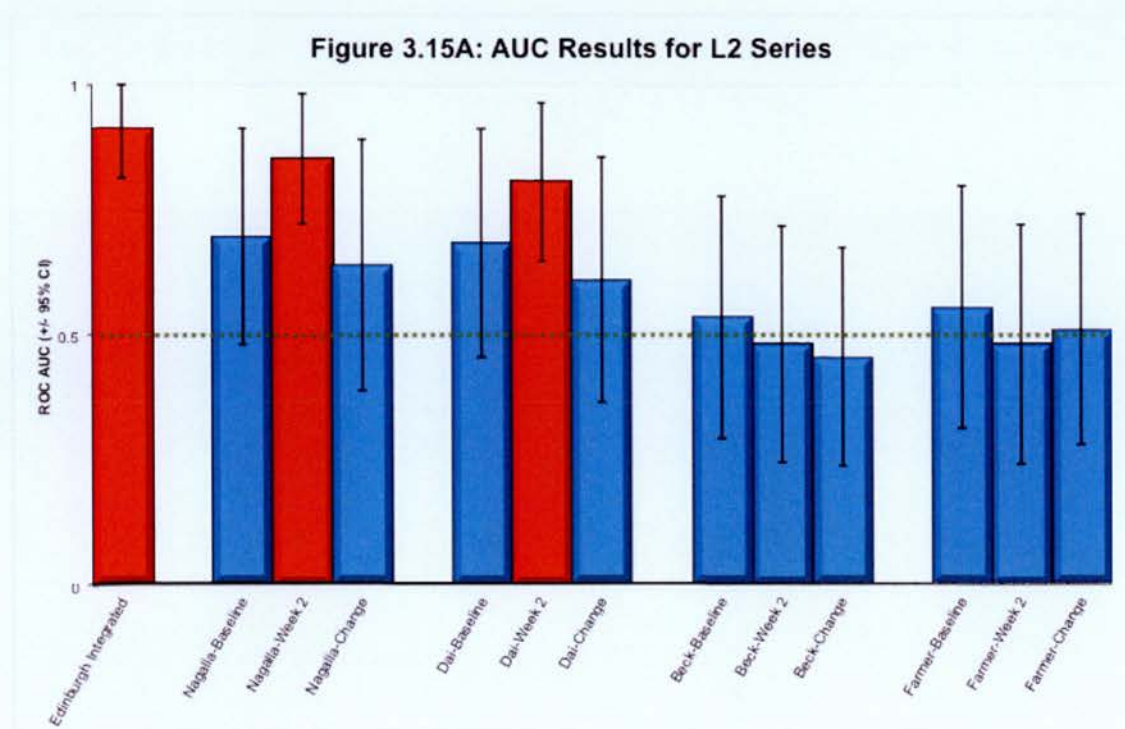


Figure 3.15B illustrates the performance of the studied signatures on the L3 series and confirms the above findings with respect to the Edinburgh signature in a new and independent dataset. Both proliferative signatures were similarly informative at the two-week time point but only the Nagalla signature was significantly informative with respect to change in gene expression. Interestingly, these results also demonstrate that stromal signatures may be informative with respect to change in gene expression.

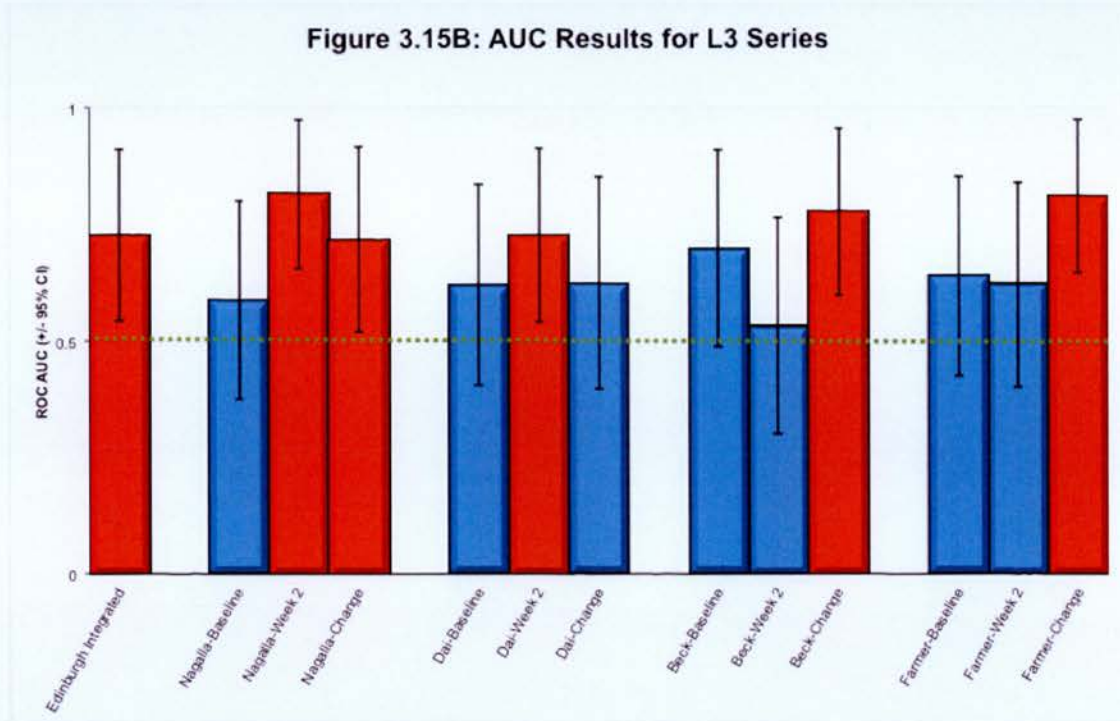
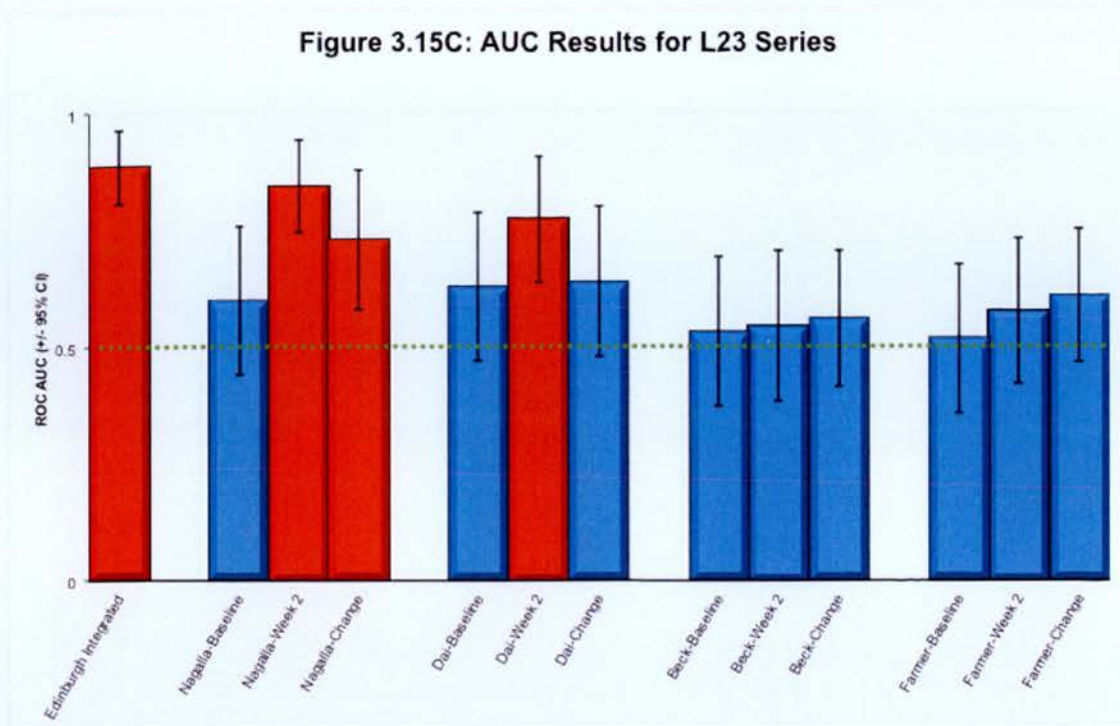


Figure 3.15C compares signature performance on the extended L23 series. As this includes the L2 dataset, it would be anticipated that the Edinburgh signature remains highly informative of treatment response, whilst the proliferative signatures perform similarly to that seen in the L2 and L3 datasets. However, in this larger series, the significance of the stromal signatures is lost.

Figure 3.15C: AUC Results for L23 Series



3.16 Quantitative PCR Validation of Microarray Data

Quantitative PCR was used for the validation of microarray results using genes selected from samples of the L2 dataset, as previously published [508]. When gene expression changes between baseline and the two-week samples were stratified for their magnitude of change (>1.5 fold), consistent direction of change and the statistical significance of that change, a total of 143 genes were identified of which 91 were down-regulated and 52 up-regulated by the Letrozole therapy. Only nine genes were common to all three stratifications.

For validation purposes, three up-regulated and three down-regulated genes were selected for qPCR analysis based on their presence in at least two of the three stratifications described above. These were COLEC12, MMP2 and DCN (up-regulated) and CCNB1, CDC2 and NUSAP1 (down-regulated). Primer sequences for the qPCR analysis were designed for each gene as described in Section 3.7.2.

The results of these comparative analyses are shown in Table 3.16A (and were published in Miller *et al.*^[508]). The data show that the direction, amplitude and significance of the changes were comparable between qPCR and microarray. The latter was measured using the Wilcoxon signed rank test, a non-parametric analysis of repeated measurements on a single sample; the P-values are given in the table. The correlation between the microarray and qPCR quantification of gene expression changes was positive and highly significant when compared using Spearman's rank test. The ρ value represents the correlation coefficient where a value of 1 reflects maximal correlation and zero represents no correlation. P-values for the significance of the correlation are also given. These results therefore validated the use of microarray analysis for the quantitative evaluation of gene expression in this dataset. This additional verification step is not often undertaken or reported in other publications of this nature.

Gene	Microarray			Quantitative PCR			Correlation	
	Direction of Change	Mean Fold Change (95% CI)	Wilcoxon paired test (p)	Direction of Change	Mean Fold Change (95% CI)	Wilcoxon paired test (p)	Rho (ρ)	P
CCNB1	↓	1.9 (1.6-2.1)	2×10^{-8}	↓	2.8 (1.8-3.7)	2×10^{-7}	0.55	7×10^{-6}
CDC2	↓	2.2 (1.8-2.5)	1×10^{-8}	↓	3.8 (2.8-4.8)	7×10^{-8}	0.60	7×10^{-7}
NUSAP1	↓	2.3 (2.0-2.6)	4×10^{-10}	↓	3.2 (2.5-3.8)	2×10^{-8}	0.41	2×10^{-3}
COLEC12	↑	1.6 (1.5-1.8)	6×10^{-9}	↑	2.6 (1.5-3.7)	2×10^{-8}	0.52	4×10^{-5}
MMP2	↑	1.7 (1.4-2.0)	9×10^{-6}	↑	3.2 (2.2-4.3)	1×10^{-6}	0.65	3×10^{-8}
DCN	↑	1.7 (1.4-2.0)	2×10^{-6}	↑	3.0 (1.4-4.6)	5×10^{-5}	0.44	5×10^{-4}

Table 3.16A: Comparative Analysis of Microarray and qPCR on Selected Genes.

3.17 Changes in Cyclin and CDK Expression during Neoadjuvant Letrozole Treatment

The AIs are known to reduce cellular proliferation in the majority of breast cancers and this reduction is associated with a clinical response to treatment. For this reason, a focused sub-analysis was carried out to investigate the molecular events underlying the anti-proliferative effect of AIs, specifically Letrozole, in the neoadjuvant setting. The expression of selected genes representative of the main phases of the cell cycle was measured using qPCR, as described above, in the L2 dataset samples, collected at baseline and at 2-weeks and 3-months of treatment.

The selected genes included Cyclins D1, A2, B1, CDKs 1, 2, 4 and the NUSAP1 gene (a mitotic spindle associated protein). In addition, proliferation was assessed by immunohistochemical staining for Ki67 and expressed as a percentage of cells staining positively. Changes in the expression of the Cyclins/CDKs were compared to changes in Ki67 positivity and to clinical response, classified as tumour shrinkage >50% at 3 months.

At the time of this sub-analysis, sequential biopsies were available for 62 tumours of the L2 series (42 for the CDK4 assay), clinical response was available in 56 of these cases and Ki67 change was assessed in 53 cases. The results of this analysis are summarised in Table 3.17A and in Figure 3.17A.

Marker	% Reduction		Correlation to Ki67		Association with response	
	14 days	3 Months	14 Days	3 Months	14 Days	3 Months
	Mean Reduction (CI, n)		Pearson correlation p (r, n)		T-test p (n responders / non-responders)	
Ki67	61% (47-74, 43)	73% (63-83, 53)	NA	NA	0.05 (29 / 9)	0.05 (32 / 11)
NUSAP1	46% (30-63, 56)	49% (30-69, 55)	<0.001 (0.75, 38)	<0.001 (0.50, 48)	0.007 (32 / 14)	0.01 (32 / 11)
Cyclin B1	42% (32-52, 55)	51% (39-64, 56)	0.004 (0.46, 38)	<0.001 (0.55, 48)	0.001 (32 / 13)	0.03 (32 / 12)
Cyclin A2	42% (27-58, 56)	51% (37-65, 56)	<0.001 (0.69, 38)	<0.001 (0.56, 48)	0.01 (32 / 14)	0.01 (32 / 12)
Cyclin D1	35% (26-44, 56)	45% (35-54, 55)	0.101 (0.27, 38)	0.015 (0.35, 47)	0.10 (32 / 14)	0.47 (31 / 12)
CDK1	48% (34-61, 55)	50% (32-68, 56)	<0.001 (0.64, 38)	<0.001 (0.51, 48)	0.01 (31 / 14)	0.03 (32 / 12)
CDK2	16% (5-27, 56)	14% (3-25, 56)	0.006 (0.44, 38)	0.019 (0.34, 48)	0.006 (32 / 14)	0.02 (32 / 12)
CDK4	6% ([-4]-16, 41)	16% (5-26, 42)	0.796 (-0.05, 29)	0.118 (0.26, 37)	0.08 (27 / 8)	0.45 (27 / 7)

Table 3.17A: Changes in Selected Cell-cycle Associated Genes During Treatment with Letrozole

Values in the table show the following:

- % of reduction: mean reduction (CI of mean, n);
- Correlation to Ki67: Pearson product-moment correlation, p (r, n);
- Association with response: t-test for difference between responders and non-responders, p (numbers of responders / non-responders)

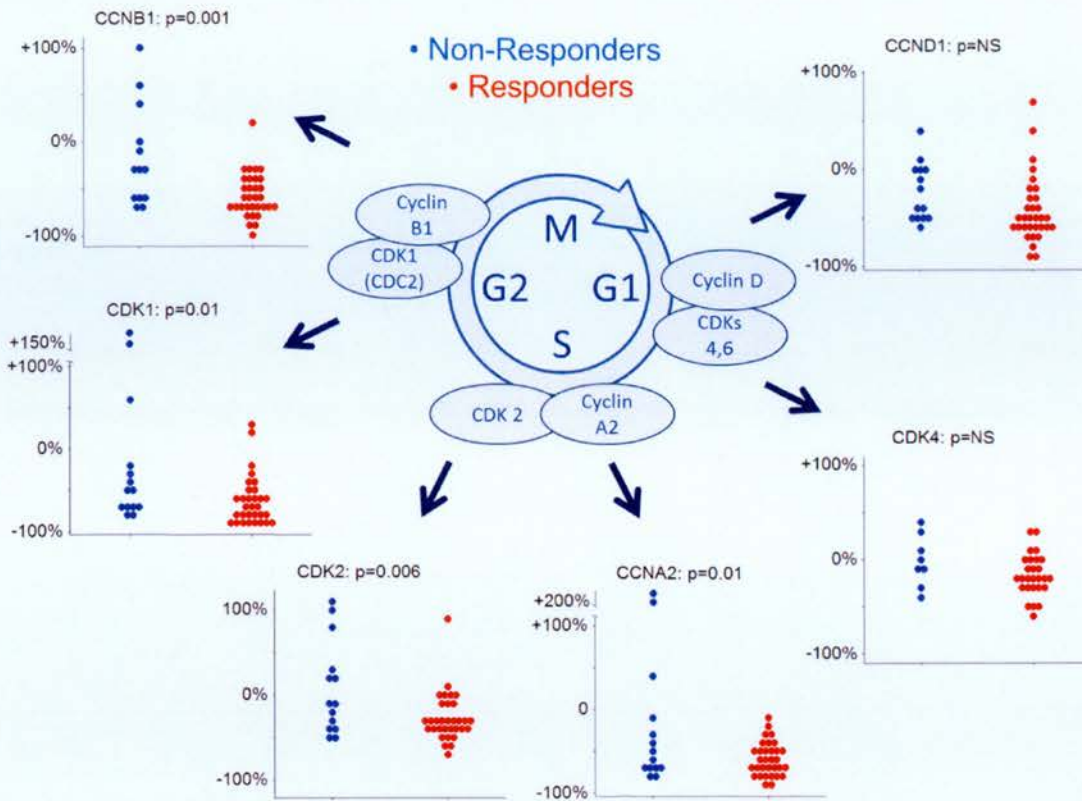


Figure 3.17A: Associations of the Change in Selected Genes with Clinical Response. The y-axes in each plot show the percentage change in mRNA expression after 14 days of treatment (red for responders, blue for non-responders); p-values show the significance of the t-test for the difference between the responders and non-responders.

These results demonstrate that two weeks of Letrozole therapy strongly decreased mRNA expression of Cyclins B1, A2, D1, CDK1 and NUSAP1 but not CDK 2 and 4 where the changes observed were only modest (or even non significant). The mean expression levels decreased by the two week time point and were maintained at a similar level after 3 months of treatment. Significant correlations between the change in Ki67 and changes in Cyclins B1/CDK1 and Cyclin A2/CDK2 were observed, demonstrating the importance of transcriptional regulation of the S and G2-M Cyclins/CDKs for the anti-proliferative effect of neoadjuvant Letrozole. The lack of correlation between Cyclin D/CDK4 and Ki67 or clinical response could reflect the importance of post-transcriptional steps in Cyclin D/CDK4 activation.

Changes in the expression of NUSAP1 and Cyclin B1 significantly correlated with clinical response and Ki67 positivity, especially after two weeks of treatment. This sub-analysis therefore highlights the potential utility of NUSAP1 and Cyclin B1 as anti-proliferative response markers during neo-adjuvant treatment with Letrozole, although they do not completely separate responders from non-responders, indicating the need for other additional markers.

3.18 Ki67 and Other Factors Predicting Survival after Neoadjuvant Endocrine Therapy

This sub-analysis aimed to determine which factors, including pre and post treatment Ki67, are independent predictors of breast cancer specific survival (BCSS), defined as the time from diagnosis to death from a breast cancer related cause, following neoadjuvant AI treatment. To this end all patients with large operable or locally advanced oestrogen receptor positive (ER+) breast cancer who were treated with 3 months of Letrozole, Anastrozole or Exemestane were included. In contrast, the dataset for the microarray studies incorporated only those patients treated with Letrozole in order to eliminate possible variation in gene expression related to treatment with the three different third-generation AIs.

The major effect of the AI's is through inhibition of tumour cell proliferation [297, 331-333, 542]. It follows therefore that the evaluation of tumour proliferation through measurement of Ki67 may allow an early and reliable evaluation of response and long-term outcome following treatment [332, 542, 543]. Cell cycle complete response (CCCR), defined as Ki67 values of <1% [331] [544] is indicative of maximal cell cycle response. It is achieved more commonly than complete pathological response in patients treated with neoadjuvant endocrine therapy. Two-week Ki67 levels have been shown to predict disease free recurrence but there are no accurate data correlating Ki67 with breast cancer specific survival [545].

3.18.1 Patients and Study Design

The study population consisted of 201 women. After surgery, all responding patients continued on an adjuvant AI for five years. Inclusion criteria were as described above for the Letrozole audit series.

All core biopsy specimens taken pre-treatment (baseline) or following three months of AI therapy were fixed in 10% neutral buffered formalin for 24 to 48 hours prior to processing and embedding by the local pathologist in paraffin wax blocks. Wax sections of 3-4 μ m were prepared for Ki67 staining. Ki-67 was determined by measuring the percentage of cells stained using the MIB-1 antibody.

In each case, the tumour size, T stage, pathological node status, grade, ER level (Allred Score), Ki67 at baseline and at three months were recorded as was tumour response, based on 3D ultrasound measurements. The pathology department at the Western General Hospital performed the oestrogen receptor immunohistochemical analysis using the Novocastra 6F11 antibody.

Objective clinical tumour response was calculated based on WHO criteria (Section 3.3.5). Patients were followed annually in clinic throughout the course of their adjuvant treatment. All follow up data for breast cancer relapse, all cause mortality and, breast cancer specific mortality was recorded.

3.18.2 Statistical Analysis

Logistic regression was performed to determine whether any of the variables recorded were predictive of tumour response, whether or not a patient underwent surgery and, if so the nature of their surgery. Survival rates were estimated by the Kaplan-Meier technique and compared using the log-rank test. All survival rates quoted are breast cancer specific unless otherwise stated. Proportional hazards analysis was performed to determine any independent prognostic factors.

3.18.3 Results

Although 201 patients were recruited to the study, not all data points were collected for each patient and therefore analysis was carried out for those 153 patients for whom all data collection was complete. The mean age of the patient population was 74.9 years (range 52-93). The demographics of the study population and the details of tumour pathology are shown in Table 3.18A.

Table 3.18A: Patient Demographics & Tumour Pathology

Endocrine treatment	Number (%)
Anastrozole	23 (15.0)
Exemestane	10 (6.5)
Letrozole	120 (78.4)
Tumour Stage	
T1	5 (3.3)
T2	102 (66.7)
T3	8 (5.2)
T4	38 (24.8)
Tumour Grade	
1	22 (14.4)
2	101 (66.0)
3	30 (19.6)
ER	
<=6	12 (7.8)
7	40 (26.1)
8	101 (66.0)
Surgical treatment to breast	
Breast conserving surgery	116 (75.8)
Mastectomy	26 (17.0)
No surgery	11 (7.2)
Pathological Node Status	
Node negative	68 (44.4)
1-3 nodes involved	43 (28.1)
4+ nodes involved	19 (12.4)
No nodal surgery	23 (15.0)

Response

The percentage reduction in tumour volume as measured by 3D ultrasound following 3 months of neoadjuvant AI treatment is shown in Table 3.18B. 103 patients (66.7%) had a partial response (at least 50% reduction in tumour volume) whilst only 4 patients had progressive disease (no response). Logistic regression shows that ER ($p < 0.0001$) was the main predictor of whether or not the patient achieved a partial response, with a higher ER correlating to better response.

Table 3.18B: Tumour Response at 3 months According to Percentage Reduction in Tumour Volume on Ultrasound.

Response	Number of Patients (%)
≥75%	53 (35)
≥50 - <75%	50 (32)
≥25% - <50%	33 (22)
≥-25% - <25%	13 (8)
<-25%	4 (3)

Operability

Logistic regression analysis showed that tumour Grade ($p=0.0014$) and US response ($p=0.0180$) were the only predictors of whether or not a patient received surgery for their breast tumour. Age ($p<0.0001$) was the main predictor of whether a mastectomy or wide local excision was performed.

Survival

Survival was measured from the date of initiation of AI therapy. One patient was lost to follow up at 35 months and another two at 42 months. For the remainder, the median follow-up was 59 months with a minimum of 38 months. Fifty-eight patients died, 35 with breast cancer present. Overall survival in this predominantly elderly population was 68.7%.

In the univariate analysis, survival was significantly associated with T stage ($p=0.0315$) and the number of involved nodes ($p=0.0005$) but not tumour size, grade or ER level.

Patients with T1-3 breast cancers had better a survival rate than those with locally advanced T4 tumours ($p=0.0315$). The five-year survival rate for patients with T1-3 tumours was 86.6% compared with 60.5% for those with T4 tumours.

Node negative patients did well following surgery with a 5-year survival rate of 97.2%. As expected, increasing nodal involvement equated to poorer survival; 70.5% for those with 1-3 nodes involved and 72.5% for those with 4 or more nodes involved nodes (Figure 3.18A).

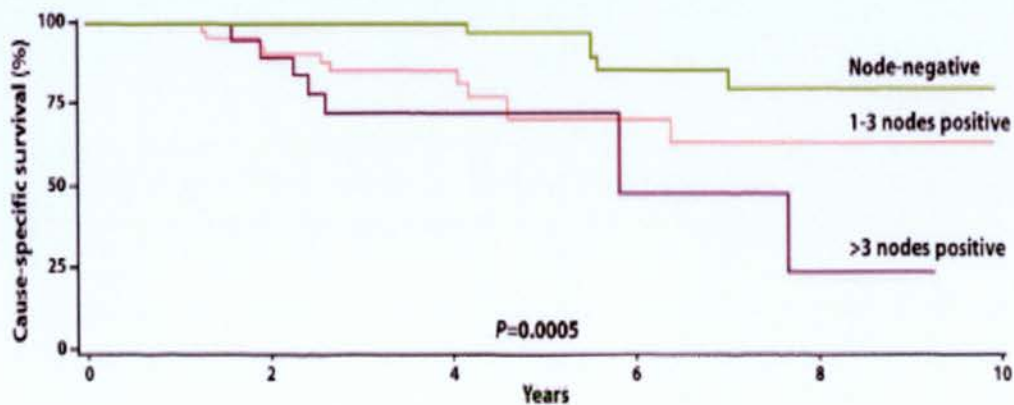


Figure 3.18A: Cause Specific Survival by Node Status at Surgery

Ki67 expression at diagnosis ($p=0.0355$), at 3 months ($p=0.0318$) and the percentage change in Ki67 between these two time points ($p=0.0027$), were also significantly associated with breast cancer mortality whereas US measured tumour response was not.

For the analysis of Ki67 measured at diagnosis, three months and in terms of change in Ki67 over the 3 months, patients were divided into three groups of approximately equal sizes, rather than using quartiles, due to the small numbers. Patients with a Ki67 of less than 10.5% had similar survival in the first 5 years to those with a Ki67 of between 10.5% and 21% whilst those with a Ki67 of 21% or more had a poorer survival (Table 3.18C, Figure 3.18B).

Table 3.18C: Ki67 at Diagnosis

Ki67 at Diagnosis	Number	5-year Survival (%)
<10.5%	50	89.7
>10.5 to <21%	50	89.7
>21%	50	66.8
P value		0.0355

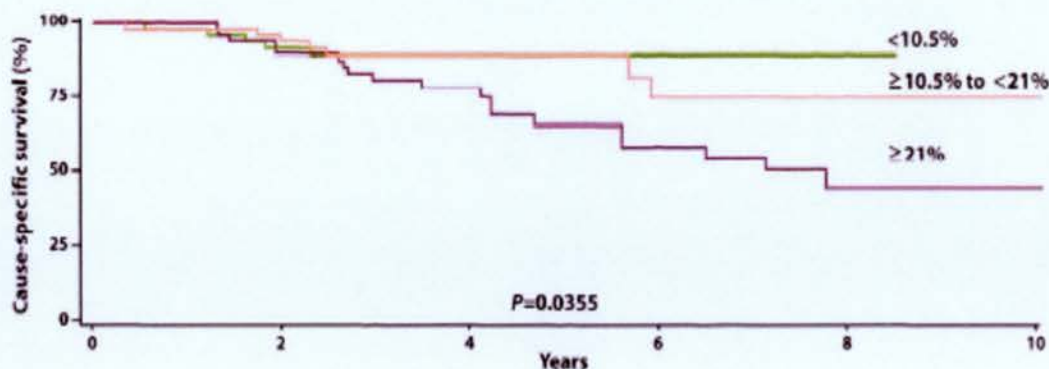


Figure 3.18B: Cause Specific Survival by Ki67 at Diagnosis

Values of Ki67 at 3 months were divided into residual proliferating values of greater than 5 %, 1-5% or less than 1 %. The 1% value has been widely used to define complete cell cycle response (CCCR). Two thirds of treated patients had residual tumour proliferation of less than 5% at three months and one half of these had values of less than 1% (CCCR) (Table 3.18D). Figure 3.18C shows that those patients with low rates of residual proliferation (<1%) following treatment had better survival outcomes, particularly compared to those with > 5% residual proliferation ($p= 0.0318$).

