Late Recurrence in ER+ Breast Cancer

Refining risk prediction beyond 5 years

Dr Juliet Richman

Ralph Lauren Centre for Breast Cancer Research

Royal Marsden Hospital

London

Supervisor: Professor Mitch Dowsett

Table of Contents

Acknowledgements 4

Dedication 5

Statement of conjoint work 6

Glossary of abbreviations 7

Abstract 8

1. Introduction 8

1.1 Significance of the estrogen receptor in breast cancer 9

1.2 Epidemiology of ER+ breast cancer 10

1.3 Management of early ER+ breast cancer 11

1.4 Recurrence in ER+ breast cancer 12

1.5 Extended adjuvant endocrine therapy 13

1.6 Biology of late recurrence 15

1.7 Predicting risk of late recurrence 18

1.8 Practical implications of late recurrence 32

1.9 Some unanswered questions: 34

1.10 Hypothesis and Aims 37

2. Testing validity of CTS5 38

2.1 Introduction 39

2.2 Aims 41

2.3 Methods 42

2.4 Results 46

2.5 Discussion 60

3. Clinical Utility of CTS5 68

3.1 Introduction 68

3.2 Aim 69

3.3 Methods 70

3.4 Results 73

3.5 Discussion 78

4. Validity of CTS5 in a high stage population 84

4.1 Introduction 84

4.2 Aims 88

4.3 Methods 89

4.4 Results 92

4.5 Discussion ... 106

5. Enhancement of CTS5 with molecular biomarkers ... 115

5.1 Introduction ... 115

5.2 Aims ... 117

5.3 Methods ... 118

5.4 Results ... 129

5.5 Discussion ... 142

6. Very late recurrence tissue study ... 147

6.1 Introduction ... 147

6.2 Aim ... 149

6.3 Methods ... 149

6.4 Results ... 165

6.5 Discussion ... 183

7. Overall conclusions and fulfilment of aims ... 191

7.1 Conclusions from CTS5 validation in RMH ... 191

7.2 Conclusions from clinical impact study ... 192

7.3 Conclusions from CTS5 validation in high stage women ... 193

7.4 Conclusions from IHC biomarker study ... 194

7.5 Conclusions from very late recurrence tissue study ... 195

7.6 Overall aims ... 196

7.7 Overarching Hypothesis ... 198

7.8 Concluding remarks ... 198

8. References ... 199

9. Appendix ... 211

9.1 IHC staining protocol/SOP for ER, PgR, Ki67, CK7, CK19 and p63 ... 211

9.2 FFPE Allprep Extraction Protocol ... 215

9.3 Nanostring 360 panel – individual genes ... 218

9.4 Illumina DNA sequencing panel of genes ... 228

9.5 Full copy number alteration results ... 229

9.6 Mutation results by gene ... 231

9.7 Gene expression results ... 232

9.8 Gene Set Enrichment Results ... 234

Acknowledgements

This work would not have been possible without the contributions of so many wonderful people. I would like to thank all the staff of the Ralph Lauren Centre for Breast Cancer Research at the Royal Marsden Hospital. I was privileged to work within such a pool of talent and I benefitted hugely from the kindness and patience of the many skilled staff of this department. Particular thanks go to Richard Buus, Ben Haynes, Elena Lopez-Knowles and Gene Schuster who never tired of my questions and were always willing to assist and guide me. I am also very grateful to Margaret Hills and to Simone Detre for their expertise in IHC.

Thanks go to the Royal Marsden Breast Unit – specifically Professor Stephen Johnston, Dr Alistair Ring, Dr Alicia Okines and Dr Mark Allen who provided so much of the clinical context of my work and regular clinical contact which was so important to me during my time out of training.

I am indebted to Ivana Sestak of the Centre for Preventative Medicine at Queen Mary University and to Lucy Kilburn from the Clinical Trials and Statistics Unit at the ICR for their respective input to the statistics that went into this research. Their tireless work and teaching was a fundamental part of this thesis.

A special mention must go to my supervisor Professor Mitch Dowsett for his endless support, encouragement and assistance over the past 4 years. His guidance and our regular catch ups were both enjoyable and invaluable to this process. As his final MD(Res) student I was well aware of the calibre of my predecessors and am so proud to have produced a thesis to sit on his bookcase alongside those of many revered oncology minds.

Thanks go to my parents and husband Aryeh who have provided unconditional love, encouragement and cups of tea. Their support over the past 4 years (and before) has been instrumental in this accomplishment and would not have been achievable without them.

Lastly, this research was made possible by a generous grant from the Cridlan Ross Smith charitable trust. I am so grateful to the trustees for enabling me to carry out this work.

Dedication

I would like to dedicate this thesis to my three boys: Yoni, Ilan and Avidan. Yoni you are my eloquent, strong-willed daredevil and you are so much braver than me. Ilan you are my imaginative, witty side-kick with the biggest heart I know and Avidan you are my cuddly little monkey with a smile to light up a room. All three of you never cease to amaze me and I am so proud to be your Mum.

Being only young children, I cannot expect you to read this thesis just yet, but it is my hope that one day you can learn something from it. This research has shown me the importance of questions, how good research is always based on a relevant question and the best research answers those questions and generates new ones. Yoni, Ilan and Avidan, I am always blown away by the fantastic questions you ask and how you have developed a thirst for knowledge at such a young age. You have taught me that we should never assume we know everything and never be afraid to ask again. You have inspired me to always ask questions and delve a bit deeper and I hope it makes me a better doctor and oncologist. It is my wish for you that you never tire of asking questions and that you always love learning new things.

The Institute of Cancer Research

STATEMENT OF CONJOINT WORK

Candidate: **Dr Juliet Richman**

Title: **Late recurrence in estrogen receptor positive early breast cancer – refining risk prediction beyond 5 years**

Statistical analysis for the test of validation of the CTS5 in the RMH cohort was overseen by Ivana Sestak from the Centre of Preventative Medicine, Queen Mary University, London.

The test of validation of the CTS5 model in the TACT clinical trial population and the effect of additional molecular biomarkers (chapters 4 and 5) was conducted in collaboration with colleagues from the Clinical Trials and Statistics Unit (CTSU) within the ICR where the patient data is held. In particular, Lucy Kilburn (principal statistician) conducted all statistical analyses for these two studies.

The exploratory tissue study into very late recurrence (chapter 6) involved the use of complex bio-informatic processes and analyses. These were conducted under the supervision of Dr Gene Schuster and Dr Richard Buus. The DNA sequencing was undertaken by Ioannis Assiotis and Nik Matthews in the Tumour Profiling Unit in the ICR.

Glossary of frequently used abbreviations

ER – Estrogen receptor

ER+/- - Estrogen receptor positive/negative

PgR – Progesterone receptor

PgR+/- - Progesterone receptor positive/negative

HER2 – Herceptin receptor 2

HER2+/- Herceptin receptor positive/negative

AI – Aromatase inhibitor

CTS – Clinical treatment score

CTS5 – Clinical treatment score at 5 years

IHC – Immunohistochemistry

IHC4 – Immunohistochemistry 4 model

DR – Distant recurrence

HR – Hazard ratio

CI – Confidence interval

DFS – Disease free survival

OS – Overall survival

CNA – Copy number alteration

RMH – Royal Marsden Hospital

EPR – Electronic patient record

AUC – Area under the curve

ROC – Receiver operator characteristic

MDT – Multidisciplinary team

CCR – Committee for clinical research

TMA – Tissue microarray

FFPE – Formalin fixed paraffin embedded

Abstract

ER positive (ER+) breast cancer has the potential to recur many years following diagnosis. Predicting who is most at risk of recurrence and when this may occur is complex. Risk prediction incorporates clinical, pathological and molecular modelling, has the potential to alter management and markedly affects the survivorship of women treated for early breast cancer. The research detailed in this thesis was aimed at testing and developing risk prediction methods for women with early breast cancer focussing purely on their prognostic potential for late recurrence.

Using retrospective validation studies and prospective utility studies, an existing clinical model was tested in cohorts of women with high stage disease and also women outside of clinical trials. This model was further tested for its utility in a real time MDT meeting setting. Immunohistochemical biomarkers were explored as potential 'add-ons' to clinical models in order to improve precision. Lastly a molecular study involving DNA and RNA sequencing, aimed at discovering biomarkers for very late recurrence, was conducted.

As a result of this work, the prognostic value of existing risk prediction models has been confirmed in a real-life cohort and its applicability broadened to include high stage post-menopausal women. The prognostic value of molecular biomarkers for late recurrence in different patient populations has been explored and questions regarding extrapolation of models in different populations have been addressed. Finally, this work identified clinical and molecular candidate biomarkers for very late recurrence including DNA mutations, high RNA expression of proliferation and cell cycle genes and baseline tumour indices such as size and lymph node burden. These would be suitable for ongoing research and development.

1. Introduction

Breast cancer is divided into sub-types based on the presence of proteins within the tumour cell. These receptors are the estrogen, progesterone receptors and HER2, which have an impact on prognosis and determine the course of clinical management. The most common subtype is estrogen receptor (ER) positive (ER+) breast cancer, which accounts for more than 80% of early breast cancer diagnoses.

1.1 Significance of the estrogen receptor in breast cancer

Knowledge of the connection between circulating estrogen and breast cancer dates back to the late 19th century when oophorectomy for breast cancer was first performed based on the theory that "secretions of the ovary favoured the growth of the cancer"[1]. This secretion was later identified as being estrogen[2] and subsequently its association with epithelial proliferation within breast tissue was described [3]. In the first half of the 20th century it was unknown that not all breast cancer was driven by estrogen and therefore a large proportion of surgical oophorectomy was performed with no chance of response.

The ER was first characterised in the 1960s where it was shown that binding of estrogen to the receptor in target tissues was a potentially reversible process that brought about a series of downstream proliferation events [4,5]. In the 1970s it was recognised that breast tumours which were rich in ER responded better to surgical endocrine ablation which led to development of the first ER assays,. These were initially ligand-binding assays, followed by enzyme immunoassays and then later immunohistochemical assays. Research efforts then switched to the search for a pharmacological anti-estrogen of which Tamoxifen (first developed as contraceptive agent) was the first that showed successful results in breast cancer regression by blocking the ER receptor[6]. Initially, treatment was administered regardless of ER status. Later, further endocrine therapies were developed, in particular aromatase inhibitors (AI) which function by preventing the conversion of peripherally produced androgens into estrogen. These two classes of drugs, and other endocrine therapies are known to be highly effective in both early and metastatic breast cancer and are used throughout the world.

In recent years, mechanisms surrounding the function of the ER have been extensively investigated. It is now known that ligand binding to the ER brings about a cascade of events, both directly within the cell nucleus and through multiple pathways including the ras/raf/MAPK pathway and the PI3K/AKT pathway [7]. This brings about transcription of genes which promote progression through the cell cycle and proliferation and by blocking these pathways with endocrine agents, cancer progression is reduced. The most significant discovery in the past decade has been that of the intrinsic subtypes of breast cancer. This is a molecular classification that categorises breast cancer into one of four subtypes by levels of expression of a panel of genes [8,9]. The reference subtype which the cancer most closely resembles in terms of its gene expression is the subtype to which it will be assigned. Two of these subtypes: luminal A and luminal B, correspond to the ER+ cancers and a key distinction between the two is that luminal B cancers have higher expression of proliferation genes in comparison to luminal A.

1.2 Epidemiology of ER+ breast cancer

In the UK there are approximately 55,200 new cases of invasive breast cancer in women every year[10]. Approximately 80% of these will be ER+[11,12]. Incidence is highest in women over the age of 50 with the highest proportion of cases occurring in the 60-69 year old age group [13]. Since the advent of routine breast screening in the UK in the late eighties, incidence has climbed year on year and is predicted to continue rising to >71,000 new cases in the year 2035 [14]. Risk of breast cancer is linked to an overall increased lifetime exposure to estrogen such as early onset menarche and late menopause and is thought to be caused by modifiable risk factors such as obesity and alcohol use in approximately 23% of cases[15]. The majority of breast cancers are diagnosed at an early stage when treatment is potentially curative [16]. This has been facilitated by the advent of screening which is now responsible for 31% of breast cancer diagnoses [17].

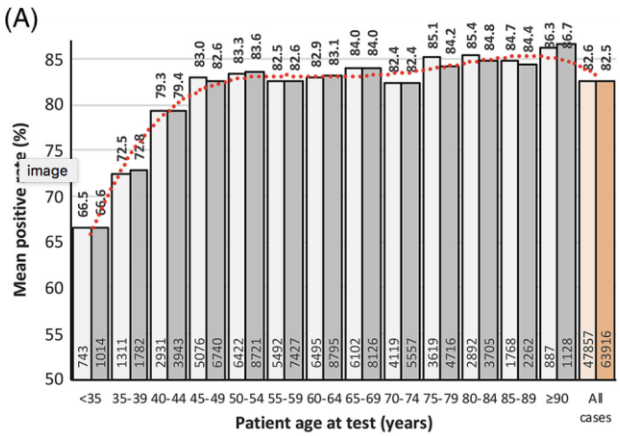


Figure 1.1 Credit Dodson et al 2018[12]
Distribution of ER-positive rates by patient age at test. The final categorical status has been used to define ER positivity. Data are shown for UK and for all patients. A polynomial trend-line (red dotted-line) is displayed for the all patient data set. Figures above the columns are proportions of tumours, those inside the base of the columns are counts of ER-positive tumours

1.3 Management of early ER+ breast cancer

In all subtypes of breast cancer, the mainstay of treatment of disease is surgery and radiotherapy followed by adjuvant medical therapy. Specifically for ER+ breast cancer, local treatment is followed by systemic endocrine therapy for at least 5 years (+/- chemotherapy) which aims to eliminate residual microscopic disease.

For pre-menopausal women, the drug used is tamoxifen which blocks the ER and prevents ligand binding. Tamoxifen has been shown to reduce recurrence rates by about half over 5 years of treatment in comparison to no treatment and reduces mortality by a third over 15 years compared with no treatment [18]. A large, collaborative meta-analysis established the benefit of tamoxifen and also the imperative role of ER testing as no benefit was seen in tumours classified as ER poor [19,20]. This meta-analysis established that 5 years of tamoxifen is better than 1 or 2 and that its effect was regardless of chemotherapy treatment and held true in both node positive and node negative women.

For post-menopausal women, the drug of choice is usually an AI as the main source of endogenous estrogen in these women is peripheral fat and not the ovaries. AIs reduce risk of recurrence by 30% compared to tamoxifen [21,22] and also confer a survival advantage with a 20% relative reduction in mortality compared with tamoxifen and an inferred 50%

relative reduction when compared with no endocrine therapy[21]. These systemic treatments coupled with early detection as a result of screening mean that approximately two thirds of women diagnosed with any type of breast cancer (not just ER+) will survive for 20 years or more[23]. In an epidemiological study of >430,000 women diagnosed with breast cancer between 1973–2000, cause of death was analysed. Amongst those with ER+ cancers, death was breast cancer-related in only 18% of women with tumours <2cm and 31% in those with tumours >2cm[24]. Therefore, the majority of women eventually died of causes unrelated to breast cancer.

1.4 Recurrence in ER+ breast cancer

A proportion of women treated for early breast cancer will go on to experience recurrence. Recurrences can either be local, loco-regional or distant. It is the latter that is the focus of much research as these are incurable and account for the vast majority of the 11,400 deaths per year in the UK from breast cancer[23]. Distant recurrences come about when, prior to surgery, cells separate from the tumour and travel through the blood or lymphatics to distant sites. If these cells are not ablated by systemic treatments they can eventually propagate and establish macrometastases.

Risk of both early and late recurrence is associated with features of the cancer such as large size and nodal involvement (Figure 1.2) which will be explored more fully later in this chapter. Owing to the ongoing risk of recurrence of ER breast cancer over many years, the cumulative risk of recurrence is appreciable even for those with low stage cancers. Distant recurrence in ER+ breast cancer occurs in approximately 20% of women over 10 years with only half of these occurring in years 0-5[25]. Unlike many other cancers, including ER-negative (ER–) breast cancer, the recurrence risk of ER+ breast cancer remains stable beyond 5 years after diagnosis[25,26]. This means that 4,000–5,000 distant recurrences beyond 5 years can be expected amongst the ~46,000 new cases of ER+ breast cancer diagnosed in the UK in 2015[12,14] (Figure 1.1). Owing to the projected increase in incidence of invasive breast cancer over the next decade[10,27], late recurrence is, and will remain, a considerable health problem. In this thesis, I define late recurrence as recurrence beyond 5 years from diagnosis.

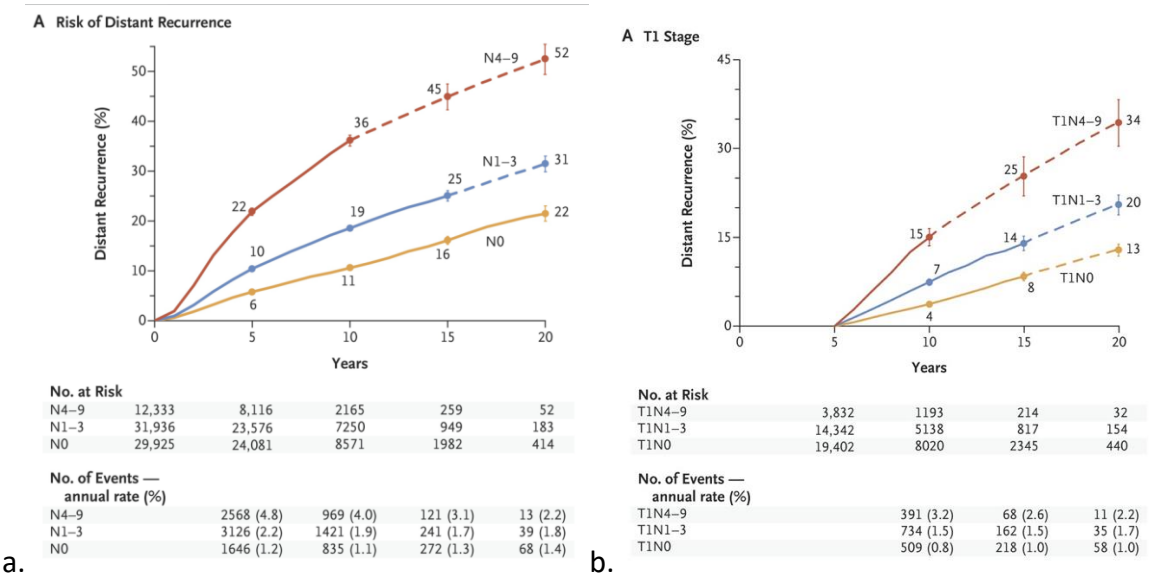


Figure 1.2 Credit Pan et al 2017 [25]
Risk of distant recurrence a. from year 0 stratified by nodal status and b. for T1 stage cancers from year 5 stratified by nodal status

1.5 Extended adjuvant endocrine therapy

Endocrine therapy is recommended at diagnosis and for at least 5 years in virtually all women with ER+ primary breast cancer owing to the survival benefits derived from this therapeutic modality as discussed above. The duration of endocrine therapy has been the subject of several clinical trials as it aims to address the question of whether continued endocrine therapy could reduce late recurrences.

1.5.1 Tamoxifen beyond 5 years

Since the early 2000s, the recommendation has been to cease tamoxifen treatment at 5 years. In the NSABP-B14 study[28], patients with early-stage ER+ breast cancer who, following 5 years of tamoxifen were re-randomised to receive 5 years of placebo, had a significant *improvement* in disease-free survival (DFS) over those who received tamoxifen (82% versus 72%). A trend (not significant) towards improved overall survival (OS) was also observed in the same group (94% versus 91%)[28].

Nevertheless, data from two studies with much larger cohorts[29,30] subsequently showed a benefit of tamoxifen beyond 5 years. In the ATLAS study[29], an absolute reduction in disease recurrence (21.4% versus 25.1%) and reduced breast cancer-related mortality (12.2% versus 15%) at 10 years follow up were observed in patients receiving an additional 5 years of tamoxifen compared with those who stopped treatment at 5 years. The difference between the extended tamoxifen and placebo arms became more apparent after 10 years of follow up, highlighting the importance of long-term follow-up in studies involving use of adjuvant therapy. The incidence of pulmonary embolism and endometrial cancer was significantly higher among patients in the extended therapy group than in the discontinuation group. These findings were consistent with those from the UK-based aTTom study[30], in which a significant recurrence benefit was seen for the extended therapy arm as well as a trend, although not significant, for improved OS. As with the ATLAS study, incidence of endometrial cancer was significantly greater in the extended tamoxifen arm.

1.5.2 Aromatase inhibitors beyond 5 years

In the pivotal MA.17 study[31], women who had received 5 years of adjuvant tamoxifen were randomised to receive either the AI letrozole or placebo for up to 5 years. Letrozole treatment was unequivocally associated with an absolute reduction in the risk of recurrence of 4.6% compared with placebo. Women receiving letrozole experienced significantly more hot flushes and arthralgia/myalgia than those receiving placebo. A significantly greater incidence of new-onset osteoporosis was observed in the letrozole arm compared with placebo (8.1% versus 6%) but this difference did not translate to an increased rate of bone fractures. The aromatase inhibitor exemestane was tested in a study with a similar design, NSABP-B33[32], in which therapy extension with an aromatase inhibitor was associated with improvements in 4-year relapse-free survival (96% versus 94%; $P = 0.004$) compared with placebo. The incidence of toxicities was similar to that observed in MA.17[11].

The benefits of extending a 5-year course of upfront, adjuvant AI appear to be less substantial. In the MA.17R study[33], 1,918 women without disease recurrence after 5 years of therapy with letrozole were randomised to receive either letrozole or placebo for a

further 5 years. An absolute reduction in disease recurrence of 3.3% was seen in the letrozole group but this difference was weighted heavily in favour of a reduction in the incidence of contralateral secondary cancers. Patients in the letrozole group had a higher incidence of osteoporosis and skeletal-related adverse events; the incidence of other toxicities was unremarkable. In the DATA study[34], a borderline significant improvement in 5-year DFS (83.1% versus 79.4%) but no improvement in OS (90.8% versus 90.4%) were observed, with 6 years versus 3 years of anastrozole following 2–3 years tamoxifen respectively. Nevertheless, in post-hoc analyses, a subgroup of patients with both ER and progesterone receptor-positive disease with lymph node involvement in the 6-year treatment group had improved outcomes. In the IDEAL study[35], 5 years of treatment with letrozole was not superior to 2.5 years of treatment, after an initial 5 years of endocrine therapy; no prognostic subgroups or biomarkers were identified in this study. Furthermore, the analysis of DFS events in IDEAL showed that the main differences in outcomes between treatment arms were related to secondary breast cancers and not to distant disease.

Taken together, these results suggest that women who have received 5 years of tamoxifen and then switch to an aromatase inhibitor are more likely to derive benefit from extended endocrine therapy. In comparison, for those who have received 5 years of therapy with an aromatase inhibitor the benefits appear modest. Crucially, little in the way of biomarkers for responders have emerged from these studies and so appropriate means of patient selection for extended therapy is still required. Endocrine therapy can be gruelling treatment for patients to endure and toxicity must be considered in decisions regarding treatment extension.

1.6 Biology of late recurrence

1.6.1 Cellular quiescence and tumour dormancy

The potentially long period that precedes distant recurrence in ER+ breast cancer is often referred to as a period of latency or tumour dormancy. The term dormancy sometimes refers to reversible cellular quiescence and, in other cases, to tumour mass dormancy[36-38].

Cellular quiescence is the maintenance of cell cycle arrest in individual cells which have metastatic potential. This is characterised by a major reduction or complete absence of proliferation. Tumour mass dormancy is conceptualised as an equilibrium between a proliferating cellular population and another population with a higher cellular death rate. Both types of dormancy are likely to co-exist with each one doing so as part of the same biological continuum, with both requiring certain triggers to bring about re-emergence of proliferation. The phenomenon of long-term dormancy is not seen in ER negative (ER-) breast cancer and one particular xenograft study has highlighted this distinction. In that study, ER+ cancer cells injected into mice remained dormant for several weeks, with very low Ki67, until stimulated with exogenous estrogen to trigger macrometastases to develop[39]. Conversely, mice injected with ER- cells developed macrometastatic deposits much more quickly. These findings reflect the unique behaviour of ER+ breast cancer cells and suggest a mechanism by which dormant micrometastatic disease could 'awaken' from dormancy in response to an increase in circulating estrogen brought about by endocrine therapy cessation.

1.6.2 Mechanisms behind dormancy and its reversal

The mechanisms that underlie dormancy and its reversal involve complex interactions between tumour cells and their microenvironment[40-44]. Other key components are cancer stem cells[45], the tumour vasculature[46] and the host immune system[47-49]. Regulation of angiogenesis has also been explored as a determinant of tumour dormancy either by inducing cellular quiescence or by other means[46,50], and the role of tumorigenic cancer stem cells has also been implicated [51].

What triggers the exit from dormancy, however, is still unknown and there is limited preclinical and translational studies that attempt to answer this question. A combination of the ongoing acquisition of mutations, along with other environmental factors (such as physiological stress and aging) is likely to tip the balance towards exit from dormancy over time. There are suggestions that non-cancer related inflammation or chronic blood loss could play a role[52] as well as conditions associated with physiological stress but robust experimental data is not available.

The study of the mechanisms underlying both types of tumour dormancy is hampered by difficulties in developing cell lines and mouse xenograft models that accurately reflect the long latent period that can precede late recurring ER+ breast cancer and also, by the length of time needed to reach the end point in these studies. Another feature that has not been fully reproduced yet in experimental breast cancer models — and is particularly relevant to ER+ breast cancer — is the effect of estrogen deprivation imposed by adjuvant endocrine treatment. The differential response to adjuvant endocrine therapy will be a key determinant in the behaviour of a micrometastasis which, in some patients, will be fully ablated and in others will respond but then relapse with a variable duration of latency in between.

1.6.3 Natural history of treated early breast cancer

As a result of the dormancy phenomenon, the potential outcomes for a patient treated with 5 years endocrine therapy and the timing thereof are numerous (Figure 1.3). Early recurrences are usually as a result of innate resistance to endocrine therapy or early acquisition of resistance. Late recurrences could arise due to innate resistance in very slowly proliferating cells but could also be due to acquired resistance in an initially sensitive cancer or due to resurgence of circulating estrogen following cessation of endocrine therapy that re-awakens previously dormant cells. This myriad of possibilities will be determined by biological properties of the cancer itself as well as the host and the complex interplay of these factors with the effects of treatment. It is therefore quite possible that early and late recurrence can differ in their biology and may be subject to different prognostic influences.

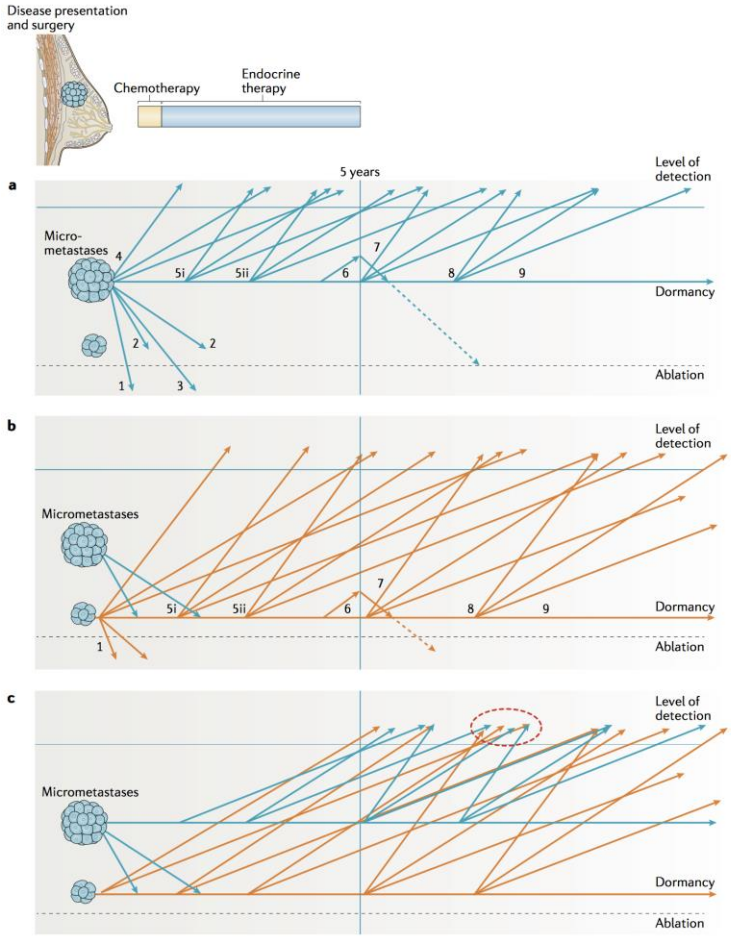


Figure 1.3 Credit Richman et al 2018 [53]
In this diagram, the slopes of the blue lines represent the growth rate of the metastases. **a** | Behaviour of a large-sized micrometastasis. Different outcomes can occur: cure owing to complete response to chemotherapy (1) or endocrine therapy (3), partial response to chemotherapy and endocrine therapy (2), innate resistance to endocrine therapy (4), stable dormant disease followed by early-stage (5i) or late-stage (5ii) acquired therapeutic resistance, beginning of disease recurrence with subsequent disease regression as a withdrawal response to adjuvant tamoxifen treatment (6), which would lead to disease to become dormant or be fully ablated, emergence of distant disease upon cessation of endocrine therapy (7) or at a later stage of follow-up monitoring (8), or establishment of stable, long-term dormancy (9). **b** | Behaviour of a small-sized micrometastatic lesion, including large lesions that have been reduced in size by adjuvant therapy. The outcomes in this situation have been described in panel a, but the emergence of recurrent disease is delayed in comparison with that scenario because of the smaller size of the metastatic lesions at the time of surgery. **c**| Multiple factors leading to the outgrowth of metastatic lesions and overt late-stage recurrence. In the example in the red dotted circle, six recurrent lesions emerge at a similar time (~7 years after surgery). The dotted blue and orange lines denote a patient in which a withdrawal response might lead to tumour ablation.

1.7 Predicting risk of late recurrence

Risk prediction of recurrence in cancer is a fundamental component of management. In terms of late recurrence, accurate risk prediction enables clinicians and patients to make informed choices about surveillance and secondary prevention in the context of treated early disease. I have discussed above that biomarkers for response to extended endocrine

therapy are lacking and therefore prognostic biomarkers and models are required to elucidate this further.

1.7.1 Individual clinicopathological predictors

Clinical and pathological predictors of late recurrence, often termed 'classical variables', have been extensively investigated. Classical variables include tumour size, tumour grade, patient age, and degree of lymph node involvement or presence of lymphovascular invasion. The burden of micrometastatic disease and its proliferation rate are likely to affect the risk and timing of disease recurrence; thus, clinico-pathological variables that inform on disease burden and proliferation should be good predictors of the risk of both early and late recurrence.

Lymph node burden is consistently shown to be associated with the risk of late recurrence[25,26,54,55]. Tumour size has also been shown to be independently associated with late recurrence and breast cancer specific survival in some studies[55-57], including a meta-analysis involving >60,000 women who were alive and free from distant recurrence at 5 years after diagnosis[25]. One of the most notable findings of the latter study was that the cumulative risk of recurrence among women with the lowest stage tumours (T1N0) was as high as 13% over 15 years. Tumour grade has not consistently been shown to independently impact on late recurrence[55].

As will be discussed in more detail below, increasing age (in a linear fashion) has been seen to be a significant prognostic factor for late recurrence[58]. This study included only post-menopausal women and so an effect of younger age would not have been seen. In contrast, another study that included both pre- and post-menopausal women found that the effect of age on late recurrence is not linear but instead had two peaks. In this study women <40 and >60 were at significantly greater risk of late recurrence compared to those in the 40-60 age group[57]. This reflects the well-known pattern with recurrence from year 0 and features in the Predict model[59]. This finding was also reflected in the most recent Oxford Overview study [25]. This showed that the under 35s were at greatest risk of late recurrence and then the next most at risk group was the 64-75 year age group. Interestingly however,

the risk of the under 35s was not as great as for the same age group when considering recurrences from year 0.

Younger women with breast cancer sometimes have biologically very aggressive cancers which, if they recur, may be more likely to do so early. Therefore, it is plausible that when early recurrences in those younger women are excluded, the relationship of late recurrence with age becomes more linear. An in-silico analysis of women within the METABRIC database demonstrated this and showed that from year 5 onwards risk of recurrence was more linear, however numbers of women under 40 were low (Figure 1.4).

Possible explanations for a dichotomous effect of age could be that younger women have more aggressive tumours but, as mentioned above, one might expect this to favour early recurrences rather than late. There have also been suggestions that younger women suffer more from endocrine therapy toxicity which affects their compliance and therefore puts them at greater risk of recurrence. Older women on the other hand may be at greater risk of recurrence due to under-treatment with fewer older women receiving chemotherapy than younger women. Age will also be confounded by menopausal status and the two are hard to separate. As will be addressed below, the circulating plasma of a pre-menopausal woman will have a different hormone profile to that of a post-menopausal woman. Therefore, variations in hormone receptor expression in breast cancer cells between the two groups (pre- and post-menopausal) are likely to reflect these differences in circulating hormones and could have distinct effects on the behaviour of cancer cells.

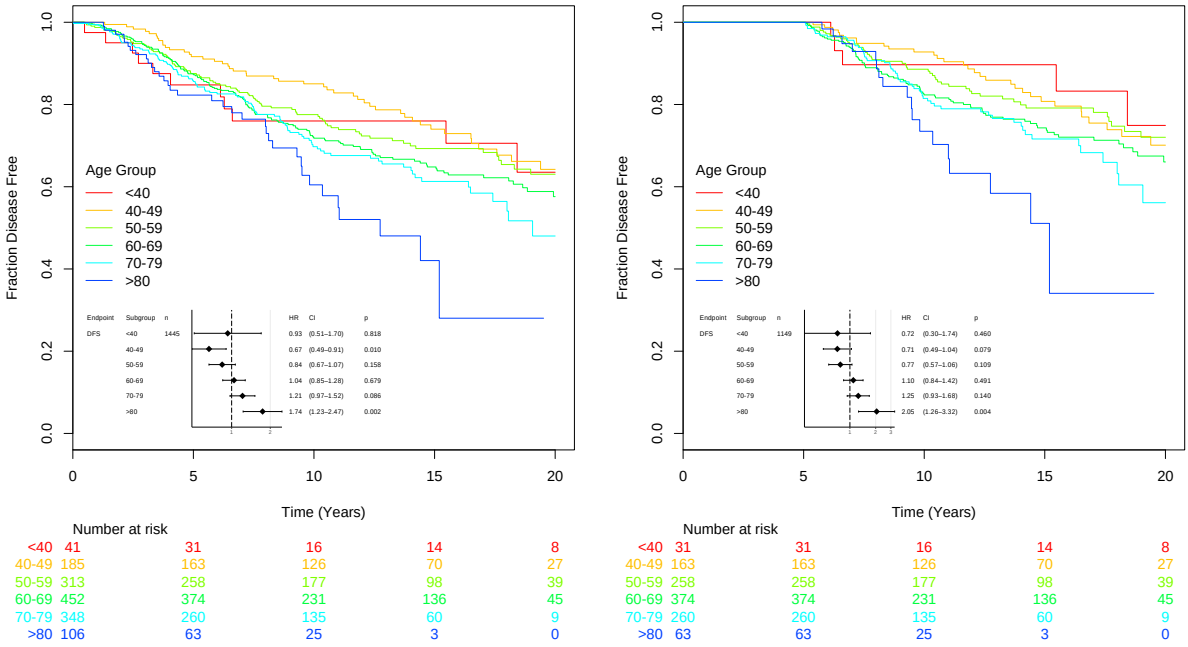


Figure 1.4 Association of age and distant recurrence in METABRIC. Left: shown from time 0 and right: shown from year 5.

1.7.2 Clinico-pathological prognostic models

Incorporation of variables into prognostic models is widely used in oncology and allows for weighting of the different components depending on their relative contribution to risk of an outcome. Prognostic models usually improve the performance of risk estimation based on single variables.

Variables such as tumour size, grade and lymph node burden have been incorporated into several prognostic models, such as the Nottingham Prognostic Index and NHS Predict, which are both designed to be used at the time of diagnosis to guide decisions on adjuvant endocrine therapy and/or chemotherapy. These methods are attractive because they represent an accessible, immediate and cost-free approach to risk prognostication that can be used in most routine clinical settings worldwide.

The clinical treatment score at 5 years (CTS5) has been specifically developed to predict risk of late recurrence in women who have not had distant disease recurrence 5 years after diagnosis and have completed 5 years of endocrine therapy[60]. This score was based on the

same principle as the clinical treatment score at diagnosis (termed CTS) which itself was a model of clinical variables developed to be used alongside the immunohistochemistry 4 (IHC4) for early recurrence[61]. CTS5 was developed on the patients from the ATAC study[62] who were all post-menopausal women with ER+ breast cancer, randomised after surgery to receive either anastrozole or tamoxifen or both. There is minimal strong evidence that has emerged from extended endocrine therapy trials for predictive biomarkers and certainly, given the modest recurrence benefit in some studies, extended endocrine therapy is not warranted for all women. The intended use of CTS5 is for guiding decisions regarding extended endocrine therapy based on the level of risk of late recurrence and the perceived benefit of continuing treatment.

The algorithm was built based on multivariate analysis of several parameters prognostic for late recurrence. The model combines patient age at diagnosis, lymph node burden (in categories), tumour size (thresholded at 30mm) and tumour grade into a mathematical formula.

Consistent with the findings of previous studies[26,54,55], having a high burden of lymph-node involvement is a strong predictor of late recurrence, hence its heavy weighting in the CTS5, indicating the prognostic relevance of this factor in the late-disease setting. Similarly tumour size features as a strong prognostic indicator.

The algorithm was then validated in the patients from the BIG 1-98 study who were postmenopausal women randomised after surgery to receive either 5 years of single agent anastrozole or tamoxifen or sequential tamoxifen and anastrozole for 2 and 3 years respectively. In the validation dataset, increasing CTS5 score had a positive correlation with risk of recurrence. Additionally, the study cohort was divided into risk groups using CTS5, the lowest of which had a 3.6% risk of distant recurrence at years 5–10. On the basis of this validation, CTS5 was re-calibrated using the combined data from the 2 trials. The final CTS5 algorithm is now freely available.

$$\mathbf{CTS5 = 0.438 \times nodes + 0.988 \times (0.093 \times size - 0.001 \times size^2 + 0.375 \times grade + 0.017 \times age)}$$

The authors of this study concluded that owing to their very low risk of late recurrence, women who fell into the low risk category had such a minimal potential for benefit from continuing endocrine therapy that it is appropriate for them to stop at 5 years. Further validation of the re-calibrated CTS5 is required and has not been attempted to date. This is one of the main objectives of this research.

Since its development, CTS5 has been tested in additional patient cohorts. In the IDEAL study dataset[35] of 1800 postmenopausal women, it produced three distinct risk categories but the calibration of observed versus expected events overestimated the number of expected distant recurrences. It is likely these results are affected by the fact that this study of extended endocrine therapy compared 2.5 additional years with 5 additional years and therefore all patients received at least 7.5 years letrozole, which would account for the much lower event rate in the study overall. CTS5 has been tested in a cohort of 349 node positive women and classified 29% as low risk with a 5-10 year distant recurrence risk of 4.2%[63].

Since this model has been derived from data from only post-menopausal women, any prognostic features unique to pre-menopausal women will not have been captured. Tumour size, grade and lymph node burden are likely to be important prognostic factors in pre-menopausal women but could be to different degrees as for post-menopausal women. The effect of age, as discussed previously, may differ and it is uncertain whether age, as a linear predictor of late recurrence, holds true for pre-menopausal women in the context of late recurrence.

1.7.3 Individual molecular biomarkers

Molecular biomarkers, in this context, can be defined as any biological feature of a cancer cell whose presence or burden is associated with outcome. Knowledge of cellular receptors is fundamental to the management of breast cancer and these receptors also function as biomarkers. Quantification of these biomarkers in cancer cells using different methods has been studied in terms of their relationship with recurrence outcomes.

As mentioned earlier assays for ER, PgR and HER2 have progressed in their precision and now measurement using IHC is the gold-standard for clinical purposes. This technique involves using antibodies to detect a target antigen (receptor) in the breast cancer tissue. The antibody is bound to a chromogen which stains the cells expressing that antigen. Scoring of these receptors is usually conducted using semi-quantitative measures such as the Allred score which considers degree and intensity of staining. Alternatively, percentage of positively stained cells can be used either alone or in combination with intensity of staining to produce an H score.

More recently methods of mRNA quantification of gene biomarkers have led to more precise measurement of ER and PgR as well as many other cancer related genes and these have been employed in many research studies. The added scope of gene expression quantification enables a more comprehensive analysis of the plethora of endocrine dependent and proliferative pathways within a cancer cell than can be gained by IHC analysis.

1.7.4 Biomarkers measured by IHC

1.7.4.1 ER and PgR

Studies have shown that high ER confers a favourable prognosis with regards to late recurrence in women who have received 5 years of endocrine therapy[64]. Additionally, there is data to suggest that in ER+ patients treated with tamoxifen, PgR>10% was associated with a better response to treatment compared with PgR<10% and reduced late recurrences[65]. For PgR<10% tamoxifen effect is seen only in years 0-5. In a study of women over the age of 60, lower ER positive percentage was significantly associated with late recurrence in comparison to non-recurrent controls[56]. This study demonstrated a higher proportion of PgR positive cancers among non-recurrent controls in comparison to late recurrence cases but numbers within this subgroup were low. In other studies, some including pre-menopausal women, despite levels of ER and PgR showing univariate significance for late recurrence, this has not borne out in multivariate analysis[55,57]. Two large overviews of adjuvant studies have failed to find a significant prognostic effect of quantitative ER and PgR measurement on distant recurrence at ten years beyond just

"positive" or "negative"[18,66]. However a large number of trials in Oxford Overview meta-analyses used older methods such as ligand binding assays or enzyme immunoassays to measure receptor expression[18] and it has been concluded that continuous versus categorical measurement of ER and PgR were more likely to find associations with outcome[66].

1.7.4.2 Ki67 Index

Ki67 is a protein marker of cellular proliferation. It is present in rapidly dividing cells and is often sparse or absent in quiescent cells. The information it provides regarding proliferation will have overlap with pathological grade as both parameters reflect mitotic rate. Several large meta-analyses have examined the prognostic role of Ki67 for recurrence from year 0[67,68] however evidence for its independent prognostic value has been inconsistent and studies are beset with marked heterogeneity making comparisons difficult. This, coupled with difficulties of reproducibility in IHC staining and scoring has meant that its measurement has not been established in routine practice. Ki67, however, has been shown to be of prognostic value for late recurrence in studies of patients who have completed 5 years of endocrine therapy however, fails to retain significance in multivariate analysis[55,69]. However, in a recent study of purely lobular breast cancers, Ki67 index was shown to improve the prognostic value of CTS5 and risk stratify the cohort into three groups, the lowest of which had a 1.9% risk of recurrence in years 5-10[70]. Additionally, Ki67 showed significant prognostic value within the IHC4 – a model for recurrence in years 0-10[61].

Overall, it would appear that there is prognostic value of levels of ER PgR and Ki67 for predicting late recurrence, but study findings have been inconsistent and different methods of ER and PgR measurement are hard to compare. It is certainly the case that more quantitative measurement is necessary to detect differences but the prognostic impact of these biomarkers seen for early recurrence is less clear for late recurrence. The impact of menopausal status on a woman's circulating hormone profile will impact on the hormone receptor expression of their breast cancer and studies have shown that the roles of PgR and Ki67 differ between pre- and post-menopausal women[71,72].

1.7.5 Individual biomarkers measured by gene expression quantification

Gene expression is the process by which the genetic code, as set out by the DNA, is converted into mRNA and then into a functional protein. Cells of a cancer will differentially express genes depending on their function so even if a particular DNA sequence is present in the cell, it may not be expressed. Therefore to measure gene expression and gain an idea of the functional process important in that cell, it is important to quantify RNA.

Several studies using molecular gene expression quantification have been conducted retrospectively in cohorts of patients with ER+ breast cancer, some of whom received endocrine therapy for 5 years.[73-78]. In one such study, tamoxifen treated breast tumours were classified according to the average expression levels of 12 genes encoding proliferation genes, and four ER-associated genes based on the gene signature Oncotype DX (*ESR1*, *PGR*, *SCUBE2*, and *BCL2*)[73]. Among patients with highly proliferative tumours, a high ER-associated gene expression score was indicative of improved DFS 0–5 years after diagnosis but of worse DFS 5–10 years after diagnosis, a result that was consistent across patients regardless of whether they had lymph-node involvement.

A similar analysis of genes and gene modules from the Oncotype DX signature in patients from the TransATAC study[74] revealed that high expression of ER-related genes was indicative of a favourable prognosis in years 0–5 but not in years 5–10 — *ESR1* and *SCUBE2* were the genes whose prognostic relevance changed the most in this regard[74]. Furthermore, high levels of *PgR* expression, a favourable prognostic feature in the early recurrence setting, lost prognostic value in the late recurrence setting, a finding that was also reported in the EBCTCG meta-analysis[25] in which Ki67 retained a moderate level of prognostic relevance for late recurrence.

The results of these studies of individual molecular biomarkers measured by quantitative mRNA methods[73,74] suggest that high ER-related gene expression confers a good prognosis in years 0-5 in women treated with endocrine therapy. A possible explanation for this could be that dormant tumour cells with high ER expression are well controlled by endocrine therapy but this control is relinquished on cessation of treatment. This leads to a

resurgence of circulating estrogen which would then trigger the greatest level of proliferation in high ER expressing cells. Late recurrence, therefore, becomes more likely in these cancers. This pattern, however was not reported in a similar study which found no prognostic significance of *ER* mRNA levels for early recurrence but a did show a positive association with late recurrence[79].

It is undeniable that gene expression quantification of *ER*, ER-related genes and proliferation genes provides an increased level of precision than IHC methods. ER signalling and its complex downstream effect is dependent on hundreds of genes whose expression cannot be captured purely by measurement of ER and PgR by IHC. Similarly, Ki67 is not the only determinant of proliferation and so the additional coverage of a large number of genes made possible by gene expression quantification would be expected to increase the prognostic information available. However, gene expression quantification comes at a high financial cost which will preclude its use in many parts of the UK and the world.

1.7.6 Molecular models

As with the clinicopathological models, molecular data have been combined into risk prediction models for recurrence in ER+ breast cancer. One such model is the IHC4 model discussed below. Additionally, several molecular tests enable quantitation of the levels of the RNA from multiple genes in a primary tumour sample, and have shown prognostic value for the evaluation of risk of distant recurrence in years 0–10 after diagnosis. Several of them have shown value for predicting late recurrence.

1.7.6.1 IHC4

Routinely measured ER, PgR along with HER2 and Ki67, using IHC have been integrated in the IHC4 score, which has been validated as a prognostic model for risk of recurrence at 0–10 years[61]. This model incorporates quantitatively scored values for ER, PgR, Ki-67 and HER2 status and combines this with patient age, tumour size, grade, lymph node burden and type of endocrine therapy to produce a risk of recurrence in years 0-10. In combining the receptor values with clinico-pathological variables, the model is able to capture a wide range of data that is independently prognostic for recurrence. It was initially developed on

the patients from the transATAC study and was shown to have prognostic value in the early recurrence setting. Importantly, this quantitatively measured model has a superior prognostic value compared to categorical IHC based subtyping[80].

For late recurrence however, IHC4 has been reported to not have prognostic value superior to that derived from standard clinical variables[55,81]. This difference can be explained by the strong weight (and direction thereof) that the ER component has in the IHC4 algorithm. The shift in prognostic importance of ER-related genes after 5 years (discussed above) would be expected to result in the differential weightings allocated to the IHC4 components which would then be no longer appropriate for late recurrence.

1.7.6.2 PAM50

The PAM50 score combines the expression levels of 50 genes and tumour size, which are used to define the intrinsic subtypes of breast cancer, and information from a set of proliferation-related genes for prognostic purposes[82]. PAM50 has prognostic value for the prediction of recurrence risk between years 0–10[83], with similar performance for the prediction of risk at years 0–5 and 5–10[55,84]. An analysis of 2 combined datasets demonstrated that PAM50 had prognostic value in the late recurrence setting, independent of the clinical treatment score (CTS)[85].

1.7.6.3 EndoPredict

Endopredict is a 12-gene expression panel that is combined with tumour size and nodal status to generate an EPClin score[79]. EPClin has good prognostic value for the prediction of disease recurrence 0–10 years after diagnosis[79,86], and of late recurrence in two large translational studies[78,86].

1.7.6.4 Breast Cancer Index

The Breast Cancer Index (BCI) is a molecular score that incorporates the ratio of *HOXB13/IL17BR* levels and a proliferation-related 5-gene module referred to as Molecular Grade Index [87] . It has demonstrated independent prognostic value for late recurrence[88] and has shown prognostic value in lymph node positive[81] and negative cancers[89]. *HOXB13/IL17BR* ratio has also demonstrated predictive value for responses to extended

letrozole therapy, albeit in a small, case control study. In patients involved in the MA17 study[11], a high ratio in the placebo group was independently associated with a significantly increased risk of recurrence and also conferred greater recurrence free survival benefit from extended letrozole compared to low ratio)[90]. BCI is the only test to claim predictive value for extended endocrine therapy.

1.7.6.5 Comparisons between CTS5 and BCI

A further study has investigated the concordance between CTS5 and BCI in a small cohort of women with low stage ER+ cancers, 14% of whom were pre-menopausal [91]. In this study the intermediate and high risk CTS5 categories were combined. The two models showed moderate levels of concordance. 24% of CTS5 low patients reclassified as BCI high and 26% CTS5 high reclassified as BCI low [92]. Many re-classifications were for patients that fell in the 5-10% risk (intermediate) CTS5 category.

Taken together, despite not being developed specifically for the evaluation of the risk of late recurrence, PAM50, EPClin and BCI all have good prognostic value in this setting. It is undeniable that clinical and molecular variables provide independent prognostic information and therefore the most accurate risk prediction models will be those that incorporate both elements. All three tests enable clinically valid risk groups to be defined, the lowest of which includes women for whom extended endocrine therapy might not be warranted. Further validation in higher stage women and in pre-menopausal women would broaden the scope of women in whom these signatures can be applied.

1.7.6.6 Late recurrence and dormancy signatures

Non-commercial signatures specific to late recurrence and dormancy have been developed and tested in addition to the commercially available gene expression models described above. These models, which have demonstrated an association with metastasis free survival, incorporate genes involved in processes such as cellular quiescence and dormancy[50] and epithelial mesenchymal plasticity[93]. How such signatures perform in comparison with the established commercial signatures remains to be established. However, in a large study of time-dependent gene-expression signatures for the prediction

of late recurrence, a specific 5–10-year signature did not improve risk estimation above that provided by a 0–10-year model[94].

1.7.7 DNA aberrations as biomarkers of late recurrence

1.7.7.1 Mutations

DNA mutations in the form of point mutations and translocations are common in all tumour types and accumulate over time. These alterations arise in proliferative cells and therefore, in a dormant but not fully quiescent tumour, mutations can still accumulate. When DNA alterations affect a driver gene, the situation can lead to the emergence of a dominant clone. Such an alteration could facilitate the conversion of a small, dormant micrometastatic deposit into an overt metastatic lesion, and therefore lead to late recurrence. In ER+ breast cancer, the most common mutations are those affecting the *PIK3CA* and *TP53* genes[95,96].

Molecular studies on paired primary and distant recurrence specimens have led to the conclusion that whilst the genomic landscape of the two tumours are similar, heterogeneity exists in subclonal mutations that have arisen after micrometastatic colonies have seeded at distant sites[95].

Mutations in *ESR1* that confer constitutive activation of the mutated ER protein independent of ligand binding have been detected in ~20% of tumour samples derived from endocrine therapy treated patients with metastatic ER+ breast cancer[97], and in ~40% of circulating tumour DNA (ctDNA) samples from patients with metastatic ER+ breast cancer receiving fulvestrant[98]. Studies have shown that *ESR1* mutations are much less frequently associated with adjuvant use of AI as opposed to AI treatment in the metastatic setting, and never associated with tamoxifen use without AI[99]. A retrospective study involving 42 women treated for at least 2 years with aromatase inhibitors in the adjuvant setting and who subsequently had disease recurrence showed that only two patients had *ESR1* mutations that were detectable in plasma ctDNA at the time of recurrence[100].

1.7.7.2 Copy number alteration

Most genetic alterations driving the progression of breast cancer appear to arise from copy number alterations (CNA) i.e. deletion and amplifications. Their contribution to metastatic dissemination has been characterised in several detailed studies, albeit with modest-sized cohorts[101,102]. CNA leading to enhanced ER modulated cell cycle progression has been associated with poor response to neo-adjuvant AI[103]. Additionally, CNA have been combined with gene expression data into a molecular signature for recurrence prognostication from time of diagnosis[102]. Internal validation has demonstrated the ability of the score associated with this signature to separate the study cohort into three distinct risk groups, the lowest of which had a distant recurrence rate of 11% over 7 years (at a median follow-up duration of 99 months)[102]. Most recently a molecular analysis of copy number alteration and gene expression profiling was applied to over 1900 patients from the METABRIC dataset[104]. This analysis revealed that there were 4 clusters into which 26% of the ER+ patients grouped. These clusters were associated with a higher risk of late recurrence and were enriched for key copy number alterations, some of which are known to be driver alterations. Amplifications of regions of chromosomes 8, 11 and 17 led to the overexpression of genes such as *RPS6KB1*, *ZNF703*, *MYC*, *FGF3* and *CCND1* which are all known oncogenes some of which are potentially therapeutically targetable.

1.7.8 Minimal residual disease in ER+ breast cancer

The proportion of women with micrometastatic or minimal residual disease after surgery is estimated at 40%. Currently, following surgery, patients are recommended to receive adjuvant systemic endocrine therapy and/or chemotherapy based on baseline risk factors, owing to the lack of a clinically available method for measuring minimal residual disease. The same situation occurs in patients who have received endocrine therapy for 5 years with regards to deciding to extend treatment or not. Methods of quantifying minimal residual disease such as circulating tumour cells (CTCs) and circulating tumour DNA (ctDNA) are being developed having shown promise for monitoring of metastatic disease[98,105,106].

1.7.8.1 Circulating tumour cells and circulating tumour DNA

CTCs present in plasma of women who have completed 5 years of endocrine therapy are associated with a 13-fold increase in the risk of late recurrence compared to undetectable circulating tumour cells in a small study with relatively short follow up [107]. To further establish the prognostic value of CTCs in late recurrence, longer follow-up durations, as well as serial measurements during treatment with endocrine therapy and after cessation, would be useful to identify those women with MRD. Similarly with ctDNA, prognostic value for early recurrence has been shown when used in the neoadjuvant[108] and adjuvant settings[109]. Longer follow-up durations and prospective study designs are required to assess the value of ctDNA monitoring in guiding adjuvant therapy and extrapolate its relevance for late recurrence.

1.8 Practical implications of late recurrence

1.8.1 Survivorship of patients

The enduring risk of recurrence in ER+ breast cancer creates a different survivorship experience in comparison to survivors of other cancers where risk of recurrence reduces markedly after 5 years. For patients themselves, managing a potential lifelong risk of recurrence presents both physical and psychological burden. Patients frequently cite endocrine therapy as a daily reminder of their cancer that they would much rather forget about and, therefore, committing to a further 5 years could have important psychological as well as physical implications, but also uncertain levels of benefit.

1.8.2 Extended endocrine therapy

As discussed at the start of this chapter, the role of endocrine therapy in the first 5 years is undisputed and its effectiveness unquestionable. In contrast, management strategies beyond 5 years are far from clear and this includes both surveillance of breast cancer survivors and the question of extended endocrine therapy. Therefore, risk prediction beyond 5 years is of paramount importance in clarifying the way forward in terms of management.

1.8.3 Implications for risk prediction

Risk prediction models need to be broadly applicable, accessible and sufficiently precise. In achieving this, compromise will be necessary. This will require robust model building, validation of these models in different cohorts of women and the championing of methods that do not preclude their use outside of specialist centres in the developed world. Risk prediction models need to capture as much information as possible whilst retaining a simplicity that allows it to be reproduced and utilised widely across hospital based and primary care settings.

One crucial step in risk prediction model development is to test for their clinical validity in a variety of relevant populations and also for their utility in changing decisions and improving long-term outcomes for patients. By carrying out tests of validation in different cohorts, one can see whether a tool that performs well in its development dataset extrapolates to an independent dataset. Testing clinical utility will show whether models are likely to have a significant impact on routine practice and will be situation specific. Without these vital studies, it cannot be known how well the tool will perform in day-to-day practice. Even amongst prognostic models that perform well in a variety of populations, it may still not show clinical utility by resulting in some sort of change in practice that ultimately benefits patients. It is these two key steps which demonstrate the clinical relevance of these models and their place in long term management.

1.8.4 Implications for research

Since risk of recurrence persists to 20 years, follow up within a study of late recurrence needs to be at least this long in order to be certain a woman is recurrence free. This is practically very seldom possible in translational research and to conduct this in a retrospective fashion would require long term storage of archival tissue. Prospective trials of this nature will be expensive to run, require large numbers of participants and will not yield results for many years. A retrospective study would ameliorate some of these

disadvantages, but the trade-off is that patients with adequate follow up will not be up to date and will therefore not necessarily reflect modern practice.

The greater awareness of the continuous risk of late recurrence in patients with ER+ breast cancer that has unfolded over the past few years has provided multiple new, previously unappreciated challenges for research, clinical management and for survivors of breast cancer. The research that I have undertaken aims to develop and strengthen the evidence for risk prediction models for late recurrence so that they may become embedded into routine practice for the long-term management of ER+ breast cancer survivors.

1.9 Some unanswered questions:

1.9.1 Can CTS5 be extrapolated to a non clinical-trial population and pre-menopausal women?

The CTS5 model was developed on a purely post-menopausal, clinical trial cohort. This means that it was a highly selected population and women with significant co-morbidities or frailty would have been prevented from entering. This is routine practice with clinical trials but can lead to unrepresentative trial cohorts. Further to this, prognostic value of CTS5 for pre-menopausal women and women with HER2 positive (HER2+) breast cancer remains to be seen. It is likely that of the CTS5 variables, the most likely variation in pre-menopausal women will be the effect of age and whether this linear predictor is appropriate for the younger age group.

1.9.2 Can CTS5 be extrapolated to higher stage patients?

Many prognostic tools (both clinical and molecular) have been trained and validated on low risk clinical-trial populations of patients. It remains to be established whether CTS5 can be extrapolated into higher stage patients. In higher stage women the thresholding of tumour size at 30mm may no longer be appropriate and additionally a greater proportion of women receiving chemotherapy may have a bearing on the performance of CTS5.

If it demonstrates clinical validity in higher stage patients, then it will be a more widely applicable tool. This will enable informed recommendations to be made regarding extended endocrine therapy in a larger number of women than is currently possible and in this way more women can be confident in their management recommendations.

1.9.3 Does CTS5 have clinical utility for aiding extended endocrine therapy decision making?

Clinical utility is the ability of a prognostic test to alter the course of management in a way that improves outcomes or reduces toxicity for patients or reduces cost for the healthcare provider. In the context of late recurrence, it is best tested in a prospective study of extended endocrine therapy with patients stratified according to CTS5. This method enables researchers to determine the predictive value of CTS5 for benefit from extended therapy. Establishing this in an extended adjuvant setting is problematic owing to the long follow-up times needed to be able to definitively answer this question. An alternative approach is to carry out a decision impact study where treatment decisions are made before and after implementing the prognostic model and analysing the impact the model had on treatment decisions. This type of study would not provide information of predictive benefit of CTS5 but would give a more immediate indication of likely utility.

Generating evidence for clinical utility will expand its use both alone and in combination with other models. If CTS5 were to be widely employed, then it would reduce clinician to clinician and centre to centre variability in treatment recommendations.

1.9.4 Are there any molecular alternatives to gene expression profiles for late recurrence and could this improve the CTS5?

It is discussed above that clinical and molecular predictors of recurrence provide independent prognostic value and combining them in the IHC4+CTS improved the prognostic value of either component alone for early recurrence[61]. Therefore, in a similar way, combining CTS5 (which captures no molecular variables) with a molecular model

specific to late recurrence may improve prognostic of CTS5. This has been demonstrated in the superior performance of gene expression models that incorporate clinical features over those that do not. One alternative to gene expression profiling would be to develop an IHC based prognostic model (similar to IHC4) measuring molecular biomarker expression at the protein level. Comparative studies have shown that quantitative IHC scoring for the routinely measured IHC biomarkers (ER, PgR, and Ki67) has a high concordance with mRNA transcript demonstrating that changes in gene expression should translate to protein receptor expression [110].

If the individual biomarkers demonstrate prognostic value for late recurrence, training an IHC model specifically for late recurrence could provide a molecular risk prediction tool that captures different information to CTS5 and therefore improves it. Additionally, such a model would be low cost and relatively accessible and could further inform on extended endocrine therapy recommendations.

1.9.5 Can we identify any biomarkers for very late recurrence beyond 10 years from diagnosis?

Extending endocrine therapy to ten years is one option to control for the enduring risk of recurrence but at the ten-year mark, could there be a role for extending therapy even longer? Currently there is no clinical trial data to support such an approach but there are clinicians who advocate ongoing treatment for women without significant toxicity who they feel remain at high risk of recurrence. For those tolerating treatment then this could be reasonable. However, protracted therapy places a considerable burden on patients and healthcare resources for questionable benefit. It would therefore be useful to explore any prognostic biomarkers for recurrence beyond 10 years to allow safe cessation of endocrine therapy at that point.

1.10 Hypothesis and Aims

1.10.1 Overall Hypothesis

Predicting risk of late recurrence can be improved by development of existing clinical and molecular biomarkers and prognostic models.

1.10.2 Specific Aims

The first 3 aims are intended to provide evidence that could potentially be incorporated into clinical practice almost immediately. The fourth aim is for an exploratory study proposed to provide preliminary data on the molecular biology of primary ER+ breast cancers that recur beyond 10 years.

1.10.2.1 Test validity of CTS5 in additional patient cohorts.

This will ascertain whether CTS5 can be extrapolated to pre-menopausal and higher stage patients and also to unselected patients not part of a clinical trial. The CTS5 will be retrospectively applied to a higher stage clinical trial cohort and an unselected RMH patient cohort to test whether it accurately predicts late recurrence. If proven, this will establish the validity of CTS5 within multiple patient cohorts and increase the number of women in whom this prognostic tool could be applied.

1.10.2.2 Test the clinical utility of CTS5.

This will enable us to gauge how use of CTS5 impacts on patient outcomes. A prospective decision impact study will aim to establish whether CTS5 impacts on clinician treatment decisions regarding extended endocrine therapy beyond 5 years. It also aims to evaluate how use of the CTS5 could alter routine practice.

1.10.2.3 Develop an IHC based prognostic model for use in combination with CTS5.

This will involve training an IHC model specifically for late recurrence using tissue from a clinical trial with over 10 years of follow up. The prognostic value of individual variables will be explored and, on this basis, a model built. It will be combined with CTS5 with the specific question of whether it improves the prognostic performance of CTS5 alone. An IHC based

model will enable incorporation of molecular biomarkers known to be important for late recurrence into a prognostic algorithm. Such a model could potentially be a low cost, accessible alternative to genomic expression models.

1.10.2.4 Explore genomic differences in primary tumours between women with very late recurrence and women with no recurrence.

This work aims to discover differences between very late recurrent and non-recurrent ER+ primary breast cancers. It will utilise archival tissue with long follow up periods to ensure non-recurrent controls are truly non-recurrent. The study will explore RNA expression and DNA aberrations to identify biomarkers for very late recurrence. If such biomarkers are identified, this may provide a hypothesis for further research into risk stratification for patients beyond ten years from diagnosis given their ongoing potential for recurrence.

2. Testing validity of CTS5

2.1 Introduction

CTS5, as discussed in chapter 1, is a prognostic tool for distant recurrence in years 5-10 from diagnosis following curative treatment for early ER+ breast cancer. To date, it is the only tool of its kind that was developed specifically for risk prediction of late recurrence. Its intended use is at 5 years to aid decisions regarding extending endocrine therapy to 10 years in view of data suggesting this approach for some women[29-31,33]. The tool combines routinely collected clinico-pathological data (age at diagnosis, tumour size, tumour grade and lymph node burden) into a mathematical algorithm to produce a risk of distant recurrence from years 5-10. This algorithm assigns co-efficients to each variable that reflect their relative contribution to the outcome.

CTS5 was trained on the data from women who were alive and distant recurrence-free at 5 years in the ATAC study[62] and tested on the BIG1-98[111] study dataset. Based on this test of validation, a re-calibrated CTS5 was derived from the combined datasets of the two studies. This version of CTS5 is recommended for use and is available online[112]. Both datasets were large adjuvant endocrine therapy studies in post-menopausal women with early ER+ breast cancer. CTS5 has shown good prognostic performance with a HR of 2.07 (95% CI 1.88 to 2.28) for late distant recurrence with each unit increase in score[58]. Additionally, the tool divided the combined cohort into three distinct risk groups, the lowest of which contained 41% of the trial population who had a median 3.6% risk of distant recurrence in years 5-10. This large, combined dataset provides robust evidence for the prognostic value of CTS5. On the basis of this data, it has been proposed that those who fall into a CTS5 'low-risk' category, can safely avoid extended endocrine therapy[58].

Women in the ATAC and BIG 1-98 trials had relatively low stage disease and were all post-menopausal. Furthermore, populations of clinical trial patients can behave differently to 'real-life' patients who are clinically unselected and have a wider range of co-morbidities and backgrounds. It is not known how the score performs in a cohort of unselected, non-trial patients as well as pre-menopausal patients and consequently how widely the CTS5 can be applied.

2.1.1 Defining validity

Testing validity of a prognostic model is an imperative part of model development as, by definition, a model will perform well in the dataset in which it was developed, but these results may be over-optimistic. Tests of clinical validity, in the context of prognostic models, aim to determine whether the performance of a prognostic model in a separate dataset, matches or comes close to matching its performance in the development dataset. The degree of closeness of performance required to a deem a model clinically valid will be situation based and will be defined differently across groups of clinicians and different disease states.

In this study it is the re-calibrated CTS5 (based on the combined ATAC and BIG 1-98 datasets) that will be tested.

Validation can be broadly assessed in 3 ways. The first is by internal validation which is commonly done by sample splitting. This method involves dividing the initial development cohort into 2. One half is used to develop the model and the other half is then used to validate. This gives quick results and would be a good indicator of whether it will be worthwhile taking the model forward to further tests of validation. However, this method does not test the model in a cohort that is different from the development set as one half is likely to be similar to the other.

The second type of validation, temporal validation, involves testing a model on a subsequent cohort within the same centre. Lastly, external validation would be the testing of the model in a completely independent dataset which would need to be appropriately similar but could explore differences from the development set. Such validation exercises use retrospective datasets and therefore take advantage of long follow up which is important for studies of early breast cancer. This is the method used to validate the original CTS5(ATAC) – by utilising the BIG -198 dataset.

Within a test of validation, certain statistical analyses are important. In particular, calibration and discriminatory performance of the model should be assessed[113]. Calibration is the degree to which the outcomes in the validation set match the expected outcomes. The expected outcomes are generated using each individual's event probability calculated

from their risk score from the original model. Calibration can be assessed by plotting observed versus predicted events with a χ^2 test to determine difference between the two. Discrimination is the ability of a test to correctly identify someone with an outcome event over someone without. This is best demonstrated by plotting a receiver operator characteristics (ROC) curve. This plots sensitivity against 1-specificity i.e. true positives against false positives. The area under the curve, also known as the c-statistic gives an overall value of discriminative ability of the model. Typically, prognostic models have a c-index of 0.6-0.85[114]. The 21 gene recurrence score (Onctoype DX) has shown an AUC, for recurrence free survival, of 0.69 in a validation study of 465 women with ER+, HER2 negative (HER2-) breast cancer[115] and AUC, for overall survival, of 0.60 in a validation study of the national cancer database in the US[116].

Calibration and discrimination can be tested across the whole cohort and with the cohort divided into clinically relevant subgroups. Additionally, tests of validation aim to test the separation of a cohort into risk groups. This is particularly relevant for CTS5 where the risk groups are intended to guide management decisions regarding extended endocrine therapy. The Kaplan Meier survival method is useful within a study of validation as it can incorporate the element of time into the validation exercise which ROC curves cannot. If risk groups are generated in a development set, then the same parameters should be applied in the test set i.e. for CTS5 the cut-offs established in the original combined ATAC/BIG 1-98 datasets for the three risk groups must be employed for the validation dataset. A test of validity would be whether proportions of events in each group was similar across the two datasets and if separation of outcome measures across the risk groups was maintained in the test set.

2.2 Aims

To test the validity of CTS5 within an unselected, consecutive cohort of women with early ER+ breast cancer who remained distant recurrence free at 5 years from diagnosis.

More specifically:

To establish whether CTS5 has a similar prognostic performance within this independent dataset compared with that in the combined ATAC and BIG 1-98 datasets.

To establish whether CTS5 predicts, with sufficient accuracy for clinical use, late recurrences in pre-and post-menopausal women who have reached 5 years from diagnosis.

To establish whether CTS5 adequately divides the cohort into clinically useful risk groups.

In order to achieve these aims, I conducted a large retrospective cohort study of women treated for ER+ early breast cancer in a single centre. The cohort was selected to determine whether CTS5 is applicable in women who are clinically unselected and in pre-menopausal women.

2.3 Methods

2.3.1 Patient cohort

This retrospective, external test of validation was conducted in a single centre and was approved by the Royal Marsden Hospital (RMH) committee for clinical research (CCR). Patients were identified through an initial search conducted by the Royal Marsden Hospital (RMH) Information team. The following selection criteria were used: female, diagnosed with ER+ breast cancer in years 2000-2007, underwent surgery with curative intent at the Royal Marsden Hospital. These dates were chosen as by 2000 the recommendation for patients with ER+ disease to receive 5 years of endocrine therapy was universal. The upper time point of 2007 was selected so that all patients would be at least 10 years from diagnosis. Patients were excluded if they had non-invasive cancer on histopathological analysis, less than 5 years of documented follow up from diagnosis, distant recurrence within 5 years of diagnosis, death due to causes other than breast cancer within 5 years of diagnosis, ER-negative disease and lack of CTS5 input data. Patients who died due to non-breast cancer causes beyond 5 years of diagnosis were censored at the date of death.