Table 3.18D: Ki67 at 3 Months

Ki67 at 3 months	No of Patients (%)	5-year survival (%)
<1	47 (30.7)	86.9
>=1-<5	47 (30.7)	86.7
>=5	56 (36.6)	70.8
P value		0.0318

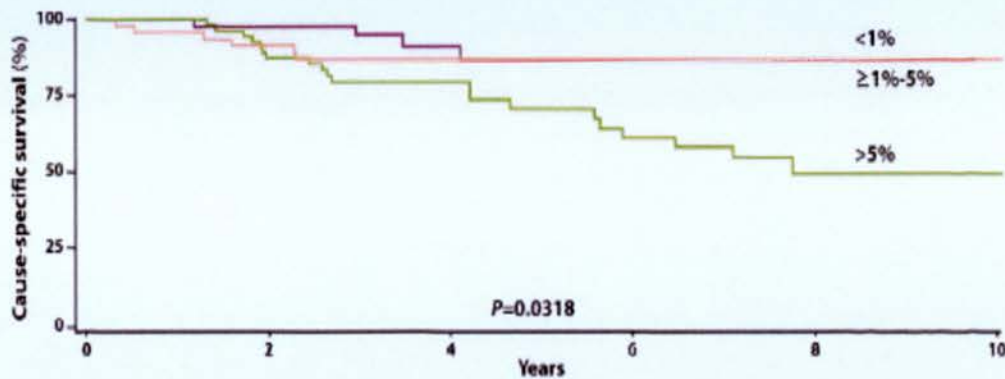


Figure 3.18C: Ki67 at 3 Months and Cause Specific Survival

The percentage fall in Ki67 observed in patients at three months was stratified according to predetermined cut-offs considered to represent significant changes in proliferation. The results in Table 3.18E show that around two thirds of patients had a greater than 40% reduction in Ki67 following 3 months of treatment and more than half of these showed an over 80% fall in proliferation. A third had lesser reduction (<40%), no change or an increase in Ki67 at 3 months. Those patients with falls in proliferation (>80%) following AI treatment were found to have a better breast cancer specific survival than those with lesser reductions (40-80%) and, those with the least change (<40%) had the poorest survival outcomes $p=0.0027$ (Figure 3.18D).

Table 3.18E: Percentage Reduction in Ki67 at 3 Months:

% Reduction in Ki67	No of Patients (%)	5-year survival (%)
≥80	51 (35.3)	86.8
>=40-<80	45 (29.4)	80.9
<40	54 (36.6)	42.7
P value		0.0027

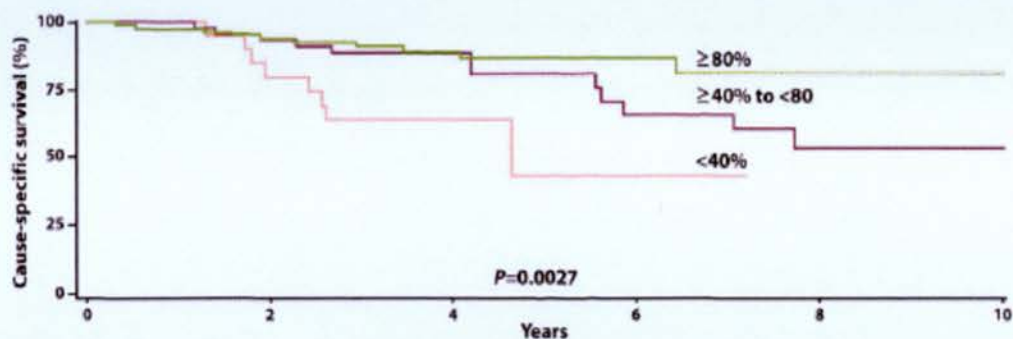


Figure 3.18D: Percentage Change in Ki67 and Cause Specific Survival

Multivariate Analysis

Proportional hazards analysis was initially performed using only those variables available at diagnosis, i.e. year of treatment, age, clinical size, T stage, ER, grade, Ki67 and which AI was used. Data on all these variables was available for 152 patients. The analysis showed that T stage ($p=0.0188$) and baseline Ki67 ($p = 0.0205$) were significant independent prognostic factors for survival.

When adding in those variables not available until later, including US response, type of surgery, number of involved nodes, Ki67 at 3 months, percentage reduction in Ki67 and whether radiotherapy was given post-operatively, only 126 patients were left for analysis. The number of involved nodes ($p=0.0007$), percentage reduction in Ki67 ($p=0.0029$) and grade ($p=0.0383$) were the only significant independent prognostic factors for BCSS.

4. Discussion – PART 1

The Present Study

The over-riding purpose of any cancer classification system is to help plan effective treatment and provide an accurate prognosis for each patient. In addition, classification can help to standardise cancer descriptions for research purposes and communication. Furthermore it provides an important starting point for the investigation of the mechanisms underlying cancer development, metastasis and response or resistance to therapeutic intervention.

The current histopathological classifications of breast cancer, including hormone receptor status and proliferative markers, such as Ki67, do not adequately reflect its diversity; consequently patients may be given inaccurate prognoses and may be subjected to ineffective and unnecessary treatment. For example, up to 60% of breast cancer patients will receive chemotherapy during the course of their treatment and yet only a minority, those with actual systemic disease (2-15%) will stand to benefit^[455, 546, 547]. Similarly, approximately 40% of patients treated with endocrine therapy will not benefit. Many patients are therefore exposed to the toxicity of these agents without a survival advantage. The implications for the healthcare economy of such unnecessary treatments are also evident, with an average cost of approximately £35,000 for each breast cancer patient. More importantly, the use of inappropriate or ineffective therapy can only serve to delay the institution of alternative, more effective treatments or surgery. In order to target treatment specifically to those who stand to benefit, better predictors of response need to be developed. This is especially true for intermediate grade tumours whose phenotype, classification and treatment are particularly challenging.

The advent of microarray technology has allowed researchers to rapidly measure the expression of many thousands of genes in a tumour sample at multiple time points before, during and after treatment and to assess changes in gene expression over time. The results have been used to derive new molecular classification systems and genetic signatures that can potentially be used for prognostic and predictive purposes.

The work described in this thesis was designed to validate the efficacy of the only published signature that is predictive for neo-adjuvant endocrine response in breast cancer. To place this signature in context, its performance has been compared to that of four other signatures that are potentially associated with neo-adjuvant response: two proliferative and two stromal gene expression signatures. Despite there being a large number of genetic signatures published for adjuvant breast cancer treatment, to date, very little has been published for neoadjuvant endocrine response prediction. The validation and comparison of the signatures has been carried out on a previously published dataset, a newly generated dataset and an extended dataset incorporating both of these.

The work carried out directly by the author included: the recruitment of patients into the Edinburgh Letrozole audit series and the processing and design of ethical applications and patient consent forms; core biopsy sampling; clinical measurement of tumour response; documentation of patient outcomes and side-effects and establishment and maintenance of a revised database for recording this information. In the laboratory setting, the author also performed RNA extraction and quality control prior to microarray analysis. Subsequently, the author ran quantitative PCR analyses to validate the gene expression profiles identified on microarray and carried out a focused sub-analysis of genes associated with different phases of the cell cycle. In addition, the author collected patient and tumour demographics including Ki67 staining and compared these to the available long-term outcome data. Ki67 immunohistochemistry was performed by the pathology department. RNA microarray was performed at specialised core facilities and analysis of the microarray data was carried out with the help of a bioinformatician using complex algorithms, the principles of which were discussed in the methods section.

Several other original research projects have utilised the data and samples collected by the author, but these are beyond the scope of this thesis.

4.1 Findings of the Present Study

4.1.1 Validation of the Edinburgh Signature

The Letrozole audit series (L2) was previously used to assess gene expression changes following two weeks of neoadjuvant Letrozole therapy and identified 143 genes whose expression was altered during treatment in one of four different patterns. The ontology of the down-regulated genes reflected cell cycle progression and mitosis in particular, whilst up-regulated genes were associated with organ development, ECM regulation and inflammatory response ^[508].

A follow-up study from our unit correlated the changes in gene expression profiles of patients treated with two weeks of neoadjuvant Letrozole to the clinical response measured after three months of treatment and thereby derived the predictive signature validated in the present study.

Here, for the first time, the Edinburgh signature has been validated against an independent (L3) and an extended (L23) dataset. Furthermore, the L3 validation set demonstrates that the predictive power of the signature is still present when different microarray platforms are used. The validation also legitimises the use of this signature as a benchmark for comparison with the proliferative and stromal signatures used.

The AUC analysis of the Edinburgh signature on the L2 dataset demonstrates a very high sensitivity and specificity, as would be expected from a signature being tested on the dataset from which it was derived. Nevertheless good predictive power is maintained across all three datasets.

Ontological analysis of the genes within the signature previously indicated that ribosomal proteins showed higher baseline expression levels that then fell significantly on treatment in responsive tumours. In non-responding cases, increased expression during early treatment was observed. This is likely to reflect the suppression of protein synthesis in responsive tumours and the escape from such suppression in resistant tumours ^[510].

4.1.2 Performance of the Proliferative Signatures

The decision to use two proliferative signatures in the present study was made because proliferation genes constitute a major part of adjuvant prognostic signatures ^[522] and also because we have shown earlier that cell cycle genes were the most down-regulated genes early during neo-adjuvant endocrine treatment ^[508]. Due to the lack of published predictive signatures in the neoadjuvant setting, both of the chosen signatures had been derived in the adjuvant setting. Both were independently validated and contained only genes associated with proliferation.

In this study, whilst neither proliferation signature is predictive of response at baseline, both are significantly predictive at the two-week time point across all datasets and to a similar degree to the Edinburgh signature. This indicates that the level of expression of proliferation-associated genes, whilst not sufficiently discriminating between responders and non-responders at baseline, becomes so after two weeks of Letrozole treatment. This observation is important because it demonstrates the potential to identify non-responders early during the course of their treatment, so that alternative therapy can be considered. The predictive power of the proliferation signatures is consistent with other methods of measuring tumour proliferation^[539]. Since the levels of cellular proliferation in responsive tumours should drop significantly in responders but less so, or not at all, in non-responders, it follows that the difference in proliferation gene levels at two weeks is likely to be more informative.

On the other hand, it might then be expected that measurement of the *change* in proliferative gene expression levels over the two weeks would also be significantly informative of response. In the present study however, only the Nagalla signature demonstrates that the degree of change in proliferative gene expression between baseline and two weeks is significantly predictive of response to treatment and only in the L3 and L23 datasets. The reasons for the lack of significance in the L2 dataset are not entirely clear, but may possibly relate to differences in signature translation for the Affymetrix platform or to the statistical power of the study.

4.1.3 Performance of the Stromal Signatures

The stromal signatures used here, were selected for similar reasons to the two proliferative ones mentioned above: firstly, stromal rearrangement has been associated with the biology of neo-adjuvant endocrine treatment ^[309, 326, 327] and secondly, we previously demonstrated that stromal rearrangement genes were amongst those most up-regulated after short-term treatment with Letrozole ^[508].

Contrary to our expectations, stromal signatures showed only a modest association with response. The results from the present study demonstrate that neither stromal signature was informative at baseline or at two weeks when these time-points were considered in isolation. However, when considering the *change* in gene expression between these two time points, both signatures are significantly informative on the L3 dataset only, where the predictive power is similar to that of the proliferative and Edinburgh signatures. Once again, the reason why this finding should only be significant on the L3 dataset and not the L2 dataset is not clear. Nevertheless, this result does suggest that the *change* in stromal gene expression during the course of treatment is more relevant to response prediction than the absolute levels of stromal gene expression. This fact may reflect the complexity of stromal rearrangement and the delayed timescale of events when compared to proliferation.

In contrast to studies on tumour epithelial cells, the majority of published data suggest that gene expression and functional changes observed in the tumour stroma are not due to genetically abnormal stromal cells in breast cancer.

Using laser capture micro-dissection to separate tumour stroma from epithelium, Ma *et al.* also demonstrated the importance of change in levels of stromal gene expression in breast cancers. Specifically, the most dramatic changes in gene expression in the stromal component were observed during the transition from normal breast tissue to DCIS, as opposed to the transition from DCIS to invasive carcinoma. The latter was nonetheless associated with up-regulation of several ECM degrading proteases ^[513]. These observations appear to be compatible with those of the present study in that they both suggest that change in stromal gene expression is most informative. No other studies have specifically looked at the stromal response to aromatase inhibitors in the neoadjuvant setting.

When interpreting the stromal signatures' expression it is important to recognise that such signatures may be quite diverse. Specifically, the Farmer signature includes different types of collagens, including types 1 and 3, which would be consistent with "Desmoid Type" or "scar formation" as described by Beck *et al.* [524]. In contrast, Beck describes their signature, which includes genes associated with epithelial-stromal interaction and basal membrane *etc.* as a "Fibromatosis Type". Biologically, the stromal microenvironment may include many and varied functions, such as mechanical support (e.g. collagens and extracellular matrix), adhesion (e.g. decorin, fibronectin), local immune response and nutrition *etc.* Thus biologically different sets of genes could be reported as "stromal signatures" and the specific gene composition must be taken into account when comparing different stromal signatures.

4.1.4 Expression of the Cyclins

Aberrations in cell cycle machinery are amongst the most common mechanisms underlying cancer development, as described in the introduction. Thus far however, most of the research in this area has focused on the G0-G1-S transition. For instance, Cyclin D is amplified in up to 15% of breast cancers [548] and the expression of aberrant forms of Cyclin E is strongly associated with poor outcome [549]. In contrast, the results from the present study have highlighted the importance of transcriptional changes in the S2-M phase regulators - Cyclin B and CDK1. One possible interpretation of this finding is that mutations affecting cell cycle regulation at the G0-G1 level may act through changes to the downstream expression of G2 and M phase genes.

4.1.5 Changes in Ki67 Expression as a Predictor of Breast Cancer-Specific Survival During Neoadjuvant AI Treatment.

This sub-analysis looked at potential clinical, pathological and, biological predictors of outcome in breast cancer patients treated with neoadjuvant AI endocrine therapy.

The clinical factors found to bear prognostic significance in relation to breast cancer specific survival (BCSS), were pathological node status ($p= 0.0007$) and tumour grade ($p= 0.0336$). Previous studies in patients receiving neoadjuvant chemotherapy [550] or endocrine therapy [551-553] for locally advanced breast cancer have shown that tumour grade and surgical node status were significantly predictive of disease free and overall survival in multivariate analyses.

Furthermore, it has been suggested that a poor early clinical response to neoadjuvant AIs may be predictive of early breast cancer recurrence in the long-term and could be used as an indicator for alternative adjuvant therapy [554]. Whilst this may be true, there is a lack of

precision in clinical measurements, which may also lack the sensitivity and rapidity of change needed to detect a response to treatment at an early stage. In this sub-analysis, clinical response was not found to be an important predictor of breast cancer specific survival.

With respect to biological markers that allow earlier and more accurate evaluation of long-term outcome, the IMPACT study produced convincing evidence that a change in Ki67 value after neoadjuvant AI treatment was predictive of disease free survival ^[555]. That study investigated 158 patients randomised to three months of neo-adjuvant treatment with either Anastrozole or Tamoxifen or both in combination. The findings were that: (i) Suppression of Ki67 at 2 and 12 weeks was significantly greater with Anastrozole than Tamoxifen ($p=0.004$ and $p<0.001$ respectively) and, (ii) Ki67 predicted the recurrence free survival of the much larger but equivalent adjuvant ATAC trial ($p= 0.004$) ^[556-558]. Further analyses have since suggested that the absolute values of Ki67 following short-term neoadjuvant AI therapy (2 weeks) are similarly predictive of recurrence free survival ($p=0.004$) whereas baseline Ki67 values are not ^[542].

Ellis has further explored the importance of the absolute post treatment Ki67 values in patients treated with sixteen weeks of neoadjuvant Tamoxifen or Letrozole. Cell Cycle Complete Response (CCCR, or Ki67 <1%) was an independent prognostic factor for both relapse free and overall survival ^[559], suggesting that patients who have effectively switched off tumour proliferation at 16 weeks will do better in the long-term.

The present study differs from those discussed in two regards. Firstly, breast cancer specific survival has been used as an outcome measure to give a more accurate reflection of deaths from the disease. Secondly, proliferation was measured following three months of treatment, as opposed to two weeks. Previous studies have indicated that whilst the majority of ER positive breast tumours treated with an AI reduce proliferation by 2 weeks and maintain this up to the 3 month time point, a few patients will demonstrate a secondary rise in proliferation after 2 weeks, possibly indicating early resistance ^[555]. Three-month values were therefore considered to be more reflective of actual response to neoadjuvant treatment.

This present sub-analysis though small, furthers our knowledge with regard to molecular prognostication. Firstly, the data support the concept that Ki67 expression after three months of endocrine treatment predicts long-term outcome and in particular death from breast cancer. Secondly, three-month Ki67 values might better reflect survival outcome compared to the values at two weeks, since some tumours will develop resistance between these time points. Thirdly, the findings may help in the management of individual patients receiving adjuvant endocrine therapy by indicating the need for alternative therapy when the AI fails to switch off or significantly reduce tumour proliferation.

If confirmed, the quantification of changes in Ki67 during endocrine therapy potentially has major implications for molecular prognostication by allowing the biological detection of response or resistance earlier than clinical indicators. It could therefore inform decisions on future treatment for individual patients and may aid in the evaluation of novel therapies used in combination with the AI's to circumvent resistance. The use of an mTOR inhibitor Everolimus in combination with AIs is currently the subject of investigation in to the management of endocrine resistant breast cancers in the BOLERO trial and other studies [560]. Specifically BOLERO 4 is looking at the safety and efficacy of combining Everolimus with Letrozole in ER positive, HER 2 negative metastatic breast cancer patients (see <http://clinicaltrials.gov/show/NCT01698918>).

4.2 Comments on Methodology

4.2.1 Use of the Neoadjuvant Treatment Setting

The mechanisms of tumour response and resistance to endocrine therapy need to be elucidated so that they can be targeted or bypassed respectively. To this end, researchers can focus their investigations either on patients with early disease treated with surgery and adjuvant therapy, patients with metastatic disease, or on those treated with neoadjuvant therapy.

The principle drawback of adjuvant studies is that the primary tumour has already been excised and cannot therefore be used to assess response directly. Instead, other measures of response such as disease free and overall survival must be employed, necessitating lengthy follow-up in large cohorts of patients. Adjuvant studies may also be confounded by occult spread of disease at the time of diagnosis and must therefore incorporate untreated control groups, which in many circumstances may be unethical. In addition, the markers of response are usually measured in the excised primary tumour whilst the study end points are dependant upon the behaviour and progression of micro-metastatic disease.

The neoadjuvant setting on the other hand, lends itself to the investigation of various breast cancer therapies for a number of reasons. Firstly, it relates to whole tumours in their natural biological environment. Secondly, the primary tumour under investigation has not yet been excised and so multiple, sequential tissue samples can be obtained from the same primary tumour before, during and after treatment. Thirdly, the findings regarding tumour response in the neoadjuvant setting can be extrapolated to inform treatment decisions in the adjuvant setting. Overall, many fewer cases and shorter follow up are required for neoadjuvant studies.

Whilst it is possible to use *in vitro* cell culture or animal models to study tumour biology directly, both have obvious and significant limitations and the results obtained from such studies may not be clinically applicable. Attempts have been made to identify chemotherapy predictive signatures using *in-vitro* models ^[462] but the validity of these has subsequently been disputed on a number of levels; some have been withdrawn and such studies are no longer considered to be reliable ^[561-565]. The possible reasons for this include the fact that cell lines represent only one component of the original tumour cell population, specifically selected for growth in culture. It is unlikely that the gene expression profiles of such cultures adequately reflect those of whole tumours. Additionally, since the mechanisms of tumour proliferation and resistance may involve interactions between different cell types within the tumour, data derived from a single cell line are likely to be fundamentally flawed and not clinically relevant ^[562]. Animal studies have several advantages but the results are rarely borne out in humans.

Patients undergoing neoadjuvant treatments for breast cancer are therefore the perfect subjects for direct studies of tumour biology and response.

4.2.2 Timings of Treatment and Sample Acquisition

It has been shown that neoadjuvant Tamoxifen can reduce tumour size as early as one month after initiation of treatment, although the majority of patients respond at 2-3 months and a further few show a delayed response at 3-6 months. During the latter time frame however, a few of the earlier responding tumours become resistant to treatment and increase in size ^[566]. In terms of proliferative response, which is significantly correlated with both clinical and pathological responses ^[166], a decrease in Ki67 occurs in the majority of tumours treated with Letrozole at 10-14 days and up to 80% by 3 months ^[166, 330]. On this basis and borrowing from experience with cytotoxic neoadjuvant therapy, three months of neoadjuvant Letrozole had been chosen as standard treatment for the Edinburgh patients and is widely accepted practice because it allows sufficient time for tumour response whilst avoiding prolonged treatment in non-responders, thereby minimising potential delays in prescribing surgery or alternative drug therapies. The collection of core biopsy samples pre-treatment and at two weeks, allows time for the gene expression changes to occur whilst still permitting detection of these changes as early as possible so that appropriate treatment decisions can be made. These time points were therefore chosen in this study.

4.2.3 Tumour Response Assessment

Three-dimensional ultrasound was used as the principal clinical marker of response because this modality has previously been shown to correlate with actual tumour size on excision better than calliper measurement or mammography ^[535]. Additionally, since neoadjuvant therapy is primarily used to shrink tumours prior to surgical excision, it was felt to be a more significant end point than the assessment of pathological or proliferative responses.

For the purposes of the present study, it was necessary to define and separate patients who clearly responded to Letrozole from those who clearly did not. This was achieved by excluding an intermediate response group as defined in the methods section. The criteria for defining each of these groups was deliberately very stringent to minimise the risk of cross-contamination of the responder / non-responder groups.

4.2.4 Sample Collection

The use of core biopsies in this study provided sufficient high quality tissue for RNA analysis without significantly reducing the size of the tumour. This was important since tumour volume was measured as the primary end-point. Multiple samples were taken to ensure sufficient material for analysis and the samples were also confirmed to contain at least 20% malignant cells. Nevertheless, the actual tumour content may have varied from one sample to the next and this could have influenced subsequent analyses. Furthermore, core biopsies taken from tumours with marked heterogeneity may not necessarily have collected sample material representative of the tumour as a whole. To minimise these variables, ultrasound guided biopsies were performed whenever tumour size and location made accurate localisation difficult.

Micro-dissection of the tissue samples in order to enrich the tumour content and maximise RNA yields was not performed, because of our wish to include the stroma and any other non-malignant components that might effect gene expression, in their natural proportions.

4.2.5 Signature Selection

Since there is only one published predictive signature for use in the neoadjuvant endocrine setting, it was necessary to select suitable prognostic and predictive signatures from the adjuvant setting for comparison in this study. The selection of these signatures was somewhat arbitrary but attempted to include those that clearly reflected the biology of endocrine therapy. The signatures were also selected for their 'purity' in terms of proliferative or stromal genes so that biological interpretation of the results could be made more easily. The Edinburgh signature, which was derived from the L2 dataset and is the only published neoadjuvant endocrine response signature, was validated against the independent L3 and the extended L23 datasets. In this way, the performance of the other signatures could be compared to that of a signature that was optimised for this treatment modality.

4.2.6 Microarray Analysis

Unlike other forms of gene expression profiling, such as quantitative PCR, microarray technology allows the rapid assessment of thousands of genes, equating to the entire human genome on a single array. Such technology has identified genes whose expression levels are altered by neoadjuvant endocrine therapy ^[507-509]. Despite some concerns, the general reliability and reproducibility as well as the quantitative nature of microarray technology have been clearly established ^[567]. The present study included validation of gene expression by quantitative PCR, which confirmed that microarray data correlates well with gene expression levels as measured by alternative means.

Microarray can be performed on formalin-fixed, paraffin embedded or fresh frozen samples. The latter are associated with less nucleic acid degradation and were therefore used in this study to provide good quality RNA for qPCR and microarray analysis. The facilities for the proper handling and processing of fresh frozen tissue are readily available at the Edinburgh Breast Unit.

The Affymetrix HG-U133A system was used for the first 137 patients of the Edinburgh Letrozole Audit series, following which the Illumina® HT-12 system was used. The latter not only became locally available but is also considered to be technically superior for a number of reasons that have already been discussed. Nonetheless both platforms are widely used in microarray research making comparisons relatively straightforward.

The interpretation of the vast amount of data produced by microarray is extremely challenging and attempts must be made to remove all background 'noise' without discarding any of the important and relevant information. All necessary means of data pre-processing and elimination of bias were carried out on the data produced in this study. In addition, the use of multiple signatures, testing on multiple datasets and the utilisation of different array platforms served to reinforce the validity of the conclusions drawn.

4.2.7 Multi-Signature Comparisons

This study is the first attempt to compare and validate multiple proliferative and stromal gene signatures in the neoadjuvant setting. A previous study compared the prognostic power of multiple prognostic models, including gene expression signatures and somewhat surprisingly demonstrated that the expression of a single proliferation gene AURKA, performed as well as multi-gene signatures ^[568]. The authors suggested that one reason for this might be the 'over-fitting' of data mining methods, which can occur when training and validation datasets overlap. To avoid over-fitting in this study we first validated the Edinburgh signature on a completely new series of patients and secondly, measured gene expressions using an alternative microarray platform.

4. Discussion – PART 2

The Clinical Utility of the Molecular Subtypes and Multi-Gene Signatures

4.3 The Shortcomings of Current Classifications

Historically, breast cancers with similar histopathological characteristics were noted to behave similarly leading to the notion that these classification systems were adequate. The true heterogeneity of the disease however was underestimated and it soon became clear that invasive ductal carcinomas showed highly variable outcomes and response to treatment despite being classified together into relatively few histopathological groups ^[569]. However, the practical means to sub-classify breast cancers did not exist until the advent of advanced molecular biological techniques. Since then, new subtypes of breast cancer have been identified that show distinct and more consistent behavioural characteristics despite often having identical histopathological features ^[446, 447, 449]. Nevertheless, there is still some controversy regarding the exact number and definition of these subtypes.

Histopathological assessments are also subject to potential technical errors leading to misclassification. It has been shown for example, that up to 20% of ER, PR and HER2 quantifications in the clinical setting may be inaccurate with implications for therapeutic strategy in those patients ^[570, 571]. Similarly, there is a lack of absolute concordance between clinically validated histopathological and microarray classification in some cases, for example in tumour HER2 status ^[444, 572-575]. Such discrepancies can also result in misclassifications or the misinterpretation of published research datasets. Although yet to be proven, it would seem likely that genetic classification systems would be more rigorous, reliable and reproducible, whilst promising more accurate prognostication and individualised treatment. There are however several confounders of this assumption, which are discussed below.

4.4 The Potential and Limitations of Molecular Subtyping

The advent of gene expression profiling, which promised to revolutionise breast cancer diagnosis and treatment, was met with an initial wave of enthusiasm. As with many new technologies however, the difficulties and limitations have gradually become more apparent ^[576, 577].