Another initial exclusion criterion of the study was receiving more than 5 years of endocrine therapy. The reason behind this was that in the original ATAC and BIG 1-98 studies, patients were treated with 5 years of endocrine therapy only and owing to the absence of data at the time for benefit of extending this treatment time to 10 years, very few women would have continued their endocrine therapy beyond this point. Additionally, by including those who took endocrine therapy for longer than 5 years, the treatment effect will mean that the event rate of the whole cohort will be lower than expected. This could then affect the apparent performance of the model in comparison to the training set. However, as the data collection progressed it became clear that there was an appreciable proportion of women who continued endocrine therapy beyond 5 years. These tended to be younger women or those who had higher risk features of their primary cancer. It seemed likely that to exclude all these cases would enrich my dataset with much lower risk women than is representative of a 'real life' situation and would also adversely affect the population size and event number. For this reason, the exclusion criteria for the study were modified to include women who received more than 5 years of endocrine therapy and secondary analyses were performed for those who received less than or equal to 60 months' endocrine therapy.

2.3.2 Data Collection and Analysis

For each patient identified within the initial IT search, the EPR was manually scrutinised. This involved searching each patient individually and manually locating the necessary baseline, input and outcome data from within the EPR and entering it into a separate database. The following data was collected for eligible patients: date of birth, date of diagnosis, menopausal status at diagnosis (if not recorded, women were presumed to be post-menopausal if aged 51 or older), pathological tumour size, histopathological grade, pathological nodal status, ER status, HER2 status, type and duration of endocrine therapy, occurrence and dates of distant recurrence, local/contralateral recurrence, date of last follow up and date and cause of death.

The CTS5 value was calculated for each patient using the published algorithm for the combined ATAC and BIG 1-98 datasets[58].

CTS5=0.438 x nodes + 0.988 x (0.093 x size - 0.001 x size^2 + 0.375 x grade + 0.017 x age)

In this model, node negative = 0, 1 positive node = 1, 2-3 positive nodes = 2, 4-9 positive nodes = 3 and >9 positive nodes = 4. Histological grade 1 = 0, grade 2 = 1 and grade 3 = 2. Tumour size is continuous and is capped at 30mm. The expected risk of recurrence for each individual was determined from the equation relating risk to CTS5 which is the logarithm of the baseline cumulative hazard function; calculated using the Breslow method with a correction factor for each sub-group analysed. This estimates the risk between 5 and 10 years. Expected risk was also modified for those women who recurred in years 5-10 or those who died of non-breast cancer causes in order to take into account the shorter follow-up these women had.

Where patients received neo-adjuvant chemotherapy or endocrine therapy, the pre-treatment and pathological tumour indices for size, grade and lymph node burden were compared and the higher value was used. Where patients were discharged after 5 years or were part of an open access follow-up program and hence not seen in a clinic (as is often the case in treated breast cancer in the UK), continued adherence to endocrine therapy

was assumed to 5 years unless there was documentation of early cessation or continuation beyond that point. For patients with less than 10 years of documented follow up, it was assumed that they were recurrence-free at the time of data collection.

2.3.3 Statistical Considerations

The primary population for validation was post-menopausal women and the secondary population was pre-menopausal women. The primary endpoint of the study was distant recurrence in years 5-10 and the secondary endpoint was distant recurrence in years 5 onwards. Distant recurrence was defined as metastatic disease at a distant site and therefore women with local or loco-regional relapse or contralateral new primary events in years 0-5 were not excluded. Local, loco-regional or contra-lateral recurrences in years 5+ were not treated as events. Patients with second primary cancers in years 0-5 were not excluded from the analysis. Death in years 5+ from other causes without distant recurrence was a censoring event. At the time of data collection, all included women had had the potential for 10 years of follow up and therefore, as stated above, those without documented follow up (i.e. a clinic visit or telephone call) were assumed to be recurrence free and were not censored at date of last follow up.

Pre-planned subgroup analyses were performed for chemotherapy-treated versus chemotherapy-naïve patients. Initially it was intended to analyse the HER2+ versus HER2- subgroups however owing to low patient numbers for the HER2+ sub-group and event numbers this analysis was not performed. Post-hoc sub-group analyses were performed for women who took less than or equal to 60 months of endocrine therapy versus those who took more than 60 months.

2.3.4 Statistical Analysis

For each of the two endpoints, prognostic value of the continuous CTS5 score was calculated using hazard ratios (HR) for a unit increase in the score along with 95% confidence intervals and likelihood ratio χ^2 for distant recurrence. Log-rank P values were calculated to determine significance. Patients were stratified into risk groups using the

CTS5 cut-offs (<3.13, 3.13-3.86 and >3.86) established in the original combined datasets which sought to define a low, intermediate and high-risk group with a <5%, 5-10% and >10% 5-10 year DR risk respectively. Kaplan Meier survival method was used to obtain the time to event curves for the three risk categories based on the observed distant recurrence time and events. The 5-10 year distant recurrence risks (%) were calculated with corresponding 95% confidence interval. HR between low/intermediate and low/high-risk groups were calculated with corresponding 95% confidence intervals to demonstrate discrimination between risk groups. Discrimination was further evaluated using ROC.

Calibration of the CTS5 risk model was tested by comparing observed (O) DR events to expected (E) primary endpoint events with the dataset divided into quintiles and into CTS5 risk groups. The expected events were calculated based on the baseline hazard for distant recurrence in the combined ATAC/BIG 1-98 datasets against which the model's performance is being calibrated. For those women who recurred in years 5-10 and those who died of non-breast cancer causes in years 5-10, their 5-10 year baseline hazard was adjusted according to the amount of follow up beyond 5 years that they had. The overall performance of the CTS5 is shown by likelihood ratio Chi square test and observed versus expected ratio with corresponding 95% confidence interval.

2.4 Results

2.4.1 Patient cohorts

A total of 3598 women who had undergone surgery with curative intent for early breast cancer at RMH between 2000-2007 were selected. After exclusions (Figure 2.1), a total of 2428 women were included in the analysis.

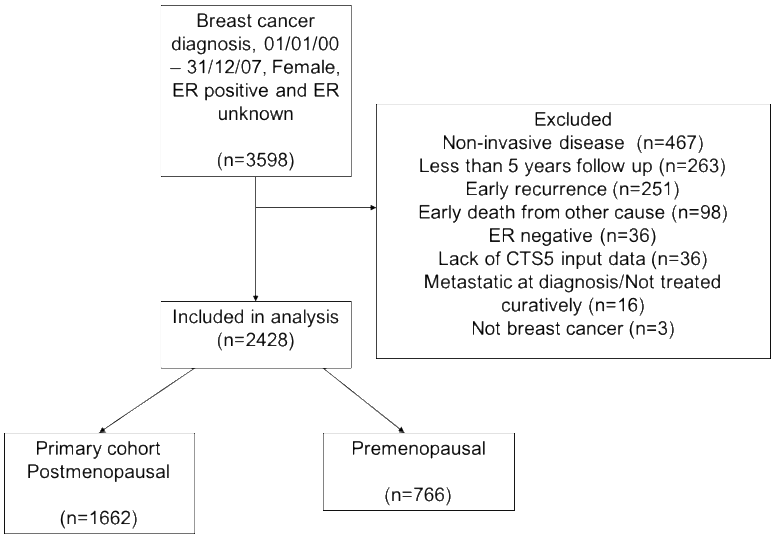


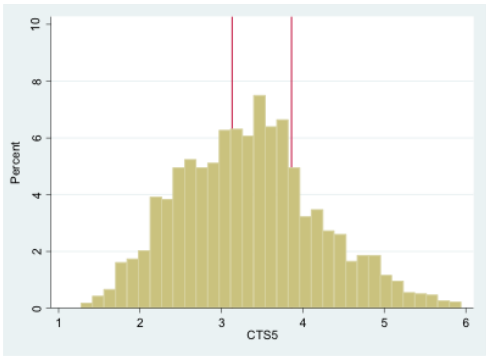
Figure 2.1 Consort diagram

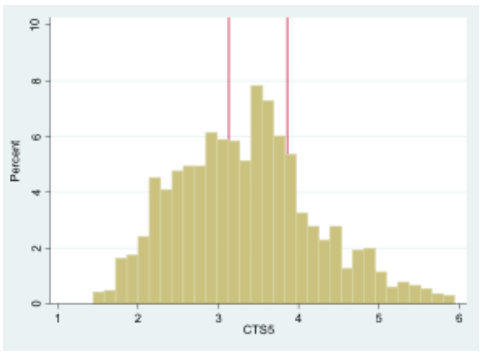
The cohort included 1662 post-menopausal women and 766 pre-menopausal women. Baseline characteristics are detailed in Table 2.1. All women had ER+ tumours and 9% of tumours were known to be HER2+. PgR was positive for 48% of women however, PgR status was unknown in over half of women. Most women received between 4 and 6 years of endocrine therapy with a small proportion taking less than 4 years or more than 6 years. Median calculated CTS5 score was 3.33 (IQR 2.71-3.85) for the post-menopausal cohort and 3.31 (IQR 2.69-3.85) for the pre-menopausal cohort and these equate to a 5-10 year risk of distant recurrence of 6.0% and 5.9% respectively which was statistically not different (P=0.46). There was a normal distribution of scores among the two populations (Figure 2.2).

	Pre-menopausal (N=766)	**Post-menopausal (N=1662)**	**All (N=2428)**
Median age (IQR)	46 (41-49)	62.5 (57-70)	57 (49-66)
Receptor Status			
ER positive	766 (100%)	1662 (100%)	2428 (100%)
PR positive	417 (54.4%)	765 (46.0%)	1182 (48.7%)

PgR negative	49 (6.4%)	173 (10.4%)	222 (9.1%)
PgR unknown	300 (39.2%)	724 (43.6%)	1024 (42.2%)
HER2 positive	97 (12.7%)	125 (7.5%)	222 (9.1%)
HER2 negative	612 (79.9%)	1383 (83.2%)	1995 (82.2%)
HER2 unknown	57 (7.4%)	154 (9.3%)	211 (8.7%)
Tumour size (mm), median (IQR)*	22 (14-33)	18 (12-27)	19 (12-33)
Grade			
Well	105 (13.7%)	356 (21.4%)	461 (19.0%)
Intermediate	405 (52.9%)	919 (55.3%)	1324 (54.5%)
Poor	256 (33.4%)	387 (23.3%)	643 (26.5%)
Nodes			
Negative	459 (59.9%)	1211 (72.9%)	1670 (68.8%)
1 positive	133 (17.4%)	201 (12.1%)	334 (13.8%)
2-3 positive	96 (12.5%)	134 (8.1%)	230 (9.5%)
4-9 positive	64 (8.4%)	77 (4.6%)	141 (5.8%)
9+ positive	14 (1.8%)	39 (2.4%)	53 (2.2%)
Chemotherapy	540 (70.5%)	468 (28.2%)	1008 (41.5%)
Adjuvant endocrine therapy			
< 12 months	57 (7.4%)	70 (4.2%)	127 (5.2%)
12-48 months	46 (6.0%)	55 (3.3%)	101 (4.2%)
48-72 months	481 (62.8%)	1318 (79.3%)	1799 (74.1)
>60 months	194 (25.3%)	257 (15.5%)	451(18.6%)
Tamoxifen only	409 (53.4%)	577 (34.7%)	986 (40.6%)
AI only	60 (7.8%)	504 (30.3%)	564 (23.2%)
Switch (Tam to AI)	263 (34.3%)	546 (32.9%)	809 (33.3%)
Unknown	34 (4.4%)	35 (2.1%)	69 (2.8%)

Table 2.1 Baseline characteristics of the pre- and post-menopausal populations.

a. 

b. 

c.

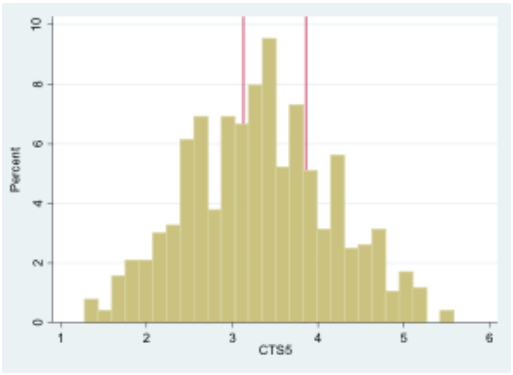


Figure 2.2 Distribution of CTS5 scores a. for the whole population, b. for post-menopausal women and c. for pre-menopausal women.

Over the follow-up period, 149 (9%) DRs were recorded in post-menopausal women, with 99 (6%) breast cancer deaths. In pre-menopausal women, 94 (12%) DRs were recorded with 48 (6%) breast cancer deaths. DR in years 5-10 were 107 (6%) and 51 (7%) for post- and pre-menopausal women respectively. Median time to DR for the post-menopausal cohort was 8.3 years and for the pre-menopausal cohort, 10.7 years. Of women with known HER2+ cancer, there were 8 DR. Median follow-up overall was 13.4 years from diagnosis and all patients had reached 10 years from diagnosis.

	Premenopausal (N=766)	**Postmenopausal (N=1662)**	**All (N=2428)**
DR in years 5-10	51 (6.7%)	107 (6.4%)	158 (6.5%)
DR after 5 years (including beyond 10 years)	94 (12.3%)	149 (9%)	243 (10.0%)
Deaths after 5 years	68 (8.9%)	388 (23.4%)	456 (18.8%)
Breast Cancer Deaths	48 (70.6%)	99 (25.2%)	147 (32.2%)
Other	8 (11.8%)	91 (23.5%)	99 (21.7%)
Unknown	12 (17.7%)	198 (51%)	210 (46.1%)
Median follow-up (yr from diagnosis), (IQR)	13.8 (11.9-16.0)	13.2 (11.2-15.6)	13.4 (11.4-15.7)

Table 2.2 Endpoint events in the pre- and post-menopausal populations

2.4.2 Performance of CTS5 for events in years 5-10 in the post-menopausal cohort

CTS5 was prognostic for late distant recurrence (DR) in post-menopausal women with a HR of 2.18 (1.78-2.67), p<0.0001, LR-χ^2=53.54) for an increase in one CTS5 unit (Table 2.3). Prognostic value of CTS5(ATAC) in the BIG 1-98 validation study showed HR 2.07 (1.88-2.28), LR-χ^2=212.1. Using the predefined cut-off points, CTS5 split the cohort into three

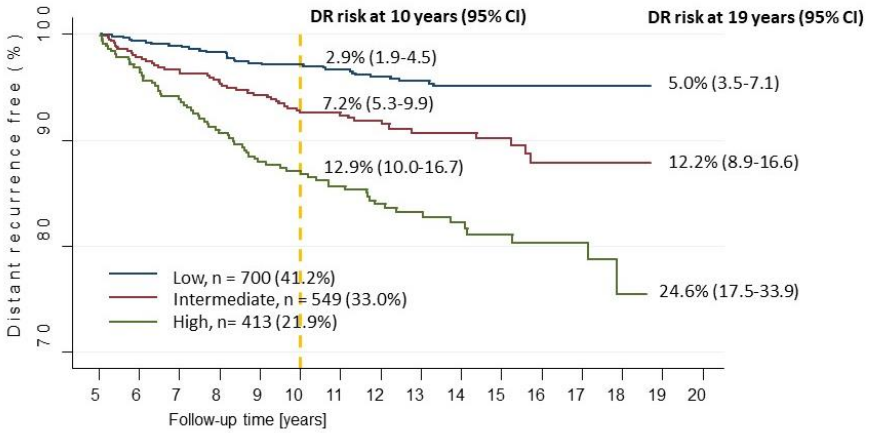
distinct risk groups (Figure 2.3a). HR for distant recurrence for intermediate and high-risk women was 2.28 (1.32-3.93), p=0.003 and 3.81 (2.27-6.41), p<0.0001 respectively, compared to those categorised into the low risk group (Table 2.3). This is in comparison to CTS5(ATAC) in the BIG 1-98 validation set where HR for DR between low and intermediate and low and high risk groups was 2.19 (1.61 – 2.98) and 5.33 (4.02 – 7.07) respectively.

The low risk category in this RMH cohort comprised 42% of the cohort whose risk of distant recurrence at 10 years was 2.9% (1.9-4.5) . This figure is comparable to that of the ATAC and BIG 1-98 datasets where 42% and 43% of the cohorts respectively fell into the low risk category with a DR risk of 2.5% and 3.6% respectively (Figure 2.4).

CTS5 remained prognostic regardless of chemotherapy treatment, with a HR of 2.08 (1.42-3.05), p<0.0001, LR-χ^2=14.12 for chemotherapy-treated and for untreated patients HR=2.42 (1.79-3.26), p=<0.0001, LR-χ^2=31.06).

Table 2.3 Prognostic performance of CTS for post- and pre-menopausal women for distant recurrence in years 5-10 and in years 5+. N numbers in brackets indicate number of women for that category in postmenopausal/premenopausal subgroups e.g. CTS5 continuous, no chemotherapy = 1194 post=menopausal women and 226 pre-menopausal women.

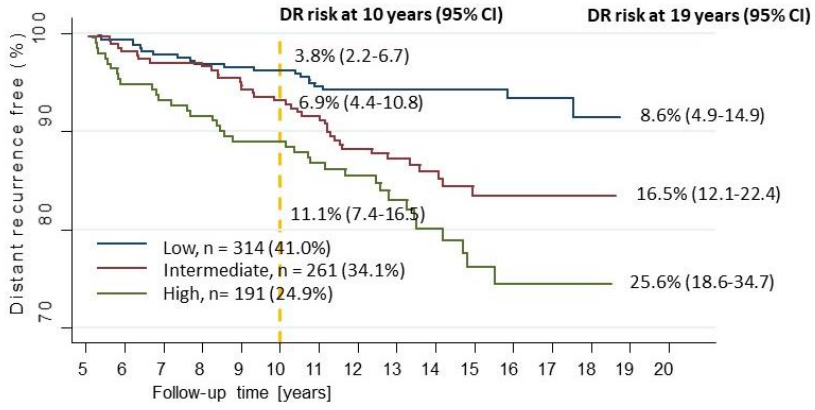
	Years 5-10				**Years 5+**			
	Postmenopausal women (N=1662)		**Premenopausal women (N=766)**		**Postmenopausal women (N=1662)**		**Premenopausal women (N=766)**	
	HR (95% CI)	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**
CTS5 continuous (n=1662)	2.18 (1.78-2.67)	<0.0001	1.84 (1.32-2.56)	<0.0001	2.08 (1.74-2.47)	<0.0001	1.89 (1.48-2.41)	<0.0001
CTS5 continuous No chemotherapy (n=1194/226)	2.42 (1.79-3.26)	<0.0001	1.65 (0.75-3.66)	1.42	2.27 (1.74-2.96)	<0.0001	2.42 (1.39-4.23)	0.002
CTS5 continuous Chemotherapy (n=468/540)	2.08 (2.03-3.2)	<0.0001	2.21 (1.42-3.41)	<0.0001	1.82 (1.33-2.48)	<0.0001	1.85 (1.35-2.55)	<0.0001
CTS5 risk groups								
Low (n=700/341)	Reference		Reference		Reference		Reference	
Intermediate (n=549/261)	2.51 (1.46-4.33)	0.001	1.82 (0.88-3.78)	0.11	2.25 (1.43-3.52)	<0.0001	2.34 (1.36-4.04)	0.002
High (n=413/191)	4.67 (2.78-7.85)	<0.0001	3.04 (1.50-6.19)	0.002	4.34 (2.84-6.64)	<0.0001	3.49 (2.03-6.02)	<0.0001
CTS5 continuous </= 60 months ET (n=1405/572)	2.55 (2.03-3.20)	<0.0001	1.83 (1.22-2.74)	0.003	2.29 (1.88-2.79)	<0.0001	1.81 (1.35-2.41)	<0.0001



Number at risk

Low	700	688	676	661	647	644	602	508	413	328	258	181	111	43
Intermediate	549	513	498	480	457	436	396	323	259	202	148	91	55	21
High	413	390	363	339	319	304	284	228	185	155	119	80	52	21

a.



Number at risk

Low	314	312	306	303	302	300	285	239	205	170	128	99	63	29
Inter	261	256	252	251	245	241	227	189	151	117	85	61	41	19
High	191	179	176	173	166	166	151	125	93	74	52	30	18	11

Low Intermediate High

b

Figure 2.3 Kaplan Meier curve of all DR events for a. postmenopausal women, b. premenopausal women (vertical yellow dashed lines to demarcate 5-10 year cut-off). Risk of DR at two different time points is given for each CTS5 risk group with corresponding 95% CI. Numbers in each CTS5 risk category (low, intermediate and high) are given with percentage of whole pre or postmenopausal sub-group.

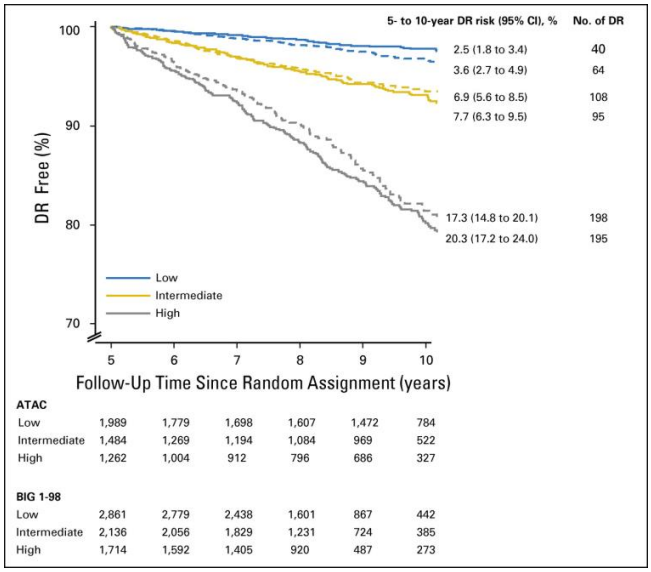


Figure 2.4 Credit Dowsett et al 2018 [58] - DR events for CTS5 risk groups in the ATAC (solid lines) and BIG 1-98 (dotted line) datasets.

When split into quintiles by increasing score, CTS5 calibrated well as the observed versus expected distant recurrence events were similar in all except the top quintile where CTS5 overestimated the number of expected events (χ^2= 7.09) (Table 2.4 and Figure 2.5a). The expected mean risk in the top quintile was 19.4% A similar picture was seen when dividing the cohort by CTS5 risk group. The χ^2 for difference between observed and expected events was statistically significant only in the high-risk group (6.8) and the overall observed versus expected ratio was 0.82 (0.67-0.99). p=0.041 (Figure 2.5b). The c-statistic for area under the ROC curve was 0.67 in comparison to that of the ATAC/BIG 1-98 dataset where the c-statistic was 0.74 (Figure 2.6).

	Cohort by quintile						Cohort by risk group				
Post-menopausal by quintile (all)	**Quintile**	**Observed**	**Expected**	**Chi2**	**N**	**Pre-menopausal by quintile (all)**	**Quintile**	**Observed**	**Expected**	**Chi2**	**N**
	1	6	6.9	0.12	333		1	5	3.1	1.16	157
	2	10	12.5	0.49	333		2	7	5.8	0.25	150
	3	27	18.5	3.91	332		3	8	9.21	0.16	153
	4	21	28.1	1.80	332		4	11	14.01	0.65	153
	5	43	64.4	7.09	332		5	20	28.18	2.37	153
	TOTAL	**107**	**130.36**	**4.19**			**TOTAL**	**51**	**60.3**		
Post-menopausal by risk group (all)	Low	20	20.9	0.04	700	**Pre-menopausal by risk group (all)**	Low	12	9.3	0.82	314
	Intermediate	37	37.4	0.004	549		Intermediate	18	18.5	0.01	261
	High	50	72.1	6.75	413		High	21	32.6	4.11	191
Post-menopausal by quintile (</=60 months ET)	1	4	5.6	0.44	281	**Pre-menopausal by quintile (</=60 months ET)**	1	4	2.1	1.8	115
	2	10	9.4	0.03	281		2	4	3.9	0.002	115
	3	21	14.1	3.33	282		3	7	6	0.17	114
	4	20	21.5	0.11	280		4	8	8.7	0.06	114
	5	39	42.7	0.31	281		5	13	17.5	1.14	114
	TOTAL	**94**	**93.31**	**0.01**			**TOTAL**	**36**	**38.2**	**0.13**	
Post-menopausal by risk group (</=60 months ET)	Low	18	19.2	0.08	652	**Pre-menopausal by risk group (</=60 months ET)**	Low	11	8	1.13	275
	Intermediate	37	31.4	1.00	471		Intermediate	12	13.4	0.14	189
	High	39	42.8	0.34	282		High	13	16.8	0.86	108

Table 2.4 Calibration of observed versus expected events for post- and pre-menopausal entire cohorts and post- and pre-menopausal cohorts of women who received </= 60 month endocrine therapy

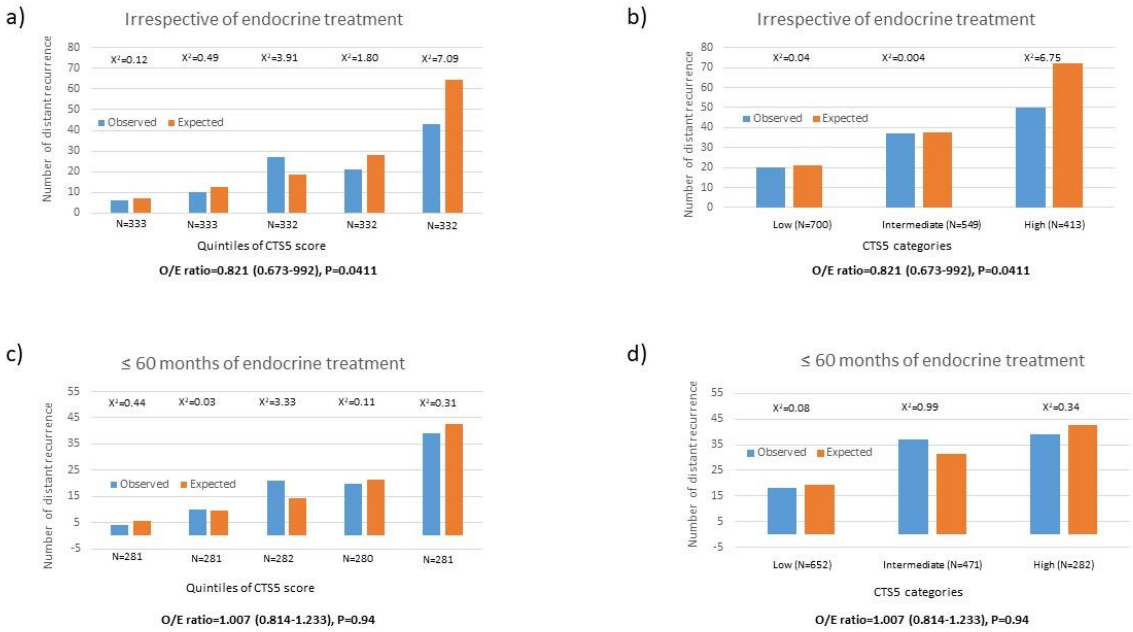


Figure 2.5 Histograms for observed versus expected events in postmenopausal women according to a) quintiles and irrespective length of endocrine therapy use, b) risk groups and irrespective length of endocrine therapy use, c) quintiles and 60 months or less of endocrine therapy use, and d) risk groups and 60 months or less of endocrine therapy use.

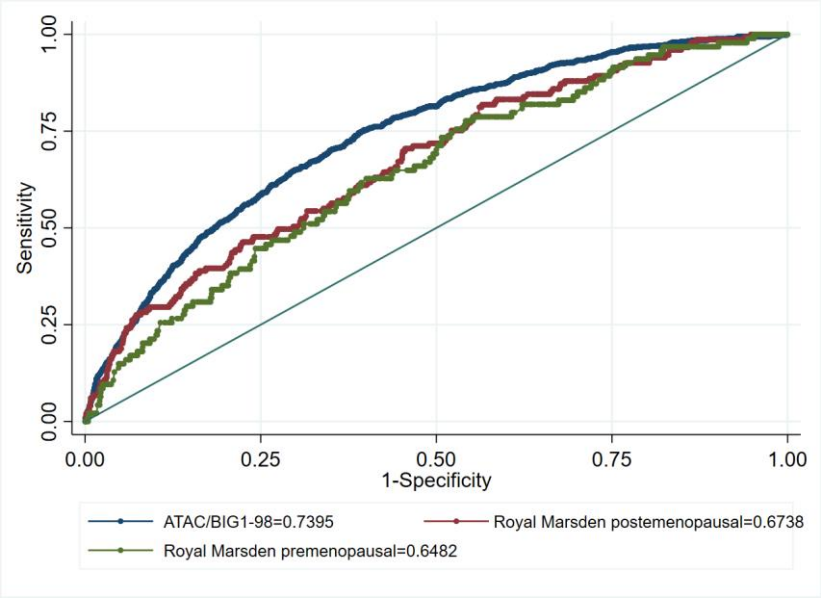


Figure 2.6 ROC curve for pre- and post-menopausal women and for ATAC/BIG1-98 datasets

2.4.3 Performance of CTS5 for events in years 5-10 in the pre-menopausal cohort

CTS5 was prognostic for late distant recurrence in pre-menopausal women with a HR for distant recurrence of 1.84 (1.32-2.56), p<0.0001, LR-χ^2=13.25 for a unit increase in CTS5 value (Table 2.3). For women treated with chemotherapy (n=540), CTS5 was significantly prognostic: HR= 2.21 (1.42-3.41), p<0.0001, LR-χ^2=12.50. However, for those who had not received chemotherapy (n=226), CTS5 score was not associated with the prediction of late distant recurrence, HR=1.65 (0.75-3.66), p=1.42.

When divided into the pre-defined risk groups, HR for distant recurrence for the high-risk group was 3.04 (1.50-6.19), p=0.002) compared to those categorised as low risk (Table 2.3). The low-risk group had a median risk of distant recurrence at 10 years of 3.8% (2.2-6.7) however, no significant separation of low and intermediate risk curves was observed (HR= 1.82 (0.88-3.78), p=0.11) and considerable overlap was seen in confidence intervals for HRs for distant recurrence (Table 2.3 and Figure 2.3b). Calibration of CTS5 in the pre-menopausal cohort showed that observed versus expected events for all quintiles was similar although a statistically significant difference was seen for the CTS5 high risk group (χ^2=4.1) (Table 2.4 and Figure 2.7a,b). The expected mean risk in the top quintile was 18.4%. The overall O/E ratio was 0.85 (95%CI 0.63-1.11), p=0.23. Discrimination by ROC curve was similarto the post-menopausal cohort with a c-index of 0.65 (Figure 2.6).

2.4.4 Prognostic value of CTS5 for events in years 5 and beyond in the post-menopausal cohort

When analysing all distant recurrence events post 5 years, continuous CTS5 had HR 2.08 (1.74-2.47), p<0.0001; LR-χ^2=64.30 for the postmenopausal cohort (Table 2.3). When divided into risk groups, significant risk stratification was observed with a HR for distant recurrence 4.34 (2.84-6.64), p<0.0001) and 2.25 (1.43-3.52) p<0.0001) for the high and intermediate groups respectively when compared to the low risk group (Table 2.3). Women in the low risk group had a risk of distant recurrence at 19 years of 5% (3.5-7.1) (Figure 2.3a).

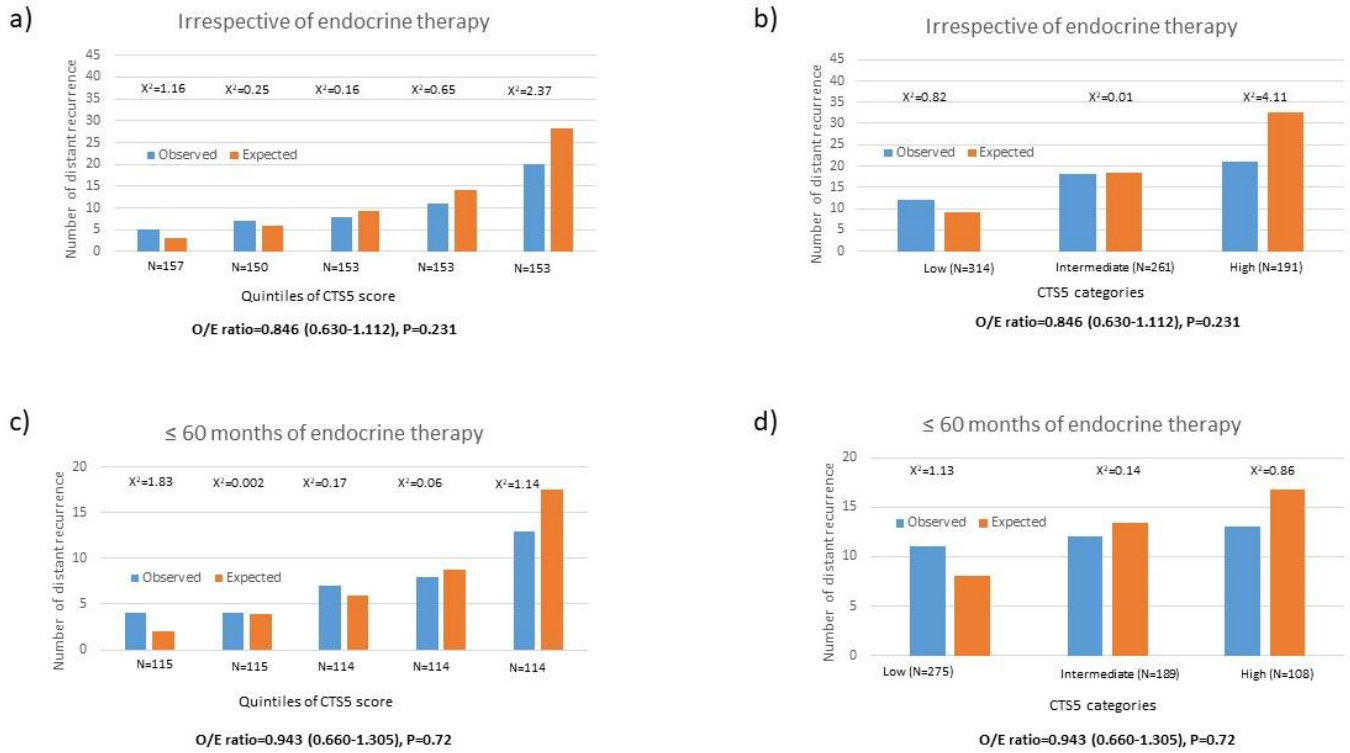


Figure 2.7 Histograms for observed versus expected events in pre-menopausal women according to a) quintiles and irrespective length of endocrine therapy use, b) risk groups and irrespective length of endocrine therapy use, c) quintiles and 60 months or less of endocrine therapy use, and d) risk groups and 60 months or less of endocrine therapy use.

2.4.5 Prognostic value of CTS5 for events in years 5 and beyond in the pre-menopausal cohort

For pre-menopausal women, continuous CTS5 had a HR for distant recurrence of 1.89 (1.48-2.41), p<0.0001; LR-χ^2=25.22 (Table 2.3), this effect was seen regardless of chemotherapy use. When divided into risk categories, HR was 3.49 (2.03-6.02), p<0.0001 and 2.34 (1.36-4.04), p=0.002 for the high and intermediate risk groups respectively when compared to the low risk group (Table 2.3). Women who were categorised into the low-risk group had a distant recurrence risk at 19 years of 8.6% (4.9-14.9) (Figure 2.3b).

2.4.6 Exploratory sub-group analysis in women who received no more than 60 months endocrine therapy for events in years 5-10.

During data collection, it became apparent that an appreciable proportion of women took endocrine therapy beyond 5 years. To exclude these women would have biased the dataset towards a lower stage population and therefore subgroup analyses were performed based on duration of endocrine therapy.