The initial approach for the identification of the molecular subtypes was based upon hierarchical clustering analysis, which is somewhat subjective and cannot be prospectively employed for the classification of individual tumour samples ^[454, 578-580]. Single sample predictors (SSP) were therefore developed which compare the expression profile of a given tumour sample with the average expression profile of each molecular subtype ^[449, 457, 573]. In the last decade, three such SSPs have been published ^[449, 457, 573]. One of these uses quantitative PCR to classify samples into molecular subtypes whilst the others employ gene expression profiling techniques. ^[573]

Subtle variations in defining the average gene expression profile of each molecular subtype can however change the classification of samples when using SSPs ^[453, 581, 582]. For example, in a cohort of 295 breast cancers that were classified into molecular subtypes using different SSPs, a concordance of only 64% was obtained. Furthermore, the classification of tumours into the two subtypes of ER positive breast cancer, luminal A (good prognosis) or B (poor prognosis), appears to be strongly dependent on the SSP used ^[453] and on the expression of proliferation genes within the sample ^[449, 457, 463, 474, 569, 573]. In any case, the level of proliferation gene expression most likely forms a continuum making such a binary division somewhat arbitrary, akin to the histological division into low, intermediate and high-grade tumours ^[444, 473, 476].

Haibe-Kains *et al.* have attempted to address some of the potential limitations of SSPs with their Subtype Classification Models (SCM). The latter are based upon sets of genes whose expressions is correlated with ER, HER 2 and Aurora Kinase A, a proliferation module chosen to discriminate between high and low proliferation tumours [473, 477]. The same group however recently demonstrated that quantitative assessment of these three genes alone is sufficient for the classification of the clinically relevant breast cancer subtypes and that the addition of more genes to a classification model may be of little value whilst impeding clinical utility [583].

The intrinsic subtype classification first proposed by Perou [446, 447] is now considered to be an evolving classification [584] and important discrepancies still exist with respect to the number [450-452], definition [445, 453, 454, 585] and the prognostic and predictive value of each subtype. The useful information they provide beyond that of ER, PR, HER2 and proliferation has yet to be established [455, 585] and standardisation of the definitions and methods for molecular classification has yet to be produced and validated [442, 444, 453, 455]. The evolution of intrinsic classification is evident with the recent addition of three further subtypes of ER negative cancers: Interferon rich [454, 457], Molecular apocrine [451, 452, 586] and Claudin low [450, 587]. Although some of the intrinsic subtypes suggest a cell of origin (e.g. luminal or basal), this concept is likely to be erroneous [447, 449, 588-592]. Additionally, it has been suggested that 'Normal breast-like' cancers may simply represent an artefact of frozen tissue procurement, rather than a distinct molecular subtype [444, 450, 453, 457, 573, 593].

PAM50 is the most recent intrinsic subtyping assay and has potential for clinical application [573, 574]. It comprises a set of 50 genes identified from a total of nearly two thousand intrinsic genes associated in proliferation, ER regulation, HER2, basal and myoepithelial characteristics. PAM50 assays are reported in terms of a risk of relapse score (ROR), which applies to node negative patients who are not receiving adjuvant systemic therapy [573]. It has been argued however, that despite their potential [574], PAM50 and other "intrinsic subtype" based genetic classifications are not yet suitable replacements for the traditional grade, hormone receptor, Ki67 and HER2 based prognostic models [442, 444, 453, 455].

4.5 The Promise of Gene Signatures in Prognostication and Prediction

Molecular subtyping promises to refine the currently used clinico-pathological classification systems, but the promise of individualised treatment planning and prognosis depends upon the development of predictive and prognostic genetic signatures. Since the description of the first molecular subtyping model, numerous signatures, which aim to identify key gene expression patterns associated with prognosis or prediction of response to a variety of systemic treatments, have been published ^[442-444]. In drawing from studies of gene expression signatures, it is necessary to consider the finding that prognostic signatures can be readily generated from random gene sets and that these may perform as well as published gene expression signatures ^[594]. Whilst this indicates that vast numbers of genes are associated with proliferation and therefore correlated with cancer prognosis, the important point is that the genes within any particular signature cannot be taken to be directly involved in the mechanism of that cancer.

4.5.1 Prognostic Signatures

The first generation prognostic signatures contain surprisingly few genes in common, perhaps reflecting the vast numbers of genes that are involved in the regulation of cellular proliferation and their downstream pathways ^[595, 596]. Despite this, meta-analyses confirm good concordance in their ability to classify patients into groups sharing a similar prognosis. The signatures also confirm that ER positive and ER negative breast cancers are indeed governed by distinct biological processes ^[476, 477] and that the latter are associated with a poor prognosis in the majority of cases ^[157, 459, 597]. Changes in the expression of proliferation genes also distinguish between good and poor prognosis subtypes ^[473, 474, 598], with higher expression being naturally associated with poorer prognosis ^[446, 447, 449, 599]. It is not surprising therefore that these signatures can also be predictive of response to chemotherapeutic agents, which target proliferation ^[442-445, 476]. Nevertheless, whilst these signatures provide relevant information for prognostication, they are yet to be incorporated into routine clinical practice and none are supported by level I evidence for their prognostic power ^[442, 444, 455]. Two trials addressing this issue (MINDACT and TAILORx) are ongoing.

The fully accrued TAILORx trial (Trial Assigning Individualised Options for Treatment) is evaluating Oncotype DX® in ER positive, HER2 negative, node negative disease to determine the need for chemotherapy in addition to endocrine therapy ^[482]. Oncotype DX® is a commercially available qPCR based analysis of 21 genes (16 cancer related and 5 reference) designed for application in ER positive, node negative breast cancers receiving adjuvant Tamoxifen, which was shown to outperform clinico-pathological factors for predicting 10-year distant recurrence ^[485]. Based on the expression of genes associated with proliferation, ER and HER2 signalling, and invasion, the test result is presented as a continuous recurrence score RS ranging from 0-100 and equating to the risk of 10 year disease recurrence ^[485]. The continuously variable range of ER positivity may be better reflected in the Oncotype DX® assay. In the trial, patients with an RS score below 10 were given hormone therapy alone; those with RS 11-25 were randomly assigned to receive hormone therapy with or without chemotherapy and those with an RS over 25 received standard hormonal and chemotherapy. When completed, this study should determine which patients in the intermediate risk group have sufficiently low risk of recurrence to avoid the need for chemotherapy ^[482].

Just as proliferation gene expression signatures are important in ER positive disease, it appears that stromal and immune response gene signatures may have more relevance in ER negative disease. Recently, second generation signatures specific for each cancer subtype have been reported in studies of the tumour microenvironment and host immune response ^[442, 494, 497, 498]. These have been shown to be potential prognosticators in ER negative and triple negative breast cancers ^[494, 600, 601]. Wang and colleagues identified a 76-gene 'activated fibroblast' signature that predicted poorer prognosis in terms of distant metastases free survival, which was independent of age, tumour size, grade and ER status. The authors proposed that this signature could be used to predict patients who are unlikely to benefit from adjuvant chemotherapy ^[458]. Similarly, Finak identified a 26-gene stromal prognostic signature that was independent of tumour ER or HER2 status ^[497]. Whilst these results are promising more research and evidence is required in support of their observations.

There are a variety of reasons why the first generation signatures have not become clinically useful tools. Firstly, as we have seen, the ability to determine prognosis corresponds to the level of expression of oestrogen dependant, proliferation and cell cycle related genes ^[473, 477]. Therefore, it may be the case that the signatures simply represent a sophisticated proliferation assay akin to more standardised histological grading systems and thus have little more to add ^[602]. Furthermore, proliferation appears to be prognostic in ER positive but not ER negative disease and so the signatures cannot be applied to the latter ^[458, 459, 477, 597].

Secondly, the classification and treatment of breast cancers at either end of the spectrum of prognosis are more easily determined than that of intermediate grade tumours. The latter not only display intermediate phenotype and survival but are also subject to the greatest degree of intra and inter-observer variation during classification, making prognostication and treatment decisions extremely difficult ^[603]. Generally, the first generation signatures cannot adequately subdivide intermediate tumours into subgroups that can direct treatment planning. The application of Oncotype DX® does 'push' some intermediate grade tumours into a low risk group that do not require chemotherapy and others into a high-risk group that do.

In one study, treatment recommendations based on standard pathological grading were changed for 31% of patients after considering their Oncotype DX® scores and in 22.5% of cases chemotherapy was omitted altogether ^[604]. However 40-60% of patients who are classified as intermediate (Grade II) using standard histopathological means are still allocated into the intermediate group based on their Oncotype DX® RS scores. Clinical decision-making is therefore not made much easier in the very group for whom prognostication and management are already difficult ^[605]. In order to improve clinically useful sub-classification of this group of tumours, better-defined predictors and prognostic markers are required. To this end, Haibe-Kains *et al.* developed a 97-gene 'GGI' score with the purpose of sub-classifying ER positive, intermediate grade II tumours into groups with distinct clinical outcomes that are closely related to those of high or low grade tumours ^[467, 473]. As with other prognostic signatures, GGI also has predictive capabilities with GGI grade III deriving more benefit from chemotherapy than grade I ^[470].

Comparison of the GGI to other 70-gene and 76-gene prognostic signatures for their ability to predict distant metastasis-free survival in primary, untreated node negative breast cancer demonstrated concordance of 68-88% between signatures - the 70-gene and GGI signatures had the highest concordance. Thus, despite limited overlap of their genes, these signatures have similar prognostic performance ^[603].

A third reason for the lack of clinical utility of first generation signatures is the fact that many of them were developed to predict short term disease recurrence (<5 years) and have shown reduced prognostic significance after 5-10 years of follow-up, indicating that they may only be useful in predicting early but not late disease recurrence ^[459, 478]. Patients requiring extended therapy due to a high risk of late disease relapse may therefore be inappropriately reassured and have beneficial treatment discontinued.

Fourthly, as with histopathological markers, technical artefact can diminish the accuracy and therefore the efficacy of the first generation signatures. For example, the allocation of patients into prognostic subgroups may be dependant upon the percentage of neoplastic versus non neoplastic (normal breast, stromal and inflammatory) cells in a given tissue sample. The first generation signatures were themselves derived from tissues of varying neoplastic content. Indeed, evidence suggests that the percentage of neoplastic cells in a sample does have a substantial impact on the final gene expression profile and on the ability to derive biologically meaningful data ^[606]. On the other hand, the high levels of standardisation of some gene expression assays such as Oncotype DX® and MammaPrint® are perhaps more reproducible than the current assays for ER, PR and HER2 ^[481, 607, 608]. The vast majority of gene expression based assays are however, not sufficiently standardised to achieve high-level concordance between laboratories.

Finally, most of the signatures developed thus far, have been subjected to very limited independent validation. A variety of different microarray platforms and statistical tools have been used to analyse the data and there is little overlap between the genes of each signature. All of these factors make comparison between studies and validation of individual studies extremely difficult ^[603].

4.5.2 The Transition from Prognostic to Predictive Signatures

Although many prognostic markers have been shown to have predictive value, efforts are underway to find biomarkers capable of specifically predicting a patient's response or resistance to a particular treatment so that individual they can be managed accordingly. In fact, there is currently a strategic shift away from prescribing chemotherapy on the basis of prognosis to doing so on the basis of prediction of response. As a mark of this transition, a number of attempts have been made to identify predictive components within the previously published prognostic signatures.

However, one of the difficulties inherent in studies of adjuvant treatments is the need to monitor survival endpoints that require prolonged follow-up and large patient numbers to achieve significance.

The clinico-pathological Pre-operative Endocrine Prognostic Index (PEPI) score combines tumour size, nodal status, ER and Ki67 indices to give a total relapse score ^[155, 315]. This can be used to distinguish between patients with low risk of relapse who can be treated with endocrine therapy alone from high-risk patients who will require additional chemotherapy. Specifically, patients with pathological stage 0 or I and a PEPI score of 0 at the time of surgery had a sufficiently low rate of relapse to suggest that further adjuvant therapy was not required. In contrast, patients with high pathological stage and a PEPI score of 3, required additional adjuvant treatment. Whilst PEPI is not a genetic predictive marker, patients with luminal-A tumours identified by PAM50 are associated with better prognosis PEPI scores than those with luminal B tumours and a better outcome when treated with endocrine therapy alone ^[301].

Over the last decade, many groups have defined multiple gene prognostic signatures with the aim of identifying patients with a good enough prognosis to allow them to forego chemotherapy ^[158, 458, 467, 485, 486, 609-611] ^[443, 444]. The first of these was the 70-gene signature, which could be applied to patients who had not received systemic therapy, in order to identify those who were unlikely to develop distant disease within 5 years ^[158]. This was shown to be an independent and better predictor of outcome and response to adjuvant chemotherapy than standard clinico-pathological variables ^[157, 158, 458, 467, 485, 486, 609-611], leading to the development of the commercially available version of the 70-gene signature – MammaPrint®. Other examples include the 76-gene signature ^[458, 612] and the Genomic Grade Index ^[442-444, 467, 470, 613-616]. Whilst *Oncotype DX*® RS has been shown to be predictive of benefit from the addition of chemotherapy to Tamoxifen ^[501, 617], the predictive power of MammaPrint® ^[615, 616] and GGI ^[470] have however only been tested in retrospective datasets in patients receiving multi-drug chemotherapy.

In an exploratory study evaluating 323 different gene signatures in 550 node negative patients who did not receive adjuvant systemic therapy, it was suggested that prognostic predictions can be made in ER positive but not ER negative or HER2 positive disease ^[598]. However other studies have identified immune regulated signatures that might predict survival within HER2 or triple negative tumours, although within these, the good prognosis group still demonstrated a high incidence of relapse ^[618, 619]. With respect to ER negative breast cancers, a retrospective microarray analysis has identified an increased stromal meta-gene expression that predicted resistance to FEC chemotherapy and therefore treatment failure ^[496].

In ER positive node positive disease the low risk groups as predicted by MammaPrint®, Oncotype DX® and PAM50 have relatively high recurrence rates nonetheless (over 10% at 8-10 years) making denial of adjuvant chemotherapy inappropriate^[157, 488, 489, 574, 620]. However, in a subset of patients with 1-3 positive nodes the Oncotype DX® and MammaPrint® low risk groups have a relapse rate similar to that of node negative disease when treated with hormone therapy alone (<10% at 5-10 years)^[488, 489, 621].

The ongoing prospective phase III multi-centre randomised MINDACT trial (**M**icroarray **I**n **N**ode **N**egative and 1-3 positive lymph-node **D**isease may **A**void **C**hemotherapy) currently aims to test whether patients with a high risk of recurrent disease according to clinico-pathological features (Adjuvant! Online) can be spared chemotherapy if their MammaPrint® result indicates relatively low risk^[487, 622]. This is again indicative of the trend of using a predictive utility as a prognostic test.

4.5.3 Predictive Signatures in the Adjuvant Endocrine Therapy Setting

Only a few gene expression signatures have been designed specifically to predict endocrine response in ER positive tumours and none have yet been shown to confer benefit if implemented in the clinical management of breast cancer patients. ER status is of course, the main indicator for the institution of endocrine therapy. Whilst ER negative tumours almost never respond, ER positive tumours respond in 50-70% of cases^[623]. The fact that the remaining 30-50% of ER positive tumours do not respond presents an urgent incentive to find a means of early categorisation of these tumours based upon their likely treatment response. Whilst PR and HER2 status provide some additional predictive information, they are not sufficiently reliable indicators of endocrine response.

A number of endocrine predictive signatures have therefore been developed. For example, Jansen *et al.* identified and validated a 44-gene signature that was predictive of response to adjuvant Tamoxifen in advanced ER positive disease. The signature included oestrogen regulated genes, apoptosis, ECM formation and immune response genes^[506]. Furthermore, their signature performed significantly better than traditional predictive factors although the authors acknowledged that further research was required to validate the use of this or similar signatures in directing the need to institute alternative adjuvant therapy in non-responders^[506]. Similarly, PAM50 has been validated in an independent cohort of ER positive patients treated with adjuvant Tamoxifen and again outperformed standard clinico-pathological variables including Ki67, PR and histological grade^[574]. The ubiquitous Oncotype DX® is also independently predictive in postmenopausal ER positive patients treated with Anastrozole^[488].

The Sensitivity to Endocrine Therapy (SET) index is based on 165 genes that are co-expressed with the ER gene. The test categorises ER positive tumours into groups that predict low, intermediate or high response to any endocrine therapy ^[479]. Effectively, this signature aims to reflect the activity of oestrogen signalling pathways in ER +ve tumours, therefore highlighting those where the receptor is functional. Unlike other multi-gene signatures that essentially evaluate proliferation in ER positive tumours, the SET index appears to be predictive of response to endocrine therapy independently of the inherent prognosis of the tumour. SET also identifies a subgroup of tumours with excellent prognosis in a Tamoxifen treatment group (node negative and high SET index) and in a chemo-endocrine treatment group (high and intermediate SET index) ^[479].

The Breast Cancer Index (BCI) assesses the likelihood of distant recurrence in ER positive, node negative patients. In validation studies, the BCI stratified nearly half of such patients treated with Tamoxifen into a low risk group for disease recurrence ^[611, 624]. The BCI also reflects the promising strategy of combining mutually complementary signatures. Unlike the other commercialised tests, BCI has been developed from the combination of two indices: a HOXB13: IL17BR (H-I) gene ratio, which predicts metastases in ER positive, Tamoxifen treated patients ^[486] and a five proliferation gene index (MGI) which discriminates grade I from grade III disease in a similar way to GGI ^[611]. Two distinct approaches to combining the HI ratio and MGI have been investigated recently in ER positive, node negative patients and were more or less equivalent in their stratification of the 10-year risk of breast cancer related death ^[625].

A comparison of BCI, Oncotype DX® and the immunohistochemical prognostic test ICH4 for their ability to predict long term risk of recurrence in ER positive patients found that only the BCI accurately identified patients at risk beyond 5 years of treatment with either Anastrozole or Tamoxifen, whilst all three could predict earlier risk. Beyond 5 years, BCI was clearly able to distinguish those 40% of patients with a significant risk of recurrence. This of course offers the prospect of targeting prolonged therapy to those most likely to benefit ^[626].

4.5.4 Predictive Signatures in the Neoadjuvant Chemotherapy Setting

In the neoadjuvant chemotherapy setting, the elimination of detectable disease at the primary tumour site (pathological complete response, pCR) is often used as a surrogate for systemic response to chemotherapy agents, although it is only achieved in 20-30% of patients ^[627]. Pathological complete response rates to Anthracycline and Taxane based chemotherapy are around 7% for luminal A, 17% for luminal B, 36% for HER2 and 43% for basal-like subtypes ^[628]. Overall, the majority of responders are non-luminal A tumours, which helps to explain why gene expression signatures that discriminate luminal A from other intrinsic subtypes can also predict pCR ^[470, 501, 616].

Straver *et al.* assessed the predictive value of MammaPrint® in the neoadjuvant chemotherapy setting on 167 patients using pCR as a measure of chemosensitivity. All patients had invasive breast cancer greater than 3cm in diameter. None of the good MammaPrint® prognosis group achieved pCR, whereas 20% of the poor prognosis group, which included all triple negative cancers, did. Amongst the non triple-negative group, the response was significantly associated with the MammaPrint® classification ^[616]. These findings are in keeping with the expectation that the good prognosis groups detected by MammaPrint have lower proliferation rates and are therefore less sensitive to chemotherapy. They nonetheless indicate the potential for MammaPrint® to be used as a predictive marker of response to neoadjuvant chemotherapy.

Similarly, Oncotype DX® has also been correlated with pCR in patients receiving neoadjuvant chemotherapy and was found to have predictive value for benefit from chemotherapy in patients with ER positive, lymph node negative disease ^[501]. GGI has also been shown to be a predictor of chemotherapy response in the neoadjuvant setting in both ER negative and positive patients ^[470].

Several groups have attempted to identify chemotherapy response signatures by comparing gene expression profiles in high and low responsive tumours ^[499, 502-504]. The majority of these studies were done using biopsies taken before neoadjuvant chemotherapy was initiated. In this way, MD Anderson developed a 30-gene signature (DLDA-30), subsequently validated, that predicts response to T-FAC chemotherapy better than a combination of age, grade and ER status (sensitivity 92% versus 61%) ^[502]. However this finding has since been disputed in a further independent study ^[629].

4.5.5 Predictive Signatures in the Neoadjuvant Endocrine Therapy Setting

To date, the only endocrine neoadjuvant predictive signature that has been independently validated is the Edinburgh 205-gene signature used in the present study.

Concurrently with the publications from the Edinburgh group described above, Mackay *et al.* measured the molecular response before and after two weeks of treatment with either Letrozole or Anastrozole and identified 421 genes including the classic oestrogen-response genes, whose expression was affected by the treatment. Proliferation genes were down-regulated and ECM associated genes up-regulated^[507]. In this study, the changes in gene expression profiles were not correlated to response to the aromatase inhibitors, but instead related to changes in Ki-67, which could be considered as a response surrogate.

Mello-Grand *et al.* identified a 54-gene signature from a small group (13 matched samples) of ER positive patients treated with Anastrozole for three months, which was predictive of sensitivity to neoadjuvant Anastrozole in all of the patients. They also observed that Anastrozole treatment was associated with a decreased expression of cellular proliferation genes and an increased expression of inflammatory genes. In addition, enrichment of cell cycle inhibition and induction of immune response genes occurred in responders^[521]. No independent validation of this signature has yet been made.

More recently, the genomic grade index (GGI) has been shown to be a useful predictor of response to neoadjuvant Letrozole in post-menopausal ER positive breast cancer patients. Women with low GGI were shown to be more likely to respond to three months of treatment than those with high GGI. Importantly, low proliferation following two weeks of Letrozole therapy, as measured by GGI and Ki67, was associated with tumour response in GGI high-grade but not GGI low-grade tumours, suggesting that the early assessment of proliferation may be useful in predicting response, particularly in high-grade ER positive breast cancer^[539].

To date, there appear to be no studies focused upon stromal gene expression changes as predictive markers of response to endocrine treatment in either the adjuvant or neoadjuvant settings. However, the local inflammatory component reported by Mello-Grand *et al.* may potentially be associated with some aspects of the stromal microenvironment, as discussed earlier. Of note, both Mackay's^[507] and Miller's^[508] groups reported changes in extracellular matrix gene expression during neoadjuvant endocrine treatment and Farmer's group reported the predictive capacity of a stromal signature with respect to neoadjuvant cytotoxic chemotherapy.

4.5.6 Combining Markers to Improve Power

Whilst clinico-pathological variables have been shown to add prognostic information independent of that offered by the first generation signatures^[442-444], when these variables are standardised for histological grade, protein receptor status and Ki67, little additional information seems to be provided by the signatures^[455, 630]. Nevertheless, *Oncotype DX*® has been promoted as an adjunct to clinico-pathological classification in evaluating the prognosis of ER positive, node negative patients and in predicting their response to chemotherapy^[631, 632]. It has also been shown to be independently prognostic of 10-year distant recurrence in ER positive, postmenopausal women receiving Anastrozole^[488].

Just as the combination of standard clinico-pathological and genetic markers of prediction and prognosis have, in some cases, been shown to improve their prognostic or predictive power, it is possible that combinations of multi-gene signatures may be similarly useful, particularly when they are associated with different cellular pathways^[339]. Mefford *et al.* have recently discussed the potential advantages of combining proliferation and stromal based signatures with clinico-pathological measures to add prognostic information in breast cancer patients^[633]. They demonstrated that the incorporation of stromal cell signatures significantly improves the performance of proliferation based first generation signatures in the adjuvant setting.

Another example of the successful implementation of this strategy was the Breast Cancer Index (BCI) described earlier.

Generally however, it remains to be seen whether combinations of gene signatures with or without standard histopathological markers can significantly improve prognostication and prediction. For the time being, it has been suggested that signatures should currently be used only as ancillary tools rather than as a substitute for clinico-pathological staging^[442-444].

5. Future Studies

5.1 Pertaining to the Present Study

There are several avenues for further investigation that could be applied to the data collected for the current study. These include:

- (i) Analysis of the microarray data obtained from the final, three-month samples to determine if the changes evident after two weeks of Letrozole treatment are maintained or modified by further treatment.
- (ii) Identifying the most significantly changed genes between responders and non-responders from each of the chosen signatures and then attempting to construct a novel, refined signature from these that has improved predictive power.
- (iii) Following on from this, it would be interesting to look at the possibility of combining the most informative genes from the proliferation and stromal signatures to determine whether such a combination is a better predictor of response to AI therapy than either alone.
- (iv) Investigating the ontology of these most significantly changed genes to see whether this can cast light on the mechanisms underlying resistance to aromatase inhibitors or identify new targets for breast cancer treatment. Such an investigation must consider the comments of Venet *et al.* discussed above ^[594].
- (v) Verifying or refuting the stromal signature result observed in this study using other published stromal signatures. It may also be worthwhile correlating the stromal gene expression of the samples to their stromal content at baseline to ensure that stromal signatures are not simply reflective of relatively high stromal content.
- (vi) Similarly, attempting to correlate the present results with the measurement of the ratio of tumour epithelium and stroma for each sample at baseline.
- (vii) Expanding the response groups by taking account of long-term follow-up data, to determine whether some members of the intermediate response group can be redefined as responders or non-responders and then to determine if this could have been predicted at two weeks using the gene signatures under investigation.
- (viii) Assessing the efficacy of the predictive signatures with respect to pathological tumour response (e.g. central scar formation, mitoses) at three months and not just tumour size on USS.

- (ix) Comparing the predictive power of the chosen signatures with that of standard baseline clinico-pathological markers for correlation and to determine if additional power is obtained by combining the two.
- (x) Correlating response prediction based on the chosen signatures with long-term patient outcome data.

5.2 General Considerations for Future Studies

The present study highlights some of the difficulties experienced when trying to compare data derived from different studies, carried out on different populations defined by different standards and under different conditions. No two studies are quite alike, with each research group attempting to carve out their own niche of expertise. Nevertheless, reliable and comparable results will only come from large, well-designed studies with a high degree of standardisation in the definitions and methods used. Ideally, standardisation should include:

- i) Study inclusion and exclusion criteria
- ii) Sampling time-points and technique.
- iii) Classification of the ratio of tumour to stroma ratio in biopsy samples to allow better comparison between published results and to help distinguish the relative roles of these components.
- iv) Tighter definitions of endpoints such as pCR and reduction in tumour volume.
- v) Better definition of response cut-off points.
- vi) Microarray inter-platform translation methods
- vii) Microarray data processing
- viii) Statistical analysis and reporting

With this in mind, there are a number of questions that need to be answered by future studies. One of the most pressing is how best to accurately and reliably re-define intermediate grade tumours, as defined by a variety of classification systems, into responders or non-responders to a given treatment. This is the most difficult group of patients on whom to plan treatment and determine prognosis and is consequently the most likely to be treated ineffectively, subjecting them to unnecessary physical side effects and wasting valuable resources. Furthermore the sensitivity and specificity of genetic signatures are not perfect. Even with the most promising predictive signatures, including those used in this study, there are individual cases with poor predictive scores who are nonetheless clearly responding to treatment.

It seems likely that the next generation of prognostic and predictive multi-gene signatures will only be applicable to specific drugs and subtypes of breast cancers since the nature of these subtypes and their mechanisms of resistance to different drugs appear to be fundamentally different. Several groups are already investigating prognostic and predictive signatures for each subgroup independently. However, most available studies are carried out on broader groups of breast cancer patients for whom the use of such pooled data will have a tendency to discriminate between the subtypes but have less power to predict response within a subtype. Thus the predictive value of a potential biomarker for a single subtype would be diluted by the presence of other subtypes in such studies.

A further target for future studies is to improve prognostication and prediction for ER negative disease. Many cell cycle and proliferative genes are predictive of outcome for ER positive disease but relatively few genes predict outcome of ER negative cancers. Recent studies have indicated that ER negative disease expresses immune response genes that provide prognostic information but further research is certainly needed because the 5 year relapse rate, even in patients given a good prognosis, is approximately 20%, too high to justify withholding adjuvant chemotherapy ^[445, 475, 477, 494].

The current range of available treatment options is of course limited and there may seem to be little point in being able to sub-classify breast cancers to a greater degree than the number of treatment options available. For example, if aromatase inhibitors are the only treatment option for two putative subtypes of breast cancer, what is the purpose of being able to distinguish between these two groups? The answer is two-fold. Firstly, clinicians would be able to better inform the patients of their chances of response and secondly, attempting to do so will help to identify groups for whom the currently available treatments are inadequate. The genetic profiles of these groups of patients could then be investigated further with the aim of identifying novel targets and strategies for treatments or means of circumventing their mechanisms of drug resistance.

Finally, an intelligent combination of transcriptional signatures such as those studied in this project, with data on DNA and protein levels will help to provide biologically driven integrative approaches for future tumour classifications and response prediction.

6. Conclusions

Breast cancer is a predominant cause of cancer deaths worldwide and especially in the West. The management of breast cancer is undertaken with a combination of surgery, and systemic therapies. Choosing the most appropriate therapy for an individual patient and providing them with a prognosis, is currently an imprecise science based upon the clinico-pathological features of their cancer. Many patients do not respond to treatment as anticipated and many others are given unnecessary treatments from which they will derive no benefit. In either scenario, the personal wellbeing of the patient and the financial wellbeing of healthcare systems are put under considerable strain.

The work presented in this thesis, is an attempt to assess and validate the efficacy of five multi-gene signatures for their ability to predict response to neoadjuvant endocrine therapy treatment in ER positive breast cancer patients. By observing changes in gene expression at an early stage of treatment, it is hoped that patients who are predicted to respond, will continue their treatment assured in the knowledge that it is likely to be effective. The remainder can be converted to alternative, more effective treatments at the earliest opportunity. The aims and objectives of this project have been fully achieved.

Patients treated with neo-adjuvant Letrozole have been continuously enrolled into the study to generate a new dataset with which to evaluate the capacity for response prediction using various transcriptional signatures.

Amongst these was the previously published 205-gene Edinburgh neo-adjuvant endocrine response predictive signature. For the first time, this signature was tested against an independent dataset and was confirmed to be significantly associated with response. Moreover, this validation was performed in a platform-independent manner.

Subsequently, the performance of the Edinburgh signature was compared to that of four other recently published signatures: two associated with proliferation (Nagalla 2013, Dai 2005) and two signatures associated with stromal rearrangement (Farmer 2009, Beck 2008). Furthermore, the proliferation and stromal signatures were compared to each other.

In order to diversify the comparison populations, the analysis was expanded to include a previously published dataset, thereby providing results for both datasets as well as a combined dataset.

Of those studied, only the Nagalla 2013 proliferation signature demonstrated performance similar to that of the Edinburgh neoadjuvant signature.

Both proliferation signatures demonstrated consistent, statistically significant association with neo-adjuvant response in all datasets after 2 weeks of treatment but not at pre-treatment baseline. The change in proliferation gene expression between these two time-points was associated with response for only one signature (Nagalla 2013) and then in only one of the studied datasets.

The stromal signatures were not associated with neo-adjuvant response to Letrozole either before or at two weeks of treatment. However, changes in gene expressions in both stromal signatures showed association with response in the newly collected dataset only and not in the previously published dataset.

In light of the strong performance of proliferation signatures, we conducted two extended sub-analyses, focused on selected proliferation genes and markers and using alternative laboratory methods: qPCR and IHC.

Quantitative PCR was used to measure the expression of genes associated with different phases of the cell cycle. The G2-M-phase transcriptional changes (NUSAP1 and Cyclin B1) were significantly correlated with clinical response, which was not the case for genes involved in the G0-G1 transition (Cyclins D1, CDK 4). This sub-analysis highlighted the potential utility of NUSAP1 and Cyclin B1 as anti-proliferative response markers during neo-adjuvant treatment with Letrozole.

Ki67 IHC staining was related to long-term outcome and was found to be significantly predictive both in respect to the absolute expression levels at baseline and three months and with respect to the change in these levels during treatment.

Taken together our results confirm the utility of transcriptional markers for predicting response to neo-adjuvant endocrine treatment and provide a new dataset to compare alternative signatures and markers.

Nevertheless, the currently available multi-gene signatures and molecular markers are not sufficiently reliable to justify withholding therapy in the clinical environment. Therefore it seems likely that traditional histopathological classification systems, combined with clinical staging will remain the mainstay of prognostication and prediction for a while longer. Further characterisation of the genes whose expression is altered during treatment and the molecular pathways they contribute to, should lead towards a better understanding of the disease and new opportunities for intervention.

However, considerable challenges lie ahead.

7. References

1. Ferlay, J., et al., *GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide*, 2010, International Agency for Research on Cancer. : Lyon, France. Available from <http://globocan.iarc.fr>.
2. Westlake, S. and N. Cooper, *Cancer incidence and mortality: trends in the United Kingdom and constituent countries, 1993 to 2004*. *Health Stat Q*, 2008(38): p. 33-46.
3. *Global Statistics (2008)* Cancer Research UK Statistical Information Team, 2011, using data from GLOBOCAN 2008 v1.2. IARC: Lyon, France. Available at <http://globocan.iarc.fr> and <http://infocancerresearch.uk.org>.
4. *Cancer Incidence and Mortality by Cancer Network.*, National Cancer Intelligence Network, 2008.: UK.
5. Statistics., U.K., Office for National Statistics. 2011.Data available at <http://www.ons.gov.uk/ons>.
6. *What is the Lifetime Risk of Developing Cancer?: The effect of Adjusting for Multiple Primaries*, in *Br J Cancer* 2011, Cancer Research UK Statistical Information Team. Using data from Br J Cancer. Available at <http://info.cancerresearchuk.org>.
7. *Breast Cancer Incidence and Mortality Statistics*, ISD Scotland, September 2011: Scotland. Available at <http://isdscotland.org>.
8. Berry, D.A. and P.M. Ravdin, *Breast cancer trends: a marriage between clinical trial evidence and epidemiology*. *J Natl Cancer Inst*, 2007. **99**(15): p. 1139-41.
9. Brewster, D.H., et al., *Declining breast cancer incidence and decreased HRT use*. *Lancet*, 2009. **373**(9662): p. 459-60; author reply 461.
10. Chlebowski, R.T., et al., *Breast cancer after use of estrogen plus progestin in postmenopausal women*. *N Engl J Med*, 2009. **360**(6): p. 573-87.
11. Glass, A.G., et al., *Breast cancer incidence, 1980-2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status*. *J Natl Cancer Inst*, 2007. **99**(15): p. 1152-61.
12. Kumle, M., *Declining breast cancer incidence and decreased HRT use*. *Lancet*, 2008. **372**(9639): p. 608-10.
13. Ravdin, P.M., et al., *The decrease in breast-cancer incidence in 2003 in the United States*. *N Engl J Med*, 2007. **356**(16): p. 1670-4.
14. Rossouw, J.E., et al., *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial*. *JAMA*, 2002. **288**(3): p. 321-33.
15. *Mortality Statistics: Deaths registered in 2010.*, Office for National Statistics: National Statistics, London.
16. *Deaths Time Series Data, Deaths in Scotland 2010*, General Register Office for Scotland 2010.
17. Autier, P., et al., *Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database*. *BMJ*, 2010. **341**: p. c3620.
18. Kingsmore, D., et al., *Specialisation and breast cancer survival in the screening era*. *Br J Cancer*, 2003. **88**(11): p. 1708-12.
19. Lavelle, K., et al., *Non-standard Management of Breast Cancer Increases with Age in the UK: a Population Based Cohort of Women Older than or Equal to 65 years*. *Br J Cancer*, 2007. **96**(8): p. 1197-203.
20. *The Second All Breast Cancer Report*, 2011, National Cancer Intelligence Network.
21. Stapelkamp, C., et al., *Predictors of early death in female patients with breast cancer in the UK: a cohort study*. *BMJ Open*, 2011. **1**(2): p. 247.
22. *Breast Cancer Survival Statistics*, Cancer Research UK: London. Available at <http://cancerresearchuk.org>.
23. Chia, K., et al, *Do Younger Female Breast Cancer Patients Have A Poorer Prognosis? Results From A Population Based Survival Analysis*. *Int J Cancer*, 2004. **108**(5): p. 761-765.
24. Mathew, A., M. Pandey, and B. Rajan, *Do younger women with non-metastatic and non-inflammatory breast carcinoma have poor prognosis?* *World J Surg Oncol*, 2004. **2**: p. 2.
25. Coleman, M.P., et al, *Cancer Survival Trends in England and Wales, 1971-1995: Deprivation and NHS Region*, 1999, The Stationery Office London.

26. Sasieni, P.D., et al., *What is the lifetime risk of developing cancer?: The effect of adjusting for multiple primaries.* [submitted] 2011.
27. Murray, E. and J.M. Dixon, *Breast Carcinoma: Stages I and II*, in *General Surgery. Principles and International Practice.* , K.I. Bland, Buchler M., Csendes A., Sarr M., Garden O., Wong J, Editor. 2009, Springer-Verlag London Ltd: London. p. 1523-1540.
28. Key, T.J., P.K. Verkasalo, and E. Banks, *Epidemiology of breast cancer.* *Lancet Oncol*, 2001. **2**(3): p. 133-40.
29. Garcia-Closas, M., et al., *Established breast cancer risk factors by clinically important tumour characteristics.* *Br J Cancer*, 2006. **95**(1): p. 123-9.
30. Tanner, J.M., *Trend towards earlier menarche in London, Oslo, Copenhagen, the Netherlands and Hungary.* *Nature*, 1973. **243**(5402): p. 95-6.
31. Koprowski, C., et al., *Diet, body size and menarche in a multiethnic cohort.* *Br J Cancer*, 1999. **79**(11-12): p. 1907-11.
32. *Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease.* *Lancet*, 2002. **360**; **9328**: p. 187-95.
33. Ewertz, M., et al., *Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries.* *Int J Cancer*, 1990. **46**(4): p. 597-603.
34. Ma, H., et al., *Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies.* *Breast Cancer Res*, 2006. **8**(4): p. R43.
35. *Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer.* *Lancet*, 1997. **350**; **9084**: p. 1047-59.
36. Titus-Ernstoff, L., et al., *Long-term cancer risk in women given diethylstilbestrol (DES) during pregnancy.* *Br J Cancer*, 2001. **84**(1): p. 126-33.
37. *Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies.* *Lancet*, 1996. **347**: p. 1713-27.
38. Beral, V., *Breast cancer and hormone-replacement therapy in the Million Women Study.* *Lancet*, 2003. **362**(9382): p. 419-27.
39. Ross, R.K., et al., *Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin.* *J Natl Cancer Inst*, 2000. **92**(4): p. 328-32.
40. Parkin, D.M., *Cancers attributable to reproductive factors in the UK in 2010.* *Br J Cancer*, 2011. **105** Suppl 2: p. S73-6.
41. Key, T., et al., *Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies.* *J Natl Cancer Inst*, 2002. **94**(8): p. 606-16.
42. Gunter, M.J., et al., *Insulin, insulin-like growth factor-1, endogenous estradiol, and risk of colorectal cancer in postmenopausal women.* *Cancer Res*, 2008. **68**(1): p. 329-37.
43. Key, T.J., et al., *Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies.* *Lancet Oncol*, 2010. **11**(6): p. 530-42.
44. Larsson, S.C., C.S. Mantzoros, and A. Wolk, *Diabetes mellitus and risk of breast cancer: a meta-analysis.* *Int J Cancer*, 2007. **121**(4): p. 856-62.
45. Tworoger, S.S., et al., *A prospective study of plasma prolactin concentrations and risk of premenopausal and postmenopausal breast cancer.* *J Clin Oncol*, 2007. **25**(12): p. 1482-8.
46. Eliassen, A.H., et al., *Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women.* *J Natl Cancer Inst*, 2006. **98**(19): p. 1406-15.
47. Kaaks, R., et al., *Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC).* *J Natl Cancer Inst*, 2005. **97**(10): p. 755-65.
48. Xue, F. and K.B. Michels, *Intrauterine factors and risk of breast cancer: a systematic review and meta-analysis of current evidence.* *Lancet Oncol*, 2007. **8**(12): p. 1088-100.

49. Parkin, D.M., L. Boyd, and L.C. Walker, 16. *The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010*. Br J Cancer, 2011. **105 Suppl 2**: p. S77-81.
50. Key, T.J., et al., *Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women*. J Natl Cancer Inst, 2003. **95**(16): p. 1218-26.
51. Reeves, G.K., et al., *Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study*. BMJ, 2007. **335**(7630): p. 1134.
52. Deapen, D., et al., *Rapidly rising breast cancer incidence rates among Asian-American women*. Int J Cancer, 2002. **99**(5): p. 747-50.
53. Bingham, S.A., et al., *Are imprecise methods obscuring a relation between fat and breast cancer?* Lancet, 2003. **362**(9379): p. 212-4.
54. Boyd, N.F., et al., *Dietary fat and breast cancer risk revisited: a meta-analysis of the published literature*. Br J Cancer, 2003. **89**(9): p. 1672-85.
55. Thiebaut, A.C., et al., *Dietary fat and postmenopausal invasive breast cancer in the National Institutes of Health-AARP Diet and Health Study cohort*. J Natl Cancer Inst, 2007. **99**(6): p. 451-62.
56. Dong, J.Y. and L.Q. Qin, *Soy isoflavones consumption and risk of breast cancer incidence or recurrence: a meta-analysis of prospective studies*. Breast Cancer Res Treat, 2011. **125**(2): p. 315-23.
57. Velentzis, L.S., et al., *Lignans and breast cancer risk in pre- and post-menopausal women: meta-analyses of observational studies*. Br J Cancer, 2009. **100**(9): p. 1492-8.
58. Allen, N.E., et al., *Moderate alcohol intake and cancer incidence in women*. J Natl Cancer Inst, 2009. **101**(5): p. 296-305.
59. Baan, R., K. Straif, and Y. Grosse, *Carconogenity of alcoholic beverages*. Lancet Oncol, 2007. **8**: p. 292-3.
60. Hamajima, N., et al., *Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease*. Br J Cancer, 2002. **87**(11): p. 1234-45.
61. Key, J., et al., *Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues*. Cancer Causes Control, 2006. **17**(6): p. 759-70.
62. Parkin, D.M., *Cancers attributable to consumption of alcohol in the UK in 2010*. Br J Cancer, 2011. **105 Suppl 2**: p. S14-8.
63. Rinaldi, S., et al., *Relationship of alcohol intake and sex steroid concentrations in blood in pre- and post-menopausal women: the European Prospective Investigation into Cancer and Nutrition*. Cancer Causes Control, 2006. **17**(8): p. 1033-43.
64. Lahmann, P.H., et al., *Physical activity and breast cancer risk: the European Prospective Investigation into Cancer and Nutrition*. Cancer Epidemiol Biomarkers Prev, 2007. **16**(1): p. 36-42.
65. Monninkhof, E.M., et al., *Physical activity and breast cancer: a systematic review*. Epidemiology, 2007. **18**(1): p. 137-57.
66. Chan, M.F., et al., *Usual physical activity and endogenous sex hormones in postmenopausal women: the European prospective investigation into cancer-norfolk population study*. Cancer Epidemiol Biomarkers Prev, 2007. **16**(5): p. 900-5.
67. Parkin, D.M., *Cancers attributable to inadequate physical exercise in the UK in 2010*. Br J Cancer, 2011. **105 Suppl 2**: p. S38-41.
68. Kakizaki, M., et al., *Sleep duration and the risk of breast cancer: the Ohsaki Cohort Study*. Br J Cancer, 2008. **99**(9): p. 1502-5.
69. Megdal, S.P., et al., *Nightwork and breast cancer risk: a systematic review and meta-analysis* Eur J Cancer, 2005. **41**: p. 2023-32.
70. Schernhammer, E.S. and S.E. Hankinson, *Urinary melatonin levels and postmenopausal breast cancer risk in the Nurses' Health Study cohort*. Cancer Epidemiol Biomarkers Prev, 2009. **18**(1): p. 74-9.
71. Verkasalo, P.K., et al., *Sleep duration and breast cancer: a prospective cohort study*. Cancer Res, 2005. **65**(20): p. 9595-600.
72. Gonzalez-Perez, A., L.A. Garcia Rodriguez, and R. Lopez-Ridaura, *Effects of non-steroidal anti-inflammatory drugs on cancer sites other than the colon and rectum: a meta-analysis*. BMC Cancer, 2003. **3**: p. 28.

73. Khuder, S.A. and A.B. Mutgi, *Breast cancer and NSAID use: a meta-analysis*. Br J Cancer, 2001. **84**(9): p. 1188-92.
74. Mangiapane, S., M. Biettner, and P. Schlattman, *Aspirin use and breast cancer risk: a meta-analysis and meta-regression of observational studies from 2001 to 2005*. Pharmacoepidemiol Drug Saf 2008. **17**: p. 115-24.
75. Takkouche, B., C. Regueira-Mendez, and M. Etminan, *Breast cancer and use of nonsteroidal anti-inflammatory drugs: a meta-analysis*. J Natl Cancer Inst 2008. **100**: p. 1439-47.
76. Largent, J.A., et al., *Hypertension, antihypertensive medication use, and breast cancer risk in the California Teachers Study cohort* Cancer Causes Control, 2010. **21**(10): p. 1615-24.
77. Cogliano, V.J., et al., *Preventable exposures associated with human cancers*. J Natl Cancer Inst, 2011. **103**(24): p. 1827-39.
78. Luo, J., et al., *Association of active and passive smoking with risk of breast cancer among postmenopausal women: a prospective cohort study*. BMJ, 2011. **342**: p. d1016.
79. Xue, F., et al., *Cigarette smoking and the incidence of breast cancer*. Arch Intern Med, 2011. **171**(2): p. 125-33.
80. Hankinson, S.E. and D.J. Hunter, *Breast Cancer Textbook of Cancer Epidemiology*. 2002, New York: New York. Oxford University Press. 301-37.
81. John, E.M., et al., *Medical radiation exposure and breast cancer risk: findings from the Breast Cancer Family Registry*. Int J Cancer, 2007. **121**(2): p. 386-94.
82. Alm El-Din, M.A., et al., *Breast cancer after treatment of Hodgkin's lymphoma: risk factors that really matter*. Int J Radiat Oncol Biol Phys, 2009. **73**(1): p. 69-74.
83. Gold, D.G., J.P. Neglia, and K.E. Dusenbery, *Second neoplasms after megavoltage radiation for pediatric tumors*. Cancer, 2003. **97**(10): p. 2588-96.
84. Kenney, L.B., et al., *Breast cancer after childhood cancer: a report from the Childhood Cancer Survivor Study*. Ann Intern Med, 2004. **141**(8): p. 590-7.
85. Taylor, A.J., et al., *Risk of breast cancer in female survivors of childhood Hodgkin's disease in Britain: a population-based study*. Int J Cancer, 2007. **120**(2): p. 384-91.
86. Berrington de Gonzalez, A., et al., *Second solid cancers after radiotherapy for breast cancer in SEER cancer registries*. Br J Cancer, 2010. **102**(1): p. 220-6.
87. Berrington de Gonzalez, A. and S. Darby, *Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries*. Lancet, 2004. **363**(9406): p. 345-51.
88. Parkin, D.M. and S.C. Darby, *Cancers in 2010 attributable to ionising radiation exposure in the UK*. Br J Cancer, 2011. **105** Suppl 2: p. S57-65.
89. Green, J., et al., *Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk*. Lancet Oncol, 2011. **12**(8): p. 785-94.
90. Van den Brandt, P.A., et al., *Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk*. Am J Epidemiol, 2000. **152**(6): p. 514-27.
91. *Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease*. Lancet, 2001. **358**(9291): p. 1389-99.
92. Pharoah, P.D., et al., *Family history and the risk of breast cancer: a systematic review and meta-analysis*. Int J Cancer, 1997. **71**(5): p. 800-9.
93. Ford, D., et al., *Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium*. Am J Hum Genet, 1998. **62**(3): p. 676-89.
94. Antoniou, A., et al., *Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies*. Am J Hum Genet, 2003. **72**(5): p. 1117-30.
95. Peto, J., N. Collins, and R. Barfott, et al *Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer*. J Natl Cancer Inst, 1999. **91**: p. 943-9.
96. Turnbull, C. and N. Rahman, *Genetic predisposition to breast cancer: past, present, and future*. Annu Rev Genomics Hum Genet, 2008. **9**: p. 321-45.
97. Hartmann, L.C., et al., *Benign breast disease and the risk of breast cancer*. N Engl J Med, 2005. **353**(3): p. 229-37.

98. Kerlikowske, K., et al., *Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy*. J Natl Cancer Inst, 2003. **95**(22): p. 1692-702.
99. Levi, F., et al., *Cancer risk after radiotherapy for breast cancer*. Br J Cancer, 2006. **95**(3): p. 390-2.
100. Rubino, C., et al., *Relation of risk of contralateral breast cancer to the interval since the first primary tumour*. Br J Cancer, 2010. **102**(1): p. 213-9.
101. Soerjomataram, I., et al., *Risks of second primary breast and urogenital cancer following female breast cancer in the south of The Netherlands, 1972-2001*. Eur J Cancer, 2005. **41**(15): p. 2331-7.
102. Volk, N. and V. Pompe-Kirn, *Second primary cancers in breast cancer patients in Slovenia*. Cancer Causes Control, 1997. **8**(5): p. 764-70.
103. Oza, A.M. and N.F. Boyd, *Mammographic parenchymal patterns: a marker of breast cancer risk*. Epidemiol Rev, 1993. **15**(1): p. 196-208.
104. Tamimi, R.M., et al., *Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women*. J Natl Cancer Inst, 2007. **99**(15): p. 1178-87.
105. McCormack, V.A. and I. Dos Santos Silva, *Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis*. Cancer Epidemiol Biomarkers Prev, 2006. **15**(6): p. 1159-69.
106. Berg, W.A., et al., *Diagnostic Accuracy of Mammography, Clinical Examination, US and MR Imaging in Pre-operative Assessment of Breast Cancer*. Radiology, 2004. **233**: p. 830-49.
107. Dixon, J.M., *Indications and Techniques of Breast Biopsy*. Curr Prac Surg, 1993. **5**: p. 142-8.
108. Dixon, J.M. and R.E. Mansel, *ABC of breast diseases. Symptoms assessment and guidelines for referral*. BMJ, 1994. **309**(6956): p. 722-6.
109. Sainsbury, J.R., et al., *ABC of breast diseases. Breast cancer*. BMJ, 1994. **309**(6962): p. 1150-3.
110. Evans, A., *Ductal carcinoma in situ (DCIS): Are we over detecting it?* Breast Cancer Research, 2004. **6** (suppl 1).
111. Afonso, N. and D. Bouwman, *Lobular carcinoma in situ*. Eur J Cancer Prev, 2008. **17**(4): p. 312-6.
112. Devilee, P., A. Fatteneh, and F.A. Tavassoli, *Tumours of The Breast and Female Genital Organs*, 2003, World Health Organisation (WHO): Oxford.
113. Anderson, W.F., et al., *Epidemiology of inflammatory breast cancer (IBC)*. Breast Dis, 2005. **22**: p. 9-23.
114. Bloom, H. and W.W. Richardson, *Histological Grading and Prognosis in Breast Cancer; A Study of 1409 cases of which 359 have been followed for 15 years*. British Journal of Cancer, 1957. **11**(3): p. 359-377.
115. Elston, C. and I. Ellis, *Pathologic Prognostic Factors in Breast Cancer. The Value of Histological Grades in Breast Cancer. Experiences of a Large Study with Longterm Follow-Up*. Histopathology, 1991. **19**: p. 403-10.
116. Miller, W.R., *Oestrogens and breast cancer: biological considerations*. Br Med Bull, 1991. **47**(2): p. 470-83.
117. Awada, A., I. Bozovic Spasojevic, and L. Chow, *New therapies in HER2-positive breast cancer: a major step towards a cure of the disease?* Cancer Treat Rev, 2012. **38**(5): p. 494-504.
118. Goncalves, A., et al., *[Triple-negative breast cancer: histoclinical and molecular features, therapeutic management and perspectives]*. Bull Cancer, 2013. **100**(5): p. 453-64.
119. *AJCC Cancer Staging Manual, 7th Edition*.
120. Dixon, J.M., *A Companion to Specialist Surgical Practice: Breast Surgery*. 3rd Edition ed. 2006, London: Elsevier/ W.B Saunders.
121. Kuenen-Boumeester, V., et al., *Ultrasound-guided fine needle aspiration cytology of axillary lymph nodes in breast cancer patients. A preoperative staging procedure*. Eur J Cancer, 2003. **39**(2): p. 170-4.
122. Chetty, U., et al., *Management of the axilla in operable breast cancer treated by breast conservation: a randomized clinical trial*. Edinburgh Breast Unit. Br J Surg, 2000. **87**(2): p. 163-9.

123. Hayes, D.F., B. Trock, and A.L. Harris, *Assessing the clinical impact of prognostic factors: when is "statistically significant" clinically useful?* Breast Cancer Res Treat, 1998. **52**(1-3): p. 305-19.
124. Blamey, R.W., et al., *Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999.* Eur J Cancer, 2007. **43**(10): p. 1548-55.
125. Veronesi, U., et al., *Prognostic Significance of number and level of axillary node metastases in breast cancer.* Breast, 1993. **3**: p. 224-228
126. Singletary, S.E., et al., *Revision of the American Joint Committee on Cancer staging system for breast cancer.* J Clin Oncol, 2002. **20**(17): p. 3628-36.
127. Olivetto, I.A., et al., *Long-term survival of patients with supraclavicular metastases at diagnosis of breast cancer.* J Clin Oncol, 2003. **21**(5): p. 851-4.
128. Leone, B.A., et al., *Stage IV breast cancer: clinical course and survival of patients with osseous versus extraosseous metastases at initial diagnosis. The GOCS (Grupo Oncologico Cooperativo del Sur) experience.* Am J Clin Oncol, 1988. **11**(6): p. 618-22.
129. Ellis, I.O., et al., *Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up.* Histopathology, 1992. **20**(6): p. 479-89.
130. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.* Histopathology, 1991. **19**(5): p. 403-10.
131. Elston, C.W., I.O. Ellis, and S.E. Pinder, *Pathological prognostic factors in breast cancer.* Crit Rev Oncol Hematol, 1999. **31**(3): p. 209-23.
132. Singletary, S.E., *Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy.* Am J Surg, 2002. **184**(5): p. 383-93.
133. El-Gohary, Y.M., et al., *Prognostic significance of intratumoral and peritumoral lymphatic density and blood vessel density in invasive breast carcinomas.* Am J Clin Pathol, 2008. **129**(4): p. 578-86.
134. Dixon, J.M., *ABC of Breast Diseases 4th edition*, ed. J.M. Dixon. 2012: Wiley-Blackwell.
135. Dunnwald, L.K., M.A. Rossing, and C.I. Li, *Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients.* Breast Cancer Res, 2007. **9**(1): p. R6.
136. Masuda, H., et al., *Role of epidermal growth factor receptor in breast cancer.* Breast Cancer Res Treat, 2012. **136**(2): p. 331-45.
137. Menard, S., et al., *HER2 as a prognostic factor in breast cancer.* Oncology, 2001. **61 Suppl 2**: p. 67-72.
138. Khasraw, M. and R. Bell, *Primary systemic therapy in HER2-amplified breast cancer: a clinical review.* Expert Rev Anticancer Ther, 2012. **12**(8): p. 1005-13.
139. Bouzubar, N., et al., *Ki67 immunostaining in primary breast cancer: pathological and clinical associations.* Br J Cancer, 1989. **59**(6): p. 943-7.
140. Gerdes, J., et al., *Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation.* Int J Cancer, 1983. **31**(1): p. 13-20.
141. Gerdes, J., et al., *Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67.* J Immunol, 1984. **133**(4): p. 1710-5.
142. Spyrtatos, F., et al., *Correlation between MIB-1 and other proliferation markers: clinical implications of the MIB-1 cutoff value.* Cancer, 2002. **94**(8): p. 2151-9.
143. Kontzoglou, K., et al., *Correlation between Ki67 and breast cancer prognosis.* Oncology, 2013. **84**(4): p. 219-25.
144. Urruticoechea, A., I.E. Smith, and M. Dowsett, *Proliferation marker Ki-67 in early breast cancer.* J Clin Oncol, 2005. **23**(28): p. 7212-20.
145. Van Diest, P.J., E. Van der Wall, and J.P. Baak, *Prognostic value of proliferation in invasive breast cancer: a review.* J Clin Pathol, 2004. **57**(7): p. 675-81.
146. Pietilainen, T., et al., *The important prognostic value of Ki-67 expression as determined by image analysis in breast cancer.* J Cancer Res Clin Oncol, 1996. **122**(11): p. 687-92.