1405 post-menopausal women received up to 60 months of endocrine therapy and these women had a HR for distant recurrence of 2.55 (2.03-3.20), p<0.0001) for continuous increase in CTS5 value (Table 2.3). When divided into CTS5 risk categories, time to event curves were distinctly separate with significant heterogeneity between intermediate and low (HR 3.08 (1.76-5.42), p<0.0001) and high and low (HR 5.78 (3.31-10.11), p<0.0001). Calibration of observed versus expected events for this exploratory sub-group was similar to the primary analysis but was improved for the highest quintile (Table 2.4 and Figure 2.5c,d). Numbers of observed versus expected events were not significantly different for any of the quintiles and there was an overall observed versus expected ratio of 1.01 (0.81-1.23), p=0.94. In the 257 women who had extended endocrine therapy, there were significantly fewer overall observed (13) versus expected (37.1) events (χ^2 =15.7), overall observed versus expected ratio=0.35 (0.19-0.60),p<0.001.

572 pre-menopausal women received no more than 60 months of endocrine therapy. These women had a HR for distant recurrence of 1.83 (1.22-2.74), p=0.003) for continuous increase in CTS5. When divided into CTS5 risk categories, time to event curves were similar to the primary analysis and showed the risk groups in the pre-menopausal population were less distinct than for post-menopausal women. HR for DR for was 3.25 (1.46-7.25), p=0.004) for high versus low groups and 1.59 (0.70-3.61), p=0.26 for the intermediate versus low group. CTS5 showed similar calibration to the primary analysis but was improved for the top quintile (Table 2.4 and Figure 2.7c,d). Numbers of observed versus expected events were not significantly different in any of the quintiles with an overall χ^2 of 0.28, observed versus expected ratio = 0.94 (0.66-1.30), p=0.72. In the 194 women who received more than 60 months of endocrine therapy, there were fewer observed (15) versus expected

(22) events but this was not statistically significant - (χ^2 =2.29), overall observed versus expected ratio=0.68 (0.38-1.12),p=0.13.

2.5 Discussion

2.5.1 Comparison with development set

The populations represented in this study included both post- and pre-menopausal women. The baseline tumour characteristics of these women showed trends of higher stage disease compared to the ATAC and BIG 1-98 patients. This RMH cohort had a greater proportion of node positive cancers, grade 3 cancers and patients who received chemotherapy in comparison to the ATAC dataset and also, but to a lesser degree, to the BIG 1-98 dataset. The proportion of chemotherapy treated women in ATAC and BIG 1-98 was 20% and 24% whereas in the RMH cohort this figure was 41.5%. The latter figure was much higher in pre-menopausal women 70% of whom received chemotherapy. Taking into account just post-menopausal women the percentage of patients treated with chemotherapy in the RMH cohort was 28%.

15% of post-menopausal women and 25% of pre-menopausal women received more than 5 years of endocrine therapy, whereas in the ATAC and BIG 1-98 studies, this figure is unknown but is anticipated to be very low. DR event rate beyond 5 years was 7% in ATAC and 5.5% in BIG 1-98 and in this study the rate was substantially higher at 10%. Although data on co-morbidities and performance status were not collected as part of this study, it is reasonable to assume that there are patients within the RMH cohort with comorbidities or low performance status that would have excluded them from those studies had they been a patient at that time. Altogether, this cohort is relatively similar but with important differences to the combined development set and provides a suitable cohort in which to test validity of CTS5.

2.5.2 Performance of CTS5 in post-menopausal women

The results shown above demonstrate that the CTS5 was significantly prognostic for late distant recurrence for the entire population. For the post-menopausal cohort, this was regardless of chemotherapy treatment, and the HRs were similar for events in years 5-10 and for years 5 onwards. DR risk estimates were less precise at 19 years than at 10 owing

to the lower numbers of women at risk making risk estimation at this point more uncertain. Calibration of observed versus expected events was only significantly different for the top quintile/CTS5 high risk group where CTS5 over-estimated the risk of recurrence. Translated into current clinical practice, this may lead to more high scoring women being prescribed extended endocrine therapy than needed owing to a perceived greater risk, however, as these women's observed risk was still high, they would probably warrant extended endocrine therapy regardless.

The ROC curve demonstrates reasonable discriminative ability of the score for pre- and post- menopausal women which although lower than that of the development set, is within an acceptable range and is on par with other prognostic models in breast cancer[115,116]. CTS5 stratified postmenopausal women into three distinctive risk groups, with significantly different risk of recurrence in the low, intermediate and high-risk categories. The low-risk group identified here had a very low distant recurrence risk and comprised 41% of the cohort, which translates to a substantial number of women who may be deemed able to avoid extended endocrine therapy. The time to event curve separation closely mirrors that of the development set[58] (Figure 2.8). The high risk group in ATAC had a poorer outcome than the same group in this study which may be, in part caused by the proportion of women who received prolonged endocrine therapy in this RMH study. It should be noted that in the development dataset, the original CTS5(ATAC) algorithm was used whereas in this study the published algorithm based on the combined ATAC and BIG1-98 datasets was tested. However, the changes to the co-efficients in the new algorithm were minor and the prognostic value of the two models are very similar with overlapping confidence intervals and so the comparison can be considered reliable (Figure 2.9).

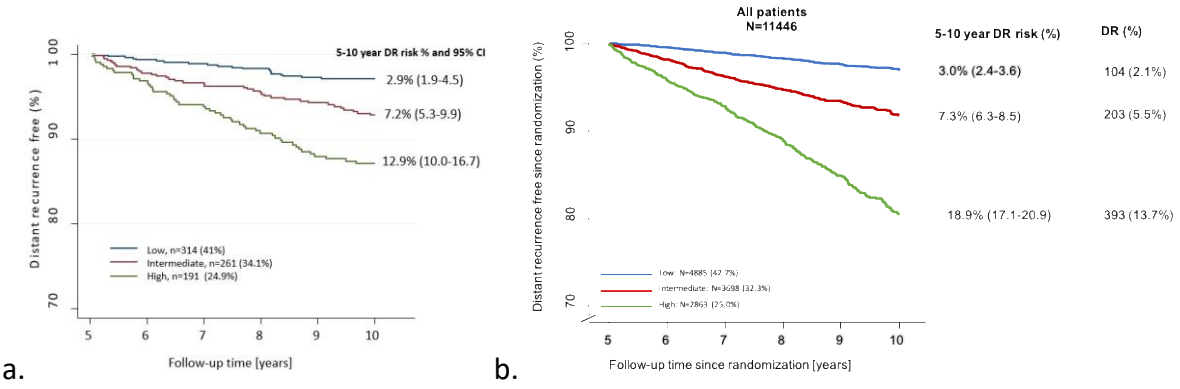


Figure 2.8 Time to event curves for low, intermediate and high risk CTS5 groups in the a. RMH dataset and b. combined ATAC and BIG1-98 datasets.

Algorithm	HR for late recurrence	p. value	95% CI
ATAC CTS5	2.940465	0.0001	2.61-3.31
ATAC/BIG1-98 CTS5	2.701031	0.0001	2.49-2.93

Figure 2.9 Comparison of HR for late recurrence in the development dataset of the original ATAC CTS5 and the combined CTS% algorithm from the ATAC and BIG 1-98 datasets.

2.5.3 Performance of CTS5 in pre-menopausal women

Prior to this study, CTS5 has only been tested in one cohort of premenopausal women within the TAILORX study[117]. The TAILORX population was of a broadly low risk and has been reported as showing that CTS5 performed poorly for women under 50 in comparison to those over 50[118]. It was anticipated that since the CTS5 was developed in a purely postmenopausal cohort, its performance for pre-menopausal women would be somewhat weaker given the inherent biological differences between pre- and post-menopausal breast cancer and also given the effect of age (discussed below).

For the pre-menopausal cohort, CTS5 was significantly prognostic for distant recurrence in years 5-10 and years 5 onwards. As with post-menopausal women, estimates of DR risk at 19 years become less certain with lower numbers of women at risk. When divided into chemotherapy and no chemotherapy subgroups, CTS5 remained prognostic only in the

chemotherapy group however this may be due to low numbers of pre-menopausal women who did not receive chemotherapy. Risk group stratification was not as distinctive as in the post-menopausal cohort. This is reflected in the overlap of the intermediate and low risk groups event curves for years 5-10 and the wide confidence intervals associated with the HRs for recurrence.

Interestingly there is an almost doubling of events for this cohort when the analysis was extended beyond 10 years. Incorporation of these later data resulted in an improvement in separation of the risk curves and significant different in HRs for recurrence between the three risk groups. It is therefore possible that the overlap seen in low and intermediate risk groups for events in years 5-10 only, may be caused by lower event numbers rather than failings of the score itself. This is something that would require confirmation in a study with a larger proportion of pre-menopausal women. However, owing to the improved distinction between risk groups that becomes evident with longer follow-up, clinicians can have some confidence in the ability of CTS5 to risk stratify pre-menopausal women. Altogether, CTS5 demonstrates prognostic ability in pre-menopausal women although, elements of validity as defined previously and usefulness in guiding management is lacking in comparison to post-menopausal women.

2.5.4 Effect of extended endocrine therapy

When considering patients who took a maximum of 60 months endocrine therapy, we see that calibration is improved for both pre- and post-menopausal women indicating the superior precision of CTS5 for women who have had only 5 years of endocrine therapy – which is the point at which it is practically applicable. In this sub-group analysis, a discrepancy is only seen in the calibration of events for women who took more than 60 months of endocrine therapy where there were significantly fewer observed than expected events. This implies that the difference in observed versus expected events in the top quintile and high-risk group for all post-menopausal women is related to the presence of women who took extended endocrine therapy (34% and 27% respectively) and therefore had fewer recurrence events. This would suggest that there is a recurrence benefit in

continuing endocrine therapy in women with high CTS5 scores. This hypothesis is important and merits testing.

2.5.5 Age and late recurrence

As discussed in chapter 1, the effect of age on breast cancer recurrence risk from time of diagnosis in all age women is not linear. Instead, women at the extremes of age are at greatest risk as seen in the 2017 EBCTCG analysis[25]. CTS5 was developed on purely post-menopausal women, few of whom would have been younger than 50. Having removed these younger women, the effect of age would become more linear which accounts for the behaviour of this variable within the model. Had CTS5 been developed on a dataset that included younger women, the coefficient for age may well have been different or possibly age may not have been a statistically significant prognostic variable. One possible reason for the better than expected performance of CTS5 for pre-menopausal women is that younger patients who have a very poor prognosis owing to aggressive disease, would be more likely to recur in the first 5 years after diagnosis. These women would not be included in any dataset for late recurrence model building and their exclusion might create a more linear relationship between age and late recurrence.

2.5.6 Strengths and limitations of the study

Strengths of this study are that it is comprised of a large cohort selected purely by virtue of time and location of their treatment, none of whom were excluded based on performance status and therefore represent a 'real-life' cohort. Additionally, median documented follow up was over 13 years and all women were at least 10 years from diagnosis demonstrating the reliability of this study in drawing conclusions about late recurrence.

Limitations of the study are that it is a retrospective analysis and therefore relies on adequate documentation which can be inconsistent. However, the EPR in this centre is robust and documentation was generally complete and of a high standard. Being a retrospective study, patients within it were treated over ten years ago and these

treatments will differ from modern practices. The analysis in this study made no adjustment for this fact which raises an important issue of whether adjustment for older data should be made.

Additionally, there was a proportion of patients with less than 10 years of documented follow up. The assumption was made that if they had no documented DR or death and had not presented back to this centre they would be recorded as being alive and distant recurrence free and not to censor them at time of last follow up. Consequently, there is a possibility that events have been missed in patients who were lost to follow up. We estimate that this number is low given that most women who experience recurrence would return to their original treating hospital for management, particularly when that hospital is a specialist centre. However, this is not always the case and some women would have moved away or chosen to go elsewhere. Missed events due to this reason could partially account for the discrepancy in observed versus expected events in the top quintile/high risk group for the overall post-menopausal population.

This study included women with local or loco-regional recurrence in years 0-5 as the aim of the study was to capture a population as close to 'real-life' as possible. Approximately 2% of luminal A breast cancers treated surgically will recur locally[119] which, will amount to a substantial number of women. To exclude these cases, therefore, would be to diminish the power of the study and applicability of the CTS5. However, it must be acknowldeged that local recurrences of ER+ breast cancer will usually be treated with further systemic therapy which will have a bearing on late recurrence risk. This treatment effect may have caused a lower than expected number of late recurrences than if these patients had been excluded. Equally, however, women who experienced local recurrence may have had a more aggressive biology making late recurrence more likely. Sub-group analysis of women who suffered local recurrence would have been informative and should be looked at in future studies of this type.

This study does not give any information regarding the value of CTS5 in predicting benefit from extended endocrine therapy. Applying the score to a cohort that has been

randomised to continue endocrine therapy to ten years versus stopping at five should therefore be conducted to generate evidence for the clinical utility of CTS5.

2.5.7 Further work

To further strengthen the evidence for validity of CTS5 it should be tested in women with higher stage disease, a larger premenopausal cohort, and in women with HER2+ disease as this study contained too few women with HER2+ cancers for any meaningful analysis. If, following testing in a larger pre-menopausal cohort, the model lacks validity in this population, re-calibration of the model for pre-menopausal women should be considered. In addition, it would be beneficial to investigate the predictive ability of CTS5 for response to extended endocrine therapy. The most efficient way of testing this would be to apply it to patients within a randomised trial of extended endocrine therapy. To this end, the trial management group of the MA17 study were approached with a view to conducting a validation study in that dataset however this was deemed not feasible due to the lack of grade data collected from patients in this study.

2.5.8 Clinical application

The evidence presented here strongly supports CTS5 as an accurate prognostic tool for predicting late distant recurrence in post-menopausal women in the absence of extending endocrine therapy. CTS5 represents a cost-free, accessible tool that could be implemented immediately. Whilst several commercially available genomic models have shown prognostic value for late DR[81,85,86,120], none of these are currently approved or funded for use in the UK for this purpose. Additionally, molecular assays, whilst adding more precision to risk estimation, carry a high cost and burden of resources which will preclude their use in many centres both in the UK and worldwide.

This study has tested validity of the re-calibrated CTS5 based on the combined ATAC and BIG 1-98 datasets. The data supports the recommendations made in the original literature that post-menopausal women who fall into the low-risk group can safely stop endocrine therapy at 5 years owing to the low chance of benefit from continuing[58]. Women who fall

into the CTS5 high category are at sufficient risk to merit continuing endocrine therapy. For those that fall into the intermediate category the management pathway is less clear and it may be the case that genomic models could further risk stratify these women. Similarly, for pre-menopausal women, those in the low-risk group have a small potential benefit from continuing endocrine therapy. However, firm conclusions regarding the clinical validity of the risk categories in this group cannot be made based on this data. This means that validation in a larger cohort would be ideal to explore this further before firm treatment recommendations can be made. This is particularly important for pre-menopausal women, many of whom would have received upfront tamoxifen and therefore stand to gain more benefit from extended endocrine therapy beyond 5 years[29,30].

The ability of prognostic models to categorise patients into risk groups is of clinical use and helps to standardise recommendations within and between centres. Specifically, for the CTS5 whose predictive value with regards to extended endocrine therapy benefit is unknown, the low-risk group is of particular interest. It must be remembered however, that cut-off values for CTS5 risk groups are arbitrary and, for all women, one must consider the absolute risk quoted by CTS5 as well as the risk category that is assigned.

Here, it is shown that CTS5 extrapolates to a clinically unselected, non-trial population, including pre-menopausal women and, in post-menopausal women, is applicable regardless of chemotherapy treatment. On the basis of this validation and the existing data, CTS5 can and should be employed for post-menopausal women treated for ER+, HER2- breast cancer who are distant recurrence free at 5 years for risk estimation of late recurrence and also as an aid for making extended endocrine therapy recommendations. Regardless of CTS5 outcome, decisions regarding extended endocrine therapy must consider toxicity profile and individual patient preference.

3. Clinical Utility of CTS5

3.1 Introduction

For a validated prognostic tool to be accepted for use in the UK and elsewhere, as well as demonstrating clinical validity, it must demonstrate clinical utility. This is the concept that use of a tool results in a change in a management decision and benefits patients. NICE (National Institute of Care and Health Excellence) take clinical utility strongly into consideration when approving prognostic tests[121]. Often this involves an assessment of the cost effectiveness of the tool and whether it adds sufficient additional information to warrant the expense. CTS5 is a cost neutral tool and as such does not have to demonstrate cost effectiveness. Nevertheless, it is still important to determine whether it has a positive impact on patient outcomes. In the case of CTS5 it would be desirable to see that using the tool results in a change in management that ultimately reduces late recurrences.

The intended use of CTS5 is to guide decision making regarding extended endocrine therapy at the 5-year mark after diagnosis and initial treatment. At this stage, women will be coming to the end of their 5 years of endocrine therapy and some may be candidates for extending that treatment to ten years. Utility of the score is therefore evaluable at this time point to assess the impact it would have on recommendations for extending endocrine therapy.

When evaluating a risk prediction model that is utilised at the 5-year mark, one must take into account the context of follow up and the forum in which this takes place. Most women, at this stage are part of a nurse-led open access follow up system without regular clinician consultations. In some, but not all centres, patients reaching the 5-year mark are discussed in a multi-disciplinary team (MDT) meeting but not in all and indeed the forum in which CTS5 would be used will vary. This means that clinical utility will be situation dependent and may be clear in some settings but not in others. For this reason, ideally, it should be tested in several different healthcare settings.

The most robust method of testing clinical utility is through a prospective study of extended endocrine therapy versus placebo in which patients enrolled are stratified by CTS5 score and randomised to either the control or experimental arms of the study. In this way, one can directly evaluate the benefit of extended therapy for different CTS5 scores or categories. A study of this kind however, would be expensive and would not yield results for some years.

Another way of testing utility is to assess the impact of a prognostic tool on current practice in a before and after type study design known as a decision impact study. This compares treatment decisions or recommendations before and after application of a prognostic tool and one can then evaluate its impact on those decisions. This is a quick and simple method of testing clinical utility and whilst it would not provide evidence of improved patient outcomes, it would give indications about the likely impact of a prognostic tool on standard practice.

Decision impact studies have been employed for assessing clinical utility of the Oncotype DX (Recurrence Score)[122-124], PAM50 ROR[125] and also the EPClin[126] and in all situations led to an overall reduction in adjuvant chemotherapy recommendations. The relatively large evidence base for clinical utility of Oncotype in comparison to other molecular models led to it being the first NICE-approved genomic model for use in guiding adjuvant chemotherapy decisions. No testing of clinical utility of CTS5 has been performed thus far.

3.2 Aim

To test, using a decision impact study, how use of CTS5 changes practise with regards to extended endocrine therapy recommendations in women who have reached 5 years on endocrine therapy without distant recurrence.

3.3 Methods

3.3.1 Patient Population

A protocol for a decision impact study was approved by the RMH CCR as a service evaluation. In RMH, women coming to the end of 5 years of follow up for early breast cancer, who are DR free are discussed in an MDT meeting. ER+ women within this group who are on endocrine therapy are discussed with a view to deciding whether to recommend continuing treatment to 10 years or to stop at 5. In May 2018, CTS5 was adopted at RMH to be used in this MDT meeting to aid these decisions for post-menopausal women with ER+, HER2- breast cancer who are taking endocrine therapy. Prior to implementation of this change, no standardised tool was being employed for this purpose and decisions were made on the basis of clinicians' consideration of individual biomarkers known to be important for late recurrence such as lymph node involvement[26,54,55] and tumour size[55].

Eligible patients were women who were post-menopausal at the time of diagnosis, who underwent surgery with curative intent for ER+ early breast cancer and who were on endocrine therapy. Exclusions were made for women with HER2+ cancer, women on whom input data for CTS5 was lacking and for women who had prematurely ceased (or never started) endocrine therapy e.g. for reasons of poor tolerance.

3.3.2 Procedures

Prior to the MDT meeting a patient list was drawn up by a senior nurse. This list was then reviewed by me (the clinician researcher) who scrutinised the hospital EPR to assess for eligibility. The following data were collected for all eligible patients: age at diagnosis, pathological tumour size, pathological tumour grade, lymph node burden and adjuvant systemic therapy (chemotherapy/endocrine therapy). Patients treated with neo-adjuvant chemotherapy or endocrine therapy were included and the greater of their neoadjuvant and excision tumour indices were input into the CTS5 model. CTS5 was calculated using the published algorithm below along with their absolute late recurrence risk in years 5-10 and

their risk categorised according the algorithm below, published in the original literature[58]. Additionally, each patient was assigned a risk category according to their CTS5 value, where CTS5 <3.13 is low risk, 3.13-3.86 is intermediate risk and >3.86 is high risk.

CTS5 = 0.438 xnodes + 0.988 x (0.093 x size - 0.001 x size^2 + 0.375 x grade + 0.017 xage)

For the purposes of the MDT meeting the calculation was made on the publicly available website – www.cts5-calculator.com and the risk curve produced on the website was shown (Figure 3.1).

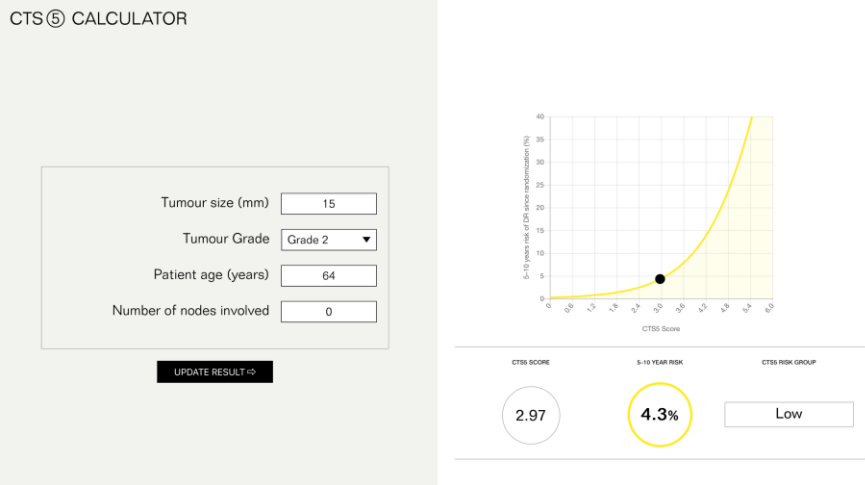


Figure 3.1 Example CTS5 data input and risk curve for individual patient

In the MDT meeting, clinicians were asked to make two recommendations for each patient, Wherever possible, recommendations had to be binary (stop or continue). The first was to be made prior to being told the CTS5 result and the second was to be made with knowledge of the CTS5 result along with the risk category and a visual representation of the patient's individual risk (Figure 3.1). The second recommendation was then communicated to the patient by letter as per standard practice.

As discussed in chapter 1, CTS5 is a prognostic and not predictive tool and as such gives no information about likely response to extended endocrine therapy based on the score value. As such, the current intended use of CTS5 is in identifying those who can be spared extended endocrine therapy by virtue of their low risk of recurrence. It was, however,

implied in the validation study in chapter 2 within the subgroup analysis where CTS5 high risk women who received more than 60 months of endocrine therapy had fewer events than those who had more. Therefore, extended endocrine therapy should be strongly considered for these women. With this in mind, the assumption was made that, for the purpose of this decision impact study, anyone who was CTS5 low-risk should be recommended (on the 2nd recommendation) to stop and women who are CTS5 high-risk should be recommended to continue – this assumption is referred to below as the "CTS5 recommendation". For those in the intermediate category no uniform recommendation can be made based on the current evidence and the management approach for these women will depend more on their specific risk, tolerance of endocrine therapy thus far and personal preferences.

3.3.3 Statistical considerations

3.3.3.1 Sample size

It was estimated that approximately 20% of patients would have a de-escalation of their treatment recommendation and 2% would have an escalation of treatment recommendation. The sample size was based on an estimate of 5% null proportion (p0) of recommendation change and 15% alternative proportion (p1) of recommendation change after seeing the CTS5 results. Based on the above figures using A'Hern's single stage method with 5% alpha level and 90% power it was decided that a total of 77 patients would be needed and an expected minimum of 8 patients with a recommendation change for CTS5 to be considered to have established clinical utility. It was decided half way through the study, owing to the ease of data collection to extend the study by a further 3 months to improve statistical power.

3.3.3.2 Endpoints

The primary endpoint was change in recommendation after knowing CTS5. Secondary endpoints were: de-escalation of recommendation after knowing CTS5, concordance between clinician recommendation 2 and CTS5 recommendation for low and high-risk women. Further descriptive secondary endpoints were: reasons for discordance between

MDT and CTS5 recommendation for low and high-risk women, and patterns of recommendations for intermediate risk women.

3.3.4 Statistical Analysis

The proportion of patients whose treatment recommendation changed (both escalation and de-escalation) as a result of CTS5 were calculated. Proportions of concordant and discordant second MDT and CTS5 recommendations were calculated and Cohen's kappa statistic was used to assess degree of concordance. Unpaired t-test was used to evaluate differences in absolute risk of late recurrence for women recommended to continue versus those recommended to stop within the CTS5 intermediate category. Fishers exact test was used to assess association between endocrine therapy agent used and recommendation 2.

3.4 Results

3.4.1 Patient Population

136 eligible patients were discussed in the MDT meeting over a period of 8 months for consideration of extended endocrine therapy. Table 3.1 shows the baseline characteristics of the cohort and how it split into the CTS5 risk categories. The median age of women in this study was 65 years with a median tumour size of 22.9mm. 58% of tumours were grade 2 and 74% were node negative. 21% had received either adjuvant or neo-adjuvant chemotherapy. An AI was used as endocrine therapy in 85% of women. Median CTS5 value was 3.38 which equates to a risk of distant recurrence of 6.3% over 5 years.

3.4.2 MDT first recommendations

Of all first recommendations, 84 (62%) were made to stop and 50 (37%) to continue. For 2 patients, no agreement was reached. In the CTS5 low-risk group, 96% of recommendations were to stop and in the CTS5 high risk group, 88% of recommendations were to continue. In the intermediate group, 60% of recommendations were to continue and 40% to stop (Figure 3.2a).

3.4.3 Changes in recommendation

For the primary endpoint, 2 (1.5%) patients' recommendations changed after clinicians were told the CTS5 result. Both changes were for escalation from stop to continue in intermediate-risk patients (Figure 3.2b). Owing to the very low number of changes, the planned statistical analysis for these figures were not performed.

Baseline Characteristic	**Total Cohort 136**
Median Age – years (range)	65 (45-89)
Median Tumour Size - mm (IQR)	22.9 (11.75-30)
Grade 1 – n (%)	36 (26.5)

Grade 2 – n (%) Grade 3 – n (%)	79 (58.1) 21 (15.4)
Node negative – n (%) 1-3 nodes positive – n (%) >3 nodes positive – n (%)	101 (74.3) 27 (19.9) 8 (5.9)
Chemotherapy treated - n (%)	28 (21)
Endocrine therapy agent – n (%) AI Tamoxifen Unknown	 115 (85) 20 (15) 1 (0.7)
Median CTS5 Score	3.38 (1.61-5.56)
CTS5 Low – n (%) CTS5 Intermediate – n (%) CTS5 High – n (%)	55 (40.4) 47 (34.5) 34 (25)
Median 5-10 year distant recurrence risk (%) CTS5 Low CTS5 Intermediate CTS5 High	 3 7 19

Table 3.1 Baseline characteristics and CTS5 values of study cohort

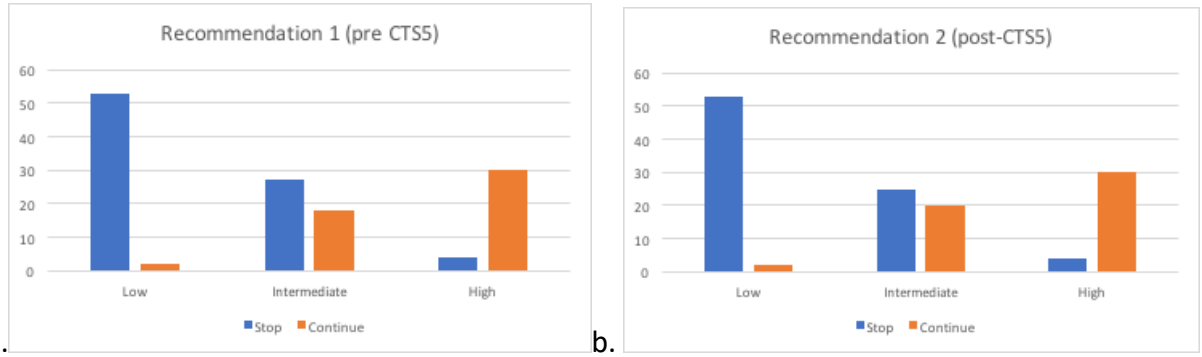


Figure 3.2 a. recommendation 1 by CTS5 risk category, b. recommendation 2 by CTS5 risk category

3.4.4 Concordance between CTS5 and MDT recommendation

Amongst patients within the CTS5 low-risk group, 53 (96%) MDT and CTS5 recommendations were concordant with a Cohen's κ statistic of 0.92. Similarly, in the CTS5

high risk group, 30 (88%) MDT and CTS5 recommendations were concordant with a Cohen's κ statistic of 0.76. (Figure 3.3).

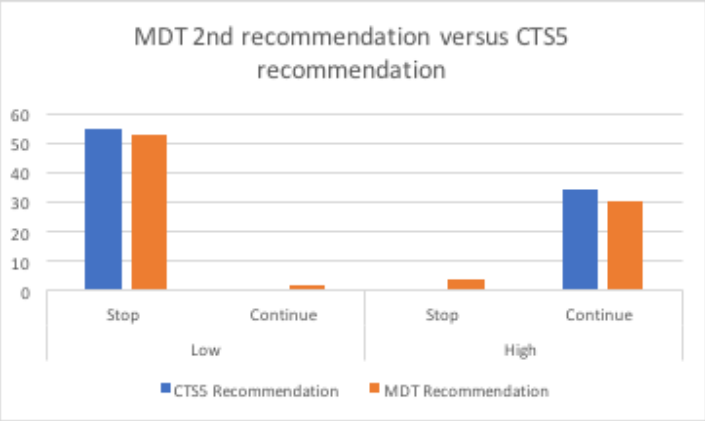


Figure 3.3 Concordance between clinician recommendation and CTS5 recommendation for women who were CTS5 low and CTS5 high risk

3.4.5 Reasons for discordance between CTS5 and MDT recommendation

Of the 'CTS5 high' patients who were recommended to stop (n=4), 3 patients had documented troublesome toxicity from treatment and 1 patient had recently been diagnosed with metastatic cancer from a non-breast primary and therefore continued adjuvant treatment for breast cancer was no longer considered appropriate. Out of the 'CTS5 low' patients who were recommended to continue (n=2), for one patient this course of action was taken owing to them both being node positive and for the other this was due their cancer being multi-focal.

3.4.6 Patterns of recommendation within the overall cohort

The heatmap (Figure 3.4) demonstrates the recommendations made for the whole cohort and suggests a basis on which these were made. One can clearly see that increasing tumour grade, size and lymph node burden were key features in the decision to continue as opposed to stop. Patient age as a single variable did not show a distinct pattern with relation to recommendation. Endocrine agent used for upfront treatment does not affect the treatment recommendation. There is clearly a higher proportion of patients who received chemotherapy to the right of the heatmap reflecting their higher stage disease and amongst these patients one can see many more recommendations to continue

endocrine therapy compared to those who did not receive chemotherapy. Tumour size and lymph nodes cluster together and to a lesser degree tumour grade.

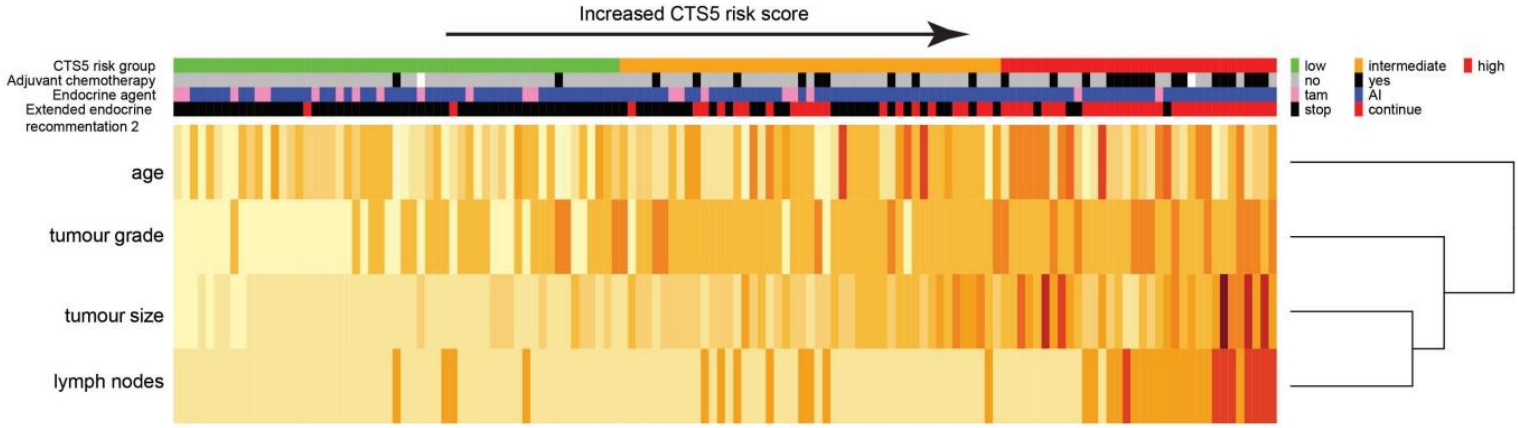


Figure 3.4 Heatmap showing CTS5 variables for each patient who are classified according to CTS5 risk group, adjuvant chemotherapy treatment, endocrine agent and the MDT recommendation 2 within this study. CTS5 variables are represented by increasing colour intensity for increasing value of that variable.

3.4.7 Patterns of recommendations for CTS5 intermediate risk group

Of the 'CTS5 intermediate' patients who were recommended to stop (n=27), reasons for this course of action were given for 8 patients. Six patients had suffered significant toxicity, 1 had developed a second cancer elsewhere and 1 patient had had a breast cancer whose ER Allred score was only 4 out of 8 on the resection specimen. For women who were recommended to continue (n=20), 2 were due to the bilateral nature of their cancer and 2 were because they were node positive and this variable swayed the decision.

A visual representation of the factors affecting recommendations can be seen in the heatmap (Figure 3.4). Of those in the intermediate group who received tamoxifen as their upfront treatment for either 2 or 5 years (n=6), half were recommended to stop. There was no significant association between tamoxifen use and MDT recommendation compared with AI use in this group. The median CTS5 calculated risk of recurrence among the intermediate risk group was 7.3%. Of those intermediate-risk patients who were recommended to stop, their mean risk was 6.9% and of those who were recommended to continue, their mean risk was 7.3%. This difference was not significant.

3.5 Discussion

3.5.1 Patient population

The population studied here had similar baseline features to the post-menopausal cohort of the validation study described in chapter 2. Some key differences, however, were

evident. Overall, this population was older, had larger tumours but lower grade and nodal status compared to the validation cohort. This cohort had a marginally higher median CTS5 value and a higher proportion of women who received chemotherapy. The vast majority of women had received upfront AI as their endocrine treatment (as expected) and therefore the potential benefit of extending to 10 years reflects that seen in studies such as MA17r[33] rather than those using upfront tamoxifen.