147. Trihia, H., et al., *Ki-67 expression in breast carcinoma: its association with grading systems, clinical parameters, and other prognostic factors--a surrogate marker?* *Cancer*, 2003. **97**(5): p. 1321-31.
148. De Azambuja, E., et al., *Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients.* *Br J Cancer*, 2007. **96**(10): p. 1504-13.
149. Stuart-Harris, R., et al., *Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients.* *Breast*, 2008. **17**(4): p. 323-34.
150. Goldhirsch, A., et al., *Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer.* *J Clin Oncol*, 2003. **21**(17): p. 3357-65.
151. Luporsi, E., et al., *Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review.* *Breast Cancer Res Treat*, 2012. **132**(3): p. 895-915.
152. Galea, M.H., et al., *The Nottingham Prognostic Index in primary breast cancer.* *Breast Cancer Res Treat*, 1992. **22**(3): p. 207-19.
153. Haybittle, J.L., et al., *A prognostic index in primary breast cancer.* *Br J Cancer*, 1982. **45**(3): p. 361-6.
154. Harvey, J.M., et al., *Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer.* *J Clin Oncol*, 1999. **17**(5): p. 1474-81.
155. Ellis, M.J., et al., *Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics.* *J Natl Cancer Inst*, 2008. **100**(19): p. 1380-8.
156. Cuzick, J., et al., *Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer.* *J Clin Oncol*, 2011. **29**(32): p. 4273-8.
157. Van de Vijver, M.J., et al., *A gene-expression signature as a predictor of survival in breast cancer.* *N Engl J Med*, 2002. **347**(25): p. 1999-2009.
158. Van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer.* *Nature*, 2002. **415**(6871): p. 530-6.
159. Forrest, A.P., et al., *Randomised controlled trial of conservation therapy for breast cancer: 6-year analysis of the Scottish trial.* *Scottish Cancer Trials Breast Group.* *Lancet*, 1996. **348**(9029): p. 708-13.
160. Veronesi, U., et al., *Twenty Year Follow Up of a Randomised Study Comparing Breast Conserving Surgery with Radical Mastectomy for Early Breast Cancer.* *N Engl J Med*, 2002. **347**: p. 11227-11232.
161. Veronesi, U., et al., *Quadrantectomy versus lumpectomy for small size breast cancer.* *Eur J Cancer*, 1990. **26**(6): p. 671-3.
162. Franceschini, G., et al., *Conservative and radical oncoplastic approaches in the surgical treatment of breast cancer.* *Eur Rev Med Pharmacol Sci*, 2008. **12**(6): p. 387-96.
163. Asgeirsson, K.S., et al., *Size of invasive breast cancer and risk of local recurrence after breast-conservation therapy.* *Eur J Cancer*, 2003. **39**(17): p. 2462-9.
164. Gage, I., et al., *Pathologic margin involvement and the risk of recurrence in patients treated with breast-conserving therapy.* *Cancer*, 1996. **78**(9): p. 1921-8.
165. Smitt, M.C., et al., *The importance of the lumpectomy surgical margin status in long-term results of breast conservation.* *Cancer*, 1995. **76**(2): p. 259-67.
166. Miller, W.R., et al., *Proliferation, steroid receptors and clinical/pathological response in breast cancer treated with letrozole.* *Br J Cancer*, 2006. **94**(7): p. 1051-6.
167. Peintinger, F., et al., *Residual specimen cellularity after neoadjuvant chemotherapy for breast cancer.* *Br J Surg*, 2008. **95**(4): p. 433-7.
168. Sasano, H., T. Suzuki, and T. Moriya, *Pathology of breast cancer following neoadjuvant therapy*, in *Endocrine Therapy in Breast Cancer*, W.R. Miller and J.M. Ingle, Editors. 2002, Marcel Dekker: New York. p. pp. 213-222.
169. Semiglazov, V.F., et al., *Phase 2 randomized trial of primary endocrine therapy versus chemotherapy in postmenopausal patients with estrogen receptor-positive breast cancer.* *Cancer*, 2007. **110**(2): p. 244-54.

170. Gendy, R.K., J.A. Able, and R.M. Rainsbury, *Impact of skin-sparing mastectomy with immediate reconstruction and breast-sparing reconstruction with miniflaps on the outcomes of oncoplastic breast surgery*. Br J Surg, 2003. **90**(4): p. 433-9.
171. Zurrida, S., et al., *The Changing Face of Mastectomy (from Mutilation to Aid to Breast Reconstruction)*. Int J Surg Oncol, 2011. **2011**: p. 980158.
172. Kuciejewska, A. and I. Smith, *Locally Advanced Breast Cancer*, in *General Surgery. Principles And International Practice*, B.M. Bland K, Csendes A, Garden O, Sarr M and Wong J, Editor. 2009, Springer Science: London. p. 1541-1549.
173. Mansel, R.E., et al., *Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial*. J Natl Cancer Inst, 2006. **98**(9): p. 599-609.
174. Ahmed, M. and M. Douek, *What is the future of axillary surgery for breast cancer?* Ecancermedicallscience, 2013. **7**: p. 319.
175. Galimberti, V. and B.F. Cole, *Axillary versus sentinel-lymph-node dissection for micrometastatic breast cancer--authors' reply*. Lancet Oncol, 2013. **14**(7): p. e251-2.
176. Galimberti, V., et al., *Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (IBCSG 23-01): a phase 3 randomised controlled trial*. Lancet Oncol, 2013. **14**(4): p. 297-305.
177. Giuliano, A.E., et al., *Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial*. JAMA, 2011. **305**(6): p. 569-75.
178. *Management of Breast Cancer in Women. A National Clinical Guideline*, 2005, Scottish Intercollegiate Guidelines Network: Edinburgh.
179. *Favourable and Unfavourable Effects on Longterm Survival of Radiotherapy for Early Breast Cancer; An Overview of the Randomised Trials*. Lancet, 2000. **355**(9217): p. 1757-70.
180. *Radiotherapy for early breast cancer*. Cochrane Database Syst Rev, 2002(2): p. CD003647.
181. Malmstrom, P., et al., *Breast conservation surgery, with and without radiotherapy, in women with lymph node-negative breast cancer: a randomised clinical trial in a population with access to public mammography screening*. Eur J Cancer, 2003. **39**(12): p. 1690-7.
182. Whelan, T.J., et al., *Does locoregional radiation therapy improve survival in breast cancer? A meta-analysis*. J Clin Oncol, 2000. **18**(6): p. 1220-9.
183. White, J.R., et al., *Long-term outcome from RTOG 9517: A phase I/II study of accelerated partial breast irradiation (APBI) with multicatheter brachytherapy (MCT) following lumpectomy for early-stage breast cancer*. Journal of Clinical Oncology, 2012. **30**(S27).
184. Chua, B., et al., *Increased use of adjuvant regional radiotherapy for node-positive breast cancer in British Columbia*. Breast J, 2004. **10**(1): p. 38-44.
185. Martijn, H., et al., *Locally advanced breast cancer: radiotherapy alone versus surgery and radiotherapy*. Radiother Oncol, 1989. **16**(1): p. 23-9.
186. Recht, A., et al., *Postmastectomy radiotherapy: clinical practice guidelines of the American Society of Clinical Oncology*. J Clin Oncol, 2001. **19**(5): p. 1539-69.
187. *Improving Outcomes for Breast Cancer: Research Evidence for the Manual Update*, 2002, National Institute for Clinical Evidence: York.
188. *Breast Cancer Services. National Overview*, 2003, NHS Quality Improvement Scotland: Edinburgh.
189. *Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group*. Lancet, 1998. **352**(9132): p. 930-42.
190. Goldhirsch, A., et al., *Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009*. Ann Oncol, 2009. **20**(8): p. 1319-29.
191. Levine, M., et al., *Randomised Trial of Intensive Cyclophosphamid, Epirubicin, and Fluorouracil Chemotherapy Compared with Cyclophosphamide, Methotrexate, and Fluorouracil in Pre-menopausal Women with Node Positive Breast Cancer*. J Clin Oncol, 1998. **16**(8): p. 2651-58.
192. Earl, H.M., et al., *NEAT: National Epirubicin Adjuvant Trial--toxicity, delivered dose intensity and quality of life*. Br J Cancer, 2008. **99**(8): p. 1226-31.

193. Hassan, M.S., et al., *Chemotherapy for breast cancer (Review)*. *Oncol Rep*, 2010. **24**(5): p. 1121-31.
194. Henderson, I.C., et al., *Improved outcomes from adding sequential Paclitaxel but not from escalating Doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer*. *J Clin Oncol*, 2003. **21**(6): p. 976-83.
195. Oakman, C., et al., *Overall survival benefit for sequential doxorubicin-docetaxel compared with concurrent doxorubicin and docetaxel in node-positive breast cancer--8-year results of the Breast International Group 02-98 phase III trial*. *Ann Oncol*, 2013. **24**(5): p. 1203-11.
196. Roche, H., et al., *Sequential adjuvant epirubicin-based and docetaxel chemotherapy for node-positive breast cancer patients: the FNCLCC PACS 01 Trial*. *J Clin Oncol*, 2006. **24**(36): p. 5664-71.
197. Bonadonna, G., M. Zambetti, and P. Valagussa, *Sequential or alternating doxorubicin and CMF regimens in breast cancer with more than three positive nodes. Ten-year results*. *JAMA*, 1995. **273**(7): p. 542-7.
198. Citron, M.L., et al., *Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741*. *J Clin Oncol*, 2003. **21**(8): p. 1431-9.
199. Charfare, H., S. Limongelli, and A.D. Purushotham, *Neoadjuvant chemotherapy in breast cancer*. *Br J Surg*, 2005. **92**(1): p. 14-23.
200. Fisher, B., et al., *Effect of Pre-operative Chemotherapy on the Outcome of Women with Operable Breast Cancer*. *J Clin Oncol*, 1998. **16**: p. 2672-85.
201. Galow, J.R., et al., *Preoperative therapy in invasive breast cancer: pathologic assessment and systemic therapy issues in operable disease*. *J Clin Oncol*, 2008. **26**(5): p. 814-9.
202. Mauri, D., N. Pavlidis, and J.P. Ioannidis, *Neoadjuvant versus adjuvant systemic treatment in breast cancer: a meta-analysis*. *J Natl Cancer Inst*, 2005. **97**(3): p. 188-94.
203. Fisher, B., et al., *Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-18*. *J Clin Oncol*, 1997. **15**(7): p. 2483-93.
204. Heys, S., et al., *Neoadjuvant Docetaxel in Breast Cancer. 3 year Survival Results from the Aberdeen Trial*. *Clin Breast Cancer*, 2002. **3 Suppl 2**: p. S69-S74.
205. Coconi, G., et al., *Randomised Trial comparing Cyclophosphamide, Methotrexate and 5-Fluorouracil (CMF) with rotational CMF, Epirubicin and Vincristine as Primary Chemotherapy in Operable Breast Carcinoma*. *Cancer*, 2002. **95**(2): p. 228-35.
206. Van der Hage, J., et al., *Preoperative Chemotherapy in Primary Operable Breast Cancer: Results from the European Organisation for Research and Treatment of Cancer Trial*. *J Clin Oncol*, 2001. **19**(22): p. 4224-37.
207. Chan, S., et al., *Prospective Randomised Trial of Docetaxel versus Doxorubicin in Patients with Metastatic Breast Cancer*. *J Clin Oncol*, 1999. **17** (8): p. 2341-54.
208. *Capecitabine in Stage IV Breast Cancer*, 2003, Cancer Care Ontario. Breast Cancer Site Group: Toronto. Available at <http://www.cancercare.on.ca>.
209. *Vinorelbine in Stage IV Cancer*, 2005, Cancer Care Ontario. Breast Cancer Disease Group: Toronto. Available at <http://www.cancercare.on.ca>.
210. Chia, S., et al., *Locally advanced and inflammatory breast cancer*. *J Clin Oncol*, 2008. **26**(5): p. 786-90.
211. Fossati, R., et al., *Cytotoxic and hormonal treatment for metastatic breast cancer: a systematic review of published randomized trials involving 31,510 women*. *J Clin Oncol*, 1998. **16**(10): p. 3439-60.
212. Lewis, R., et al., *The Clinical Effectiveness of Trastuzumab for Breast Cancer*. *Health Technol Assess*, 2002. **6**(13): p. 1-71.
213. Gianni, L., et al., *Treatment with trastuzumab for 1 year after adjuvant chemotherapy in patients with HER2-positive early breast cancer: a 4-year follow-up of a randomised controlled trial*. *Lancet Oncol*, 2011. **12**(3): p. 236-44.
214. Goldhirsch, A., et al., *2 years versus 1 year of adjuvant trastuzumab for HER2-positive breast cancer (HERA): an open-label, randomised controlled trial*. *Lancet*, 2013.

215. Piccart-Gebhart, M., et al., *Trastuzumab after Adjuvant Chemotherapy in HER2 Positive Breast Cancer*. N Eng J Med, 2005. **353**: p. 1659-72.
216. Romond, H., et al., *Trastuzumab plus Adjuvant Chemotherapy for Operable HER2 Positive Breast Cancer*. N Eng J Med, 2005. **353**: p. 1673-84.
217. Baselga, J., *Herceptin alone or in combination with chemotherapy in the treatment of HER2-positive metastatic breast cancer: pivotal trials*. Oncology, 2001. **61 Suppl 2**: p. 14-21.
218. Slamon, D., et al., *Use of Chemotherapy plus Monoclonal Antibody against HER2 for Metastatic Breast Cancer that Overexpresses HER2*. N Eng J Med, 2001. **344(11)**: p. 783-92.
219. Goldhirsch, A., et al., *2 years versus 1 year of adjuvant trastuzumab for HER2-positive breast cancer (HERA): an open-label, randomised controlled trial*. Lancet, 2013. **382(9897)**: p. 1021-8.
220. Metzger Filho, O., E.P. Winer, and I. Krop, *Pertuzumab: Optimizing HER2 blockade*. Clin Cancer Res, 2013.
221. Beatson, G.T., *On the Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a New Method of Treatment, with Illustrative Cases*. Lancet, 1896. **2**: p. 104-7.
222. Anderson, E., *The role of oestrogen and progesterone receptors in human mammary development and tumorigenesis*. Breast Cancer Res, 2002. **4(5)**: p. 197-201.
223. Russo, J. and I.H. Russo, *Breast development, hormones and cancer*. Adv Exp Med Biol, 2008. **630**: p. 52-6.
224. Anderson, E., R.B. Clarke, and A. Howell, *Estrogen responsiveness and control of normal human breast proliferation*. J Mammary Gland Biol Neoplasia, 1998. **3(1)**: p. 23-35.
225. Allred, D.C., S.K. Mohsin, and S.A. Fuqua, *Histological and biological evolution of human premalignant breast disease*. Endocr Relat Cancer, 2001. **8(1)**: p. 47-61.
226. Laron, Z., R. Pauli, and A. Pertzalan, *Clinical Evidence of the Role of Estrogens in the Development of the Breasts*. Proc R Soc Edinburgh 1989(95): p. 13-22.
227. Key, T.J. and M.C. Pike, *The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer*. Eur J Cancer Clin Oncol, 1988. **24(1)**: p. 29-43.
228. Bocchinfuso, W.P. and K.S. Korach, *Mammary gland development and tumorigenesis in estrogen receptor knockout mice*. J Mammary Gland Biol Neoplasia, 1997. **2(4)**: p. 323-34.
229. Humphreys, R., et al., *Use of the PRKO Mice to Study The Role of Progesterone in Mammary Gland Development*. J Mammary Gland Biol Neoplasia, 1997(2): p. 343-354.
230. Schairer, C., et al., *Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer risk*. JAMA, 2000. **283(4)**: p. 485-91.
231. Miller, W.R., *Background and Development of Aromatase Inhibitors in Aromatase Inhibitors*, F. BJA, Editor. 2008, Berghauser Verlag AG, Switzerland: Berlin, Germany.
232. Miller, W.R., *Endocrine treatment for breast cancers: biological rationale and current progress*. J Steroid Biochem Mol Biol, 1990. **37(4)**: p. 467-80.
233. Sommer, S. and S.A. Fuqua, *Estrogen receptor and breast cancer*. Semin Cancer Biol, 2001. **11(5)**: p. 339-52.
234. Kumar, V., et al., *Functional domains of the human estrogen receptor*. Cell, 1987. **51(6)**: p. 941-51.
235. Ribeiro, R.C., P.J. Kushner, and J.D. Baxter, *The nuclear hormone receptor gene superfamily*. Annu Rev Med, 1995. **46**: p. 443-53.
236. Klein-Hitpass, L., et al., *A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor*. Nucleic Acids Res, 1988. **16(2)**: p. 647-63.
237. Jackson, T.A., et al., *The partial agonist activity of antagonist-occupied steroid receptors is controlled by a novel hinge domain-binding coactivator L7/SPA and the corepressors N-CoR or SMRT*. Mol Endocrinol, 1997. **11(6)**: p. 693-705.
238. Enmark, E., et al., *Human estrogen receptor beta-gene structure, chromosomal localization, and expression pattern*. J Clin Endocrinol Metab, 1997. **82(12)**: p. 4258-65.

239. Hall, J.M. and D.P. McDonnell, *The estrogen receptor beta-isoform (ERbeta) of the human estrogen receptor modulates ERalpha transcriptional activity and is a key regulator of the cellular response to estrogens and antiestrogens*. *Endocrinology*, 1999. **140**(12): p. 5566-78.
240. Clarke, C.L. and R.L. Sutherland, *Progesterin regulation of cellular proliferation*. *Endocr Rev*, 1990. **11**(2): p. 266-301.
241. Dowsett, M., *Biology of Estrogen-Dependent Breast Cancer*. Aromatase Inhibitors for the Treatment of Breast Cancer, ed. E. MJ. 2008, New York: CMP Healthcare Media, Oncology Publishing Group.
242. Collins, L.C., M.L. Botero, and S.J. Schnitt, *Bimodal frequency distribution of estrogen receptor immunohistochemical staining results in breast cancer: an analysis of 825 cases*. *Am J Clin Pathol*, 2005. **123**(1): p. 16-20.
243. Miller, W., *Aromatase Inhibitors*. *Endocr Relat Cancer*, 1996. **3**: p. 65-79.
244. Love, R.R. and J. Philips, *Oophorectomy for breast cancer: history revisited*. *J Natl Cancer Inst*, 2002. **94**(19): p. 1433-4.
245. Leung, S.F., et al., *Ovarian ablation failures by radiation: a comparison of two dose schedules*. *Br J Radiol*, 1991. **64**(762): p. 537-8.
246. Prowell, T.M. and N.E. Davidson, *What is the role of ovarian ablation in the management of primary and metastatic breast cancer today?* *Oncologist*, 2004. **9**(5): p. 507-17.
247. Robertson, J.F. and R.W. Blamey, *The use of gonadotrophin-releasing hormone (GnRH) agonists in early and advanced breast cancer in pre- and perimenopausal women*. *Eur J Cancer*, 2003. **39**(7): p. 861-9.
248. Goel, S., et al., *LHRH agonists for adjuvant therapy of early breast cancer in premenopausal women*. *Cochrane Database Syst Rev*, 2009(4): p. CD004562.
249. Stearns, V., et al., *Breast cancer treatment and ovarian failure: risk factors and emerging genetic determinants*. *Nat Rev Cancer*, 2006. **6**(11): p. 886-93.
250. Lephart, E.D., *A review of brain aromatase cytochrome P450*. *Brain Res Brain Res Rev*, 1996. **22**(1): p. 1-26.
251. Longcope, C., J.H. Pratt, and S.H. Schneider, *The Interconversion and Aromatisation of Androgens by Human Adipose Tissue*. *J Clin Endocrinol Metab*, 1978(46): p. 146-52.
252. Miller, W.R. and A.P. Forrest, *Oestradiol synthesis by a human breast carcinoma*. *Lancet*, 1974. **2**(7885): p. 866-8.
253. Perel, E. and D.W. Killinger, *The Interconversion and Aromatisation of Androgens by Human Adipose Tissue*. *J Steroid Biochem* 1979(10): p. 623-7.
254. Smuk, M. and J. Schwiers, *Aromatisation of Androgens by Muscle and Adipose Tissue In-Vitro*. *J Clin Endocrinol Metab*, 1977(45): p. 1009-12.
255. Miller, W.R., *Aromatase inhibitors in the treatment of advanced breast cancer*. *Cancer Treat Rev*, 1989. **16**(2): p. 83-93.
256. Combs, D., *Recent Developments in Aromatase Inhibitors*. *Exp Opin Ther Patents* 1995(5): p. 529-34.
257. Geisler, J., *Aromatase inhibitors: from bench to bedside and back*. *Breast Cancer*, 2008. **15**(1): p. 17-26.
258. Miller, W.R., *Aromatase inhibitors and breast cancer*. *Minerva Endocrinol*, 2006. **31**(1): p. 27-46.
259. Buzdar, A.U., et al., *An overview of the pharmacology and pharmacokinetics of the newer generation aromatase inhibitors anastrozole, letrozole, and exemestane*. *Cancer*, 2002. **95**(9): p. 2006-16.
260. Dixon, J.M., et al., *Letrozole suppresses plasma estradiol and estrone sulphate more completely than anastrozole in postmenopausal women with breast cancer*. *J Clin Oncol*, 2008. **26**(10): p. 1671-6.
261. Dowsett, M., et al., *In vivo measurement of aromatase inhibition by letrozole (CGS 20267) in postmenopausal patients with breast cancer*. *Clin Cancer Res*, 1995. **1**(12): p. 1511-5.
262. Geisler, J., et al., *Influence of letrozole and anastrozole on total body aromatization and plasma estrogen levels in postmenopausal breast cancer patients evaluated in a randomized, cross-over study*. *J Clin Oncol*, 2002. **20**(3): p. 751-7.

263. Sainsbury, R., *Aromatase inhibition in the treatment of advanced breast cancer: is there a relationship between potency and clinical efficacy?* Br J Cancer, 2004. **90**(9): p. 1733-9.
264. McDonnell, D.P. and S.E. Wardell, *The molecular mechanisms underlying the pharmacological actions of ER modulators: implications for new drug discovery in breast cancer.* Curr Opin Pharmacol, 2010. **10**(6): p. 620-8.
265. Jordan, V.C., *Selective estrogen receptor modulation: a personal perspective.* Cancer Res, 2001. **61**(15): p. 5683-7.
266. Love, R.R., et al., *Effects of tamoxifen on bone mineral density in postmenopausal women with breast cancer.* N Engl J Med, 1992. **326**(13): p. 852-6.
267. Fisher, B., et al., *Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14.* J Natl Cancer Inst, 1994. **86**(7): p. 527-37.
268. Pinkerton, J.V. and S.R. Goldstein, *Endometrial safety: a key hurdle for selective estrogen receptor modulators in development.* Menopause, 2010. **17**(3): p. 642-53.
269. Shelly, W., et al., *Selective estrogen receptor modulators: an update on recent clinical findings.* Obstet Gynecol Surv, 2008. **63**(3): p. 163-81.
270. Johnston, S.J. and K.L. Cheung, *Fulvestrant - a novel endocrine therapy for breast cancer.* Curr Med Chem, 2010. **17**(10): p. 902-14.
271. AstraZeneca, *"Faslodex (fulvestrant) injection"*. Review 08/04, 2004.
272. Osborne, C.K., et al., *Double-blind, randomized trial comparing the efficacy and tolerability of fulvestrant versus anastrozole in postmenopausal women with advanced breast cancer progressing on prior endocrine therapy: results of a North American trial.* J Clin Oncol, 2002. **20**(16): p. 3386-95.
273. *Tamoxifen for early breast cancer: an overview of the randomised trials.* Early Breast Cancer Trialists' Collaborative Group. Lancet, 1998. **351**(9114): p. 1451-67.
274. O'Regan, R., V.C. Jordan, and W.J. Gradishar, *Tamoxifen and contralateral breast cancer.* J Am Coll Surg, 1999. **188**(6): p. 678-83.
275. Howell, A., et al., *Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer.* Lancet, 2005. **365**(9453): p. 60-2.
276. Jakesz, R., et al., *Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial.* Lancet, 2005. **366**(9484): p. 455-62.
277. Baum, M., et al., *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, 2002. **359**(9324): p. 2131-9.
278. Regan, M.M., et al., *Interpreting Breast International Group (BIG) 1-98: a randomized, double-blind, phase III trial comparing letrozole and tamoxifen as adjuvant endocrine therapy for postmenopausal women with hormone receptor-positive, early breast cancer.* Breast Cancer Res, 2011. **13**(3): p. 209.
279. Jonat, W. and C. Mundhenke, *The FACE trial: letrozole or anastrozole as initial adjuvant therapy?* Cancer Invest, 2007. **25**(1): p. 14-8.
280. Goss, P.E., et al., *Exemestane versus anastrozole in postmenopausal women with early breast cancer: NCIC CTG MA.27--a randomized controlled phase III trial.* J Clin Oncol, 2013. **31**(11): p. 1398-404.
281. Coombes, R.C., et al., *A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer.* N Engl J Med, 2004. **350**(11): p. 1081-92.
282. Van de Velde, C.J., et al., *Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial.* Lancet, 2011. **377**(9762): p. 321-31.
283. Goss, P.E., et al., *A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer.* N Engl J Med, 2003. **349**(19): p. 1793-802.
284. Goss, P.E., et al., *Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17.* J Natl Cancer Inst, 2005. **97**(17): p. 1262-71.
285. Robinson, A., *A review of the use of exemestane in early breast cancer.* Ther Clin Risk Manag, 2009. **5**(1): p. 91-8.