3.5.2 Utility of CTS5

On the basis that it did not meet the primary endpoint, this study does not demonstrate that CTS5 has clinical utility for making recommendations about extended endocrine therapy in this centre. The most likely reason for this outcome is that the existing basis on which clinicians within this MDT make extended endocrine therapy recommendations (prior to using CTS5) are identical to the input variables within the algorithm. Therefore, when the CTS5 outcome was revealed, the recommendation was very often concordant with that of the MDT and therefore CTS5 was not necessary to aid decisions. This lack of clinical utility for making a change in current practice may be restricted to specialist centres such as these where expertise exists for informal integration of the information included in CTS5.

3.5.3 Patterns of recommendations

3.5.3.1 Overall

Overall, the individual tumour variables were the main basis on which recommendations were made prior to knowing the CTS5. Age alone was not a strong factor. Choice of upfront endocrine agent did not impact on recommendations and those who received chemotherapy reflected their higher stage of disease.

3.5.3.2 Low and high-risk groups

Concordance in recommendations between MDT and the CTS5 treatment assumption for the low and high-risk patients was high. This suggests that a standardised approach of

stopping endocrine therapy for CTS5 low-risk women and continuing for CTS5 high-risk women is acceptable to this MDT.

3.5.3.3 Intermediate risk group

Within the intermediate group, unsurprisingly, there was much less uniformity of recommendation. There was no difference in the median absolute 5-10 year risk of recurrence between those recommended to stop versus those recommended to continue. This suggests that other factors not included in CTS5 or different weighting of the individual variables were considered when recommendation 2 was made. The initial endocrine therapy drug used did not significantly affect recommendations although patient numbers are very low for this analysis. One may expect more women who have received upfront tamoxifen to be recommended to continue on the strength of the evidence for prolonged treatment in this context, but this was not evident here. Reasons for stopping endocrine therapy in this group were often toxicity driven demonstrating that clinicians were reluctant to recommend a course of management for their patients that they would struggle with. In trials of extended endocrine therapy, adherence to 5 or 6 years of additional treatment was between 57 and 64%[29,33-35,127]. This is among populations of women who had tolerated treatment sufficiently well in the first 5 years and therefore in women struggling with toxicity, adherence would be likely to be lower still. For women who were recommended to continue, this was sometimes due to perceived high-risk factors, which are not taken into account in the CTS5 such as bilateral cancers. Another reason for recommending continuation was due to dominant consideration of a single variable such as involved lymph nodes, which already features within CTS5 and therefore is already accounted for. This practice might lead to an over-estimation of risk and unnecessary treatment.

It is mentioned in chapter 2 that risk groups are clinically useful and help standardise recommendations and reduce variability. However, caution must be exercised with being dogmatic in this approach. Here we can clearly see that for women with CTS5 intermediate risk, a universal management recommendation would be inappropriate. Instead in this group, treatment recommendations need to be more individualised depending on the patient's absolute risk and their side effect profile or perhaps further aided by use of a

genomic (or other molecular) model (Figure 3.5). This theoretical algorithm was developed by my supervisor and I and recommends an initial stratification of patients based on their CTS5 value. For those who are low risk, the potential benefit of extending endocrine therapy is so low that stopping treatment can be safely recommended. For those at high risk, extending endocrine therapy should be recommended unless significant toxicity precludes this. For these two groups, further risk stratification is unnecessary and is unlikely to be helpful. For those of intermediate risk, further risk evaluation would be helpful to guide ongoing management decisions. At the current time, this should be in the form of one of the genomic models that has shown prognostic value for late recurrence (see chapter 1). The outcome of this can then re-categorise women. Those who are re-assigned a new risk category i.e., low or high should then stop or strongly consider extended therapy respectively. For those who remain at intermediate risk, the decision should then come down to individual preference and their attitude towards risk.

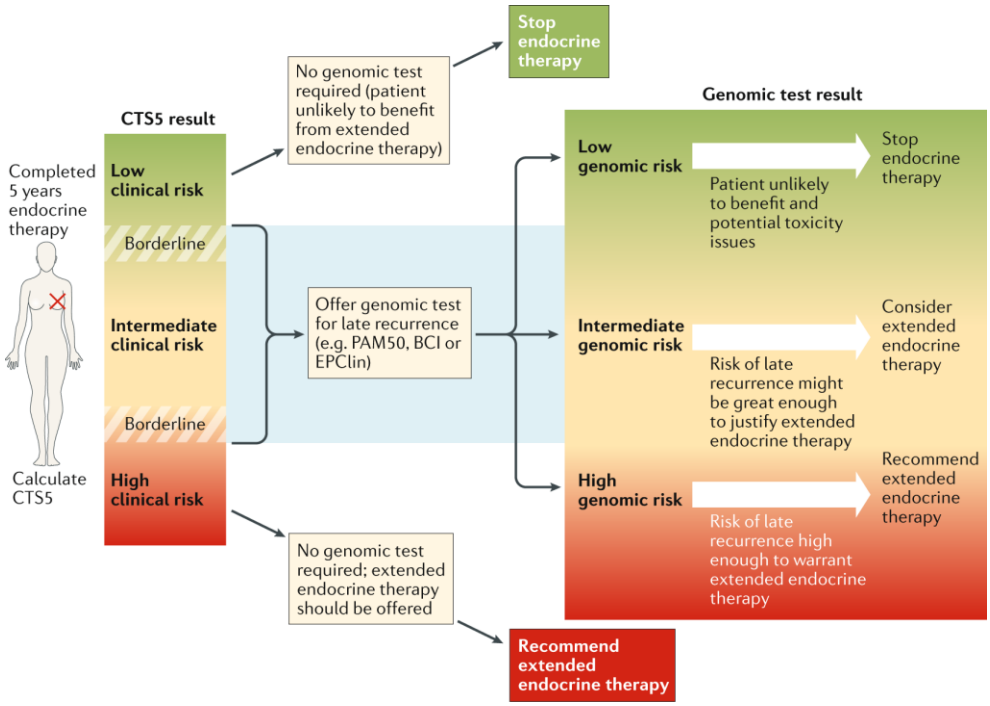


Figure 3.5 Theoretical algorithm for making extended endocrine therapy decisions at 5 years for ER+, HER2- post-menopausal women.

3.5.4 Strengths and Limitations

Strengths of this study are that it took place within the framework of an established MDT where women reaching 5 years since diagnosis are reliably identified and discussed. This should reduce the potential for selection bias. The clinicians within the MDT remained consistent removing the issue of inter-clinician variability in recommendations. Lastly, this study population is reflective of the true population of post-menopausal women with ER+ breast cancer and was similar to the validation set in chapter 2.

Limitations of the study are that it took place within a single centre and did not allow for the variations in patients, doctors and healthcare provisions in the UK and worldwide. The sample size included is relatively small given the prevalence of early ER+ breast cancer. This study assessed the impact on MDT recommendations but did not address the final course of action taken by the patient or their treatment compliance which would be useful information to procure.

3.5.5 Further work

This study took place in a tertiary referral centre with an MDT that is well versed in the literature regarding factors affecting late recurrence. Therefore, the finding that CTS5 lacks utility there is not necessarily surprising. It would be useful to test utility elsewhere in the UK and abroad as it is likely that CTS5 will prove more useful in non-specialist areas where risk factors for late recurrence are less widely known and extended endocrine therapy not commonplace. Bearing in mind the accessibility and cost-free nature of CTS5, utility in under-developed countries would be an important thing to establish. Testing utility in these areas, however, may not be as straightforward as it was in this study. The fundamental role of the MDT meeting at the five-year mark is not formalised in all centres in the same way as it is at diagnosis. Many hospitals in the UK do not have such a meeting for women reaching this milestone and it is unclear how these women are being advised. The availability of CTS5 may encourage centres to establish such MDTs. Utility could be further tested by retrospectively applying CTS5 to women within a clinical trial of extended endocrine therapy as described at the beginning of this chapter.

3.5.6 Clinical application

CTS5 is a cost-free, clinically validated tool and therefore whilst utility could not be demonstrated here, no disadvantage to its use was identified. Employing a quantitative tool for decision making at this time-point promotes consistency in recommendations and leaves less down to individual clinician and centre variability. In the absence of data for the predictive value of CTS5, the definitive correct management strategy on the basis of CTS5 results is still uncertain – particularly for the intermediate risk group. It is likely that for intermediate-risk patients, further prognostic tests will be informative and should be considered. Given the prognostic data so far, along with the accessibility of the tool, and the outcomes of this decision impact study, it is my view that CTS5 be employed for risk prognostication at 5 years and aiding extended endocrine therapy recommendations. CTS5 low-risk and CTS5 high-risk patients potentially no longer need to be discussed in an MDT meeting and instead standardised treatment recommendations could be made by one senior nurse using CTS5 (Figure 3.5) with referral into the MDT for any complex patients such as a high-risk woman with severe endocrine therapy toxicity. This would then leave just the intermediate group patients to be discussed in a formal MDT setting. This could help significantly reduce the patient list and free up time for clinicians.

For patients who warrant discussion, this should be undertaken in a formal MDT meeting at 5 years in the patient's treating centre where toxicity and co-morbidities are documented. Alternatively, CTS5 could be calculated at diagnosis and then discussed with a GP at the 5-year mark when tolerability and other patient factors are known. With the use of prognostic tools that help to standardise recommendations, clinicians must be wary of potential pitfalls of protocol-driven management and must ensure that an individualised approach is retained when making extended endocrine therapy recommendations.

4. Validity of CTS5 in a high stage population

4.1 Introduction

Chapter 2 described how CTS5 shows validity in predicting late distant recurrence in an unselected cohort of women. It was concluded that a further cohort with a larger number of pre-menopausal women would be useful to strengthen the evidence for validity in this subgroup. Additionally, it was suggested that testing CTS5 in a cohort of women with higher stage cancers would be useful as current evidence was derived from cohorts with approximately 70% node negative patients and only 20% use of adjuvant chemotherapy. To demonstrate validity in a higher stage cohort would indicate that even in a population enriched for poor prognostic features, a relatively low risk group could still be identified that might avoid extended adjuvant endocrine therapy. Lastly, testing validity of CTS5 for ER+, HER2+ breast cancer was also recommended owing to the paucity of women with HER2+ disease and events in the cohort tested in chapter 2.

The TACT trial population was selected as an optimal independent dataset, in which to test validity of CTS5 as it was comprised of women with substantially higher stage disease than

the ATAC development set and also included pre-menopausal women[128]. The study recruited ER+ and ER- women and randomised 4162 women with node positive or high risk node negative resected breast cancer to standard anthracycline based chemotherapy or anthracycline followed by docetaxel. ER+ women in this study were treated with 5 years of endocrine therapy after completion of chemotherapy. Within the overall study population, 30% of women had tumours >30mm and only 20% of women were node negative. 24% of randomised women had HER2+ cancers but it is unknown how many received HER2 directed therapy as this was not standard treatment at this time.

At a median of 62 months follow up from time of randomisation, there had been 760 distant recurrence events in total (ER+ and ER-). There was no recurrence or survival benefit for the group who received docetaxal in comparison to those who received anthracycline based chemotherapy alone (figure 1). These findings were somewhat unexpected and subsequent trials have shown the benefit of taxanes in adjuvant chemotherapy which is now commonplace[129,130].

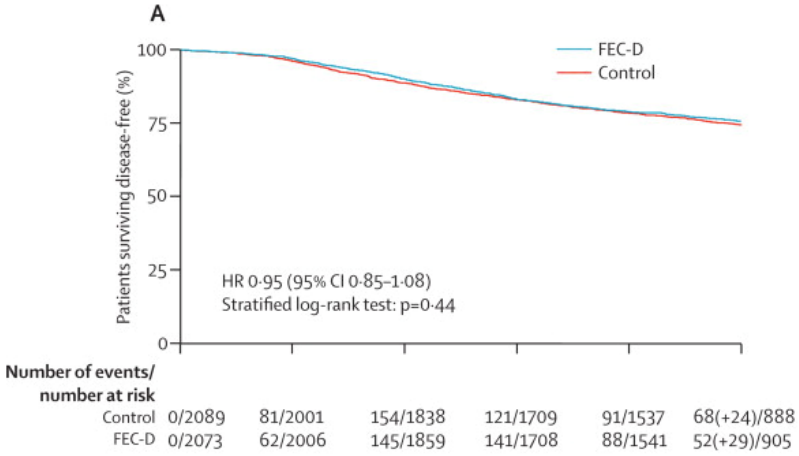


Figure 4.1 Credit: Ellis et al, 2009 [128].
Kaplan Meier plot of TACT patient showing disease free survival in each arm of the study.

Testing CTS5 on a higher stage population should allow the identification of shortcomings of the model. It may be the case that certain prognostic parameters within CTS5 that hold true in the ATAC and BIG 1-98 development datasets may not extrapolate to the TACT population. For example, in the ATAC study, only 35% of women had tumours greater than 20mm and in the TACT study, 30% of women had tumours greater than 30mm. The 30mm

tumour size threshold within the CTS5 algorithm may not be appropriate in TACT as it may classify too many larger tumours as the same size thereby not truly representing the relationship of tumour size with recurrence risk.

The issue of age in breast cancer and its impact on late recurrence has been discussed in chapters 1 and 2. The ATAC population included only post-menopausal women. As such, within model development, the effect of younger age on late recurrence risk would not have been captured. The consequence of this was the finding of a linear relationship between age and late recurrence risk. This was reflected in the BIG 1-98 validation, which also had a purely post-menopausal population. We know, however, that for recurrence from year 0, risk is increased for women younger than 35[59] and this was also seen in the large Oxford overview meta-analysis for recurrences beyond 5 years[25]. In the RMH validation, contrary to expectations, the age function did not adversely affect the calibration of the model for pre-menopausal women. This may be because the RMH pre-menopausal cohort had only 5% women under the age of 40 and therefore this small proportion did not have the power to affect the calibration. Within TACT however, where 17% of women were under 40, the linear function of age within CTS5 may not be applicable and may therefore create evidence against its validity.

Prior to the development of CTS5, many clinicians or MDTs might have erred on the side caution for women with node-positive or large cancers and recommended continued endocrine therapy for such patients. However, studies have shown that prognosticating for risk based on single variables can be inaccurate[58,85] and that women with high-risk features can still be deemed low-risk within combined prognostic models if other features are prognostically favourable. If CTS5 demonstrates validity in this population then greater confidence will be had in this tool for risk assessment of women with higher stage cancers and in spite of these features, a proportion of these women may be able to safely avoid extended endocrine therapy.

Further questions are posed by dual ER and HER2 positivity and its effect on late recurrence. Most studies of late recurrence in ER+ breast cancer exclude HER2+ women and so the question of how this subtype affects risk of late recurrence is yet to be

addressed. Approximately 10% of breast cancers are ER+/HER2+[131], which, in the UK, accounts for over 5000 cases per year. Early studies into ER+/HER2+ breast cancer have shown evidence to support the theory that these cancers show reduced responsiveness to both endocrine therapy and HER2 targeted therapy in comparison to cancers expressing only one of those biomarkers[132]. More recent studies have shown that ER+/HER2+ cancers follow a natural history of recurrence more akin to that of ER+/HER2- rather than that of ER-/HER2+[133,134]. This is supported by recent translational research demonstrating that luminal cancers (by molecular subtyping) expressing both ER and HER2 were less responsive to chemotherapy and HER2 blockade in comparison to the HER2 enriched subtype suggesting that these cancers are more ER, rather than, HER2 dependent[135].

Regarding timing of recurrence for ER+/HER2+ cancers (not treated with HER2 directed therapy), long term trial follow up data has shown that beyond 5 years, risk of recurrence in the HER2+ and HER2- subsets of ER+ cancers becomes more similar[133]. This is likely to reflect the fact that ER+/HER2+ cancers that have not recurred by the 5-year point could be more endocrine driven, rather than HER2 driven. Data on late recurrence of HER2+ cancers in women treated with HER2 targeted therapy have shown a trend (although not significant) for an increased risk of recurrence in ER+ versus ER- cancers in years 5-20. This is in contrast to a significantly greater risk of recurrence for ER-/HER2+ versus ER+/HER2+ cancers in years 0-5[136]. Interestingly, relapse free survival curves in this study shows an enduring benefit of trastuzumab to ten years in both ER+ and ER- women despite treatment ceasing after 1 year. Similar results were reported in the extended follow up of an adjuvant trial where it was seen that the benefit from trastuzumab was maintained in both hormone receptor positive and hormone receptor negative women to ten years[134].

It was, therefore, decided that the TACT study, despite having a proportion of women with HER2+ cancers, was inappropriate to investigate the significance of dual ER/HER2 positivity on late recurrence. This is because the study took place at a time when HER2 targeted therapy was not standard of care and so only a small proportion of patients would have received it. The clear benefit of trastuzumab for both ER+ and ER- cancers and duration thereof would be likely to impact on late recurrence. It was suggested that in order to

validate CTS5 in women with HER2+ breast cancer, a population who received HER2 targeted therapy would need to be used.

As explained in chapter 2, tests of validation assess the degree of closeness the performance of a model has in an independent but similar dataset to the one on which it was developed. In order to demonstrate this, certain analyses are essential to highlight distinct elements of the tool's performance in that dataset. The two most important factors to analyse are calibration (i.e. similarity between observed and expected events between the population of investigation and the original development dataset) and discrimination (i.e. ability of the model to distinguish between low and high risk patients). The statistical analysis plan for this study reflects these essential components of a test of validation.

4.2 Aims

To test the validity of CTS5 within a high-stage population (including pre-menopausal women) with early ER+ breast cancer who remained distant recurrence free at 5 years from diagnosis.

More specifically:
To test whether CTS5 has a similar prognostic performance within this independent dataset compared with that in the combined ATAC and BIG 1-98 development datasets.
To establish whether CTS5 predicts, with sufficient accuracy for clinical use, late recurrences in pre-and post-menopausal women who have reached 5 years from diagnosis.
To establish whether CTS5 adequately divides pre- and post-menopausal women into clinically useful risk groups.

4.3 Methods

4.3.1 Patient population

A proposal to test CTS5 within the TACT trial dataset was written and submitted to the TACT study trial management group. Approval was granted for the CTS5 algorithm to be applied to eligible patients. These were women with ER+ breast cancer, who had been enrolled on the study and who were alive and distant recurrence free at 5 years from randomisation. Exclusions were made for women who had received more than 72 months of endocrine therapy.

4.3.2 Statistical considerations

A statistical analysis plan was devised and analyses conducted by a senior statistician from the Clinical Trials and Statistics Unit within the Institute of Cancer Research. For each endpoint, the TACT trial population was divided into pre- and post-menopausal populations.

The statistical methods employed here were almost identical to those used in the RMH validation study (chapter 2). The main difference was in the endpoints. The primary endpoint in this study was time to distant recurrence in years 5+ following diagnosis (this was a secondary endpoint in the RMH study). Distant recurrence was defined as metastatic disease at a distant site and excluded local, loco-regional and contralateral recurrence.

Patients with local, loco-regional or contralateral recurrences in years 0-5 were not excluded from the analysis. Local, loco-regional or contra-lateral recurrences in years 5+ were not treated as events and such patients were censored at date of last follow up. Patients with distant recurrence or second primary cancers in years 0-5 were excluded from the analysis. Second primary cancers in years 5+ were not treated as events but patients were censored at the date of confirmation of the second primary cancer. Death in years 5+ from other causes without distant recurrence was a censoring event.

The first secondary endpoint was time to distant recurrence in years 5-10 (which was the primary endpoint in the BIG 1-98 validation study and the RMH study), in which events beyond 10 years are not included and these patients are censored at 10 years or date of last follow up if earlier. The other secondary endpoint was time to distant recurrence in years 5+ for HER2- subgroups.

4.3.3 Data collection and CTS5 calculation

The following data was extracted from the TACT trial database for eligible patients: date of randomisation, date of birth, menopausal status at randomisation (or following chemotherapy if not documented at diagnosis), pathological tumour size, histopathological grade, pathological nodal status, ER status, HER2 status, type and duration of endocrine therapy, dates of distant recurrence, local/contralateral recurrence, date of last follow up and date and cause of death.

CTS5 scores were derived from the model by substituting the patients' variables (age, nodal status, tumour size and grade) into the published equation:

CTS5=0.438 x nodes + 0.988 x (0.093 x size - 0.001 xsize^2 + 0.375 x grade + 0.017 x age)

The expected risk of recurrence for each individual was determined from the equation in the same way as performed for the validation study in chapter 2. Expected risk was also modified for those women who recurred in years 5-10 or those who died of non-breast cancer causes as was performed in chapter 2.

4.3.4 Statistical Analysis

For each endpoint, prognostic value of the continuous CTS5 score was calculated using hazard ratios (HR) for a unit increase in the score along with 95% confidence intervals for distant recurrence. Log-rank P values were calculated to demonstrate significance. Patients were stratified into risk groups using the CTS5 cut-offs (<3.13, 3.13-3.86 and >3.86) established in the combined ATAC and BIG 1-98 datasets. Kaplan Meier survival method was used to obtain the time to event curves for the three risk categories based on the observed distant recurrence time and events. The 5-10 year and 5 years + distant recurrence risks (%) were calculated with corresponding 95% confidence interval. HR between low/intermediate and low/high-risk groups were calculated with corresponding 95% confidence intervals to demonstrate discrimination between risk groups. Discrimination was further tested by calculation of receiver operator curves.

Calibration of the CTS5 risk model was tested by comparing observed (O) distant recurrence events to expected (E) distant recurrence events for primary endpoint events with the dataset divided into deciles and into CTS5 risk groups. The expected events were calculated using the baseline hazard for distant recurrence in the combined ATAC/BIG 1-98 datasets (shown above) against which the model's performance was calibrated. The overall performance of the CTS5 was assessed by likelihood ratio Chi square test.

4.4 Results

4.4.1 Patients

In TACT, 3122 women remained alive and DR free at 5 years from randomisation. Of these, 2257 (72.3%) had ER+ breast cancer. Exclusions were made for those who received more than 72 months of endocrine therapy (Figure 4.2). After exclusions, 2213 women remained – 1034 in the post-menopausal group and 1179 in the premenopausal group.

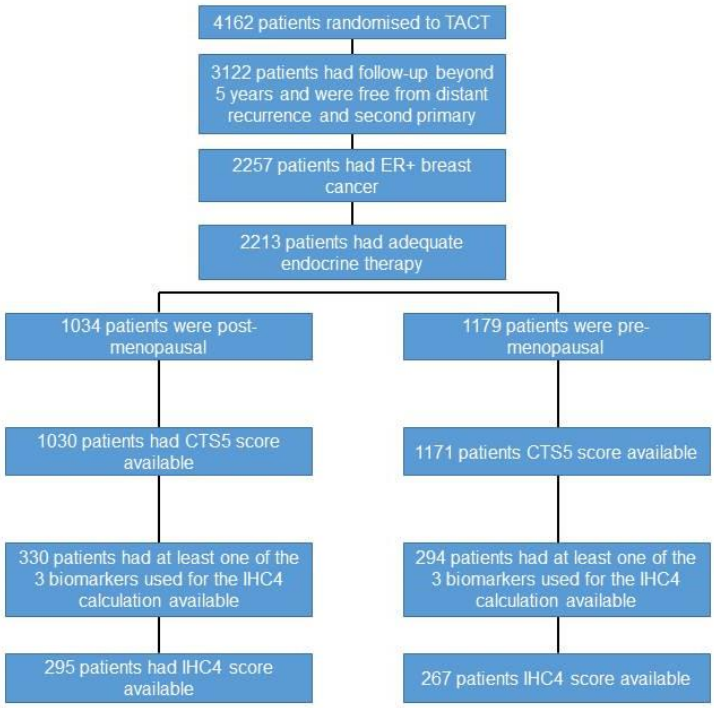


Figure 4.2 Consort diagram

4.4.2 Post-menopausal cohort

4.4.2.1 Baseline characteristics, CTS5 values and outcomes

Median age in this cohort was 54.5 years and this was similar across the risk groups. 113 (11%) women were node negative and the majority were grade 2 or 3. Median tumour size was 2.4cm. Median follow up was 12 years (Table 4.1).

CTS5 scores showed a normal distribution and median score value was 3.81 which equates to a risk of DR of 9.6% in years 5-10. 17% of the cohort were classified as low risk, 35% as intermediate risk and 48% as high risk. There were 155 DR events beyond 5 years (Table 4.2). The event rate in years 5-10 was 11.4%.

	CTS5 group			**Total (N=1030)**
	Low (N=175)	**Intermediate (N=359)**	**High (N=496)**	
Randomised treatment group				
FEC-T	89 (50.9)	186 (51.8)	240 (48.4)	515 (50.0)
Control	86 (49.1)	173 (48.2)	256 (51.6)	515 (50.0)
Age (years), median (IQR)	52.1 (48.1, 56.1)	53.9 (50.2, 58.3)	55.9 (51.9, 59.9)	54.5 (50.7, 58.7)
Age group (years)				
<40	10 (5.7)	12 (3.3)	10 (2.0)	32 (3.1)
40-49	50 (28.6)	75 (20.9)	57 (11.5)	182 (17.7)
50-59	101 (57.7)	217 (60.5)	309 (62.3)	627 (60.9)
≥60	14 (8.0)	55 (15.3)	120 (24.2)	189 (18.4)
Nodal status				
Node negative	62 (35.4)	51 (14.2)	0 (0.0)	113 (11.0)
1-3 positive nodes	108 (61.7)	259 (72.1)	168 (33.9)	535 (51.9)
≥4 positive nodes	5 (2.9)	49 (13.7)	328 (66.1)	382 (37.1)
Tumour grade				
Grade 1	45 (25.7)	45 (12.5)	13 (2.6)	103 (10.0)
Grade 2	88 (50.3)	207 (57.7)	247 (49.8)	542 (52.6)
Grade 3	42 (24.0)	107 (29.8)	236 (47.6)	385 (37.4)
Tumour size				
Median (IQR)	1.5 (1.2, 1.8)	2.2 (1.8, 2.8)	3.0 (2.3, 4.0)	2.4 (1.7, 3.5)
≤2cm	151 (86.3)	155 (43.2)	96 (19.4)	402 (39.0)
2-5cm	24 (13.7)	180 (50.1)	344 (69.4)	548 (53.2)

>5cm	0 (0.0)	24 (6.7)	56 (11.3)	80 (7.8)
>3cm	5 (2.9)	68 (18.9)	228 (46.0)	301 (29.2)
HER2 status				
HER2+	17 (9.7)	40 (11.1)	86 (17.3)	143 (13.9)
HER2-	141 (80.6)	268 (74.7)	345 (69.6)	754 (73.2)
HER2 unknown	17 (9.7)	51 (14.2)	65 (13.1)	133 (12.9)

Table 4.1 Baseline characteristics of the post-menopausal cohort

	Premenopausal (N=1179)	**Postmenopausal (N=1030)**	**All (N=2213)**
DR years 5+ (inc. events beyond 10 yrs)			
Total	198 (17%)	155 (15%)	353 (16%)
Low	32 (10%)	6 (3%)	38 (8%)
Intermediate	51 (12.%)	51 (14%)	102 (13%)
High	115 (27%)	97 (20%)	213(23%)
DR in years 5-10			
Total	147 (12.5%)	118 (11.4%)	265 (12%)
Low	26 (8.3%)	6 (3.4%)	32 (7%)
Intermediate	30 (7%)	39 (10.9%)	69 (9%)
High	91 (21.1%)	73 (14.7%)	164 (18%)
Median follow-up (years from randomisation), median (IQR)	12.0 (10.0 – 12.9)	12.0 (9.9 – 12.9)	

Table 4.2 Distant recurrence outcome data for pre- and post-menopausal women included in this analysis of CTS5. Percentages in 'Total' row are proportion of DR events out of total number of women in pre/post-menopausal subgroups. Percentages in 'Low', 'Intermediate' and 'High' rows are proportion of DR events out of no. of women in that CTS5 subgroup.

4.4.2.2 Events in years 5+

CTS5 was prognostic for late DR in post-menopausal women with a HR of 1.90 (1.52-2.38), p<0.001 for an increase in one CTS5 unit (Table 4.3). When split into three groups using the predefined cut-off points, the groups were distinct and did not overlap (Figure 4.3a). HR for distant recurrence for intermediate and high-risk women was 4.63 (1.98-10.78), p<0.001 and 6.47 (2.84-14.75), p<0.001 respectively, compared to those categorised into the low-risk group (Table 4.3). The low-risk category comprised 17% of the cohort whose risk of DR was 3.6% in years 5-10 (Figure 4.3a).

Within the post-menopausal cohort there were 754 (73%) women with known HER2- disease (with available CTS5 data). Within this sub-group, continuous CTS5 was prognostic for late recurrence in years 5 onwards, HR=2.18 (1.68-2.83), p<0.001. Median CTS5 score was 3.72 which equates to a risk of DR of 8.8% in years 5-10. CTS5 divided the HER2- sub-group into 3 risk categories in which HR = 4.29 (1.69-10.92), p=0.001, and HR = 7.04 (2.85-17.41), p<0.001 in the intermediate and high-risk groups respectively in comparison to the low-risk group (Table 4.3 and Figure 4.3b).

	Years 5+				Years 5-10			
	Postmenopausal women (N=1030)		Premenopausal women (N=1171)		Postmenopausal women (N=1030)		Premenopausal women (N=1171)	
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
CTS5 (continuous)	1.90 (1.52-2.38)	<0.001	2.03 (1.66-2.49)	<0.001	1.86 (1.44-2.4)	<0.001	2.14 (1.69-2.71)	<0.001
CTS5 HER2- (continuous) (754;783)	2.18 (1.68-2.83)	<0.001	2.26 (1.77-2.88)	<0.001	2.18 (1.61-2.95)	<0.001	2.45 (1.83-3.28)	<0.001
CTS5 risk groups (all)								
Low (175;314)	Reference		Reference		Reference		Reference	
Intermediate (359;426)	4.63 (1.98-10.78)	<0.001	1.18 (0.76-1.84)	0.48	3.39 (1.44-8.02)	0.003	0.87 (0.51-1.46)	0.59
High (496;431)	6.47 (2.84-14.75)	<0.001	3.01 (2.04-4.46)	<0.001	4.67 (2.03-10.73)	<0.001	2.87 (1.86-4.44)	<0.001
CTS5 risk groups (HER2-)								
Low (141;314)	Reference		Reference		Reference		Reference	
Intermediate (359;426)	4.29 (1.69-10.92)	0.001	1.39 (0.82-2.35)	0.22	3.01 (1.16-7.81)	0.02	0.97 (0.51-1.84)	0.92
High (496;431)	7.04 (2.85-17.41)	<0.001	3.55 (2.21-5.7)	<0.001	5.09 (2.04-12.71)	<0.001	3.44 (2.00-5.90)	<0.001

Table 4.3 Hazard ratios for distance recurrence with corresponding 95% confidence intervals and p values for continuous CTS5 score and CTS5 risk groups for the pre- and post-menopausal cohorts. Numbers in brackets in left hand column indicate number of post-menopausal women out of number of pre-menopausal women for that subgroup

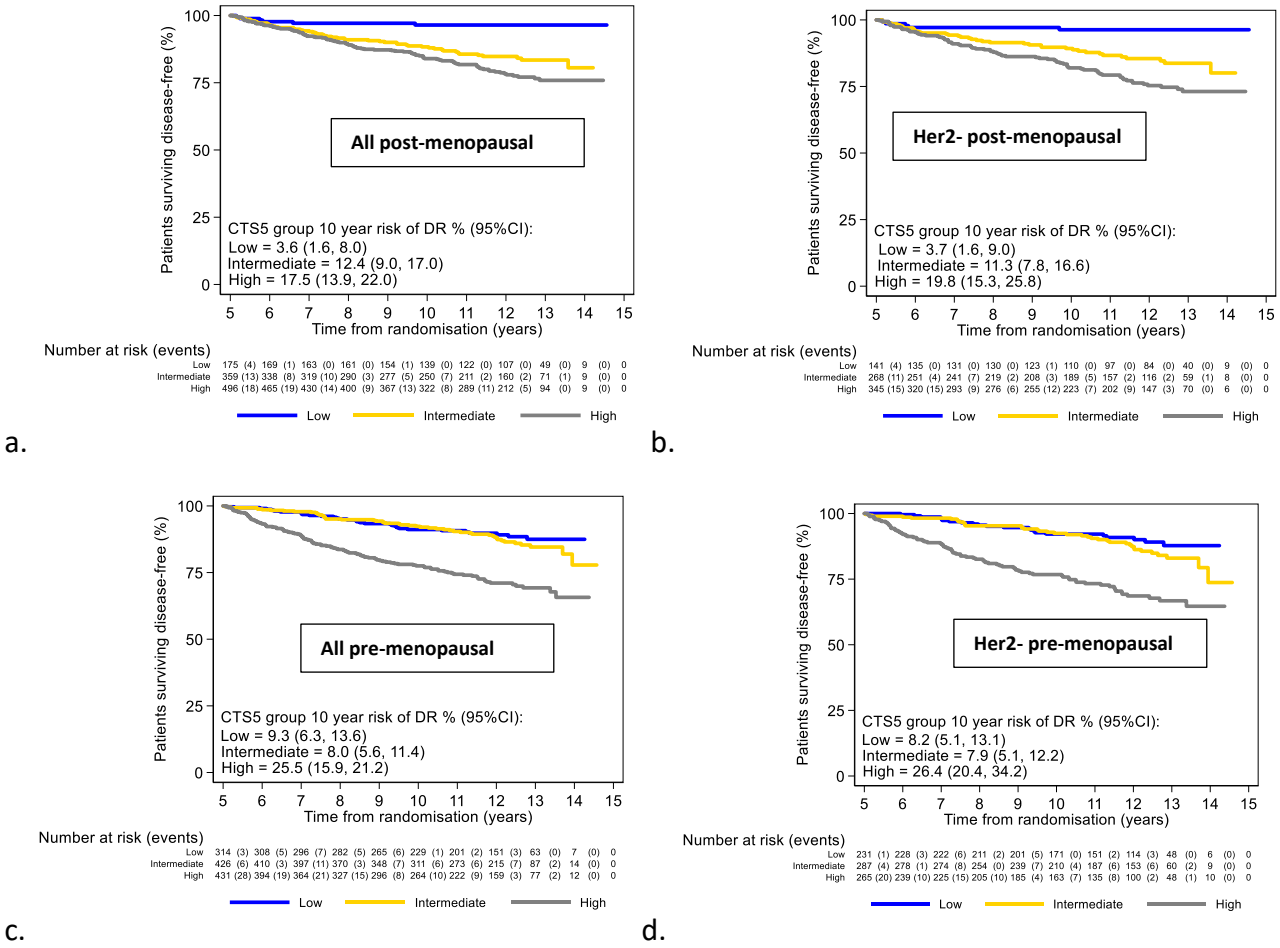


Figure 4.3 Kaplan Meier curves showing distant recurrence in years 5+ for a. all post-menopausal women, b. HER2- post-menopausal women, c. all pre-menopausal women and d. HER2- pre-menopausal women. DR risk for each CTS5 category is given with corresponding 95% CI. Note that while the KM curves show all events beyond 5 years, the DR risk for each CTS5 category is the risk at 10 years and incorporates only events up to 10 years.

4.4.2.3 Calibration in years 5-10

To compare observed and expected events with the original model, only events in years 5-10 were considered as this is the time period over which CTS5 was trained. There was a significantly higher number of observed (21) than expected (13.5) DR events in the second quintile (χ^2= 4.16) (Table 4.4, Figure 4.4a). Additionally, when dividing the cohort by risk group there was a significantly higher number of observed (39) events than expected (25.9) in the intermediate risk group (χ^2= 6.59) (Table 4.4, Figure 4.4b). The overall O/E ratio was 0.98 (0.81-1.17), p=0.86. The c-statistic for area under the receiver operator curve was 0.63.

Calibration was improved in the HER2- subgroup where there was no longer a significant difference between observed and expected events in any quintile or CTS5 risk group (Table 4.4 and Figure 4.4c and d).

	All post-menopausal patients					HER2- post-menopausal patients				
Quintile/risk group	**Observed**	**Expected**	**Number of patients**	**χ^2**	**Risk of recurrence %**	**Observed**	**Expected**	**Number of patients**	**χ^2**	**Risk of recurrence %**
1	9	8.0	206	0.12	3.9	6	4.6	151	0.47	3.0
2	21	13.5	206	4.16	6.6	14	8.9	151	2.90	5.9
3	20	20.1	205	0.00	9.8	16	13.3	151	0.54	8.8
4	29	29.2	207	0.00	14.1	19	19.8	151	0.03	13.1
5	39	49.7	206	2.31	24.1	34	36.7	150	0.19	24.3
Low	6	6.4	175	0.02	3.6	5	5.1	141	0.0	3.6
Intermediate	39	25.9	359	6.59	7.2	27	19.3	268	3.08	7.2
High	73	88.0	496	2.57	17.7	57	58.8	345	0.06	17.0

Table 4.4 Numbers of observed versus expected events for all post-menopausal women and HER2- only post-menopausal women as divided into quintiles and risk groups. χ^2 value given for difference between observed and expected events.

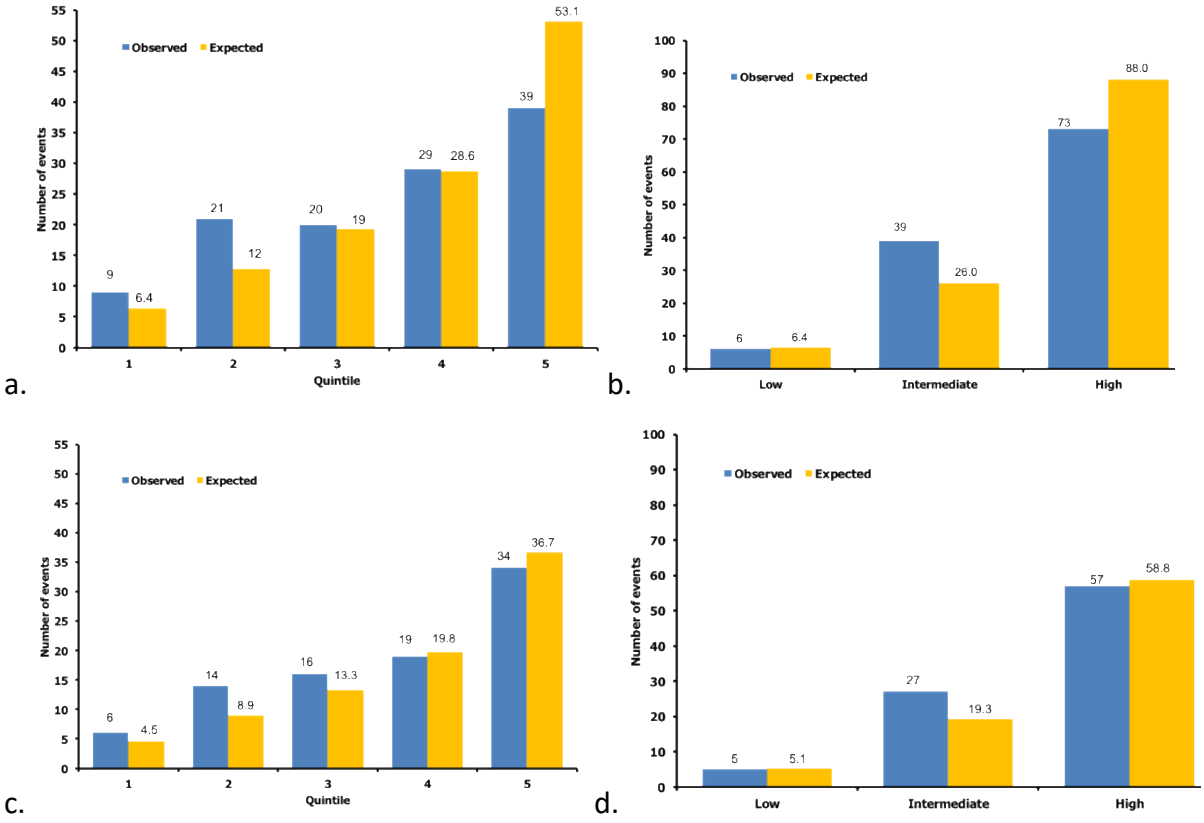


Figure 4.4 Calibration column graph of observed versus expected events in years 5-10 in post-menopausal women for a: all women by CTS5 quintile, b: all women by CTS5 risk group, c. HER2- women by CTS5 quintile, d. HER2- women by CTS5 risk group.

4.4.2.4 Events in years 5-10

CTS5 was prognostic for late DR in post-menopausal women with a HR of 1.86 (1.44-2.4) p<0.001 for an increase in one CTS5 unit (Table 4.3). When the cohort was split into CTS5 risk category, separation of the curves was clear (Figure 4.2 a and b). HR for distant recurrence for intermediate and high-risk women was 3.39 (1.44-8.02), p<0.001 and 4.67 (2.03-10.73), p<0.0001 respectively, compared to those categorised into the low-risk group (Table 4.3).

4.4.3 Pre-menopausal cohort

4.4.3.1 Baseline characteristics, CTS5 values and outcomes

Baseline characteristics of pre-menopausal women are shown in Table 4.5. Median age was 44.5. There were 282 (24.1%) women under the age of 40 and 31% of these were categorised as low risk by CTS5. There were 194 women (16.6%) with node-negative

cancers and, like the post-menopausal cohort, none of these were deemed high risk. Most cancers were grade 2 and 3. CTS5 was normally distributed in this cohort and median CTS5 value was 3.63 which equates to a 5-10 year DR risk of 8.1%. Over the entire follow up period, there were 198 DR events in the pre-menopausal cohort (Table 4.2). The event rate in years 5-10 was 12.5%.

4.4.3.2 Events in years 5+

CTS5 was prognostic for late DR in pre-menopausal women with a HR of 2.03 (1.66-2.49), p<0.001 for an increase in one CTS5 unit (Table 4.3). Using the predefined cut-off points the three risk groups markedly overlapped (Figure 4.3c), particularly evident in the low and intermediate risk curves. HR for intermediate and high-risk women was 1.18 (0.76-1.84), p=0.48) and 3.01 (2.04-4.46), p<0.001) respectively, compared to those categorised into the low-risk group (Table 4.3). The low-risk category comprised 26.6% of the cohort whose predicted risk of DR was 9.3% over years 5-10 (Figure 4.3c).

The pre-menopausal cohort was comprised of 783 known HER2- women with available data for CTS5. From these, median CTS5 value was 3.56, which equates to a risk of DR of 7.6% in years 5-10. Continuous CTS5 was prognostic for distant recurrence in years 5 onwards, HR = 2.26 (1.77-2.88), p<0.001 (Table 4.3). There was no significant difference between low and intermediate risk categories. In the high-risk group HR = 3.55 (2.21-5.7), p<0.001 in comparison to the low-risk group. The low-risk group had a predicted DR risk of 8.2% over years 5-10 (Figure 4.3d).

	CTS5 group			Total (N=1171)
	Low (N=314)	Intermediate (N=426)	High (N=431)	
Randomised treatment group				
FEC-T	141 (44.9)	217 (50.9)	224 (52.0)	582 (49.7)
Control	173 (55.1)	209 (49.1)	207 (48.0)	589 (50.3)
Age (years), median (IQR)	43.1 (38.9, 47.0)	44.4 (40.0, 47.9)	45.6 (41.6, 49.0)	44.5 (40.2, 48.2)
Age group (years)				
<40	99 (31.5)	105 (24.7)	78 (18.1)	282 (24.1)
40-49	186 (59.2)	255 (59.9)	271 (62.9)	712 (60.8)
50-59	29 (9.2)	66 (15.5)	82 (19.0)	177 (15.1)
≥60				
Nodal status				
Node negative	138 (44.0)	56 (13.2)	0 (0.0)	194 (16.6)
1-3 positive nodes	172 (54.8)	326 (76.5)	127 (29.5)	625 (53.4)
≥4 positive nodes	4 (1.3)	44 (10.3)	304 (70.5)	352 (30.1)
Tumour grade				
Grade 1	42 (13.4)	32 (7.5)	8 (1.9)	82 (7.0)
Grade 2	175 (55.7)	212 (49.8)	184 (42.7)	571 (48.8)
Grade 3	97 (30.9)	182 (42.7)	239 (55.5)	518 (44.2)
Tumour size				
Median (IQR)	1.7 (1.3, 2.1)	2.5 (2.0, 3.0)	3.2 (2.5, 4.5)	2.5 (1.9, 3.5)
≤2cm	222 (70.7)	141 (33.1)	45 (10.4)	408 (34.8)
2-5cm	87 (27.7)	266 (62.4)	306 (71.0)	659 (56.3)
>5cm	5 (1.6)	19 (4.5)	80 (18.6)	104 (8.9)
>3cm	21 (6.7)	105 (24.6)	227 (52.7)	353 (30.1)
HER2 status				
HER2-	231 (73.6)	287 (67.4)	265 (61.5)	783 (66.9)
HER2+	46 (14.7)	65 (15.3)	86 (20.0)	197 (16.8)
HER2 unknown	37 (11.8)	74 (17.4)	80 (18.6)	191 (16.3)

Table 4.5 Baseline characteristics of the pre-menopausal cohort

4.4.3.3 Calibration in years 5-10

As with the post-menopausal cohort, calibration of the model was deemed most accurately measured by using events only in years 5-10.

When split into quintiles by increasing score, there was a significantly greater number of observed than expected events for the 1st, 2nd and 4th quintiles with a χ^2 of 17.70, 11.46 and 4.39 respectively (Table 4.6 and Figure 4.5a). A similar picture was seen when dividing the cohort by CTS5 risk group in which the low and high-risk groups had a significantly greater number of observed than expected events, χ^2 = 19.35 and 7.36 respectively (Table 4.6 and Figure 4.5b). The overall O/E ratio was 1.33 (1.13-1.57), p=0.001. The c-statistic for area under the receiver operator curve was 0.65. Calibration was improved to a degree when only women with HER2- cancers were considered however a significantly greater number of observed compared to expected events was still seen in the 1st and 2nd quintiles and low and high-risk groups (Table 4.6 and Figure 4.5c and d).

4.4.3.4 Events in years 5-10

CTS5 was prognostic for late DR with a HR of 2.14 (1.69-2.71), p<0.001 for an increase in one CTS5 unit (Table 4.3). This figure was marginally improved when looking at the HER2- population only. When the cohort was split into CTS5 risk category, separation of the curves was not distinct (Figure 4.3c) and HR for distant recurrence for intermediate versus low-risk women was not significant. HR for high-risk women was 2.87 (1.86-4.44), p<0.001 compared to those categorised into the low-risk group. Similar results were seen between risk groups when looking at the HER2- only subgroup (Table 4.3 and Figure 4.3d).

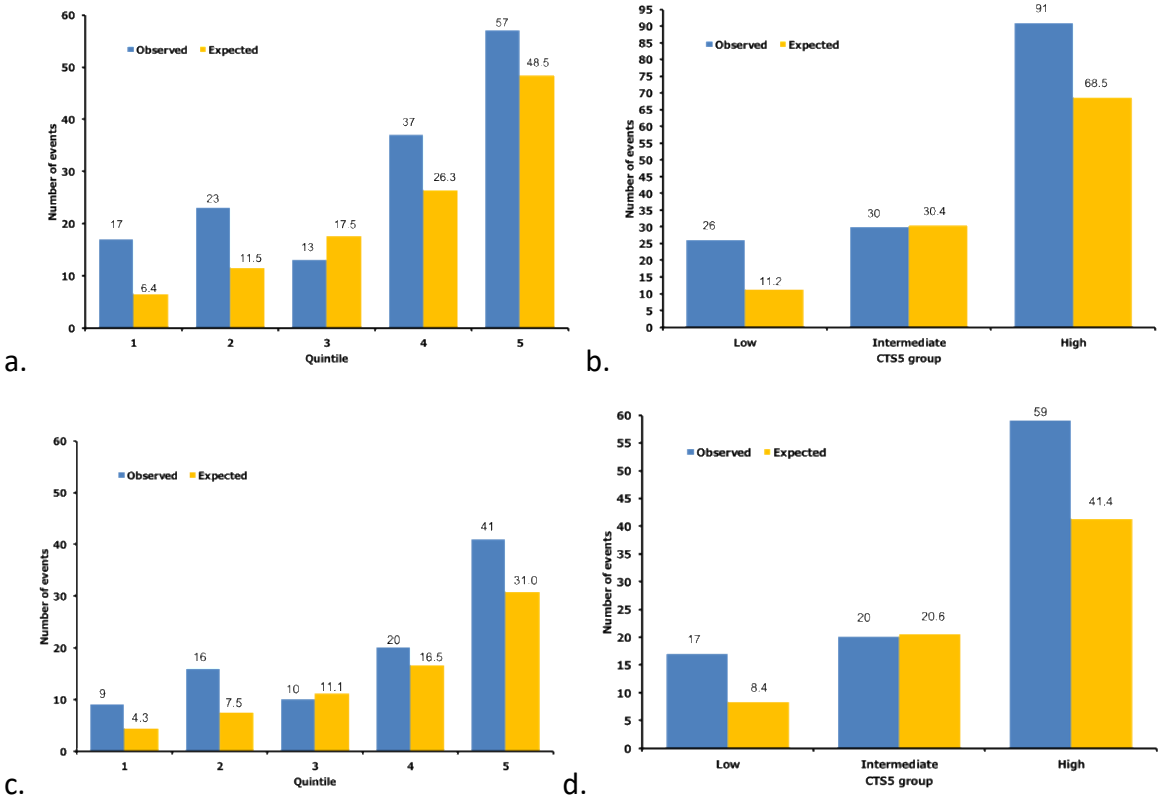


Figure 4.5 **Calibration column graph of observed versus expected events in pre-menopausal women for a: all women by CTS5 quintile, b. all women by CTS5 risk group, c. HER2- women by CTS5 quintile, HER2- women by CTS5 risk group.**

	All pre-menopausal patients					HER2- pre-menopausal patients				
Quintile/ risk group	**Observed**	**Expected**	**Number of patients**	**χ^2**	**Risk of recurrence %**	**Observed**	**Expected**	**Number of patients**	**χ^2**	**Risk of recurrence %**
1	17	6.38	234	17.70	2.7	9	4.29	156	5.16	2.8
2	23	11.51	234	11.46	4.9	16	7.46	157	9.79	4.7
3	13	17.54	234	1.17	7.5	10	11.06	156	0.10	7.1
4	37	26.26	234	4.39	11.2	20	16.52	157	0.73	10.5
5	57	48.55	235	1.47	20.7	41	30.99	157	3.24	19.7
Low	26	11.25	314	19.35	3.6	17	8.35	231	8.96	3.6
Intermediate	30	34.44	426	0.01	7.1	20	20.59	287	0.02	7.2
High	91	68.54	431	7.36	15.9	59	41.38	265	7.50	15.6

Table 4.6 Numbers of observed versus expected events for all pre-menopausal women and HER2- only pre-menopausal women as divided into quintiles and risk groups. χ^2 value given for difference between observed and expected events

4.5 Discussion

4.5.1 Post-menopausal cohort characteristics

In this sub-group, only 11% were node-negative none of whom fell into the CTS5 high risk group demonstrating the prognostic influence of this single variable. Conversely in ATAC and BIG 1-98, 68% and 61% respectively were node negative. 37% of TACT patients studied here had grade 3 cancers, a third of which were >3cm with a median tumour size of 24mm. In ATAC and BIG 1-98 only 25% and 20% of women had grade 3 disease and 10% and 11% respectively had tumours >3cm. All TACT participants received chemotherapy whereas in ATAC and BIG 1-98 this figure was 20% and 24%. Being an adjuvant chemotherapy trial that recruited in the early 2000s, TACT did not include many older patients giving a median age of postmenopausal women of 55 whereas in ATAC and BIG 1-98 it was 64 and 61.

Post-menopausal women in TACT had higher stage disease and were younger than the cohort used for the RMH validation (chapter 2). In the RMH cohort almost three quarters of women were node negative and median tumour size was 18mm (compared to 24mm in TACT) and just under a quarter had grade 3 disease. Event rate in the post-menopausal TACT validation cohort overall was 11% in years 5-10 which is 5% higher than in the RMH cohort and is to be expected for this higher stage group of women.

Thus it is clear that the baseline characteristics of the TACT trial population eligible for this validation study was substantially different from that of the original development and validation populations and fitted the aim of this study.

4.5.2 CTS5 validity for post-menopausal women

CTS5 performed well for this cohort and was significantly prognostic as a continuous score and also in the three risk categories. What is clinically important here is that risk categories were well separated in clinical outcome. The low-risk group had a DR risk of 3.6% over years 5-10, although confidence intervals cross the boundary of 5% that marks the upper limit of

risk for this group. In this higher stage population it is noteworthy that the CTS5 was still able to isolate this low risk cohort which comprised 17% of the cohort. Notably, there is no overlap between confidence intervals of distant recurrence risk between the low-risk group and the intermediate group and the Kaplan Meier curves are distinct indicating the discriminatory ability of the CTS5 to distinguish different risk women from within this population.

Calibration of observed versus expected events for the HER2- sub-group showed no significant differences. When the entire post-menopausal cohort was considered, a significant discrepancy in event numbers was seen only for those in quintile 2 and in the intermediate risk group. This is unlikely to have substantial clinical significance, and for the low-risk patients, who may be recommended to stop endocrine therapy on the basis of this data, the strength of calibration for this risk group and the 1st quintile is encouraging and supports this management approach.

Women with known HER2+ cancers accounted for 14% of the overall post-menopausal cohort analysed here and so would not have made a large impact in the overall result. Analysis of women with HER2- cancers revealed similar results in terms of distinction between risk groups and a marginally greater effect size for prognostic value of continuous CTS5 score in comparison to the whole post-menopausal cohort. This similarity in results probably reflects the fact that since most women with HER2+ cancer would not have received HER2 targeted therapy, those with more aggressive disease would have been likely to have recurred in years 0-5. Those women with HER2+ cancers who were DR free at 5 years are likely to be those with more favourable prognostic features and therefore a lower risk of late recurrence.

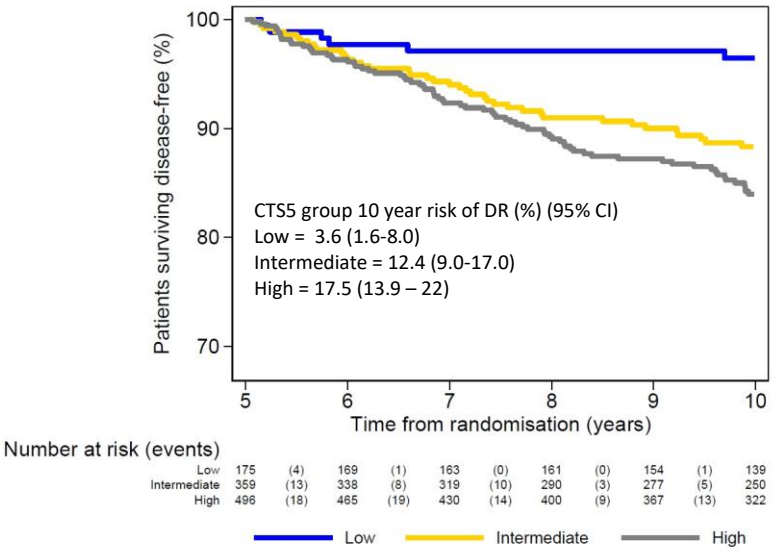
4.5.3 Comparison with ATAC/BIG 1-98 validation

This validation study had a different primary endpoint (recurrences in years 5+) compared to that within the development dataset and the RMH cohort (recurrence in years 5-10). This was thought to be appropriate for this population as this would capture recurrences beyond 10 years. By doing so this would give a more ‘real life’ picture of the prognostic

value of CTS5 and make it more relevant for long term management. This was important for this particular population with a higher anticipated DR event rate than in the development and RMH datasets. This had the effect of reducing the prognostic performance of CTS5 within the whole population and in the subgroup analyses. This occurred because the model was developed using events in years 5-10 and therefore there will be a natural loss of accuracy but even using this endpoint, CTS5 showed high prognostic value. Despite using events in years 5+ as the primary endpoint, for this study (in comparison to the RMH validation) DR risks were calculated for years 5-10 only. This was done owing to the drop off of numbers of patients at risk at the tail end of the curve which would result in a less stable analysis and wider confidence intervals if DR risk at year 15 or 20 had been calculated (as was done for the RMH validation). This lack of accuracy would have made a DR risk estimation far less useful and so it was not performed. In addition, including events beyond 10 years meant that there were more observed events for this endpoint and as such calibration of observed versus expected events would be affected. It is for this reason that calibration was only assessed for the 5-10 year event endpoint.

Magnitude of effect of continuous CTS5 on DR risk was similar in the TACT study as for CTS5(ATAC) in the BIG 1-98 validation. When comparing the prognostic value of CTS5 in women with HER2- cancers between the 2 validation studies, the effect size is identical: 2.18 (1.61-2.95) in TACT and 2.19 (1.61-2.98) in BIG 1-98. It must be remembered, as with the RMH validation in chapter 2, that in the development dataset, the original CTS5(ATAC) algorithm was used whereas in this study the published algorithm based on the combined ATAC and BIG1-98 datasets was tested. However, as discussed in chapter 2, the changes to the co-efficients in the new algorithm were minor and the prognostic value of the two models are very similar with overlapping confidence intervals and so the comparison can be considered reliable (Figure 2.9). Of clinical importance is the fact that hazard ratios between the low and intermediate risk groups are similar as depicted by the Kaplan Meier curves (Figure 4.6) and the curves continue to be distinct beyond 10 years showing the ongoing discriminative ability of CTS5. This is important given that the purpose of CTS5 is in identifying those clear low-risk women who can safely avoid extended endocrine therapy.

a.



b.

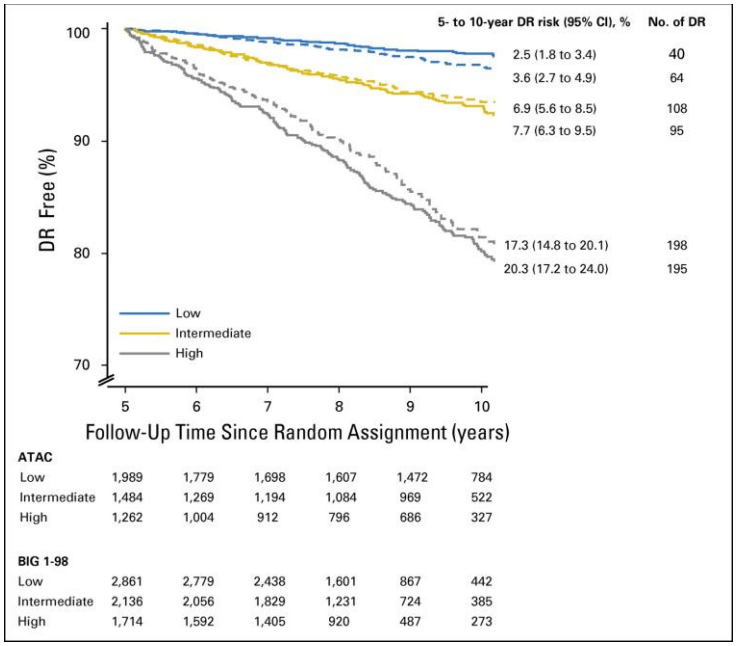


Figure 4.6 Kaplan Meier curve of DR events for a. postmenopausal women in years 5-10 in TACT, b. the ATAC and BIG 1-98 datasets (as produced by CTS5(ATAC) risk group stratification[58]).

4.5.4 Comparison with RMH validation

Despite differences in baselines features of this cohort compared to that of the RMH study, prognostic performance of continuous CTS5 in the post-menopausal cohort of TACT is similar. In TACT, a greater magnitude of difference is seen between the risk groups probably caused by a greater variation in tumour baseline characteristics between the groups. Notwithstanding all the differences between these cohorts, CTS5 highlighted a low risk group in the TACT with a distant recurrence risk of 3.6% in years 5-10 (CTS5 risk in the same group in the RMH study was 2.9%). This shows that despite the TACT study being

comprised of higher stage women, CTS5 can identify very low risk women among them. Calibration of events for low and high-risk women in this TACT analysis are similar to that of the 5-10 year events in the RMH cohort that received only 5 years of endocrine therapy.

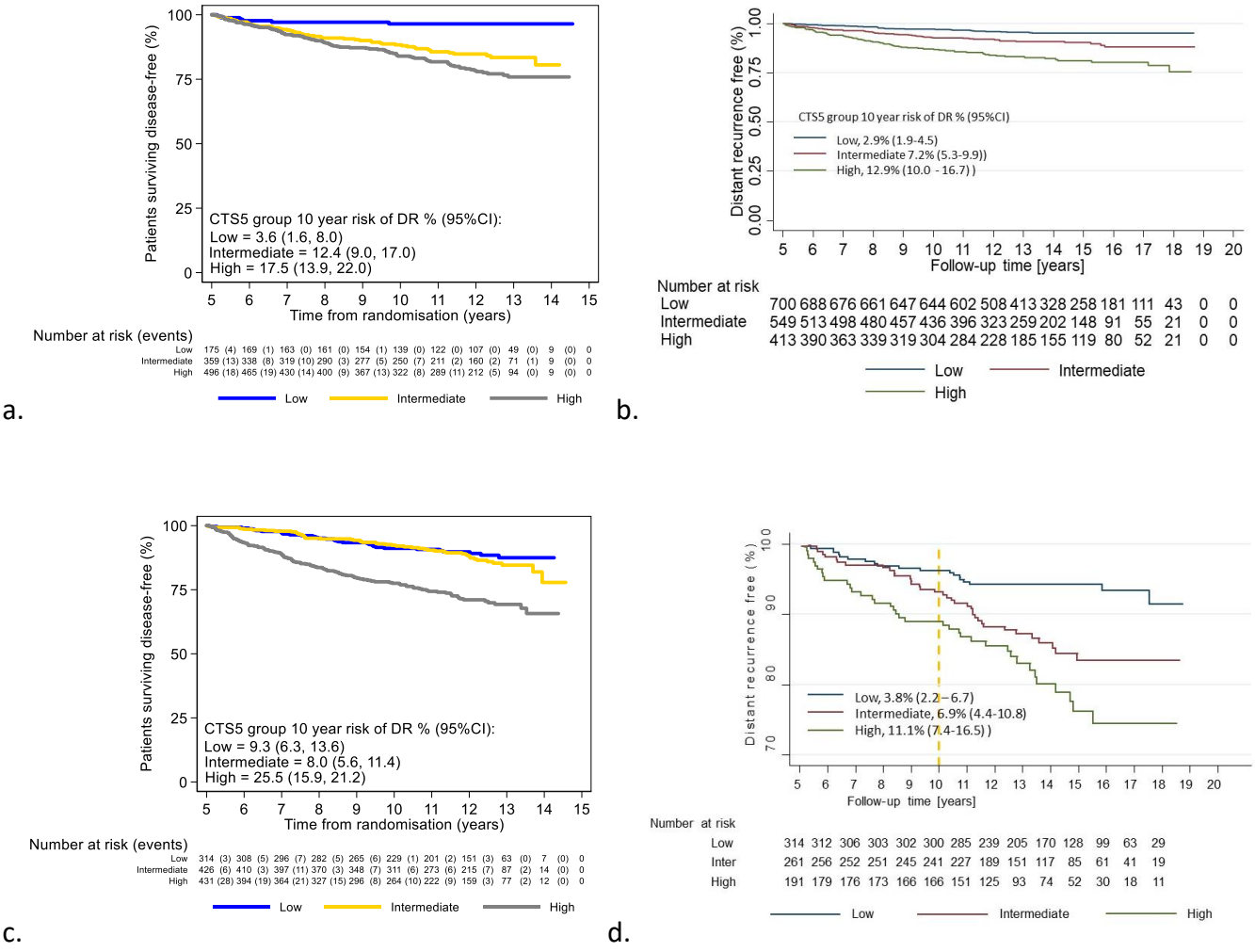


Figure 4.7 Kaplan Meier curve of DR events for a. all post-menopausal women in in TACT years 5 onwards, b. post-menopausal women in RMH (years 5 onwards) and c. pre-menopausal women in RMH (year 5 onwards) (vertical yellow dashed lines to demarcate 5-10 year cut-off) in RMH validation. CTS5 risk categories are shown with risk of distant recurrence at 10 years with corresponding 95% CI

Altogether, this data demonstrates that CTS5 can be extrapolated to a higher stage post-menopausal cohort. Similarities in performance between this dataset and the original development set as well as other tests of validation point to the robust nature of CTS5 as a prognostic model. Further, it suggests that CTS5 is an appropriate tool to be using at the 5-year mark as it divides a cohort into distinct risk groups which is helpful for clinical management. With a mean risk of recurrence of 3.6% over years 5-10 the guidance that

CTS5-low risk women can safely avoid extended endocrine therapy is applicable even in the context of women with higher stage disease.

4.5.5 Pre-menopausal cohort characteristics

Just under a quarter of pre-menopausal women were under the age of 40. Node-negative cancers comprised 17% of the cohort and, as with post-menopausal women, none of these were deemed high risk by CTS5. Again, like post-menopausal women, a third of the cohort had a tumour >3cm and just under half had grade 3 cancers.

No direct comparison with the ATAC and BIG 1-98 datasets is possible here owing to the absence of pre-menopausal women in these two studies. Comparisons with the pre-menopausal RMH validation population shows a similar median age but a higher proportion of women under 50 in the TACT study. Median tumour size was higher in TACT (25mm v 22mm) and the proportion of node negative cancers was lower (15% v 60%). Event rate in years 5-10 in this TACT study was 12.5% in comparison to 6.5% for the pre-menopausal cohort in the RMH study.

4.5.6 CTS5 validity in pre-menopausal women

This study demonstrates that CTS5 has prognostic value for late recurrence in this cohort for continuous score. However, there was a significant overlap in risk for the low and intermediate risk groups. Distinction between low and high-risk groups was clear but this is perhaps not the most clinically helpful distinction. This lack of separation of curves is seen in both the 5-10 and 5+ year endpoints and when considering HER2- only women.

Large discrepancies in observed versus expected events were seen for women in lower quintiles/risk group and those in the high-risk group. Poor calibration in the lower quintiles would be practically very problematic for this model given its main use. Under-estimation of risk in the 'CTS5 low' group could lead to women inappropriately stopping their endocrine therapy. Since nodal status, tumour size and grade are considered within the CTS5 model, if valid for pre-menopausal women, the model should accommodate this and

therefore expected events would reflect the higher risk of this sub-group. Additionally, observed and expected events within the post-menopausal cohort were not as markedly different despite these women also being of higher stage disease. One explanation could be that the tumour size threshold of 30mm had a bigger effect in the pre-menopausal cohort where a larger proportion of women had tumours >30mm than in the post-menopausal cohort suggesting that capping of tumour size is not appropriate in this population. A further reason could be that the age input into the model for this young, pre-menopausal cohort produced an inappropriately low output from the CTS5 algorithm that did not adequately reflect risk. This is supported by the observation that 26% of this sub-group were classified as low risk compared to only 17% of post-menopausal women.

Altogether, CTS5 has good prognostic value as a continuous variable for high stage pre-menopausal women however calibration was inaccurate and discrimination between low and intermediate risk women was poor and this limits its clinical usefulness. It is therefore impossible to standardise recommendations based on CTS5 groups as can be done for the low-risk group of post-menopausal women. This would have been particularly useful for pre-menopausal women (where the majority would have received tamoxifen upfront rather than an AI), the potential benefit from extended endocrine therapy based on trials of this nature[29,30,137] is greater. One must bear in mind however, that this test of validation focuses solely on high stage pre-menopausal women and therefore cannot prove that CTS5 is invalidated for all pre-menopausal women. Further tests of validation would be needed.

4.5.7 Comparison with RMH validation

In both this study and the RMH validation, CTS5 failed to produce distinct risk groups in terms of clinical outcome (Figure . In the RMH study, the event curves separated better in the 5+ years analysis than the 5-10 which was not the case in this TACT validation. Additionally, in the pre-menopausal RMH cohort, calibration of observed versus expected events was more concordant than that seen here. The reason for this is likely to be the differences in disease stage and age of women between these two validation datasets. The pre-menopausal women in TACT would have been expected to have a greater number of events than the ATAC/BIG1-98 population, thereby leading to discrepancies in calibration.

Bearing in mind this TACT study comprises a higher number of pre-menopausal women and events, the TACT validation dataset is probably more reliable for this demographic, particularly for those with high stage disease.

4.5.8 HER2 negative subgroup analysis

The primary analysis population included women with HER2+ disease as the numbers within this subgroup were low and such patients were included in the original development and the RMH validation sets. The performance of CTS5 was improved in this subgroup which is to be expected as by eliminating HER2+ cases enriches the dataset with women with ER+ disease that behaves in a more 'classical ER+' fashion compared to ER+, HER2+.

4.5.9 Strengths and Limitations

4.5.9.1 Strengths

This study sought to answer the question of whether CTS5 was valid in higher stage women including a large proportion of pre-menopausal women. The trial population had the key baseline features to address this question with an adequate number of patients and events to draw accurate conclusions.

4.5.9.2 Limitations

This study was retrospective in nature and as such problems such as recall bias and an outdated cohort must be taken into consideration. In particular, the TACT study was comprised of younger women (in both the post- and pre-menopausal cohorts) due to the tendency, at that time, to not offer chemotherapy to those over 65. If such a study were repeated today, the age of the population would undoubtedly be different.

It was an aim of this study to further test validity in pre-menopausal women owing to inconclusive results from the RMH study. With the absence of pre-menopausal women in the development and initial validation datasets, the question arises about whether testing it at all in pre-menopausal women is appropriate. One could argue that a model could never be expected to perform well in a cohort which was completely unrepresented in its

development. However, from a clinical standpoint, the applicability of CTS5 for pre-menopausal women is a valid question and many would argue that addressing it in this way is a necessary step to move forward.

It was initially intended that this study would be able to evaluate the validity of CTS5 in women with ER+/HER2+ cancer. However such an analysis was not possible given the unknowns regarding HER2 targeted therapy for women in TACT. Owing to the effect and its duration of trastuzumab on recurrence any conclusions we could draw about the validity of CTS5 in the TACT HER2+ women would now most certainly be invalid and out of date. In retrospect it perhaps would have been more appropriate to use an adjuvant study of HER2 targeted therapy in which to address this question. This is something that merits future investigation.

4.5.10 Clinical Application

For post-menopausal women, CTS5 can continue to be used to predict risk of distant recurrence beyond 5 years and can also be used to guide extended endocrine therapy recommendations. Clinicians can be confident that despite women possessing higher stage features of their cancer, these variables are included in the model and women deemed low risk by CTS5 can be recommended to stop endocrine therapy.