286. Jakesz, R., et al., *Extended adjuvant therapy with anastrozole among postmenopausal breast cancer patients: results from the randomized Austrian Breast and Colorectal Cancer Study Group Trial 6a*. J Natl Cancer Inst, 2007. **99**(24): p. 1845-53.
287. Love, R.R., et al., *Oophorectomy and tamoxifen adjuvant therapy in premenopausal Vietnamese and Chinese women with operable breast cancer*. J Clin Oncol, 2002. **20**(10): p. 2559-66.
288. Jakesz, R., et al., *Randomized adjuvant trial of tamoxifen and goserelin versus cyclophosphamide, methotrexate, and fluorouracil: evidence for the superiority of treatment with endocrine blockade in premenopausal patients with hormone-responsive breast cancer--Austrian Breast and Colorectal Cancer Study Group Trial 5*. J Clin Oncol, 2002. **20**(24): p. 4621-7.
289. Kaufmann, M., et al., *Survival analyses from the ZEBRA study. goserelin (Zoladex) versus CMF in premenopausal women with node-positive breast cancer*. Eur J Cancer, 2003. **39**(12): p. 1711-7.
290. Klijn, J.G., et al., *Combined tamoxifen and luteinizing hormone-releasing hormone (LHRH) agonist versus LHRH agonist alone in premenopausal advanced breast cancer: a meta-analysis of four randomized trials*. J Clin Oncol, 2001. **19**(2): p. 343-53.
291. Fisher, B., et al., *Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer*. J Natl Cancer Inst, 1997. **89**(22): p. 1673-82.
292. Swain, S.M., et al., *Longer therapy, iatrogenic amenorrhea, and survival in early breast cancer*. N Engl J Med, 2010. **362**(22): p. 2053-65.
293. Dixon, J.M., et al., *Outcome of Frail Elderly Patients Treated with Letrozole Alone*. Cancer Res, 2010(70 (24 Suppl)): p. Abstract P1-12-06.
294. Torrisi, R., et al., *Antitumour and biological effects of letrozole and GnRH analogue as primary therapy in premenopausal women with ER and PgR positive locally advanced operable breast cancer*. Br J Cancer, 2007. **97**(6): p. 802-8.
295. Nagaraj, G., et al., *Management of Premenopausal Women with Neoadjuvant Endocrine Therapy: A Single Institution Experience*. J Clin Oncol, 2011(29 (Suppl): Abstr e11014).
296. Eiermann, W., et al., *Preoperative treatment of postmenopausal breast cancer patients with letrozole: A randomized double-blind multicenter study*. Ann Oncol, 2001. **12**(11): p. 1527-32.
297. Smith, I.E., et al., *Neoadjuvant treatment of postmenopausal breast cancer with anastrozole, tamoxifen, or both in combination: the Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) multicenter double-blind randomized trial*. J Clin Oncol, 2005. **23**(22): p. 5108-16.
298. Cataliotti, L., et al., *Comparison of anastrozole versus tamoxifen as preoperative therapy in postmenopausal women with hormone receptor-positive breast cancer: the Pre-Operative "Arimidex" Compared to Tamoxifen (PROACT) trial*. Cancer, 2006. **106**(10): p. 2095-103.
299. Takei, H., et al., *Multicenter phase II trial of neoadjuvant exemestane for postmenopausal patients with hormone receptor-positive, operable breast cancer: Saitama Breast Cancer Clinical Study Group (SBCCSG-03)*. Breast Cancer Res Treat, 2008. **107**(1): p. 87-94.
300. Seo, J.H., Y.H. Kim, and J.S. Kim, *Meta-analysis of pre-operative aromatase inhibitor versus tamoxifen in postmenopausal woman with hormone receptor-positive breast cancer*. Cancer Chemother Pharmacol, 2009. **63**(2): p. 261-6.
301. Ellis, M.J., et al., *Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype--ACOSOG Z1031*. J Clin Oncol, 2011. **29**(17): p. 2342-9.
302. Dixon, J.M., et al., *Invasive lobular carcinoma: response to neoadjuvant letrozole therapy*. Breast Cancer Res Treat, 2011. **130**(3): p. 871-7.
303. Bear, H.D., et al., *The effect on tumor response of adding sequential preoperative docetaxel to preoperative doxorubicin and cyclophosphamide: preliminary results from National Surgical Adjuvant Breast and Bowel Project Protocol B-27*. J Clin Oncol, 2003. **21**(22): p. 4165-74.

304. Dixon, J.M., T.J. Anderson, and W.R. Miller, *Neoadjuvant endocrine therapy of breast cancer: a surgical perspective*. *Eur J Cancer*, 2002. **38**(17): p. 2214-21.
305. Dixon, J.M., et al., *Increase in response rate by prolonged treatment with neoadjuvant letrozole*. *Breast Cancer Res Treat*, 2009. **113**(1): p. 145-51.
306. Carpenter, R., et al., *A Multicentre Study to Determine the Optimum Duration of Neoadjuvant Letrozole on Tumour Regression to Permit Breast Conserving Surgery: Updated Analyses*. *J Clin Oncol*, 2010(28:670 (Suppl; abstr 670)).
307. Freedman, O.C., et al., *A randomized trial exploring the biomarker effects of neoadjuvant sequential treatment with exemestane and anastrozole in post-menopausal women with hormone receptor-positive breast cancer*. *Breast Cancer Res Treat*, 2010. **119**(1): p. 155-61.
308. Miller, W.R., et al., *Aromatase inhibitors: are there differences between steroidal and nonsteroidal aromatase inhibitors and do they matter?* *Oncologist*, 2008. **13**(8): p. 829-37.
309. Thomas, J.S., et al., *Histopathology of breast carcinoma following neoadjuvant systemic therapy: a common association between letrozole therapy and central scarring*. *Histopathology*, 2007. **51**(2): p. 219-26.
310. Sahoo, S. and S.C. Lester, *Pathology of breast carcinomas after neoadjuvant chemotherapy: an overview with recommendations on specimen processing and reporting*. *Arch Pathol Lab Med*, 2009. **133**(4): p. 633-42.
311. Huober, J., et al., *Effect of neoadjuvant anthracycline-taxane-based chemotherapy in different biological breast cancer phenotypes: overall results from the GeparTrio study*. *Breast Cancer Res Treat*, 2010. **124**(1): p. 133-40.
312. Guarneri, V., et al., *Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors*. *J Clin Oncol*, 2006. **24**(7): p. 1037-44.
313. Ring, A.E., et al., *Oestrogen receptor status, pathological complete response and prognosis in patients receiving neoadjuvant chemotherapy for early breast cancer*. *Br J Cancer*, 2004. **91**(12): p. 2012-7.
314. Ellis, M.J., et al., *Estrogen-independent proliferation is present in estrogen-receptor HER2-positive primary breast cancer after neoadjuvant letrozole*. *J Clin Oncol*, 2006. **24**(19): p. 3019-25.
315. Ellis, M.J., et al., *Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial*. *J Clin Oncol*, 2001. **19**(18): p. 3808-16.
316. Dixon, J.M., et al., *Anastrozole demonstrates clinical and biological effectiveness in oestrogen receptor-positive breast cancers, irrespective of the erbB2 status*. *Eur J Cancer*, 2004. **40**(18): p. 2742-7.
317. Young, O., et al., *Neoadjuvant Letrozole is Equally Effective in HER2 positive and Negative Breast Cancers*. *Breast Cancer Res Treat*, 2004(88 (1s)): p. S36, abs 411.
318. Larionov, A.A. and W.R. Miller, *Challenges in defining predictive markers for response to endocrine therapy in breast cancer*. *Future Oncol*, 2009. **5**(9): p. 1415-28.
319. El-Maraghi, R.H. and E.A. Eisenhauer, *Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III*. *J Clin Oncol*, 2008. **26**(8): p. 1346-54.
320. Murray, J., W.R. Miller, and J.M. Dixon, *Breast Cancer Research protocols. Neoadjuvant Endocrine Therapy Models*, in *Methods in Molecular Medicine*, S. Brooks, Harris, LA., Editor. 2005, Humana Press Inc: Totowa, New Jersey. p. 489-502.
321. Miller, A.B., et al., *Reporting results of cancer treatment*. *Cancer*, 1981. **47**(1): p. 207-14.
322. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)*. *Eur J Cancer*, 2009. **45**(2): p. 228-47.
323. Therasse, P., et al., *New guidelines to evaluate the response to treatment in solid tumors*. *European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada*. *J Natl Cancer Inst*, 2000. **92**(3): p. 205-16.

324. Therasse, P., E.A. Eisenhauer, and J. Verweij, *RECIST revisited: a review of validation studies on tumour assessment*. Eur J Cancer, 2006. **42**(8): p. 1031-9.
325. Kuerer, H.M., et al., *Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy*. J Clin Oncol, 1999. **17**(2): p. 460-9.
326. Dixon, J.M., et al., *Lessons from the use of aromatase inhibitors in the neoadjuvant setting*. Endocr Relat Cancer, 1999. **6**(2): p. 227-30.
327. Sasano, H., T. Suzuki, and T. Moriya, *Pathology of Breast Cancer following Neoadjuvant Therapy*, in *Endocrine Therapy in Breast Cancer*, W. Miller, Ingle, JN., Editor. 2002, Marcel Dekker: Ney York, USA. p. 213-222.
328. Anderson, T.J., et al., *Effect of neoadjuvant treatment with anastrozole on tumour histology in postmenopausal women with large operable breast cancer*. Br J Cancer, 2002. **87**(3): p. 334-8.
329. Miller, W.R., et al., *Pathological features of breast cancer response following neoadjuvant treatment with either letrozole or tamoxifen*. Eur J Cancer, 2003. **39**(4): p. 462-8.
330. Miller, W.R., et al., *Biological and clinical effects of aromatase inhibitors in neoadjuvant therapy*. J Steroid Biochem Mol Biol, 2001. **79**(1-5): p. 103-7.
331. Ellis, M.J., et al., *Letrozole inhibits tumor proliferation more effectively than tamoxifen independent of HER1/2 expression status*. Cancer Res, 2003. **63**(19): p. 6523-31.
332. Dowsett, M., et al., *Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer*. J Natl Cancer Inst, 2007. **99**(2): p. 167-70.
333. Dowsett, M., et al., *Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival*. Clin Cancer Res, 2005. **11**(2 Pt 2): p. 951s-8s.
334. Dowsett, M., et al., *Short term Pre-Surgical Treatment for All?* Breast Cancer Res, 2007(9 (Suppl 2), S13).
335. Miller, W.R., *Biological rationale for endocrine therapy in breast cancer*. Best Pract Res Clin Endocrinol Metab, 2004. **18**(1): p. 1-32.
336. Miller, W.R. and A.A. Larionov, *Understanding the mechanisms of aromatase inhibitor resistance*. Breast Cancer Res, 2012. **14**(1): p. 201.
337. Smith, I.E. and M. Dowsett, *Aromatase inhibitors in breast cancer*. N Engl J Med, 2003. **348**(24): p. 2431-42.
338. Miller, W.R. and A. Larionov, *Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole*. Breast Cancer Res, 2010. **12**(4): p. R52.
339. Musgrove, E.A. and R.L. Sutherland, *Biological determinants of endocrine resistance in breast cancer*. Nat Rev Cancer, 2009. **9**(9): p. 631-43.
340. Kushner, P.J., et al., *Estrogen receptor pathways to AP-1*. J Steroid Biochem Mol Biol, 2000. **74**(5): p. 311-7.
341. Gotay, C. and J. Dunn, *Adherence to long-term adjuvant hormonal therapy for breast cancer*. Expert Rev Pharmacoecon Outcomes Res, 2011. **11**(6): p. 709-15.
342. Goetz, M.P., et al., *The impact of cytochrome P450 2D6 metabolism in women receiving adjuvant tamoxifen*. Breast Cancer Res Treat, 2007. **101**(1): p. 113-21.
343. Nadji, M., *Quantitative immunohistochemistry of estrogen receptor in breast cancer: "much ado about nothing!"*. Appl Immunohistochem Mol Morphol, 2008. **16**(2): p. 105-7.
344. Van Duursen, M.B., et al., *Phytoestrogens in menopausal supplements induce ER-dependent cell proliferation and overcome breast cancer treatment in an in vitro breast cancer model*. Toxicol Appl Pharmacol, 2013. **269**(2): p. 132-40.
345. Patisaul, H.B. and W. Jefferson, *The pros and cons of phytoestrogens*. Front Neuroendocrinol, 2010. **31**(4): p. 400-19.
346. Garcia-Casado, Z., et al., *A polymorphism at the 3'-UTR region of the aromatase gene defines a subgroup of postmenopausal breast cancer patients with poor response to neoadjuvant letrozole*. BMC Cancer, 2010. **10**: p. 36.
347. Sutherland, R.L. and E.A. Musgrove, *Cyclins and breast cancer*. J Mammary Gland Biol Neoplasia, 2004. **9**(1): p. 95-104.

348. Osborne, C.K. and R. Schiff, *Mechanisms of endocrine resistance in breast cancer*. *Annu Rev Med*, 2011. **62**: p. 233-47.
349. Osborne, C.K. and R. Schiff, *Estrogen-receptor biology: continuing progress and therapeutic implications*. *J Clin Oncol*, 2005. **23**(8): p. 1616-22.
350. Girault, I., I. Bieche, and R. Lidereau, *Role of estrogen receptor alpha transcriptional coregulators in tamoxifen resistance in breast cancer*. *Maturitas*, 2006. **54**(4): p. 342-51.
351. Acconcia, F. and R. Kumar, *Signaling regulation of genomic and nongenomic functions of estrogen receptors*. *Cancer Lett*, 2006. **238**(1): p. 1-14.
352. Klinge, C.M., *Estrogen receptor interaction with estrogen response elements*. *Nucleic Acids Res*, 2001. **29**(14): p. 2905-19.
353. Bjornstrom, L. and M. Sjoberg, *Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes*. *Mol Endocrinol*, 2005. **19**(4): p. 833-42.
354. Schiff, R., et al., *Cross-talk between estrogen receptor and growth factor pathways as a molecular target for overcoming endocrine resistance*. *Clin Cancer Res*, 2004. **10**(1 Pt 2): p. 331S-6S.
355. Arnold, A. and A. Papanikolaou, *Cyclin D1 in breast cancer pathogenesis*. *J Clin Oncol*, 2005. **23**(18): p. 4215-24.
356. Barnes, D.M. and C.E. Gillett, *Cyclin D1 in breast cancer*. *Breast Cancer Res Treat*, 1998. **52**(1-3): p. 1-15.
357. Dickson, C., et al., *Amplification of chromosome band 11q13 and a role for cyclin D1 in human breast cancer*. *Cancer Lett*, 1995. **90**(1): p. 43-50.
358. Aaltonen, K., et al., *Cyclin D1 expression is associated with poor prognostic features in estrogen receptor positive breast cancer*. *Breast Cancer Res Treat*, 2009. **113**(1): p. 75-82.
359. Balasenthil, S., et al., *Estrogen receptor activation at serine 305 is sufficient to upregulate cyclin D1 in breast cancer cells*. *FEBS Lett*, 2004. **567**(2-3): p. 243-7.
360. Wilcken, N.R., et al., *Inducible overexpression of cyclin D1 in breast cancer cells reverses the growth-inhibitory effects of antiestrogens*. *Clin Cancer Res*, 1997. **3**(6): p. 849-54.
361. Bostner, J., et al., *Amplification of CCND1 and PAK1 as predictors of recurrence and tamoxifen resistance in postmenopausal breast cancer*. *Oncogene*, 2007. **26**(49): p. 6997-7005.
362. Jirstrom, K., et al., *Adverse effect of adjuvant tamoxifen in premenopausal breast cancer with cyclin D1 gene amplification*. *Cancer Res*, 2005. **65**(17): p. 8009-16.
363. Akli, S., et al., *Low-molecular-weight cyclin E can bypass letrozole-induced G1 arrest in human breast cancer cells and tumors*. *Clin Cancer Res*, 2010. **16**(4): p. 1179-90.
364. Berglund, P. and G. Landberg, *Cyclin e overexpression reduces infiltrative growth in breast cancer: yet another link between proliferation control and tumor invasion*. *Cell Cycle*, 2006. **5**(6): p. 606-9.
365. Caldon, C.E. and E.A. Musgrove, *Distinct and redundant functions of cyclin E1 and cyclin E2 in development and cancer*. *Cell Div*, 2010. **5**: p. 2.
366. Hunt, K.K. and K. Keyomarsi, *Cyclin E as a prognostic and predictive marker in breast cancer*. *Semin Cancer Biol*, 2005. **15**(4): p. 319-26.
367. Bunone, G., et al., *Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation*. *EMBO J*, 1996. **15**(9): p. 2174-83.
368. Kato, S., et al., *Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase*. *Science*, 1995. **270**(5241): p. 1491-4.
369. Schiff, R., et al., *Breast cancer endocrine resistance: how growth factor signaling and estrogen receptor coregulators modulate response*. *Clin Cancer Res*, 2003. **9**(1 Pt 2): p. 447S-54S.
370. Shou, J., et al., *Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer*. *J Natl Cancer Inst*, 2004. **96**(12): p. 926-35.
371. Wu, R.C., C.L. Smith, and B.W. O'Malley, *Transcriptional regulation by steroid receptor coactivator phosphorylation*. *Endocr Rev*, 2005. **26**(3): p. 393-9.

372. Saxena, R. and A. Dwivedi, *ErbB family receptor inhibitors as therapeutic agents in breast cancer: current status and future clinical perspective*. Med Res Rev, 2012. **32**(1): p. 166-215.
373. Olayioye, M.A., et al., *The ErbB signaling network: receptor heterodimerization in development and cancer*. EMBO J, 2000. **19**(13): p. 3159-67.
374. Witsch, E., M. Sela, and Y. Yarden, *Roles for growth factors in cancer progression*. Physiology (Bethesda), 2010. **25**(2): p. 85-101.
375. Rasmussen, B.B., et al., *Adjuvant letrozole versus tamoxifen according to centrally-assessed ERBB2 status for postmenopausal women with endocrine-responsive early breast cancer: supplementary results from the BIG 1-98 randomised trial*. Lancet Oncol, 2008. **9**(1): p. 23-8.
376. Ahnstrom, M., et al., *Role of cyclin D1 in ErbB2-positive breast cancer and tamoxifen resistance*. Breast Cancer Res Treat, 2005. **91**(2): p. 145-51.
377. Osborne, C.K., et al., *Crosstalk between estrogen receptor and growth factor receptor pathways as a cause for endocrine therapy resistance in breast cancer*. Clin Cancer Res, 2005. **11**(2 Pt 2): p. 865s-70s.
378. Yamashita, H., et al., *p53 protein accumulation predicts resistance to endocrine therapy and decreased post-relapse survival in metastatic breast cancer*. Breast Cancer Res, 2006. **8**(4): p. R48.
379. Nehra, R., et al., *BCL2 and CASP8 regulation by NF-kappaB differentially affect mitochondrial function and cell fate in antiestrogen-sensitive and -resistant breast cancer cells*. FASEB J, 2010. **24**(6): p. 2040-55.
380. Zhou, X., et al., *MicroRNA-9 as potential biomarker for breast cancer local recurrence and tumor estrogen receptor status*. PLoS One, 2012. **7**(6): p. e39011.
381. Yoshimoto, N., et al., *Distinct expressions of microRNAs that directly target estrogen receptor alpha in human breast cancer*. Breast Cancer Res Treat, 2011. **130**(1): p. 331-9.
382. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
383. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
384. Ahmad, A., et al., *Cell Cycle Regulatory Proteins in Breast Cancer: Molecular Determinants of Drug Resistance and Targets for Anti-Cancer Therapies*. Targeting New Pathways and Cell Death in Breast Cancer, ed. D.R. Aft. 2012: In Tech.
385. Butt, A.J., et al., *Cell cycle machinery: links with genesis and treatment of breast cancer*. Adv Exp Med Biol, 2008. **630**: p. 189-205.
386. Besson, A., S.F. Dowdy, and J.M. Roberts, *CDK inhibitors: cell cycle regulators and beyond*. Dev Cell, 2008. **14**(2): p. 159-69.
387. Lemmon, M.A. and J. Schlessinger, *Cell signaling by receptor tyrosine kinases*. Cell, 2010. **141**(7): p. 1117-34.
388. Hynes, N.E. and G. MacDonald, *ErbB receptors and signaling pathways in cancer*. Curr Opin Cell Biol, 2009. **21**(2): p. 177-84.
389. Perona, R., *Cell signalling: growth factors and tyrosine kinase receptors*. Clin Transl Oncol, 2006. **8**(2): p. 77-82.
390. Bhowmick, N.A., E.G. Neilson, and H.L. Moses, *Stromal fibroblasts in cancer initiation and progression*. Nature, 2004. **432**(7015): p. 332-7.
391. Cheng, N., et al., *Transforming growth factor-beta signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion*. Mol Cancer Res, 2008. **6**(10): p. 1521-33.
392. Amit, I., et al., *A module of negative feedback regulators defines growth factor signaling*. Nat Genet, 2007. **39**(4): p. 503-12.
393. Burkhart, D.L. and J. Sage, *Cellular mechanisms of tumour suppression by the retinoblastoma gene*. Nat Rev Cancer, 2008. **8**(9): p. 671-82.
394. Curto, M., et al., *Contact-dependent inhibition of EGFR signaling by Nf2/Merlin*. J Cell Biol, 2007. **177**(5): p. 893-903.
395. Adams, J.M. and S. Cory, *The Bcl-2 apoptotic switch in cancer development and therapy*. Oncogene, 2007. **26**(9): p. 1324-37.
396. Lowe, S.W., E. Cepero, and G. Evan, *Intrinsic tumour suppression*. Nature, 2004. **432**(7015): p. 307-15.

397. Apel, A., et al., *Autophagy-A double-edged sword in oncology*. Int J Cancer, 2009. **125**(5): p. 991-5.
398. Blasco, M.A., *Telomeres and human disease: ageing, cancer and beyond*. Nat Rev Genet, 2005. **6**(8): p. 611-22.
399. Hanahan, D. and J. Folkman, *Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis*. Cell, 1996. **86**(3): p. 353-64.
400. Ferrara, N., *Vascular endothelial growth factor*. Arterioscler Thromb Vasc Biol, 2009. **29**(6): p. 789-91.
401. Kazerounian, S., K.O. Yee, and J. Lawler, *Thrombospondins in cancer*. Cell Mol Life Sci, 2008. **65**(5): p. 700-12.
402. Zumsteg, A. and G. Christofori, *Corrupt policemen: inflammatory cells promote tumor angiogenesis*. Curr Opin Oncol, 2009. **21**(1): p. 60-70.
403. Talmadge, J.E. and I.J. Fidler, *AACR centennial series: the biology of cancer metastasis: historical perspective*. Cancer Res, 2010. **70**(14): p. 5649-69.
404. Klymkowsky, M.W. and P. Savagner, *Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe*. Am J Pathol, 2009. **174**(5): p. 1588-93.
405. Karnoub, A.E., et al., *Mesenchymal stem cells within tumour stroma promote breast cancer metastasis*. Nature, 2007. **449**(7162): p. 557-63.
406. Egeblad, M., E.S. Nakasone, and Z. Werb, *Tumors as organs: complex tissues that interface with the entire organism*. Dev Cell, 2010. **18**(6): p. 884-901.
407. Qian, B.Z. and J.W. Pollard, *Macrophage diversity enhances tumor progression and metastasis*. Cell, 2010. **141**(1): p. 39-51.
408. Liotta, L.A. and E.C. Kohn, *The microenvironment of the tumour-host interface*. Nature, 2001. **411**(6835): p. 375-9.
409. Conklin, M.W. and P.J. Keely, *Why the stroma matters in breast cancer: insights into breast cancer patient outcomes through the examination of stromal biomarkers*. Cell Adh Migr, 2012. **6**(3): p. 249-60.
410. Boyd, N.F., et al., *Heritability of mammographic density, a risk factor for breast cancer*. N Engl J Med, 2002. **347**(12): p. 886-94.
411. Martin, L.J. and N.F. Boyd, *Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence*. Breast Cancer Res, 2008. **10**(1): p. 201.
412. Li, T., et al., *The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer*. Cancer Epidemiol Biomarkers Prev, 2005. **14**(2): p. 343-9.
413. Ursin, G., et al., *Greatly increased occurrence of breast cancers in areas of mammographically dense tissue*. Breast Cancer Res, 2005. **7**(5): p. R605-8.
414. Provenzano, P.P., et al., *Collagen reorganization at the tumor-stromal interface facilitates local invasion*. BMC Med, 2006. **4**(1): p. 38.
415. Provenzano, P.P., et al., *Collagen density promotes mammary tumor initiation and progression*. BMC Med, 2008. **6**: p. 11.
416. Conklin, M.W., et al., *Aligned collagen is a prognostic signature for survival in human breast carcinoma*. Am J Pathol, 2011. **178**(3): p. 1221-32.
417. Orimo, A., et al., *Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion*. Cell, 2005. **121**(3): p. 335-48.
418. Wiseman, B.S. and Z. Werb, *Stromal effects on mammary gland development and breast cancer*. Science, 2002. **296**(5570): p. 1046-9.
419. Grisendi, G., et al., *Understanding tumor-stroma interplays for targeted therapies by armed mesenchymal stromal progenitors: the Mesenkillers*. Am J Cancer Res, 2011. **1**(6): p. 787-805.
420. Leivonen, M., et al., *Prognostic value of syndecan-1 expression in breast cancer*. Oncology, 2004. **67**(1): p. 11-8.
421. Beauvais, D.M. and A.C. Rapraeger, *Syndecan-1 couples the insulin-like growth factor-1 receptor to inside-out integrin activation*. J Cell Sci, 2010. **123**(Pt 21): p. 3796-807.
422. Baba, F., et al., *Syndecan-1 and syndecan-4 are overexpressed in an estrogen receptor-negative, highly proliferative breast carcinoma subtype*. Breast Cancer Res Treat, 2006. **98**(1): p. 91-8.

423. Maeda, T., J. Desouky, and A. Friedl, *Syndecan-1 expression by stromal fibroblasts promotes breast carcinoma growth in vivo and stimulates tumor angiogenesis*. *Oncogene*, 2006. **25**(9): p. 1408-12.
424. Yang, N., et al., *Syndecan-1 in breast cancer stroma fibroblasts regulates extracellular matrix fiber organization and carcinoma cell motility*. *Am J Pathol*, 2011. **178**(1): p. 325-35.
425. Watson, C.J., *Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ*. *Breast Cancer Res*, 2006. **8**(2): p. 203.
426. O'Brien, J., et al., *Alternatively activated macrophages and collagen remodeling characterize the postpartum involuting mammary gland across species*. *Am J Pathol*, 2010. **176**(3): p. 1241-55.
427. Keleher, A.J., et al., *Multidisciplinary management of breast cancer concurrent with pregnancy*. *J Am Coll Surg*, 2002. **194**(1): p. 54-64.
428. Stensheim, H., et al., *Cause-specific survival for women diagnosed with cancer during pregnancy or lactation: a registry-based cohort study*. *J Clin Oncol*, 2009. **27**(1): p. 45-51.
429. O'Brien, J., et al., *Non-steroidal anti-inflammatory drugs target the pro-tumorigenic extracellular matrix of the postpartum mammary gland*. *Int J Dev Biol*, 2011. **55**(7-9): p. 745-55.
430. Markosyan, N., et al., *Deletion of cyclooxygenase 2 in mouse mammary epithelial cells delays breast cancer onset through augmentation of type 1 immune responses in tumors*. *Carcinogenesis*, 2011. **32**(10): p. 1441-9.
431. Ristimaki, A., et al., *Prognostic significance of elevated cyclooxygenase-2 expression in breast cancer*. *Cancer Res*, 2002. **62**(3): p. 632-5.
432. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, inflammation, and cancer*. *Cell*, 2010. **140**(6): p. 883-99.
433. Gocheva, V., et al., *IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion*. *Genes Dev*, 2010. **24**(3): p. 241-55.
434. Kessenbrock, K., V. Plaks, and Z. Werb, *Matrix metalloproteinases: regulators of the tumor microenvironment*. *Cell*, 2010. **141**(1): p. 52-67.
435. Mahmoud, S.M., et al., *Tumour-infiltrating macrophages and clinical outcome in breast cancer*. *J Clin Pathol*, 2012. **65**(2): p. 159-63.
436. Bostrom, P., et al., *MMP-1 expression has an independent prognostic value in breast cancer*. *BMC Cancer*, 2011. **11**: p. 348.
437. Tetu, B., et al., *The influence of MMP-14, TIMP-2 and MMP-2 expression on breast cancer prognosis*. *Breast Cancer Res*, 2006. **8**(3): p. R28.
438. Zhang, B., et al., *[Relationship between the expression of matrix metalloproteinase-13 protein and other biomarkers, prognosis in invasive breast cancer]*. *Zhonghua Bing Li Xue Za Zhi*, 2008. **37**(7): p. 471-6.
439. Zhang, B., et al., *Tumor-derived matrix metalloproteinase-13 (MMP-13) correlates with poor prognoses of invasive breast cancer*. *BMC Cancer*, 2008. **8**: p. 83.
440. Doedens, A.L., et al., *Macrophage expression of hypoxia-inducible factor-1 alpha suppresses T-cell function and promotes tumor progression*. *Cancer Res*, 2010. **70**(19): p. 7465-75.
441. Mahmoud, S.M., et al., *The prognostic significance of B lymphocytes in invasive carcinoma of the breast*. *Breast Cancer Res Treat*, 2012. **132**(2): p. 545-53.
442. Reis-Filho, J.S., et al., *Molecular profiling: moving away from tumor philately*. *Sci Transl Med*, 2010. **2**(47): p. 47ps43.
443. Sotiriou, C. and L. Pusztai, *Gene-expression signatures in breast cancer*. *N Engl J Med*, 2009. **360**(8): p. 790-800.
444. Weigelt, B., F.L. Baehner, and J.S. Reis-Filho, *The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade*. *J Pathol*, 2010. **220**(2): p. 263-80.
445. Iwamoto, T. and L. Pusztai, *Predicting prognosis of breast cancer with gene signatures: are we lost in a sea of data?* *Genome Med*, 2010. **2**(11): p. 81.
446. Perou, C.M., et al., *Molecular portraits of human breast tumours*. *Nature*, 2000. **406**(6797): p. 747-52.

447. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
448. Gruvberger, S., et al., *Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns*. Cancer Res, 2001. **61**(16): p. 5979-84.
449. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
450. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer*. Breast Cancer Res, 2010. **12**(5): p. R68.
451. Farmer, P., et al., *Identification of molecular apocrine breast tumours by microarray analysis*. Oncogene, 2005. **24**(29): p. 4660-71.
452. Doane, A.S., et al., *An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen*. Oncogene, 2006. **25**(28): p. 3994-4008.
453. Weigelt, B., et al., *Breast cancer molecular profiling with single sample predictors: a retrospective analysis*. Lancet Oncol, 2010. **11**(4): p. 339-49.
454. Mackay, A., et al., *Microarray-based class discovery for molecular classification of breast cancer: analysis of interobserver agreement*. J Natl Cancer Inst, 2011. **103**(8): p. 662-73.
455. Weigelt, B. and J.S. Reis-Filho, *Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry*. Breast Cancer Res, 2010. **12 Suppl 4**: p. S5.
456. Sotiriou, C., et al., *Breast cancer classification and prognosis based on gene expression profiles from a population-based study*. Proc Natl Acad Sci U S A, 2003. **100**(18): p. 10393-8.
457. Hu, Z., et al., *The molecular portraits of breast tumors are conserved across microarray platforms*. BMC Genomics, 2006. **7**: p. 96.
458. Wang, Y., et al., *Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer*. Lancet, 2005. **365**(9460): p. 671-9.
459. Buyse, M., et al., *Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer*. J Natl Cancer Inst, 2006. **98**(17): p. 1183-92.
460. Desmedt, C., et al. *TRANSBIG multi-centre independent validation of the Rotterdam 76-gene prognostic signature for patients with node negative (N-) breast cancer (BC)*. in *Fifth European Breast Cancer Conference*. 2006. Nice, France.
461. Eifel, P., et al., *National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000*. J Natl Cancer Inst, 2001. **93**(13): p. 979-89.
462. Potti, A., et al., *Genomic signatures to guide the use of chemotherapeutics*. Nat Med, 2006. **12**(11): p. 1294-300.
463. Chang, H.Y., et al., *Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival*. Proc Natl Acad Sci U S A, 2005. **102**(10): p. 3738-43.
464. Chang, H.Y., et al., *Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds*. PLoS Biol, 2004. **2**(2): p. E7.
465. Glinsky, G.V., O. Berezovska, and A.B. Glinskii, *Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer*. J Clin Invest, 2005. **115**(6): p. 1503-21.
466. Glinsky, G.V., *Genomic models of metastatic cancer: functional analysis of death-from-cancer signature genes reveals aneuploid, anoikis-resistant, metastasis-enabling phenotype with altered cell cycle control and activated Polycomb Group (PcG) protein chromatin silencing pathway*. Cell Cycle, 2006. **5**(11): p. 1208-16.
467. Sotiriou, C., et al., *Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis*. J Natl Cancer Inst, 2006. **98**(4): p. 262-72.
468. Roylance, R., et al., *Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer*. Cancer Res, 1999. **59**(7): p. 1433-6.

469. Warnberg, F., et al., *Tumour markers in breast carcinoma correlate with grade rather than with invasiveness*. Br J Cancer, 2001. **85**(6): p. 869-74.
470. Liedtke, C., et al., *Genomic grade index is associated with response to chemotherapy in patients with breast cancer*. J Clin Oncol, 2009. **27**(19): p. 3185-91.
471. Dai, H., et al., *A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients*. Cancer Res, 2005. **65**(10): p. 4059-66.
472. Colombo, P.E., et al., *Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction*. Breast Cancer Res, 2011. **13**(3): p. 212.
473. Wirapati, P., et al., *Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures*. Breast Cancer Res, 2008. **10**(4): p. R65.
474. Fan, C., et al., *Concordance among gene-expression-based predictors for breast cancer*. N Engl J Med, 2006. **355**(6): p. 560-9.
475. Reyal, F., et al., *A comprehensive analysis of prognostic signatures reveals the high predictive capacity of the proliferation, immune response and RNA splicing modules in breast cancer*. Breast Cancer Res, 2008. **10**(6): p. R93.
476. Iwamoto, T., et al., *Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer*. J Natl Cancer Inst, 2011. **103**(3): p. 264-72.
477. Desmedt, C., et al., *Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes*. Clin Cancer Res, 2008. **14**(16): p. 5158-65.
478. Desmedt, C., et al., *Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series*. Clin Cancer Res, 2007. **13**(11): p. 3207-14.
479. Symmans, W.F., et al., *Genomic index of sensitivity to endocrine therapy for breast cancer*. J Clin Oncol, 2010. **28**(27): p. 4111-9.
480. Gerlinger, M. and C. Swanton, *How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine*. Br J Cancer, 2010. **103**(8): p. 1139-43.
481. Kim, C. and S. Paik, *Gene-expression-based prognostic assays for breast cancer*. Nat Rev Clin Oncol, 2010. **7**(6): p. 340-7.
482. Sparano, J.A. and S. Paik, *Development of the 21-gene assay and its application in clinical practice and clinical trials*. J Clin Oncol, 2008. **26**(5): p. 721-8.
483. Paik, S., *Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen*. Oncologist, 2007. **12**(6): p. 631-5.
484. Laronga, C., et al., *The role of the breast cancer surgeon in personalized cancer care: clinical utility of the 21-gene assay*. Am J Surg, 2012. **203**(6): p. 751-8.
485. Paik, S., et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer*. N Engl J Med, 2004. **351**(27): p. 2817-26.
486. Ma, X.J., et al., *A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen*. Cancer Cell, 2004. **5**(6): p. 607-16.
487. Cardoso, F., et al., *Clinical application of the 70-gene profile: the MINDACT trial*. J Clin Oncol, 2008. **26**(5): p. 729-35.
488. Dowsett, M., et al., *Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study*. J Clin Oncol, 2010. **28**(11): p. 1829-34.
489. Albain, K.S., et al., *Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial*. Lancet Oncol, 2010. **11**(1): p. 55-65.
490. NCCM Clinical Practice Guidelines in Oncology™. Breast Cancer online. 2011.
491. Goldhirsch, A., et al., *Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011*. Ann Oncol, 2011. **22**(8): p. 1736-47.

492. Habel, L.A., et al., *A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients*. *Breast Cancer Res*, 2006. **8**(3): p. R25.
493. Harris, L., et al., *American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer*. *J Clin Oncol*, 2007. **25**(33): p. 5287-312.
494. Teschendorff, A.E. and C. Caldas, *A robust classifier of high predictive value to identify good prognosis patients in ER-negative breast cancer*. *Breast Cancer Res*, 2008. **10**(4): p. R73.
495. Bianchini, G., et al., *Molecular anatomy of breast cancer stroma and its prognostic value in estrogen receptor-positive and -negative cancers*. *J Clin Oncol*, 2010. **28**(28): p. 4316-23.
496. Farmer, P., et al., *A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer*. *Nat Med*, 2009. **15**(1): p. 68-74.
497. Finak, G., et al., *Stromal gene expression predicts clinical outcome in breast cancer*. *Nat Med*, 2008. **14**(5): p. 518-27.
498. Bergamaschi, A., et al., *Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome*. *J Pathol*, 2008. **214**(3): p. 357-67.
499. Ayers, M., et al., *Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer*. *J Clin Oncol*, 2004. **22**(12): p. 2284-93.
500. Chang, J.C., et al., *Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer*. *Lancet*, 2003. **362**(9381): p. 362-9.
501. Gianni, L., et al., *Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer*. *J Clin Oncol*, 2005. **23**(29): p. 7265-77.
502. Hess, K.R., et al., *Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer*. *J Clin Oncol*, 2006. **24**(26): p. 4236-44.
503. Iwao-Koizumi, K., et al., *Prediction of docetaxel response in human breast cancer by gene expression profiling*. *J Clin Oncol*, 2005. **23**(3): p. 422-31.
504. Thuerigen, O., et al., *Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer*. *J Clin Oncol*, 2006. **24**(12): p. 1839-45.
505. Kao, K.J., et al., *Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization*. *BMC Cancer*, 2011. **11**: p. 143.
506. Jansen, M.P., et al., *Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling*. *J Clin Oncol*, 2005. **23**(4): p. 732-40.
507. Mackay, A., et al., *Molecular response to aromatase inhibitor treatment in primary breast cancer*. *Breast Cancer Res*, 2007. **9**(3): p. R37.
508. Miller, W.R., et al., *Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole*. *Pharmacogenet Genomics*, 2007. **17**(10): p. 813-26.
509. Cappelletti, V., et al., *Patterns and changes in gene expression following neoadjuvant anti-estrogen treatment in estrogen receptor-positive breast cancer*. *Endocr Relat Cancer*, 2008. **15**(2): p. 439-49.
510. Miller, W.R., et al., *Gene expression profiles differentiating between breast cancers clinically responsive or resistant to letrozole*. *J Clin Oncol*, 2009. **27**(9): p. 1382-7.
511. Muschler, J. and C.H. Streuli, *Cell-matrix interactions in mammary gland development and breast cancer*. *Cold Spring Harb Perspect Biol*, 2010. **2**(10): p. a003202.
512. Lewis, C.E. and J.W. Pollard, *Distinct role of macrophages in different tumor microenvironments*. *Cancer Res*, 2006. **66**(2): p. 605-12.
513. Ma, X.J., et al., *Gene expression profiling of the tumor microenvironment during breast cancer progression*. *Breast Cancer Res*, 2009. **11**(1): p. R7.
514. Mao, Y., et al., *Stromal cells in tumor microenvironment and breast cancer*. *Cancer Metastasis Rev*, 2013. **32**(1-2): p. 303-15.

515. Chen, J.L., et al., *Stromal responses among common carcinomas correlated with clinicopathologic features*. Clin Cancer Res, 2013.
516. Planche, A., et al., *Identification of prognostic molecular features in the reactive stroma of human breast and prostate cancer*. PLoS One, 2011. **6**(5): p. e18640.
517. Maffini, M.V., et al., *The stroma as a crucial target in rat mammary gland carcinogenesis*. J Cell Sci, 2004. **117**(Pt 8): p. 1495-502.
518. Bertos, N., et al., *Breast Cancer Stroma: A Predictor of Clinical Outcome and Tumour Heterogeneity*, Dec 20112, McGill University, Montreal, QC, Canada: San Antonio Breast Cancer Conference.
519. Gao, Q., et al., *Effect of aromatase inhibition on functional gene modules in estrogen receptor-positive breast cancer and their relationship with antiproliferative response*. Clin Cancer Res, 2014. **20**(9): p. 2485-94.
520. Dunbier, A.K., et al., *Molecular profiling of aromatase inhibitor-treated postmenopausal breast tumors identifies immune-related correlates of resistance*. Clin Cancer Res, 2013. **19**(10): p. 2775-86.
521. Mello-Grand, M., et al., *Gene expression profiling and prediction of response to hormonal neoadjuvant treatment with anastrozole in surgically resectable breast cancer*. Breast Cancer Res Treat, 2010. **121**(2): p. 399-411.
522. Desmedt, C. and C. Sotiriou, *Proliferation: the most prominent predictor of clinical outcome in breast cancer*. Cell Cycle, 2006. **5**(19): p. 2198-202.
523. Miller, W.R., et al., *Sequential changes in gene expression profiles in breast cancers during treatment with the aromatase inhibitor, letrozole*. Pharmacogenomics J, 2012. **12**(1): p. 10-21.
524. Beck, A.H., et al., *The fibromatosis signature defines a robust stromal response in breast carcinoma*. Lab Invest, 2008. **88**(6): p. 591-601.
525. Nagalla, S., et al., *Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis*. Genome Biol, 2013. **14**(4): p. R34.
526. Turnbull, A.K., et al., *Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis*. BMC Med Genomics, 2012. **5**: p. 35.
527. Du, P., W.A. Kibbe, and S.M. Lin, *lumi: a pipeline for processing Illumina microarray*. Bioinformatics, 2008. **24**(13): p. 1547-8.
528. Barbosa-Morais, N.L., et al., *A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data*. Nucleic Acids Res, 2010. **38**(3): p. e17.
529. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods*. Biostatistics, 2007. **8**(1): p. 118-27.
530. Pirooznia, M., et al., *SVAw - a web-based application tool for automated surrogate variable analysis of gene expression studies*. Source Code Biol Med, 2013. **8**(1): p. 8.
531. Shabalin, A.A., et al., *Merging two gene-expression studies via cross-platform normalization*. Bioinformatics, 2008. **24**(9): p. 1154-60.
532. Larionov, A., A. Krause, and W. Miller, *A standard curve based method for relative real time PCR data processing*. BMC Bioinformatics, 2005. **6**: p. 62.
533. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data*. Biostatistics, 2003. **4**(2): p. 249-64.
534. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data*. Nucleic Acids Res, 2003. **31**(4): p. e15.
535. Forouhi, P., et al., *Ultrasonography as a method of measuring breast tumour size and monitoring response to primary systemic treatment*. Br J Surg, 1994. **81**(2): p. 223-5.
536. Barnes, M., et al., *Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms*. Nucleic Acids Res, 2005. **33**(18): p. 5914-23.
537. West, R.B., et al., *Determination of stromal signatures in breast carcinoma*. PLoS Biol, 2005. **3**(6): p. e187.
538. Zheng, Q. and X.J. Wang, *GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis*. Nucleic Acids Res, 2008. **36**(Web Server issue): p. W358-63.
539. Bedard, P.L., et al., *Low residual proliferation after short-term letrozole therapy is an early predictive marker of response in high proliferative ER-positive breast cancer*. Endocr Relat Cancer, 2011. **18**(6): p. 721-30.

540. Robin, X., et al., *pROC: an open-source package for R and S+ to analyze and compare ROC curves*. BMC Bioinformatics, 2011. **12**: p. 77.
541. Lauss, M., et al., *Robust assignment of cancer subtypes from expression data using a uni-variate gene expression average as classifier*. BMC Cancer, 2010. **10**: p. 532.
542. Dowsett, M., et al., *Proliferation and apoptosis as markers of benefit in neoadjuvant endocrine therapy of breast cancer*. Clin Cancer Res, 2006. **12**(3 Pt 2): p. 1024s-1030s.
543. Dixon, J.M., et al., *Surgical issues surrounding use of aromatase inhibitors*. J Steroid Biochem Mol Biol, 2005. **95**(1-5): p. 97-103.
544. Ellis MJ., T.Y., Bhatnagar A., et al., *Cell Cycle Response after Neoadjuvant Letrozole Predicts Superior Relapse - Free and Overall Survival. Long-Term Follow-Up of the Letrozole P024 Study*. Presented ASCO, 2006.
545. Smith, I.E., et al., *Neoadjuvant treatment of postmenopausal breast cancer with anastrozole, tamoxifen, or both in combination: the Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) multicenter double-blind randomized trial*. J Clin Oncol, 2005. **23**(22): p. 5108-16.
546. Schmidt, M., et al., *Long-term outcome prediction by clinicopathological risk classification algorithms in node-negative breast cancer--comparison between Adjuvant!, St Gallen, and a novel risk algorithm used in the prospective randomized Node-Negative-Breast Cancer-3 (NNBC-3) trial*. Ann Oncol, 2009. **20**(2): p. 258-64.
547. Early Breast Cancer Trialists' Collaborative, G., *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials*. Lancet, 2005. **365**(9472): p. 1687-717.
548. Ormandy, C.J., et al., *Cyclin D1, EMS1 and 11q13 amplification in breast cancer*. Breast Cancer Res Treat, 2003. **78**(3): p. 323-35.
549. Keyomarsi, K., et al., *Cyclin E and survival in patients with breast cancer*. N Engl J Med, 2002. **347**(20): p. 1566-75.
550. McCready, D.R., et al., *The prognostic significance of lymph node metastases after preoperative chemotherapy for locally advanced breast cancer*. Arch Surg, 1989. **124**(1): p. 21-5.
551. Breast International Group 1-98 Collaborative, G., et al., *A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer*. N Engl J Med, 2005. **353**(26): p. 2747-57.
552. Kennecke, H.F., et al., *Late risk of relapse and mortality among postmenopausal women with estrogen responsive early breast cancer after 5 years of tamoxifen*. Ann Oncol, 2007. **18**(1): p. 45-51.
553. Mauriac, L., et al., *Predictors of early relapse in postmenopausal women with hormone receptor-positive breast cancer in the BIG 1-98 trial*. Ann Oncol, 2007. **18**(5): p. 859-67.
554. Lombart-Cussac, A., A. Lopez-Guerrero, and C. Fuster-Diana, *Clinical Response to four months neoadjuvant letrozole predicts early distant disease free survival in postmenopausal women with stage II-III ER/PgR positive Breast Cancer*, in ASCO 20062006.
555. Dowsett, M., et al., *Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival*. Clin Cancer Res, 2005. **11**(2 Pt 2): p. 951s-8s.
556. Howell, A., et al., *Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer*. Lancet, 2005. **365**(9453): p. 60-2.
557. Baum, M., et al., *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, 2003. **98**(9): p. 1802-10.
558. Baum, M., et al., *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, 2002. **359**(9324): p. 2131-9.
559. MJ., E., T. Y., and B. A., *Cell-Cycle response after neoadjuvant letrozole predicts superior relapse-free and overall survival. Long-term follow-up of the letrozole P024 study*. , in ASCO2006.

560. Loi, S., et al., *PIK3CA genotype and a PIK3CA mutation-related gene signature and response to everolimus and letrozole in estrogen receptor positive breast cancer*. PLoS One, 2013. **8**(1): p. e53292.
561. Baggerly, K.A., K.R. Coombes, and E.S. Neeley, *Run batch effects potentially compromise the usefulness of genomic signatures for ovarian cancer*. J Clin Oncol, 2008. **26**(7): p. 1186-7; author reply 1187-8.
562. Borst, P. and L. Wessels, *Do predictive signatures really predict response to cancer chemotherapy?* Cell Cycle, 2010. **9**(24): p. 4836-40.
563. Collingridge, D., *Expression of concern--validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial*. Lancet Oncol, 2010. **11**(9): p. 813-4.
564. Coombes, K.R., J. Wang, and K.A. Baggerly, *Microarrays: retracing steps*. Nat Med, 2007. **13**(11): p. 1276-7; author reply 1277-8.
565. Liedtke, C., et al., *Clinical evaluation of chemotherapy response predictors developed from breast cancer cell lines*. Breast Cancer Res Treat, 2010. **121**(2): p. 301-9.
566. Dixon, J.M., *Neoadjuvant Therapy: Surgical Perspectives*, ed. W.R. Miller and J.N. Ingle. 2002, New York, USA: Marcell Dekker.
567. Consortium, M., et al., *The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements*. Nat Biotechnol, 2006. **24**(9): p. 1151-61.
568. Haibe-Kains, B., et al., *A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all?* Bioinformatics, 2008. **24**(19): p. 2200-8.
569. Weigelt, B. and J.S. Reis-Filho, *Histological and molecular types of breast cancer: is there a unifying taxonomy?* Nat Rev Clin Oncol, 2009. **6**(12): p. 718-30.
570. Wolff, A.C., et al., *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer*. J Clin Oncol, 2007. **25**(1): p. 118-45.
571. Hammond, M.E., et al., *American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer*. J Clin Oncol, 2010. **28**(16): p. 2784-95.
572. Rouzier, R., et al., *Breast cancer molecular subtypes respond differently to preoperative chemotherapy*. Clin Cancer Res, 2005. **11**(16): p. 5678-85.
573. Parker, J.S., et al., *Supervised risk predictor of breast cancer based on intrinsic subtypes*. J Clin Oncol, 2009. **27**(8): p. 1160-7.
574. Nielsen, T.O., et al., *A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer*. Clin Cancer Res, 2010. **16**(21): p. 5222-32.
575. De Ronde, J.J., et al., *Concordance of clinical and molecular breast cancer subtyping in the context of preoperative chemotherapy response*. Breast Cancer Res Treat, 2010. **119**(1): p. 119-26.
576. Ransohoff, D.F., *Rules of evidence for cancer molecular-marker discovery and validation*. Nat Rev Cancer, 2004. **4**(4): p. 309-14.
577. Ioannidis, J.P., et al., *Repeatability of published microarray gene expression analyses*. Nat Genet, 2009. **41**(2): p. 149-55.
578. Pusztai, L., et al., *Molecular classification of breast cancer: limitations and potential*. Oncologist, 2006. **11**(8): p. 868-77.
579. Loi, S., et al., *Molecular forecasting of breast cancer: time to move forward with clinical testing*. J Clin Oncol, 2006. **24**(4): p. 721-2; author reply 722-3.
580. Kapp, A.V. and R. Tibshirani, *Are clusters found in one dataset present in another dataset?* Biostatistics, 2007. **8**(1): p. 9-31.
581. Haibe-Kains, B., et al., *Robustness of breast cancer molecular subtypes identification*. Ann Oncol, 2010. **21**: p. iv49-iv59.
582. Lusa, L., et al., *Challenges in projecting clustering results across gene expression-profiling datasets*. J Natl Cancer Inst, 2007. **99**(22): p. 1715-23.
583. Haibe-Kains, B., et al., *A three-gene model to robustly identify breast cancer molecular subtypes*. J Natl Cancer Inst, 2012. **104**(4): p. 311-25.
584. Perou, C.M., et al., *Clinical implementation of the intrinsic subtypes of breast cancer*. Lancet Oncol, 2010. **11**(8): p. 718-9; author reply 720-1.

585. Weigelt, B., et al., *Challenges translating breast cancer gene signatures into the clinic*. *Nat Rev Clin Oncol*, 2012. **9**(1): p. 58-64.
586. Banneau, G., et al., *Molecular apocrine differentiation is a common feature of breast cancer in patients with germline PTEN mutations*. *Breast Cancer Res*, 2010. **12**(4): p. R63.
587. Hennessy, B.T., et al., *Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics*. *Cancer Res*, 2009. **69**(10): p. 4116-24.
588. Lim, E., et al., *Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers*. *Nat Med*, 2009. **15**(8): p. 907-13.
589. Prat, A. and C.M. Perou, *Mammary development meets cancer genomics*. *Nat Med*, 2009. **15**(8): p. 842-4.
590. Sims, A.H., et al., *Origins of breast cancer subtypes and therapeutic implications*. *Nat Clin Pract Oncol*, 2007. **4**(9): p. 516-25.
591. Sorlie, T., *Introducing molecular subtyping of breast cancer into the clinic?* *J Clin Oncol*, 2009. **27**(8): p. 1153-4.
592. Stingl, J. and C. Caldas, *Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis*. *Nat Rev Cancer*, 2007. **7**(10): p. 791-9.
593. Peppercorn, J., C.M. Perou, and L.A. Carey, *Molecular subtypes in breast cancer evaluation and management: divide and conquer*. *Cancer Invest*, 2008. **26**(1): p. 1-10.
594. Venet, D., J.E. Dumont, and V. Detours, *Most random gene expression signatures are significantly associated with breast cancer outcome*. *PLoS Comput Biol*, 2011. **7**(10): p. e1002240.
595. Ein-Dor, L., et al., *Outcome signature genes in breast cancer: is there a unique set?* *Bioinformatics*, 2005. **21**(2): p. 171-8.
596. Michiels, S., S. Koscielny, and C. Hill, *Prediction of cancer outcome with microarrays: a multiple random validation strategy*. *Lancet*, 2005. **365**(9458): p. 488-92.
597. Bueno-de-Mesquita, J.M., et al., *Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER)*. *Lancet Oncol*, 2007. **8**(12): p. 1079-87.
598. Fan, C., et al., *Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures*. *BMC Med Genomics*, 2011. **4**: p. 3.
599. Perou, C.M., et al., *Distinctive gene expression patterns in human mammary epithelial cells and breast cancers*. *Proc Natl Acad Sci U S A*, 1999. **96**(16): p. 9212-7.
600. Rody, A., et al., *Identification of a clinically relevant gene signature in triple negative and basal-like breast cancer*. 2010(S5-5).
601. Yau, C., et al., *A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer*. *Breast Cancer Res*, 2010. **12**(5): p. R85.
602. Rakha, E.A., et al., *Breast cancer prognostic classification in the molecular era: the role of histological grade*. *Breast Cancer Res*, 2010. **12**(4): p. 207.
603. Haibe-Kains, B., et al., *Comparison of prognostic gene expression signatures for breast cancer*. *BMC Genomics*, 2008. **9**: p. 394.
604. Lo, S.S., et al., *Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection*. *J Clin Oncol*, 2010. **28**(10): p. 1671-6.
605. Kelly, C.M., et al., *Utility of oncotype DX risk estimates in clinically intermediate risk hormone receptor-positive, HER2-normal, grade II, lymph node-negative breast cancers*. *Cancer*, 2010. **116**(22): p. 5161-7.
606. Cleator, S.J., et al., *The effect of the stromal component of breast tumours on prediction of clinical outcome using gene expression microarray analysis*. *Breast Cancer Res*, 2006. **8**(3): p. R32.
607. Cronin, M., et al., *Analytical validation of the Oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer*. *Clin Chem*, 2007. **53**(6): p. 1084-91.

608. Glas, A.M., et al., *Converting a breast cancer microarray signature into a high-throughput diagnostic test*. BMC Genomics, 2006. **7**: p. 278.
609. Jansen, M.P., et al., *HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study*. J Clin Oncol, 2007. **25**(6): p. 662-8.
610. Ma, X.J., et al., *The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer*. J Clin Oncol, 2006. **24**(28): p. 4611-9.
611. Ma, X.J., et al., *A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer*. Clin Cancer Res, 2008. **14**(9): p. 2601-8.
612. Foekens, J.A., et al., *Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer*. J Clin Oncol, 2006. **24**(11): p. 1665-71.
613. Desmedt, C., et al., *The Gene expression Grade Index: a potential predictor of relapse for endocrine-treated breast cancer patients in the BIG 1-98 trial*. BMC Med Genomics, 2009. **2**: p. 40.
614. Loi, S., et al., *Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade*. J Clin Oncol, 2007. **25**(10): p. 1239-46.
615. Knauer, M., et al., *The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer*. Breast Cancer Res Treat, 2010. **120**(3): p. 655-61.
616. Straver, M.E., et al., *The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer*. Breast Cancer Res Treat, 2010. **119**(3): p. 551-8.
617. Paik, S., et al., *Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer*. J Clin Oncol, 2006. **24**(23): p. 3726-34.
618. Alexe, G., et al., *High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates*. Cancer Res, 2007. **67**(22): p. 10669-76.
619. Teschendorff, A.E., et al., *An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer*. Genome Biol, 2007. **8**(8): p. R157.
620. Prat, A. and e. al., 2011. J Clin Oncol, Concordance among gene-expression-based predictors for ER-positive breast cancer treated with adjuvant Tamoxifen (Abstract). **29**(suppl. (a502)).
621. Mook, S., et al., *The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study*. Breast Cancer Res Treat, 2009. **116**(2): p. 295-302.
622. Mook, S., et al., *Individualization of therapy using Mammaprint: from development to the MINDACT Trial*. Cancer Genomics Proteomics, 2007. **4**(3): p. 147-55.
623. Gonzalez-Angulo, A.M., F. Morales-Vasquez, and G.N. Hortobagyi, *Overview of resistance to systemic therapy in patients with breast cancer*. Adv Exp Med Biol, 2007. **608**: p. 1-22.
624. Jerevall, P.L., et al., *Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the Stockholm trial*. Br J Cancer, 2011. **104**(11): p. 1762-9.
625. Habel, L.A., et al., *HOXB13:IL17BR and molecular grade index and risk of breast cancer death among patients with lymph node-negative invasive disease*. Breast Cancer Res, 2013. **15**(2): p. R24.
626. Sgroi, D.C., et al., *Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population*. Lancet Oncol, 2013. **14**(11): p. 1067-76.
627. Cheang, M.C., et al., *Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype*. Clin Cancer Res, 2008. **14**(5): p. 1368-76.
628. Parker, J.S. and e. al., *Breast cancer molecular subtypes predict response to anthracycline/taxane-based chemotherapy (Abstract)*. Cancer Res., 2009. **69**((Suppl.3), a2019).