For high stage pre-menopausal women, based on data presented here, CTS5 in its current state cannot be deemed clinically valid. Whilst being prognostic of DR this is not sufficient for a prognostic model and would not be clinically useful in the setting for which it is intended. Whilst it can give an idea of DR risk, it cannot be relied upon for accurate risk prediction and certainly cannot be used to guide extended endocrine therapy decisions.

4.5.11 Future Work

Although CTS5 was not valid in the pre-menopausal cohort analysed here, it is still potentially applicable to pre-menopausal women with lower stage disease, which is the more predominant phenotype in the general population of early breast cancer sufferers. CTS5 should be evaluated in a population of pre-menopausal women to ascertain if this is

the case. If CTS5 continues to perform poorly for pre-menopausal women then efforts should be made to re-train a similar CTS5 model for pre-menopausal women in a clinical trial cohort such as that from the Soft and Text studies[138]. External validation would then be required.

Additionally, as mentioned above, testing of CTS5 in a ER+/HER2+ population would be clinically useful. Those with dual ER+/HER2+ disease present challenges for management and risk prediction. Establishing whether or not CTS5 can be used in such women, or if CTS5 were adaptable to take into account HER2 status would be a valid research question.

5. Enhancement of CTS5 with molecular biomarkers

5.1 Introduction

It has been demonstrated in chapters 2 and 4 that CTS5 has prognostic value for pre- and post-menopausal women in an unselected non-trial cohort and in a clinical trial cohort with overall higher stage disease. Further risk evaluation is unlikely to be of great value in

women who are classified as CTS5 low risk or high risk with women in these categories being recommended to stop or continue endocrine therapy respectively. This approach is in line with the findings of the clinical utility study reported in chapter 3 which saw 96% of women in the low-risk category recommended to stop endocrine therapy and 84% of women in the high risk category recommended to continue. For women in the intermediate category, however, there was far less uniformity of recommendations and it is these women, and those who are borderline, in whom further risk stratification would be useful.

CTS5 utilises only clinico-pathological variables for risk estimation of late recurrence and does not take into account molecular features of the tumour. In chapter 1, it was discussed that prognostic models such as IHC4+C, Pam50 and EPclin, that incorporate molecular as well as clinic-pathological variables, have prognostic value for recurrence. The latter two, despite being developed for recurrence from year 0, were prognostic for recurrence in years 5-10 as well[84,120]. It would therefore seem reasonable to conclude that women who fall into the CTS5 intermediate category might have their tumour sequenced in one of those assays to further evaluate their risk of recurrence in years 5-10. However, this is not currently approved by NICE nor funded by the NHS in the UK. Furthermore, access to these techniques outside of the developed world and specialist centres would be extremely limited.

In chapter 1 the prognostic value of certain molecular biomarkers for recurrence was discussed including how their prognostic significance varies depending on the methods used for quantification and how, in some cases, their prognostic significance changes in years 5-10 compared with years 0-5[55]. It was concluded that whilst IHC based methods of biomarker measurement show trends towards prognostic significance[56,64,65], these methods lack the precision and coverage of gene expression quantification and in some cases – lack the granularity to produce a range of results[18,55,57,66]. However, IHC based methods can be robust and are more widely accessible than gene expression quantification. Furthermore, studies have shown that quantitative IHC scoring for biomarkers ER, PgR, and Ki67, has a high concordance with mRNA transcript[110] demonstrating that changes in gene expression should translate to the protein expression

level. Employing more quantitative scoring of IHC detected biomarkers such as ER H scoring could be one potential method of improving precision whilst retaining the simplicity and cost-effectiveness of this technique. Additionally, the problems with inter-observer variation that comes with manual scoring involved in IHC could be ameliorated by more objective methods of automated visual analysis which, if shown to be effective, could remove those biases and variability.

The IHC4 model, which was trained and calibrated for risk estimation from year 0 alongside clinical parameters, has shown limited prognostic value for late recurrence[55,81]. Given the data discussed previously, this may be due to the weightings of the input variables within the algorithm no longer being correct rather than the relevance of the biomarkers themselves. It would therefore be reasonable to construct an IHC based model, calibrated for late recurrence, to employ alongside CTS5.

In addition, HER2 is an important prognostic biomarker which could be expected to have an impact on late recurrence in this TACT population owing to the larger proportion of HER2+ cancers than has been seen in previous CTS5 studies (chapter 2). HER2 positivity among ER+ women, as discussed in chapter 4, has been reported to have association with recurrence[139]. Benefit from trastuzumab is similar in ER+ and ER- cancers and persists to ten years or beyond[134,136]. However, it is likely that within the broad definition of HER2 amplified breast cancer, those that express a luminal molecular subtype are more endocrine sensitive and are likely to possess a natural history more in keeping with purely ER+ cancers rather than those that are HER2 enriched [135]. If HER2 status can be shown to be prognostic for late recurrence in addition to CTS5 then addition of this variable could help to improve risk stratification for ER+, HER2+ breast cancer.

5.2 Aims

1. To test the association of protein biomarkers including ER, PgR, HER2 and Ki67 with late DR in the pre- and post-menopausal, ER+ subset of the TACT population.

2. To develop a prognostic model from those factors which provide significant prognostic information to be used in conjunction with the CTS5.

3. To test the additional prognostic value of ER, PgR, HER2 and Ki67 (post-menopausal and pre-menopausal) when added to CTS5.

4. To explore the possibility of re-calibrating CTS5 for dual ER+/HER2+ women.

5. To evaluate an automated method of IHC scoring in the TACT population tissue.

5.3 Methods

5.3.1 Selection of Biomarkers

Prior to embarking on model training and development, some pilot studies were done regarding selection of biomarkers to include in the model. As described in chapter 1 and chapter 4, ER, PgR and HER2 and, to a degree, Ki67 are commonly measured biomarkers in breast cancer, and research cited in chapter 1 implies their ongoing importance in late

recurrence. This warrants their exploration as potential variables within a prognostic model for late recurrence.

The *SCUBE-2* gene is an estrogen regulated gene and encodes a tumour suppressor protein involved in signalling processes that reverse epithelial-to-mesenchymal transition (EMT). This leads to reduced levels of tumour-cell invasiveness[75]. *SCUBE-2* features in several genomic models for breast cancer recurrence. The potential of SCUBE-2 as a biomarker has been explored in a study of 156 samples from patients with breast cancer in which positive SCUBE-2 expression (measured by IHC) was independently associated with reduced distant recurrences however, only 58% of the tumours analysed were ER+, and the median follow-up duration was 44 months[76]. *SCUBE-2* RNA expression has been studied for its impact on recurrence and, along with other estrogen regulated genes, switches from being a biomarker of good prognosis to one of poor prognosis in years 0-5 and 5-10 respectively[74]. I therefore decided to explore the feasibility of incorporating this biomarker with ER, PgR and Ki67 into this IHC based study.

In order to explore this fairly novel IHC biomarker, I developed a number of research questions to evaluate whether this would be a valuable addition to the routinely tested biomarkers in breast cancer.

1. Is SCUBE-2 protein expression dynamic in response to estrogen deprivation and repletion?
2. Does SCUBE-2 expression, as measured by IHC, concord with gene expression levels measured by mRNA quantification?
3. Does SCUBE-2 expression differ between ER positive and ER negative breast cancer samples?
4. Can SCUBE-2 staining produce a wide range of staining scores to create a spectrum of SCUBE-2 positivity?
5. Does SCUBE-2 IHC expression correlate with outcome?

Experiments to address these questions are explained further in the results section of this chapter.

5.3.2 TACT patient population

Approval was sought through the trial management committee of the TACT trial to use TMAs constructed at the time of the study. Tissue from women with ER+ breast cancer who received at least 2 years of endocrine therapy and were alive and DR free at 5 years from randomisation were included . Owing to the construction of the TMAs however, individual ineligible specimens could not be excluded at the tissue staining stage and therefore scores from ineligible patients were excluded only at the data analysis stage.

5.3.3 Laboratory Methods

HER2 had already been stained and scored in a central laboratory as part of the main TACT study using IHC HercepTest and fluorescence in-situ hybridisation (FISH) for samples that were inconclusive by IHC. Three out of the four quadruplicate FFPE TMAs were selected and 5 x 4 micron sections were cut. Sections were stained for Ki67, cytokeratins (CK) 7 and 19, ER and PgR using routine IHC techniques. This process is described in detail in Appendix 9.1 but briefly, involves deparaffinising FFPE sections in xylene and graduated alcohols, low pH heat-based antigen retrieval and automated staining using the DAKO Autostainer. Counter-staining was performed using haematoxylin and sections were then dehydrated in graduated alcohol and xylene and mounted using xylene. In addition to the biomarkers described above to be assessed for prognostic association with outcome, cytokeratins 7b and 19 as well as p63 were also stained for. The cytokeratins enable identification of cancer within a specimen and p63 (a myoepithelial marker) enables a distinction to be made between invasive cancer and areas of in-situ cancer (the latter of which must be excluded from scoring). The ER, PgR and Ki67 biomarkers are hereafter referred to as the “measured biomarkers” or “measured variables”.

5.3.4 Automated visual analysis

Scoring of ER, PgR and Ki67 for both clinical and research purposes is usually done manually. This is a time-consuming process that is prone to bias and variation across centres and between individual scorers. For this project, it was decided to explore a

different approach to scoring by way of an automated visual analysis computer program. For the first 7 TMAs, two of the three stained triplicates were taken forward for scoring with the third only to be scored if there was considerable variation. For the remaining 6 TMAs one replicate of the TMA was scored. The stained slides were scanned into the computer using the Hammamatsu Nanozoomer. Scanned images were then imported into software for scoring.

The software selected for the visual analysis of biomarker staining for this study was developed by the company Visiopharm. This software uses individual applications for scoring of different biomarkers. A different application exists for each individual biomarker and they are built to recognise different cell types depending on factors such as shape and size and detects staining intensity and brightness. Each application is adaptable by the user to allow for variability in staining protocols and desired thresholds for scoring. Within this software, a programme for virtual double staining was employed for detection and scoring of the chosen biomarkers. This programme enables alignment of serial sections for scoring. The CK and p63 stained images are aligned with the image stained for the biomarker in question and the software marks out regions of interest (ROI) and areas of exclusion. Ki67, ER and PgR were then scored only in areas of invasive tumour. For those TMAs where duplicates were scored, the mean of the two scores was taken forward to data analysis.

5.3.5 Visual analysis software biomarker applications

As stated above, Visiopharm software has an application for many biomarkers in the context of the disease phenotype being studied including many commonly used cancer biomarkers. Each application uses complex computer algorithms to identify the biomarker in question.

5.3.5.1 CK7/19 and p63

Two colour deconvolution bands are used as pre-processing steps to enhance tissue and tumour regions. The classification of the tumour and non-tumour region is based on an intensity based threshold. The classifier segments the tissue into three parts: tumour, non-tumour and background corresponding to blue-, green- and clear-label, respectively. Lastly,

the tumour-, non-tumour and clear-labels are expanded, thereby closing any clear-holes in the tumour- and non-tumour labels.

5.3.5.2 Ki67, ER and PgR

The application for Ki67, ER and PgR all work within the ROI identified by the invasive tumour detection app. Within each ROI, nuclei are detected. Ki-67 is classified as positive or negative. ER and PgR are detected and classified as negative or 1+, 2+ or 3+ positive. This is done using colour deconvolution where the brown DAB stained positive nuclei are segregated from the blue and red colour of the unstained, negative nuclei. Additionally, nucleus separation is performed using a method based on shape, size and nucleus probability. For ER and PgR, positive nuclei are then subdivided into three categories depending on their staining intensities. For all three applications, nuclei that are too small are removed and then the resulting images are used to generate the output variables.

The output of the ER and PgR application is the H Score. This is calculated from the percentage of nuclei classified as 3+, 2+ and 1+ multiplied by the grade of intensity:

H Score = (percentage 3+) x 3 + (percentage 2+) x 2+ (percentage 1+)

The number of each grade of positive nuclei is also given along with the positive percentage of all positive nuclei. The output from the Ki67 application is number of negative nuclei, number of positive nuclei, total nuclei and proliferation index percentage. Any result with a total nucleus count of <50 was excluded.

ER score is shown as H score (as described above). PgR is given as percentage of positive cells. Ki67 is also scored as percentage positive as is log transformed. This is done consistently throughout the literature due to the fact that it is not normally distributed.

5.3.6 Visiopharm analysis workflow

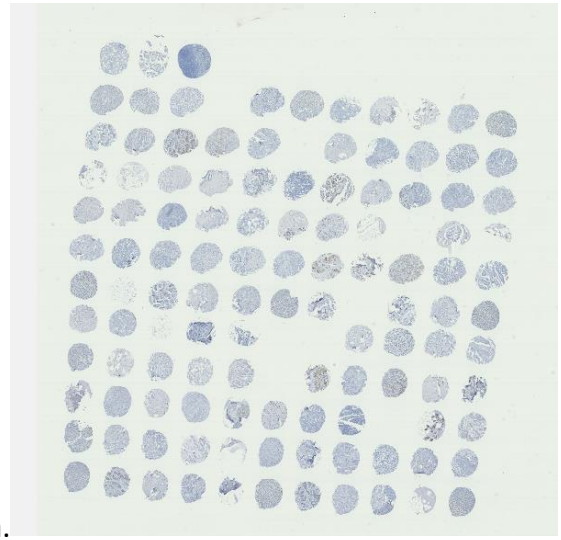
Below the steps for the visiopharm image analysis using virtual double staining for TMAs is described.

1. 3 scanned images of TMAs stained for CK7/19, p63 and exploratory biomarker were loaded into the tissue-array tab on the visiopharm software
2. An array grid was drawn over the top of the image so that each core within the image was labelled according to its position on the TMA (Figure 5.3a) and the grid was manipulated to fit as accurately over the cores on the image as possible. This was repeated for all three TMAs
3. The 3 TMAs were linked together by overlaying the three images – the exploratory biomarker on the top followed by the CK7/19 and then p63 (Figure 5.3b).
4. The linked images were de-arrayed where each core was separated.
5. The user manually checked the positional alignment of the three cores (Figure 5.3c)
6. The individual linked cores were loaded into the image analysis program on the software. The invasive and non-invasive tumour apps were run to detect epithelial cells and myoepithelium in order to define ROI and areas of exclusion (Figure 5.3d)
7. The exploratory biomarker app is run to identify presence of the biomarker within the ROI only (Figure 5.3e)

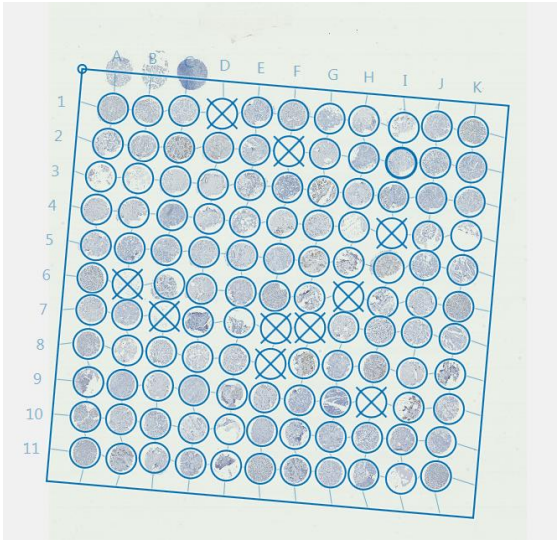
5.3.7 Biomarker application optimisation

The Visiopharm software allows a user to tune the applications and alter certain of the thresholds within the processing steps. The Visiopharm software was not routinely used in this laboratory and therefore, in order to optimise the software parameters for my staining protocols, I undertook a pilot run of the staining on a TMA from an independent source separate from that of the TACT trial TMAs. The purpose of this was to perform the staining in the same way as was planned for the TACT study TMAs and then carry out the automated visual analysis. Sections from a single TMA of ER+ early breast cancer were stained for the 5 previously mentioned biomarkers and then analysed using the in-built applications within the Visiopharm software package described above. Scoring was then reviewed and a visual impression was gained as to the sensitivity and appropriateness of the scoring in a qualitative manner. As a result of this review, changes could be made to the scoring thresholds within the processing steps for each biomarker application (see results 5.4.1).

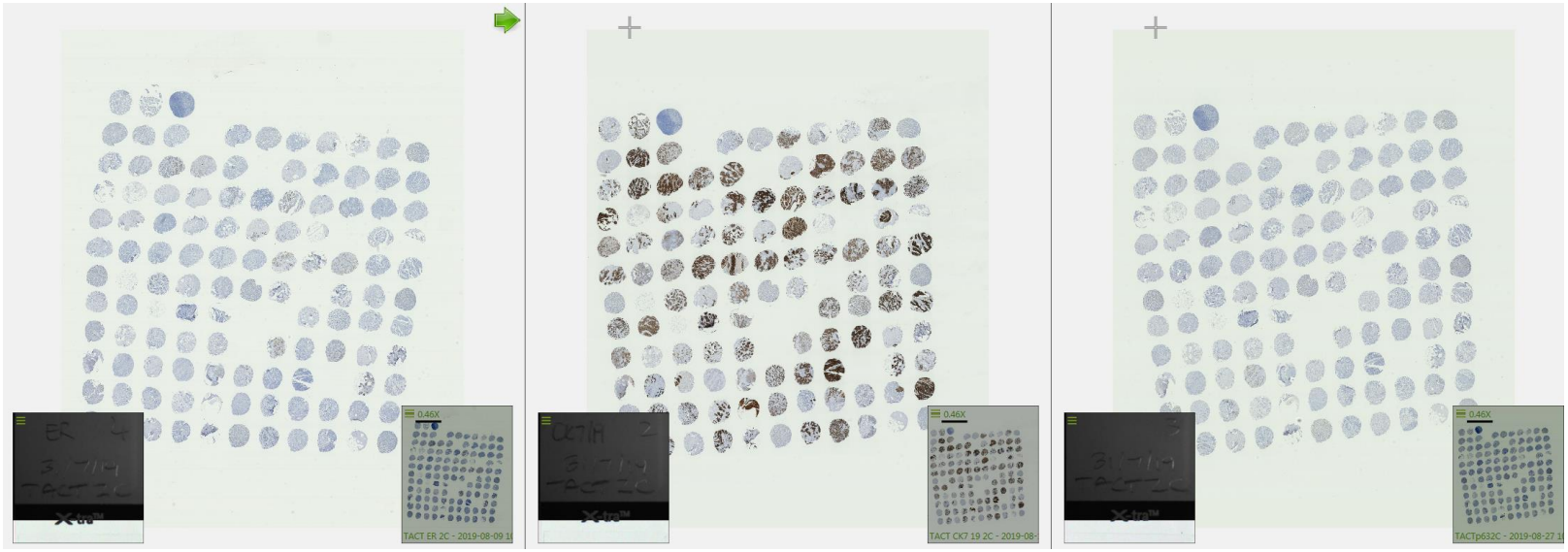
a.



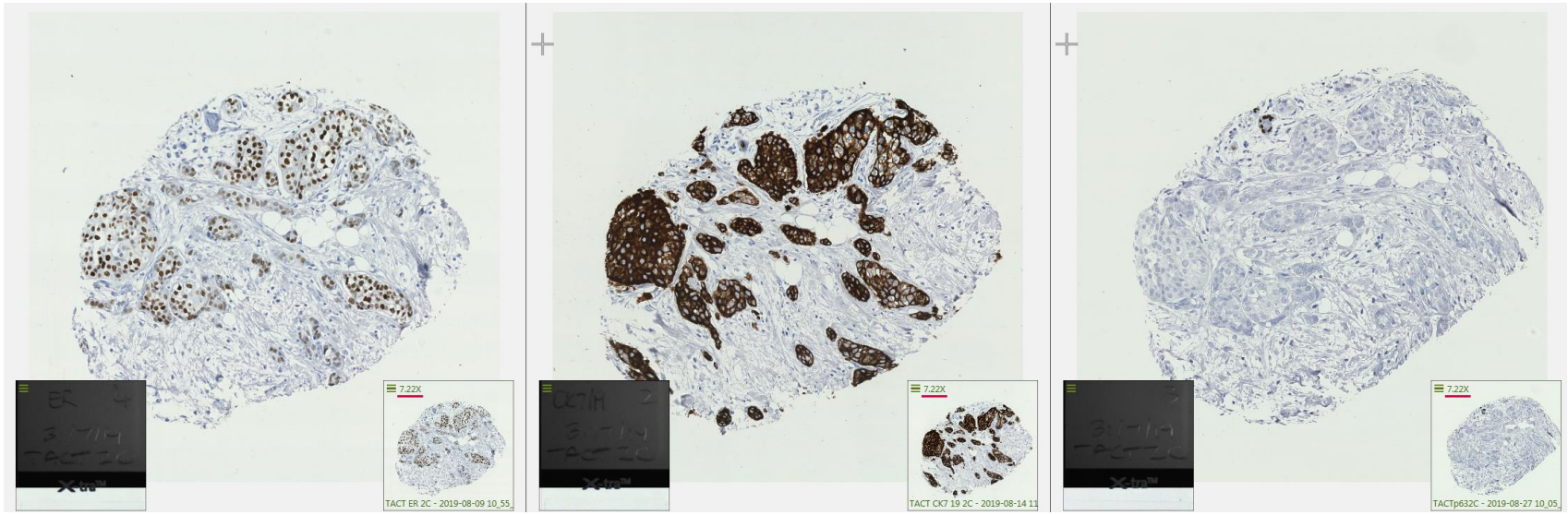
b.


A
B
D
E
F
G
H
I
J
K
1
2
3
4
5
6
7
8
9
10
11

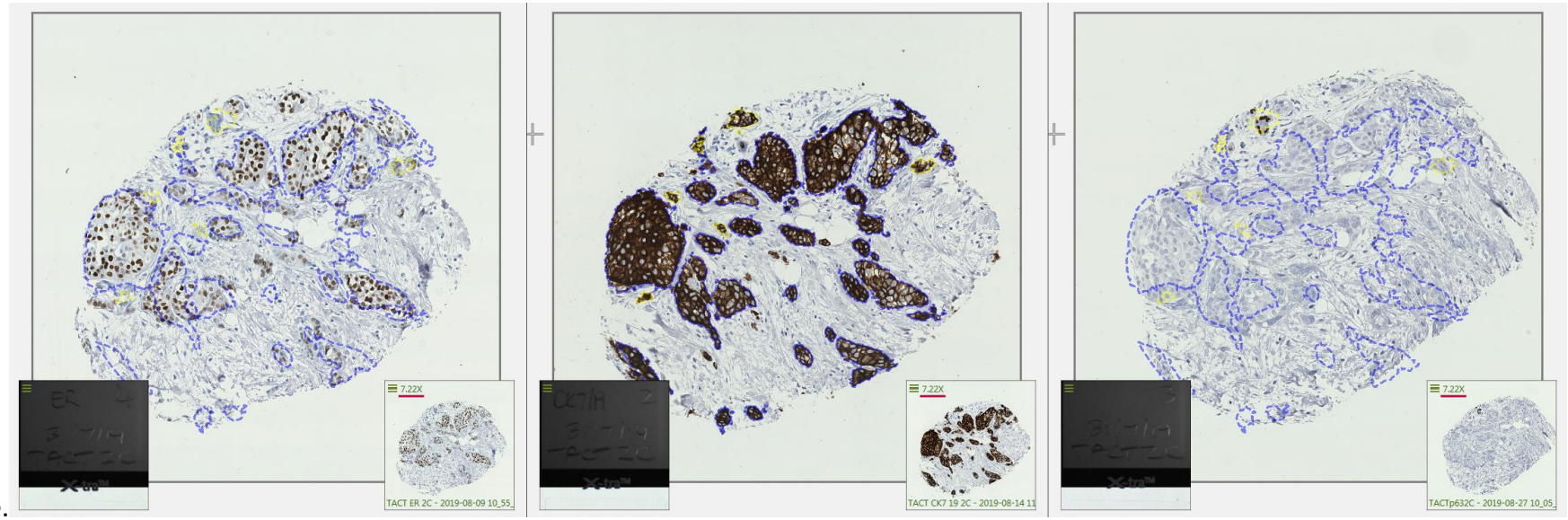
c.


0.46X
TACT ER 2C - 2019-08-09
0.46X
TACT CK7 19 2C - 2019-08-
0.46X
TACTp632C - 2019-08-27

d.


7.22X
TACT ER 2C - 2019-08-09 10_55_
7.22X
TACT CK7 19 2C - 2019-08-14 11
7.22X
TACTp632C - 2019-08-27 10_05_

e.


7.22X
TACT ER 2C - 2019-08-09 10_55_
7.22X
TACT CK7 19 2C - 2019-08-14 11
7.22X
TACTp632C - 2019-08-27 10_05_

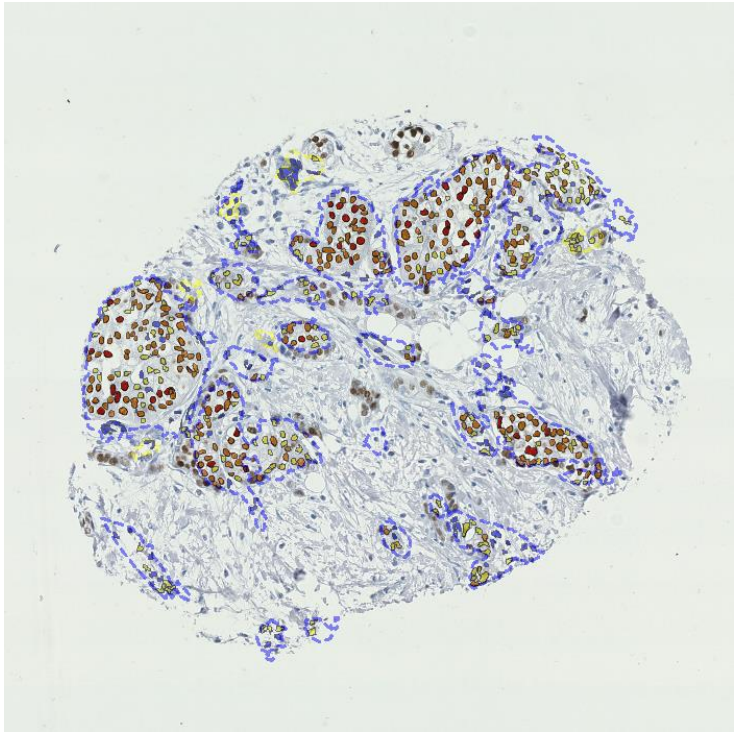
f. 

Figure 5.1 **Images generated from Visiopharm workflow: a. Scanned image of stained TMA section. b. TMA image with superimposed grid. c. 3 linked images of the same array stained for (from left to right) ER, CK7/19 and p63. d. 3 linked, individual cores (following manual checking of alignment) stained for (from left to right) ER, CK7/19 and p63. e. 3 linked cores as in image D following application of invasive (blue outline) and non-invasive tissue (yellow outline) marking, f. ER stained section following application of ER scoring (red and orange dots) only in the blue marked regions of interest and not in yellow marked regions of exclusion.**

5.3.8 Statistical Considerations

The primary endpoint of the study was time to DR in years 5+ for the whole population. DR was defined as metastatic disease at a distant site and excludes local, loco-regional and contralateral recurrence. Local, loco-regional or contralateral recurrences in years 0-5 were not excluded from the analysis. Local, loco-regional or contra-lateral recurrences in years 5+ were not treated as events and such patients were censored at last follow up. Second primary cancers in years 5+ were not treated as events and patients were censored at the date of confirmation of the second primary cancer. Death in years 5+ without distant recurrence was a censoring event.

The first secondary endpoint was time to distant recurrence (in years 5+ following diagnosis) in which patients with local, loco regional or contralateral recurrences in years 0-5 were excluded from the analysis. Local, loco-regional or contra-lateral recurrences in years 5+ were treated as events. Patients with second primary cancers in years 0-5 were excluded from the analysis. The other secondary endpoint was time to distant recurrence in years 5-10. These two endpoints were evaluated for both the whole pre- and post-menopausal populations and for the HER2- only subgroup.

The primary population for analysis was the post-menopausal cohort and the secondary population for analysis was the pre-menopausal cohort.

5.3.9 Statistical Methods

Cox regression modelling was used to show association of individual biomarkers with time to late DR. A univariate analysis was done as well as multivariate with these variables alone and in combination with CTS5. A hazard ratio for DR with corresponding 95% CI for each variable is given. Additionally, using regression modelling, a combined model of CTS5 and the biomarkers was fitted. In this model development, CTS5 was present at the start and each variable was added to the model individually. The prognostic value of each of the IHC models in addition to CTS5 compared to CTS5 alone are shown as well as change in

likelihood ratio χ^2 for both the post and pre-menopausal cohorts for each endpoint. A hazard ratio is given for risk of DR for each standard deviation increase in biomarker + CTS5 value.

5.3.10 Post Hoc Analyses

Several post-hoc analyses were performed based on initial review of the findings, Firstly, the association between the 4 biomarkers and early recurrence (years 0-5) within the post-menopausal, ER+ TACT population was analysed. For this analysis, all ER+ patients randomised within TACT who received endocrine therapy were included. Events were distant recurrences in years 0-5 (local, loco-regional and contralateral recurrences were not counted as events). Deaths in years 0-5 without distant recurrence was censored.

Additionally, it was decided that evaluation of correlations between the biomarkers and certain clinical variables would be valid. This was deemed a useful method of interrogating the IHC scoring data and establishing whether expected correlations between biomarkers were evident in this dataset. If expected correlations were not found one might be more likely to question the validity of the scoring data.

5.4 Results

5.4.1 Exploration of SCUBE-2

5.4.1.1 Question 1 – Is SCUBE-2 protein dynamic in response to estrogen deprivation and repletion?

MCF7 wild type (WT) cell lines which are estrogen dependent were used to test the hypothesis that SCUBE-2 is a dynamic protein. Four MCF7-WT cells were cultured in phenol red-free RPMI and grown to 60% confluence. They were then supplemented with 10% dextran coated charcoal (DCC) stripped serum to remove exogenous estrogen (the RPMI and DCC combination is referred to as DCC medium). Two lines were maintained in this DCC medium for 24 and 48 hours (to simulate estrogen deprivation therapy) and then RNA was extracted at those 2 time points. The other two cell lines were maintained in DCC medium and in addition treated with 10nM estradiol (E2) for 24 and 48 hours to simulate an untreated, estrogen replete cancer. RNA was then extracted at those 2 time points. It was expected that in comparison to the estrogen deprived cell lines, the expression of *SCUBE-2* would increase in the estradiol replete cell lines however the opposite was the case (Figure 5.2).

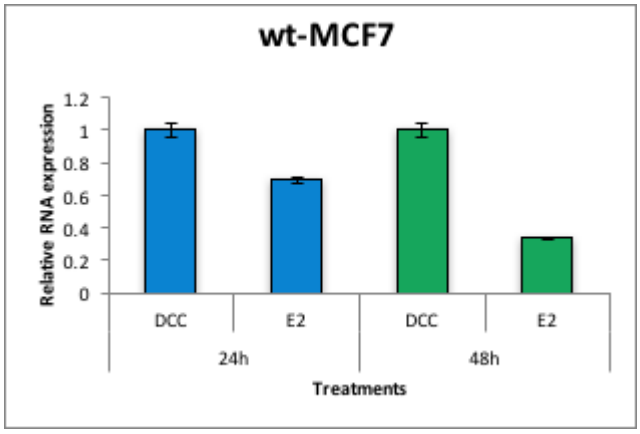


Figure 5.2 Relative RNA expression of SCUBE-2 at 24 and 48 hours incubation with either DCC or estradiol.

Simultaneous to the RNA extraction, at each time point cells were removed from all four cell line cultures to make a cell pellet which was then formalin fixed and paraffin embedded in order to test for the protein using IHC. Staining was carried out using AbCam antibody 117625 at a concentration of 1 in 1000 using a DAKO autostainer. Qualitative assessment of the staining demonstrated greater intensity and extent of staining in the estrogen

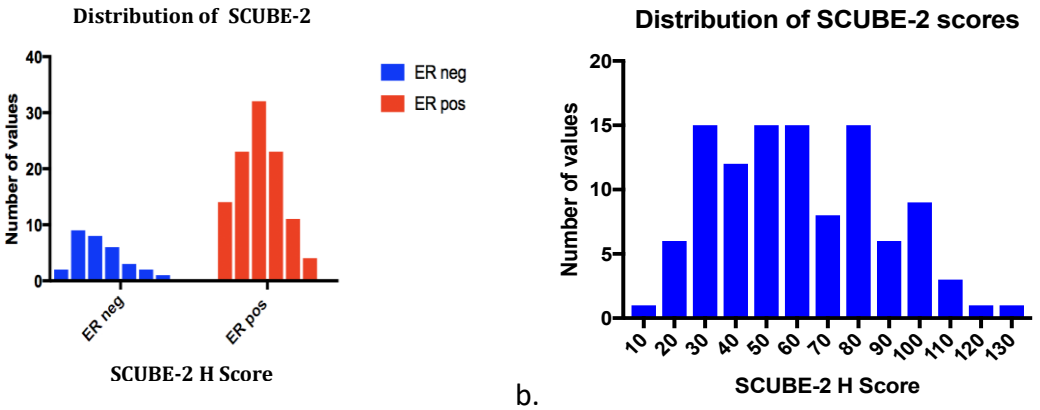
deprived cell pellet than the E2 replete pellet – concordant with the above RNA data but contrary to biologic theory.

5.4.1.2 Question 2 – Does SCUBE-2 expression, as measured by IHC, concord with gene expression levels measured by mRNA quantification?

To assess the correlation between *SCUBE-2* mRNA expression and protein expression measured by IHC, a tissue micro-array (TMA) from the trans-ATAC study was sectioned and stained for SCUBE-2 and scored using automated visual analysis software. SCUBE-2 IHC score was well correlated with ER H score (Spearman R=0.38, p<0.0001) but there was no significant correlation between SCUBE-2 H score and mRNA expression (previously quantified by this laboratory group from transATAC material).

5.4.1.3 Questions 3 and 4 – Does SCUBE-2 expression differ between ER+ and ER- breast cancer samples? Can SCUBE-2 staining produce a wide range of staining scores to create a spectrum of SCUBE-2 positivity?

To assess the association between SCUBE 2 protein expression (measured by IHC) and ER expression, the same antibody was used to test a TMA containing both ER+ and ER- human breast cancer core cuts. Sections from the TMA were stained and then SCUBE-2 H Score was calculated using Visiopharm visual analysis software. Mean SCUBE-2 H score was 60.77 in the ER+ cores versus 48.14 in the ER- cores (Figure 5.3a). Unpaired t-test showed this difference was statistically significant with a p-value of 0.02. A range of values was obtained for positively stained cores but with a paucity of higher scoring specimens (Figure 5.3a)



a. b.

Figure 5.3 Distribution of SCUBE-2 scores a. in ER- and ER+ cores and b in ER+ cores within the TMA

5.4.1.4 Question 5 – Does SCUBE-2 IHC expression correlate with outcome?

Lastly, to evaluate the prognostic value of SCUBE-2 expression by IHC, FFPE breast cancer samples from the POLAR study[140] were sectioned and stained using standard IHC methods with the same antibody and concentration as above. SCUBE-2 was scored using automated visual analysis on the Visiopharm software and an H score was given. There was no association between SCUBE-2 H score and clinical score and no association between SCUBE-2 H score and recurrence events. It was noted in this experiment that there was a considerable amount of background staining in comparison to what is usually seen with ER and PgR and this may well have led to inaccuracies in scoring using automated visual analysis.