629. Tabchy, A., et al., *Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer*. Clin Cancer Res, 2010. **16**(21): p. 5351-61.
630. Cuzick, J., et al., *Prognostic value of a combined ER, PgR, Ki67, HER2 immunohistochemical (IHC4) score and comparison with the GHI recurrence score—results from transATAC*. Cancer Res, 2009. **69**(503S-503S (Abstract)).
631. Tang, G., et al., *Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20*. Breast Cancer Res Treat, 2011. **127**(1): p. 133-42.
632. Tang, G., et al., *Recurrence risk of node- negative and ER-positive early-stage breast cancer patients by combining recurrence score, pathologic, and clinical information: A meta-analysis approach*. Proc Am Soc Clin Oncol, 2010. **28**(suppl: Abstract 509).
633. Mefford, D. and J. Mefford, *Stromal genes add prognostic information to proliferation and histoclinical markers: a basis for the next generation of breast cancer gene signatures*. PLoS One, 2012. **7**(6): p. e37646.

Appendix I: The American Joint Committee on Cancer (AJCC) and the International Union against Cancer UICC Staging Systems.

STAGE	T	N	M	COMMENTS
Stage 0	Tis	N0	M0	In situ cancer – confined to the ducts, lobules or nipple. No spread to nearby tissue. No spread to lymph nodes. No spread to distant sites.
Stage 1	T1	N0	M0	The tumour is 2 cm or less in diameter. No spread to lymph nodes. No spread to distant sites.
Stage IIA	T0	N1	M0	No tumour is found in the breast. Tumour is found in 1–3 axillary lymph nodes. No spread to distant sites.
	T1			The tumour is 2 cm or less in diameter. Tumour is found in 1–3 axillary and/or internal mammary lymph nodes. No spread to distant sites.
	T2	N0		The tumour is larger than 2 cm, but less than 5 cm in diameter. No spread to lymph nodes. No spread to distant sites.
Stage IIB	T2	N1	M0	The tumour is larger than 2 cm, but less than 5 cm in diameter. Tumour is found in 1–3 axillary and/or internal mammary lymph nodes. No spread to distant sites.
	T3	N0		The tumour is larger than 5 cm in diameter. No spread to lymph nodes. No spread to distant sites.
Stage IIIA	T0	N2	M0	No tumour is found in the breast Tumour is found in 4–9 axillary lymph nodes or in internal mammary lymph nodes. No spread to distant sites.
	T1			The tumour is 2 cm or less in diameter. Tumour is found in 4–9 axillary lymph nodes or in internal mammary lymph nodes. No spread to distant sites.
	T2			The tumour is larger than 2 cm, but not more than 5 cm in diameter. Tumour is found in 4–9 axillary lymph nodes or in internal mammary lymph nodes. No spread to distant sites.
	T3	N1 N2		The tumour is more than 5 cm in diameter. Tumour is found in 1–9 axillary nodes or in internal mammary lymph nodes. No spread to distant sites.
Stage IIIB	T4	N0 N1 N2	M0	The tumour has spread to the chest wall or skin. One of the following applies: Tumour has not spread to any lymph nodes. Tumour has spread to 1–9 axillary lymph nodes. Tumour may or may not have spread to internal mammary lymph nodes. No spread to distant sites.
Stage IIIC	Any T	N3	M0	The tumour is any size. One of the following applies: Tumour has spread to 10 or more axillary lymph nodes. Tumour has spread to 1 or more infraclavicular or supraclavicular lymph nodes. Tumour has spread to more than 3 axillary lymph nodes and to internal mammary lymph nodes. No spread to distant sites.
	T4d			Inflammatory breast cancer is classified as stage III (IIIB or IIIC), unless it has spread to distant sites or lymph nodes far from the breast, in which case it is stage IV.
Stage IV	Any T	Any N	M1	The tumour is any size. Any degree of lymph node involvement. Tumour has spread to distant sites, such as the bone, liver, lung, brain or lymph nodes far from the breast.

Appendix II: Patient Consent Forms.

Patient Consent Form Studies on Breast Disease

Please Initial Boxes

1. I have read and understood the attached information sheet.
2. I agree to undergo the procedure of core biopsy.
3. I understand that the purpose of the core biopsy is to determine the cause of my breast lump.
4. I agree to allow extra samples of my breast lump to be taken.
5. I understand how the sample will be collected and that giving this sample is voluntary and that I am free to withdraw my approval for use of the sample at any time without giving a reason and without my medical treatment or legal rights being affected.
6. I understand that the research performed on the extra sample may include research to improve our understanding of how drugs influence breast disease. I understand that if these extra samples are used for research, ethical approval will be obtained for the research project and that I will be asked to sign a separate consent form, giving permission so these extra samples can be used.
7. I agree to my General Practitioner being informed of my participation in this study.

.....
Name of Patient

.....
Date

.....
Signature

.....
Name of Person Taking Consent
(If different from researcher)

.....
Date

.....
Signature

.....
Researcher

.....
Date

.....
Signature

Letrozole Audit Consent Form

Study Number: Centre Number: Patient Identification Number:

Name of Patient:

Name of Clinician:

Please Initial Boxes

1. I confirm that I have read and understood the information sheet.
2. I understand that my participation is voluntary.
3. I agree that the tissue samples obtained during the time I am taking an aromatase inhibitor and at the time of my operation can be analysed to investigate the effects of the drugs.
4. I agree that the tissue samples can be stored for use in future studies.
5. I agree to my General Practitioner being informed of my participation in this study.

.....
Name of Patient Date Signature

.....
Name of Patient Date Signature
(To reaffirm on day of sample)

.....
Name of Person Taking Consent Date Signature
(If different from researcher)

.....
Researcher Date Signature

**Gene Expression in Aromatase Inhibitor Treated Breast Cancer
(Prospective Study)**

Study Number: Centre Number: Patient Identification Number:

Name of Patient:

Name of Clinician:

Please Initial Boxes

1. I confirm that I have read and understand the information sheet for the above study.
2. I understand that my participation is voluntary.
3. I agree that extra tissue samples and blood samples as detailed in the patient information leaflet can be taken and stored for future analysis.
4. I understand that the analyses of the tissue and blood taken will involve using new techniques for examining the genes in the cancer.
5. I agree to my General Practitioner being informed of my participation in this study.

.....
Name of Patient	Date	Signature

.....
Name of Person Taking Consent (if different from researcher)	Date	Signature

.....
Researcher	Date	Signature

**Gene Expression in Aromatase Inhibitor Treated Breast Cancer
(Retrospective Study)**

Study Number: Centre Number: Patient Identification Number:

Name of Patient:

Name of Clinician:

Please Initial Boxes

- | | |
|--|--------------------------|
| 1. I confirm that I have read and understood the information sheet for the above study. | <input type="checkbox"/> |
| 2. I understand that my participation is voluntary. | <input type="checkbox"/> |
| 3. I agree that the tissue samples obtained during the time I was taking an aromatase inhibitor and at the time of my operation can be analysed to investigate the effects of the drugs. | <input type="checkbox"/> |
| 4. I agree to have one further blood sample taken and understand the purpose of this test as explained in the patient information sheet. | <input type="checkbox"/> |
| 5. I agree that the blood sample can be stored for use in future studies. | <input type="checkbox"/> |
| 6. I agree to my General Practitioner being informed of my participation in this study. | <input type="checkbox"/> |

.....
Name of Patient Date Signature

.....
Name of Patient Date Signature
(To reaffirm on day of sample)

.....
Name of Person Taking Consent Date Signature
(If different from researcher)

.....
Researcher Date Signature

Appendix III: Patient Information Sheets

**Edinburgh Breast Unit
Surgical and Associated Services Division**

**Western General Hospital
Crewe Road South
Edinburgh
EH4 2XU**

Tel: 0131 537 1611

Patient Information

Studies on Breast Disease

You are about to have a core biopsy performed to determine the cause of your breast lump. Following injection of local anaesthetic, thin slivers of tissue obtained by a biopsy needle will be taken and sent to the pathology department to diagnose the cause of your breast lump.

The Edinburgh Breast Unit is involved in a number of research projects looking at the effect of drugs on breast lumps. We write to invite you to allow us to take extra slivers of tissue from your lump which might be used in future research studies should you receive any drug treatment.

If a definite diagnosis is not obtained from the specimens we send to the pathology department, then the stored samples will be made available to the pathology department to help diagnose the cause of your breast lump.

**Letrozole Audit
Edinburgh Breast Unit**

Patient Information Sheet

We would like to invite you to take part in a medical audit looking at the drug Letrozole. The audit will be conducted here at the Edinburgh Breast Unit, by the surgeons and medical oncologists and has been approved by a research ethics committee. Before you decide whether you wish to take part it is important for you to understand why the audit is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Background

Approximately two thirds of all breast cancers in women whose periods have stopped need the female hormone oestrogen to grow. Therefore one way of treating breast cancer is to use drugs that will deprive the cancer of oestrogen. Letrozole works by stopping the body from producing oestrogen and deprives the breast cancer of the oestrogen it requires to grow.

What are the Options for Your Treatment?

Your doctor has told you that you have breast cancer and that there are different treatments available. He/she has decided however that the best treatment for you would be hormonal therapy (Letrozole) in the first instance. This will mean that you take drugs either before you have breast surgery or it may be that you continue to take hormonal drugs without having an operation to your breast. The majority of women who have hormone sensitive tumours over 2cm in size can be made smaller using this therapy so if you do need an operation, then the amount of surgery which will follow the treatment is likely to be less extensive than would have been required without this treatment.

What is the Purpose of this Audit?

Although the majority of tumours can be made smaller using hormone therapy, we do not yet have a test which will allow us to tell within a few weeks of starting treatment whether it will be effective. At the present time therefore, we give all patients three months of treatment and during this time we carefully measure the size of the tumour monthly using clinical measurements and ultrasound.

During this study we would like to take a sample of your breast cancer 14 days after you start Letrozole treatment. Only a very small amount of tissue is needed and this can be removed with a special automatic biopsy needle. The needle test is performed after the breast skin is numbed by injecting local anaesthetic.

It is then hoped that you will continue taking hormonal treatment for 3 months. At the end of 3 months you will either have another core biopsy or have an operation.

By comparing the changes in the biopsy 10-14 days after treatment and after 3 months of treatment with the initial sample obtained at diagnosis, it should be possible for us to look at changes in your tumour over the whole of the 3-month treatment course.

Description of the Drug Treatment

If you agree to take part in this audit, you will be given Letrozole 2.5mgs to take once a day for 3 months in the first instance.

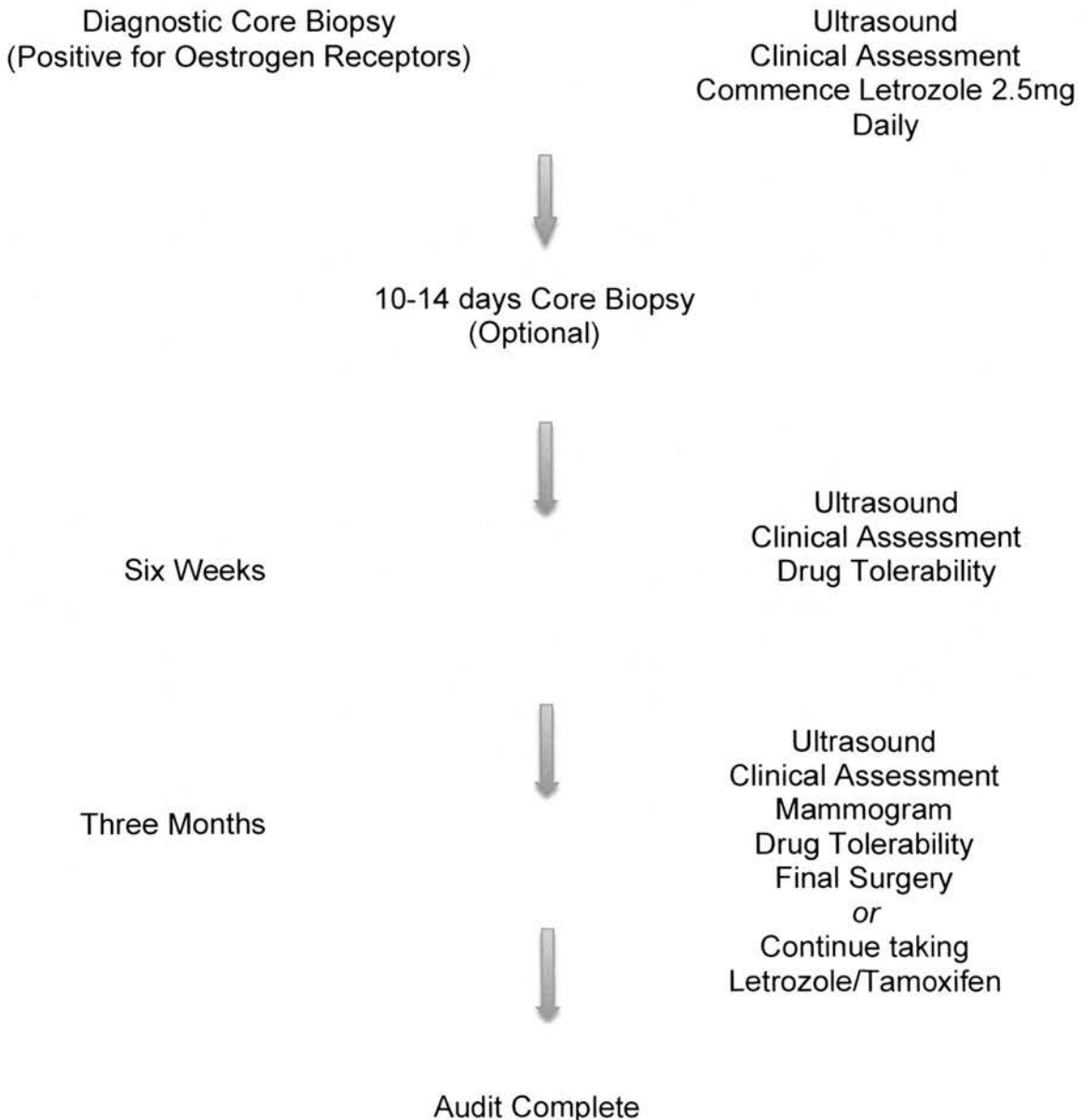
Participation in this Audit

Participation is entirely voluntary and you are free to withdraw at any time should you wish to do so. If you refuse participation or withdraw from the audit, your medical care will not be

affected in any way. Your doctor may discontinue your participation in the audit if he/she judges it to be in your best interest. Your doctor will inform you of any significant new findings related to Letrozole, which may relate to your participation in this audit. When you consent your GP will be told that you are taking part in this audit.

You should not have any additional costs as a result of taking part in this audit because all the extra tests will be performed during standard hospital visits.

Procedures During the Three Months of Treatment



Possible Benefits

This treatment should reduce the size of your tumour, which if you do have an operation should allow you to have less extensive surgery. If you do not proceed to surgery then the tablets will keep your breast cancer under control. This audit will allow us to determine whether your tumour does respond to Letrozole and will help your doctor treat your cancer more effectively.

Potential Side Effects of Letrozole

In trials in which patients were treated with Letrozole for breast cancer, side effects were generally mild to moderate and rarely ever enough to require stopping treatment.

The more common side effects that have been seen in patients treated with Letrozole 2.5mg daily are: Tiredness, hot flushes or increased sweating, changes in weight, dizziness, itching and skin rash, headache, changes in appetite, water retention, nausea and sickness, constipation or diarrhoea, stomach upset or pain.

Less common side effects are: Chest pain, viral infections, pain in the muscles, bones and joints, shortness of breath and coughing.

Do not be alarmed by this list of side effects. You may not have any of them.

You should report any discomfort or problem that you are concerned about to your doctor immediately.

What now?

If you wish to participate in this audit then please read and sign the consent form. Do not sign the form until you are sure you are happy to take part. Please note that you have the right to withdraw from this study at any time without prejudice to your legal rights or healthcare treatment.

Thank you for considering this audit. If you do have any queries, then please do not hesitate to contact us.

Contacts:

Mr J.M. Dixon
Study Investigator
Edinburgh Breast Unit
Western General Hospital
Edinburgh
Tel: 0131 537 1629

Dr Ian Kunkler
Independent Adviser
Clinical Oncologist
Dept. of Clinical Oncology
Western General Hospital
Edinburgh
Tel: 0131 537 2213

Miss Emma Murray / Miss Fiona McCaig
Clinical Research Fellows
Edinburgh Breast Unit
Western General Hospital
Edinburgh
Tel: 0131 537 3638

Miss Lorna Renshaw
Clinical Nurse Specialist
Edinburgh Breast Unit
Western General Hospital
Edinburgh
Tel: 0131 537 1000 bleep 8559

Gene Expression in Breast Cancer Treated by Aromatase Inhibitors (Prospective Study)

Patient Information Sheet

We would like to invite you to take part in a research study during your treatment with Letrozole or Anastrozole. This study will involve you giving us permission to take biopsies and blood samples for research purposes. Before you decide whether you wish to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

In some patients with breast cancer the cancer is too large or unsuitable for operation or there are medical reasons why an operation may not be the best first option of treatment. In these situations, it is useful to try to shrink down the cancer using drugs that block the hormone oestrogen being produced. From our studies it is clear that most people respond well to these drugs; others respond for a time and then the drug stops working. In a few people the drug is not so effective at shrinking the cancer. What we don't understand is why most people benefit but a few do not.

In this study we plan to look both at a sample of cancer, that will be taken from your cancer while you receive hormone blocking drugs and blood samples. The samples will be tested to see whether the changes that occur as a result of taking these drugs relate to how much the cancer shrinks. This will allow us to see if there is any way of predicting who will respond to these drugs, and importantly, tell us who might not respond.

These samples will be analysed using new techniques for examining the genes in the cancer and also the proteins the cancer produces. The human body contains thousands of genes and different proteins inside the cells and when a person develops a disease, there are changes in the genes and the proteins in the part that is diseased. Changes in genes and proteins also occur after treatment, and may allow us to tell whether or not a specific drug is having a beneficial effect.

Although this process involves analysis of genes, this is not the same as genetic testing for looking at the risk of breast cancer to the rest of your family. Instead, it is to look only at your cancer and the samples we already have stored. All of these samples are coded before we analyse them so you won't receive an individual result from these tests.

What is involved for me if I choose to take part?

You will be invited to have extra biopsies and blood tests taken and stored for research purposes. This is likely to involve one or two extra visits to hospital, at two weeks following commencement of your treatment and then again at three months. We would also make some measurement of your tumour at these visits by ultrasound. Other than this all that is required is that you read and discuss this information sheet with your friends or relatives and then make a decision on whether you are happy to give the extra samples for storage and further analysis.

Will my taking part in this study be kept confidential?

Yes. All the samples will be coded, and cannot be traced back to you except by Mr Dixon's research team in the Edinburgh Breast Unit.

What now?

If you wish to allow us to use your samples then please read and sign the consent form. Do not sign the form until you are sure you are happy to take part in the study. Please note that you have the right to withdraw from this study at any time without prejudice to your legal rights or healthcare treatment.

Thank you for considering this study. If you do have any queries then please do not hesitate to contact us.

If you wish to obtain independent advice on this study, you can contact Dr Ian Kunkler on 0131 537 2213.

Gene Expression in Aromatase Inhibitor Treated Breast Cancer (Retrospective Study)

Patient Information Sheet

You will recall during your treatment with Letrozole or Anastrozole for your breast cancer you gave us permission to take biopsies for research purposes. We would now like to ask you to take part in a research study using those biopsy samples and invite you to have one further blood test. Before you decide whether you wish to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

In some patients with breast cancer, the cancer is too large or unsuitable for operation or there are medical reasons why an operation may not be the best first option of treatment. In these situations, it is useful to try to shrink down the cancer using drugs that block the hormone oestrogen being produced. From our studies it is clear that most people respond well to these drugs. Others respond for a time and then the drug stops working. In a few people the drug is not so effective at shrinking the cancer. What we don't understand is why most people benefit but a few do not.

In this study we plan to look at the samples of cancer that have been taken from your cancer while you were taking these hormone-blocking drugs and some cells from a recent blood test. These samples will be tested to see whether the changes that occur as a result of taking these drugs relate to how much the cancer shrinks. This will allow us to see if there is any way of predicting who will respond to these drugs, and importantly, tell us who might not respond.

These samples will be analysed using new techniques for examining the genes in the cancer and blood cells. The human body contains thousands of genes and different proteins inside the cells and when a person develops a disease, there are changes in the genes and the proteins in the part that is diseased. Changes in genes and proteins also occur after treatment, and may allow us to tell whether or not a specific drug is having a beneficial effect.

Although this process involves analysis of genes, this is not the same as genetic testing for looking at the risk of breast cancer to the rest of your family. Instead, it is to look only at changes related to your cancer in the samples we already have stored and in the further blood test. All of these samples are coded before we analyse them so you won't receive an individual result from these tests.

What is involved for me if I choose to take part?

You have already had the biopsies taken and these have been stored but have not yet been tested. You are simply required to have one further blood test. This involves taking 2 teaspoons (10mls) of blood during one extra visit to hospital. All that is required from you is to read this consent form, discuss it with your friends or relatives and make a decision on whether you are happy for us to use the samples we already have stored and to have the extra blood test. If you are happy, then sign the attached consent form and return it to us in the envelope provided.

Will my taking part in this study be kept confidential?

Yes. All the samples are already coded, and cannot be traced back to you except by Mr Dixon's research team in the Edinburgh Breast Unit. It is possible that during the course of the study we will share information with other research teams across the world. Please be reassured that none of the information supplied to these teams will be traceable back to you, except by Mr Dixon's research team.

What now?

If you wish to allow us to use your samples then please read and sign the consent form. Do not sign the form until you are sure you are happy to take part in the study.

Please note that if you should decide to participate in this study we will be informing your GP as a matter of courtesy, unless you express a wish otherwise.

Travel Expenses

Any travel expenses incurred during the course of your participation in this study will be honoured in full.

Thank you for considering this study. If you do have any queries, then please do not hesitate to contact us.

If you wish to obtain independent advice on this study, you can contact Professor Ian Kunkler on 0131 537 2213.

Appendix IV: Signature Gene Lists

BECK

WISP1PDGFRL CDH11 COL3A1 DKK3 SPARC EDNRA COL6A1 COL1A2
COL5A1 COL11A1 CSPG2 FBN1 ADAM12 ADAMTS2AEBP1 ANGPTL2 COL5A2 FN1
LOX MMP11 PLAU THBS2 WNT2 COL8A1 MAFB THY1 RAB31 LUM SPON1
ITGB5 CTSK LOXL1RARRES2DPYSL3 NBL1 FBLN2BGN COL1A1 CTGF ID3INHBA
MFAP2 ROR2 SDC1 SERPINH1 COL12A1 C7orf10 GREM1 OLFML2BPOSTN
SRPX2 SULF1TWIST1 GJA1 ECM2 MN1 COL6A2 MYL9 D2S448 FHOD3
KIAA1295 LOC83468 LRRC15 ODZ4 RNF144

FARMER

DCN CSPG2 CDH11 COL3A1 FAP SERPINF1 FBN1 PDGFRL CTSK PRSS11
ASPN SPARC COL5A2 LOXL1MMP2 SPON1 SFRP4 ITGBL1 CALD1 COPZ2
MFAP2 ANGPTL2PLAU COL1A2 LRRC17 C1QTNF3 SNAI2 PCOLCE POSTN ECM2
FBLN1ADAM12 MMP11 AEBP1 PDGFRB GAS1 COL6A3 RARRES2COL6A1 C1R
NDN TGFB3 LRP1 COL10A1 DPYSL3 OLFML2B MMP14 DACT1 MGC3047 THBS2

NAGALLA

TOP2A BIRC5 KIAA0101 FOXM1 TYMS CCNB2 CDC20 UBE2C MAD2L1 PTTG1
BUB1B DLGAP5 ZWINT TRIP13 AURKA NDC80 CKS2 KIF11 NEK2 TTK
MELK CENPA CCNE2 CENPE GINS1 KIF14 CENPF KIF2C BUB1 CDKN3 RRM2
TPX2 CDK1 MKI67 NCAPH CCNA2 CCNB1 PRC1 NUSAP1 KIF4A CEP55 DTL
NCAPG HJURP KIF20A MLF1IP PBK KIF15 ASPM MCM10 CDCA3 CDCA8
KIF18B RACGAP1

DAI

STK15STK6 UBCH1D ID-GAP BUB1 PTTG2 FLJ11252 CCNB2 KNSL6
DFKZp762E1 PRC1 CENPA CDC45L STK12MAD2L1 ORC6L KIAAD165PKMYT1
E2F1 KPNA2 E2-EPF BIRC5TK1 CDC6 DJ742C19.2 CCNE2 LOC512D3 KNSL5
BLM BMD39 H2AFZ LOC51659 SNRPA1 PSMD7 YWHAZ NTN4 FST BTD

MILLER

ADSL ITM2B C3orf63 TERF1 GGH MLF1IP PPM1A SSFA2 STK19RPS10
ARPP-19 FAM128A TLE1 PIGA ARGLU1 GRB10 MOCOS RPS9 GORASP2 RPL24
RFC3 KIAA0101 BNIP1 HIVEP2 RBL2 COQ10B MAPRE2 FAM60A RPL12EIF4B TNIP1
RPL13A NDEL1 ILVBL LOC730052 /// RNPEP /// TMEM189 /// TMEM189-UBE2V1 ///
UBE2V1 DAZAP2 FLJ22222 RPS20 CHRDL1 ZNF223 BAZ2B SLC9A6 CLCC1
PER2 PAQR3 MOBKL3 MYO5A GIT1 UCRC FAM120A CDC2 TNFRSF10B DDX10
RPS29 IMPDH2 ATP5I /// MFSD7 WDR45 ZNF434 MMD INSIG1 PEX16PTGS1
ITSN2 SOS2 hCG_16001 /// hCG_2001000 /// RPL23A KIAA0265 MCM9 DLEU2 HDAC3
ARMC8 MTF2 RPL27RPS21 TMEM97 DHFR RPL37KLHL2 NUSAP1 RPS13 SMAD2
ARL17P1 DIP2C ATP6V1D ABCA1 HERC1 SETD2 RPS17 MLX DTNA EIF5B
FAM32A SSBP3 CCDC93 CMAH BLVRA SNRPF DMXL1 RPLP0 RAB33B
RUFY1 SCO2 RPS12 NBL1 FHL2 DDX28 HOXC6 RPS2 SULT1A3 /// SULT1A4
DEDD KIAA0226 LOC729362 /// RPL36A FDPS PLSCR3 STX16MAD2L1 PLSCR1 WDR1
EXPH5 ARFGAP1RNF6 ATF4 MAGED2 KIAA0564 FMR1 SAP18SIRT1 ACSL3 RPS4X
RRM2 MAGED1 SPAG7 AMPD3 GNPTAB DRD2 LOC402057 /// RPS17 JUND PARP2
LOC646912 USP22 TMEM80 PSMD14 KIF5B TK2 NCK1 CRTAP ATP5I DIO2
TNFSF10 SEC31A LIMA1 CBX1 ATP8A1 RPLP0 /// RPLP0-like RPLP1 GNA11
ANKHD1 CTBS PRPF4B CASP8 CDKN2AIP RPS6 C22orf28 LAMP2 C2orf33
CD81 FAM13A1 R3HCC1 PCGF3