Data generated from this series of experiments suggest that SCUBE-2 is related to ER but probably does not have the same close concordance with mRNA expression. The staining achieved by the AbCam antibody was reasonable but not as consistent as one would want for a prognostic biomarker. Unfortunately, the cell lines experiments showed that SCUBE-2 expression fell in response to estrogen stimulation which was the opposite of what was expected and suggests that cell line models were not useful in exploring this biomarker. The lack of association with outcome and lack of correlation with mRNA expression that resulted from these investigations of optimisation led to the conclusion that SCUBE-2 is unlikely to be a robust IHC biomarker for breast cancer recurrence. Therefore this biomarker was not pursued further within this study.

5.4.2 TMA analysis using Visiopharm software: pilot run for optimisation

A TMA containing 49 cores was stained for the 5 biomarkers and scored by the Visiopharm automated visual analysis software. On qualitative review of the scoring outcomes and comparison to manual scoring it was noted that all applications except for Ki67 and PgR were under-scoring. As a result, several alterations were made to the other applications to allow for this observed difference.

1) Non-invasive tumour detection app
    a. Lowered the threshold for staining intensity for a positive nucleus

2) Invasive tumour detection app
    a. Lowered the threshold for staining intensity for a positive nucleus
    b. Altered the processing steps which close the gaps between positively stained cells to enable more confluent ROIs to be generated
3) ER app
    a. Lowered the threshold for staining intensity of a 1+ nucleus

The amended applications were saved in the software and were then taken forward in their adapted form for use in the main study.

5.4.3 TACT TMAs

Tissue from 1412 women enrolled in the TACT study was incorporated into TMAs at the time of the running of the trial[128]. 116 women were excluded prior to scoring due to insufficient tissue. 1296 patients had their tissue scored for ER, PgR and Ki67 (as well as the CKs and p63). A further 112 patients were excluded after scoring as their core contained fewer than 50 cells. A total of 1184 samples had a score for at least 1 biomarker and then exclusions were made for ER- patients and those who had received insufficient endocrine therapy (Figure 5.4).

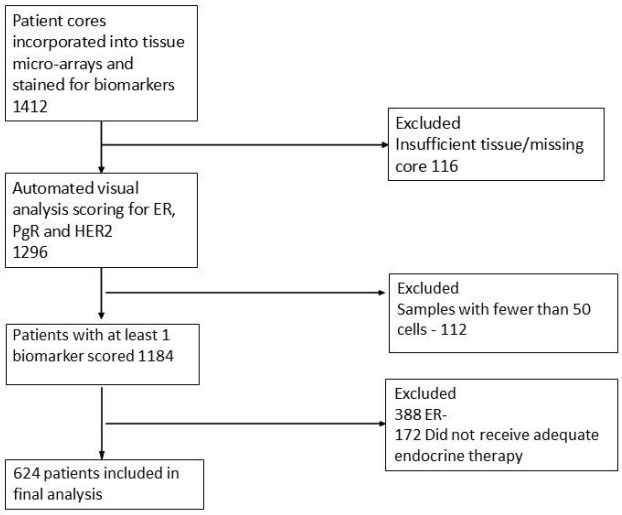


Figure 5.4 CONSORT diagram of TACT IHC study

5.4.4 Variation between TMAs

Of the 13 TMAs included in this study, 7 were scored in duplicate. TMAs A and C were compared for their degree of correlation. For Ki67, the spearman rank correlation co-efficient was 0.56 (p=<0.0001), for ER it was 0.78 (p=<0.0001) and for PgR it was 0.85 (p=<0.0001) (Figure 5.5)

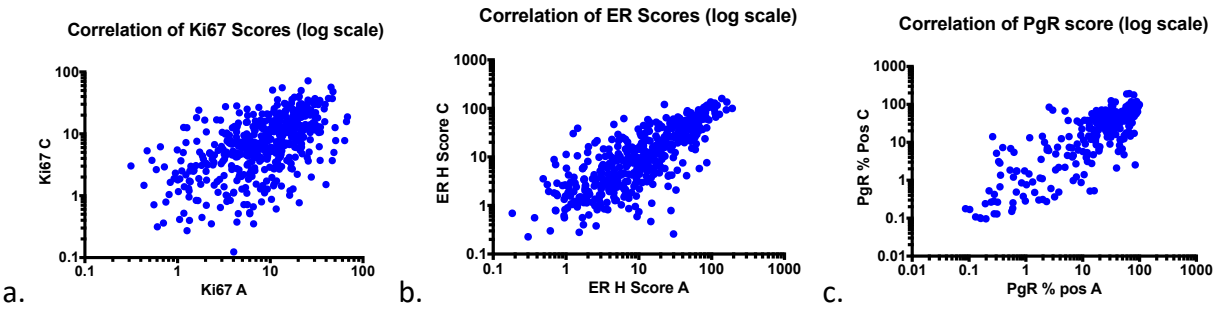


Figure 5.5 Correlation of biomarker scoring between TMA A and TMA C for the 7 TMAs scored in duplicate for a. Ki67 % positive, b. ER H score and c. PgR % positive.

5.4.5 Correlation between automated and manual scoring of Ki67

Automated visual analysis with the Visiopharm software package is not routinely used for scoring IHC stained slides in this laboratory. Therefore, as well as the optimisation of the individual applications as described above, I also undertook a comparison of manual versus automated scoring of Ki67 on a selection of cores. A total of 24 cores from different TMAs were selected on a visual basis. Selection was not random as it was intended that a broad range of Ki67 results be compared. Selected cores were manually scored by a senior scientist in the laboratory, experienced in Ki67 scoring for IHC (Figure 5.6).

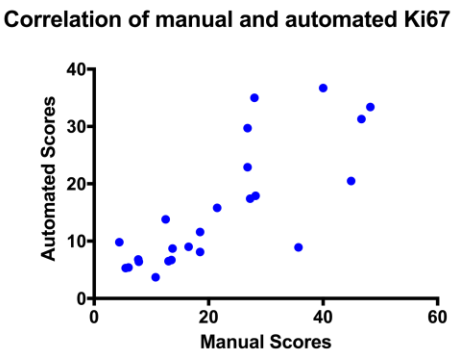


Figure 5.6 Correlation of a sample of manual and automated Ki67 scores

Overall, the automated scoring was lower than the manual scoring. This was the case for all but two of the cores. A positive correlation was shown between the two sets with a Pearson correlation coefficient of 0.7959 ($p < 0.0001$).

5.4.6 Patients in main IHC biomarker study

Biomarker expression data was generated on at least 1 of the 3 IHC biomarkers that had been newly stained in this study for 330 post-menopausal women and 294 pre-menopausal women with ER+ breast cancer. Baseline characteristics of these patients are shown in Table 5.1. Characteristics for both the pre- and post-menopausal cohorts matched those seen in the CTS5 validation study (see chapter 4). Most patients were node-positive with tumours greater than T2 stage.

	Patients with ER H score, PgR positive cells or Ki67 available		
	ER+, postmenopausal, with adequate endocrine therapy (N=330)	**ER+, premenopausal, with adequate endocrine therapy (N=294)**	**ER+, with adequate endocrine therapy total (N=624)**
Age (years), median (IQR)	55.0 (51.2, 58.6)	44.3 (40.4, 48.1)	49.9 (44.0, 55.4)
Age group (years)			
<40	5 (1.5)	67 (22.8)	72 (11.5)
40-49	55 (16.7)	186 (63.3)	241 (38.6)
50-59	207 (62.7)	41 (14.0)	248 (39.7)
≥60	63 (19.1)	0 (0.0)	63 (10.1)
Nodal status			
Node negative	31 (9.4)	57 (19.4)	88 (14.1)
1-3 positive nodes	180 (54.6)	158 (53.7)	338 (54.2)
≥4 positive nodes	119 (36.1)	79 (26.9)	198 (31.7)
Tumour grade			
Grade 1	38 (11.5)	22 (7.5)	60 (9.6)
Grade 2	186 (56.4)	150 (51.0)	336 (53.9)
Grade 3	106 (32.1)	122 (41.5)	228 (36.5)
Tumour size			
≤2cm	140 (42.4)	110 (37.4)	250 (40.1)
2-5cm	167 (50.6)	164 (55.8)	331 (53.0)
>5cm	23 (7.0)	20 (6.8)	43 (6.9)
HER2 status			
HER2+	60 (18.2)	54 (18.4)	114 (18.3)
HER2-	258 (78.2)	224 (76.2)	482 (77.2)
HER2 unknown	12 (3.6)	16 (5.4)	28 (4.5)

Table 5.1 Baseline characteristics of pre- and post-menopausal women within the IHC study

5.4.7 Biomarker scoring for the post-menopausal cohort

Scores for the three IHC biomarkers are shown in Table 5.2 and Figure 5.7. A wide range of scores was evident. HER2 results (taken from a previous translational study of the TACT trial) showed: n 1034, negative 758 (73.3%), positive = 143 (13.8%), unknown = 133 (12.9%).

	ER H-score	PgR Positive Cells	Ki67 (%)	Ln (Ki67 + 0.1)
N	318	318	320	320
Median	35.9	21.6	6.6	1.9
IQR	12.5 - 66.9	2.2 - 50.4	2.8 – 11.5	1.1, 2.5
Min-Max	0.2, 190.1	0 - 95.4	0 - 71	-2.3, 4.3

Table 5.2 IHC based scores for ER, PgR and Ki67 by automated visual analysis

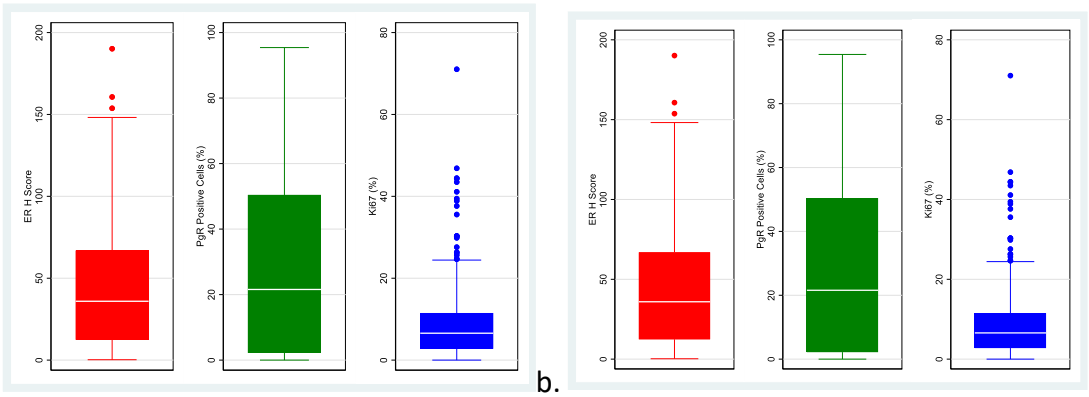


Figure 5.7 Boxplots showing distribution of the 3 continuously scored biomarkers in the a. post-menopausal population and b. pre-menopausal population. Note – here Ki67 % is not log transformed to demonstrate true distribution of the biomarker.

5.4.8 Impact on late recurrence for the post-menopausal cohort

For the primary endpoint none of the 4 biomarkers in the study (ER, PgR, Her2 and Ki67) had a significant independent association with late DR. Women with HER2- cancers were more likely to have a late recurrence compared to HER2+ women but this was not statistically significant in univariate or multivariate analysis (Table 5.3).

	Univariate analysis	Multivariate analysis (without CTS5)	Multivariate analysis (with CTS5)
ER H Score	1.00 (0.99-1.01), p=0.94	1.00 (0.99, 1.01); p=0.95	1.00 (0.99, 1.01); p=0.87
PgR % positive	1.00 (0.99 – 1.01), p=0.95	1.00 (0.99, 1.01); p=0.89	1.00 (0.99, 1.01); p=0.98
HER2 status (positive/negative)	0.64 (0.37-1.10), p=0.10	0.78 (0.35, 1.72); p=0.54	0.61 (0.27, 1.37); p=0.23
Ln (Ki67 + 0.1)	1.03 (0.81 – 1.31), p=0.81	1.08 (0.83, 1.41); p=0.55	1.01 (0.78, 1.32); p=0.93

Table 5.3 Individual IHC biomarkers within the TACT ER+ post-menopausal population and association with recurrence beyond 5 years

Further modelling tested the HR for recurrence for CTS5 adjusted for the presence of the measured variables and also looked at the impact of each of the measured variables on the prognostic power of the model (Table 5.4). CTS5, when adjusted for ER, PgR and Ki67 individually remained significantly associated with outcome and this is the case regardless of the scores for those measured biomarkers. This was reflected in the LR χ^2 value demonstrating the overall additional prognostic impact each variable had on the CTS5 value which was not significant for ER, PgR and Ki67. For HER2 status, CTS5 (adjusted for this variable) was significantly associated with outcome and HER2 status *does* impact on performance of the model. Additionally, this impacted on the overall performance of the CTS5 model reflected in the LR χ^2 value (6.48; p=0.01) (Table 5.4).

Analysis for the 5-10 year endpoint showed very similar results with HER2+ status being the only biomarker to show a trend with late distant recurrence; HR=0.52 (0.26-1.03), p=0.06 however this was not borne out in multivariate analysis. HER2+ status added to the prognostic value of CTS5: HR=0.42 (0.21-0.83), p=0.01 whereas none of the measured biomarkers impacted on the performance of CTS5. Results in the HER2- only population, were similar with no biomarkers significant for late recurrence (data not shown).

	Model with CTS5 score (continuous) and each IHC4 component separately	LR test chi2; p-value for inclusion of IHC variable into CTS5 score model
CTS5 score	2.75 (1.83, 4.11); p<0.001	0.04; p=0.83
ER H Score (continuous)	1.00 (0.99, 1.01);p=0.83	
CTS5 score	2.64 (1.78, 3.92); p<0.001	0.07; p=0.69
PR positivity (continuous)	1.00 (0.99, 1.01); p=0.69	
CTS5 score	2.09 (1.64, 2.67); p<0.001	6.48; p=0.01
HER2 (positive)	0.52 (0.30, 0.90); p=0.02	
CTS5 score	2.53 (1.72, 3.71); p<0.001	0.01; p=0.94
Ln (Ki67 +0.1)	1.01 (0.79-1.28); p=0.94	

Table 5.4 Hazard ratios for recurrence for: CTS5 adjusted for each variable (ER, PgR, HER2 and KI67) – green cells, estimate of effect of variable on the CTS5 score – orange cells and LRχ^2 for overall impact on the model for each variable for post-menopausal subgroup.

In addition, the association of all 4 biomarkers were evaluated for their impact on early recurrence. Women with HER2+ cancer were *more* likely to have an early recurrence; HR=1.68 (1.23-2.31), p=0.001 than HER2- women (contrary to the association seen for late recurrence).

Owing to the lack of prognostic association between the biomarkers, and late recurrence, no further prognostic modelling was attempted within the post-menopausal cohort.

5.4.9 Biomarker scoring for the pre-menopausal cohort

Scores for the 3 measured biomarkers are shown in Table 5.5 and Figure 5.7b. A wide range of scores was evident. HER2 results (taken from a previous translational TACT study) were: n=1179; negative=790 (67.0%); positive=198 (16.8%); unknown=191 (16.2%).

	ER H-score	PgR Positive Cells	Ki67 (%)	Ln (Ki67 + 0.1)
N	291	289	287	287
Median	20.8	34.4	6.2	1.8
IQR	7.5, 43.8	8.6, 63.9	3.3 – 11.1	1.2, 2.4
Min-Max	0.5, 160.1	0, 95.7	0 – 43.4	-2.3, 3.8

Table 5.5 IHC based scores for ER, PgR and Ki67 by automated visual analysis

5.4.10 Impact on late recurrence in the pre-menopausal cohort

As with the post-menopausal cohort, HER2+ cancers were significantly less likely to recur late compared to HER2- and this was true in both univariate and multivariate analysis. (Table 5.6). None of the other three measured biomarkers were found to be significantly associated with outcome.

	Univariable analysis	Multivariable analysis (excluding CTS5 score)	Multivariable analysis (including continuous CTS5 score)
ER H Score (continuous)	1.00 (0.99, 1.01); p=0.96	1.00 (0.99, 1.01); p=0.95	1.00 (0.99, 1.02); p=0.49
PR positivity (continuous)	1.00 (0.99, 1.01); p=0.34	1.00 (0.98, 1.00); p=0.14	1.00 (0.98, 1.01); p=0.42
HER2 (positive)	0.92 (0.62, 1.37); p=0.69	0.17 (0.04, 0.73); p=0.02	0.16 (0.04, 0.66); p=0.01
Ln (Ki67+0.1)	0.98 (0.71, 1.36); p=0.92	1.05 (0.75, 1.47); p=0.79	0.93 (0.66, 1.32); p=0.70

Table 5.6 Individual IHC biomarkers within the TACT ER+ pre-menopausal population and association with recurrence beyond 5 years

CTS5 score (adjusted for the measured variables) was significantly associated with outcome and this was regardless of ER, Pgr and Ki67 (Table 5.7). HR for DR by CTS5 value was impacted by HER2 score (HR 0.79 (0.53,1.18);p=0.02 but, unlike for post-menopausal women, this did not significantly impact on the prognostic value of the model (LR $\chi^2$1.36; p=0.24). When the analysis was performed just on the HER2- women the outcome was similar with no measured biomarkers significant for late recurrence (data not shown).

	Model with CTS5 score (continuous) and each IHC4 component separately	LR test chi2; p-value for inclusion of IHC4 component into CTS5 score model
CTS5 score	2.94 (1.84, 4.69); p<0.001	0.53; p=0.47
ER H Score (continuous)	1.00 (0.99, 1.01);p=0.45	
CTS5 score	2.87 (1.79, 4.61); p<0.001	0.00; p=0.98
PR positivity (continuous)	1.00 (0.99, 1.01); p=0.98	
CTS5 score	2.17 (1.74, 2.71); p<0.001	1.36; p=0.24
HER2 (positive)	0.79 (0.53, 1.18); p=0.02	
CTS5 score	2.84 (1.75, 4.61); p<0.001	0.50; p=0.48
Ln (Ki67+0.1)	0.89 (0.64, 1.23); p=0.47	

Table 5.7 Hazard ratios for recurrence for: CTS5 adjusted for each variable (ER, PgR, HER2 and KI67) – green cells, estimate of effect of variable on the CTS5 score – orange cells and LRχ^2 for overall impact on the model for each variable for pre-menopausal subgroup.

For the 5-10 year endpoint, women with HER2+ cancer were less likely to have a recurrence than HER2-, HR = 0.13 (0.02-0.95), p=0.05 in a multivariable analysis including CTS5. However, in model development, HER2 status did not add significant prognostic value to CTS5. Otherwise, as with the 5+ endpoint, none of the 3 measured biomarkers were significantly associated with outcome nor did they add to the prognostic value of CTS5.

As no prognostic association between the IHC biomarkers and late recurrence was found, no further prognostic modelling, beyond that demonstrated above was attempted within the pre-menopausal cohort.

5.4.11 Correlations of clinical and IHC biomarkers in TACT

To evaluate whether the results obtained were consistent with expectations, correlations between scores of the different biomarkers and also between biomarkers and tumour grade were assessed. Within the combined pre- and post-menopausal populations of the TACT study, ER H score and PgR% showed a weak, negative correlation with Ki67% (Spearman R = -0.09, p=0.004 and -0.21, p<0.001 respectively) (Figure 5.8a and b). Tumour

grade was weakly positively associated with Ki67 (Spearman R=0.22, p<0.001) and moderately negatively correlated with PgR% (Figure 5.8 c and d).

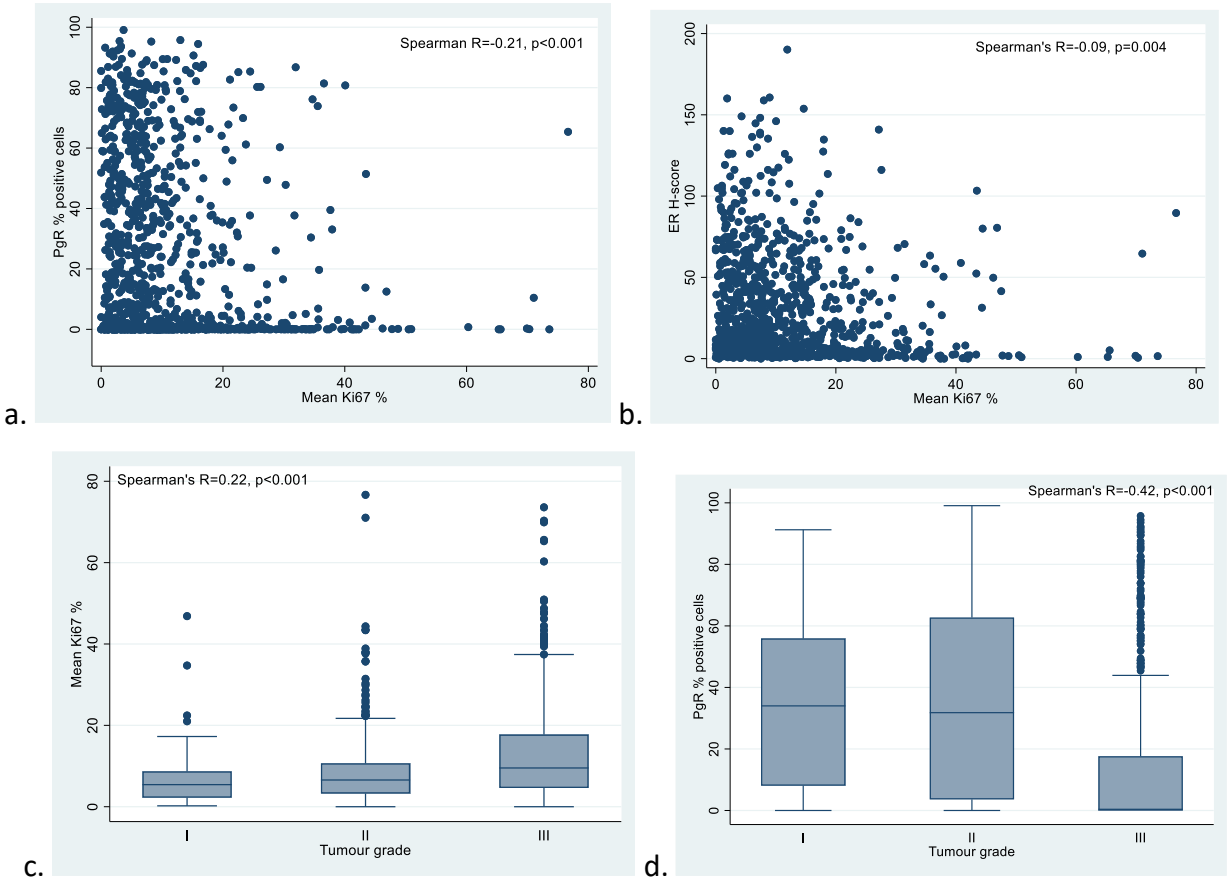


Figure 5.8 Correlations between clinical and IHC biomarkers, a. between PgR% and Ki67, b. between ER H score and Ki67, c. between Ki67 and tumour grade and d. between PgR% and tumour grade.

5.5 Discussion

5.5.1 Sample size

The original sample size was large, however, there was a substantial attrition rate. 1152 cores were stained and scored for at least one biomarker but after exclusions approximately half the patients were lost. This impeded the power of the study owing to reduced event numbers, particularly when the population was further divided into pre-and post-menopausal cohorts.

5.5.2 IHC biomarkers in TACT population

On review of the biomarker scores for the ER+ population, the median ER scores were reduced in comparison to what one would expect. However, these low scores were across the whole set of TMAs and therefore despite lower absolute scores, an overall trend should not have been affected. Additionally, median Ki67 values were lower than expected compared to other populations of ER+ women[141]. In this ER+ high stage population there are likely to be a substantial proportion of women with luminal B type cancers in whom Ki67 scores would be expected to be higher[142]. However, as with the ER scoring, the Ki67 scores tended to be globally low, most likely due to use of automated image analysis, and so trends in Ki67 levels should still have been detectable and therefore should not have affected an overall prognostic effect on recurrence if one existed.

5.5.3 Prognostic value of IHC biomarkers for late recurrence

None of the measured biomarkers, that are known to have shown prognostic value for late recurrence in other translational studies[56,64,65,69] were prognostic for late recurrence in this TACT study for either the post or pre-menopausal cohorts. This was also the case when only women with HER2- cancers were considered. In addition, the measured biomarkers showed no association with early recurrence. This is surprising given the known prognostic value of PgR and Ki67 in ER+ disease, in addition to clinical variables, for early recurrence and prognostic value of the IHC4 model for risk estimation from year 0[61].

The baseline features of the TACT population itself must be considered the primary reason for these findings. Women in TACT were, younger, of higher stage and had all received chemotherapy. This is in stark contrast to the general population of early breast cancer sufferers and also from the datasets in which most prognostic biomarkers have been studied to date[61,64,65] such as transATAC where women were all post-menopausal and chemotherapy treated patients were excluded. Such differences in clinical stage, age and chemotherapy treatment are likely to account for the absence of prognostic association between these biomarkers and both early and late recurrence. The main conclusion to be drawn from this study is that in higher stage women, there may be a prognostic relationship between biomarker expression and late recurrence but this relationship will be modest in strength. As this study was small and consisted of *only* higher stage women it is possible that a prognostic association exists between these biomarkers within higher stage cancers that the study did not have the power to detect. However, even given the caveats described such as high attrition rate and differences in biomarker scoring using automated analysis, a major prognostic effect would have been detectable using methodologies employed in this study.

When designing this study, it was unknown how these variables would behave in this distinct population. A biomarker study, such as the one conducted here should be undertaken in a lower stage, representative cohort of early breast cancer sufferers and may well demonstrate stronger associations with late recurrence. Indeed, in an exploratory analysis of the 989 post-menopausal women from the transATAC dataset, Ki67 had independent prognostic association with late recurrence (HR=1.43 (1.14-1.80), p=0.002. In a prognostic model using the same 4 biomarkers as analysed in this study, an IHC model had prognostic value for late recurrence (HR=1.48 (1.20-1.82), p=0.0003 and this significantly improved the prognostic value of CTS5 ΔLR-χ^2 4.47, p=0.035 (unpublished).

It must be remembered that, in the literature, associations between IHC measured biomarkers and recurrence (both early and late) have been inconsistent and that reliability of biomarker measurement has been improved by molecular methods. It could therefore

be argued that the inherent imprecision in IHC methods despite using quantitative scoring, hampers its use for prognostic purposes in the late recurrence setting.

5.5.4 Prognostic value of HER2

HER2 status was the only biomarker to show a significant association with late recurrence in the post- and pre-menopausal populations. For both the 5+ and 5-10 years endpoint, risk of late recurrence was greater for HER2- cancers in comparison to HER2+. This is consistent with what we know about the biology of HER2+ cancers which are more aggressive if not treated with HER2 directed therapy. In this study, where the majority of women with HER2+ cancer had not received HER2 targeted therapy, it is likely that the women that remained recurrence free at 5 years were those with good prognosis cancers and therefore had an inherently lower risk of recurrence beyond 5 years. The fact that HR for recurrence by CTS5 score was impacted by HER2 status in both pre- and post-menopausal women, suggests that a CTS5-type model should be re-fitted for a HER2+ population. However, it must be borne in mind that this analysis, which was made on women with HER2+ cancers that were not treated with HER2 directed therapy, is outdated and would need to be further explored in an up to date HER2+ cohort.

5.5.5 Quality of the IHC data

TMA staining was performed using established protocols. Biomarker scoring employed novel methods for this laboratory as described above. For this reason, validation exercises were performed to optimise scoring parameters within the Visiopharm software including a pilot study using an ER+ TMA from a different source. In addition, comparisons were drawn between scores from the duplicate TMAs to assess heterogeneity and between automated and visual scoring of Ki67. Results from these comparisons were reassuring in terms of consistency of biomarker expression between duplicate TMAs and consistency of scoring between the two methods. However, comparisons of scoring methods was limited to Ki67 and were not undertaken for ER or PgR.

Negative correlations were found between PgR and Ki67, PgR and grade and ER and Ki67 and ER which is consistent with existing literature[141,142]. Whilst the correlation co-efficients were only moderate, these findings (supported by existing data) give reassurance that the tissue staining and scoring itself was robust and correlates as expected with clinical data and between biomarkers themselves.

Altogether, it can be concluded that the methodologies of staining and scoring were robust and thoroughly explored prior to commencement of the main study. Further optimisation of a known proportion of ER+ cores within the tissue for all 3 measured biomarkers would have improved these methodologies even more but were not possible given the confines of time within this study.

5.5.6 Automated visual analysis

The use of automated visual analysis software is intended to improve precision and consistency and reduce time taken to score IHC stained sections. The software programme used in this study employs sophisticated programming in order to achieve a high level of accuracy when it scores the images and enables adaptation of these applications in order to allow for variation in staining between labs or between projects. This software was, however, sometimes complicated to use and required a high level of familiarity in order to be at ease with the multiple processes involved.

Additionally, the time-saving benefit of automated visual analysis was reduced somewhat by additional processes that were more time consuming such as manually checking the alignment of TMA cores.

5.5.7 Strengths and Limitations

As explained above the tissue being studied here was of good quality. The methods of staining were robust and are reproducible in other laboratories worldwide at low cost, even without the use of automated visual analysis.

This study was hampered by a higher than expected attrition from the number of stained TMAs down to the number of women eligible for analysis. Lower patient numbers and hence lower event numbers means that smaller effect sizes of variables on outcome become much more difficult to detect.

The process of automated scoring could have benefitted from further optimisation steps to ensure its consistency and comparability with standard manual scoring. Bearing this in mind, coupled with the fact that the Visiopharm software was, at times, complicated to use, simpler methods such as the Cognition Master© software might be a viable option as this has shown consistent results in scoring compared to manual methods for Ki67[143].

5.5.8 Clinical Application

Owing to the fact that no prognostic model could be developed to improve CTS5 using IHC measured biomarkers, the current situation with regards to risk estimation for late recurrence is unchanged. CTS5 remains the only validated prognostic tool in widespread use that enables accurate risk estimation for distant recurrence in years 5-10. In situations where CTS5 outcome does not point to a clear management path (intermediate risk) or in a complex patient where management decisions are not straight-forward, the only further option is to consider genomic tests to further risk stratify these women (see chapter 2). In the UK, these tests are not funded by the NHS for this purpose although over the next 10 years, patients who have undergone these tests at baseline to guide chemotherapy decisions may be able to use the outcome to guide risk management beyond 5 years (where models validated for late recurrence such as Pam50 or EPclin have been used).

Automated visual analysis, whilst not being perfect in this study, offers a credible option for scoring IHC stained tissue in a consistent and reproducible way that reduces workload in laboratories and increases consistency of scoring. Such methods will undoubtedly evolve over the coming years to incorporate elements of artificial intelligence to further increase precision and reliability of biomarker measurement. It is likely and appropriate that these methods be embraced to enable progress in laboratory practices.

5.5.9 Future Work

IHC methods of biomarker detection and scoring are vastly cheaper and more accessible worldwide than genomic methods so it would be worthwhile to explore this research question in other datasets, particularly given the outcomes of the preliminary trans-ATAC analysis mentioned above. Creating a model where IHC data can be combined with clinical data to improve risk estimation, particularly for those with intermediate or borderline CTS5 scores would benefit many women reaching 5 years on endocrine therapy. Importantly, this may help further risk stratify intermediate risk women in whom a management strategy is unclear.

In addition, training a model using a more recent dataset where HER2+ women received anti-HER2 therapy would benefit those with ER+/HER2+ cancers where management decisions and risk prediction is complex and harder to estimate.

As mentioned above, more widespread use of automated methods of IHC biomarker scoring is foreseeable and would have the potential to both improve laboratory practice and increase efficiency. Validation studies of these methods against current gold standards and subsequent incorporation into translational studies should be pursued to bring these techniques into routine use.

6. Very late recurrence tissue study

6.1 Introduction

ER+ breast cancer, treated with 5 years of endocrine therapy, has an almost linear risk of recurrence up to 20 years from diagnosis[25]. Prognostic features of late recurrence are well characterised[25,26,54,55] and, as has been discussed earlier in this thesis, prognostic models in the form of clinico-pathologic tools and molecular assays exist for risk prediction over this time period[58,78,120]. However, what is far less explored is the distinction between features of a very late recurring and a non-recurring primary. Features such as tumour size and lymph node involvement may be less helpful at this stage, but discriminators could still exist at a molecular level. Currently at 5 years from diagnosis, owing to recent clinical trial data[29-31,33], women have the option to continue their endocrine therapy for a further 5 years. Currently, this decision is made purely on the basis of clinico-pathologic features of the primary tumour and patient factors. However, with a potentially life-long risk of recurrence, even after completing 5 years extended endocrine therapy there remains a need to distinguish between those who will and will not recur.

It would therefore be of interest to determine whether a distinction exists between the molecular features of a primary cancer that recurs very late versus one that does not. If found, this may establish an evidence base upon which to further develop risk stratification of patients who have reached 10 years disease free in order to monitor and manage them appropriately for their level of risk.

Part of the difficulty of studying very late recurrence is the protracted follow-up time that is required to determine whether a woman is truly recurrence free. To study this retrospectively requires tissue that has been stored for many years which is often not done. Additionally, in order to power a prospective study of very late recurrence, a large sample size would be needed owing to the relative rarity of the outcome event.

As mentioned in chapter 1, alterations in gene expression[73,74] along with mutations and copy number alterations (CNA) in certain key genes are known to be associated with recurrences and long term survival in ER+ breast cancer however very little is known about the implications these have for recurrences beyond 10 years. Mechanisms driving dormancy and emergence from it undoubtedly play a part in the natural history of very late recurrence and these could include interactions between the tumour and its micro-

environment, stem-ness of existing micrometastatic disease and ER signalling pathways[53,144]. As demonstrated in Figure 1.3, a very late recurrence occurring many years after cessation of all systemic therapy is likely to arise from micrometastatic cells that have been maintained in a state of equilibrium for many years and have then re-emerged due to certain triggers[144]. An enhanced understanding of the molecular features that associate with very late recurrence may assist in decisions on active management of patients who are clinically disease-free over a long period and improve our understanding of the determinants of disease progression.

6.2 Aim

To determine whether a difference exists in gene expression or DNA alteration in primary tumours between patients with very late recurrence compared to no recurrence.

RMH retains all archival diagnostic material, which provides a valuable opportunity to study tissue, and patients for whom there is long follow up. A retrospective case control study was undertaken of RMH female patients diagnosed with early ER+ breast cancer between 20 and 30 years ago. Specifically, molecular analysis was conducted of the primary tumours of women who recurred very late and women who had not recurred in order to draw conclusions regarding the differences in molecular features of the two groups.

6.3 Methods

A protocol was written with input from statisticians and bio-informaticians, setting out the aims, endpoints, methodologies and proposed analyses for the study. An iRAS application was concurrently completed. The protocol was submitted to and approved by the RMH CCR (CCR no. 4984). HRA approval was sought and granted from HRA and Health Care Research Wales. Research ethics approval was sought and granted from the Health and

Social Care Research Ethics Committee (REC ref 18/NI/0227). This study is referred to in this thesis as the 'very late recurrence study' or 'VLR'.

6.3.1 Patient population

An information search for eligible patients was carried out. Inclusion criteria were: female, diagnosed with early breast cancer between 1988-1998, underwent curative surgery at RMH, follow up of at least 10 years (20 years for controls). Exclusion criteria were ER-disease, metastatic disease within 6 months of diagnosis, patient not treated curatively, non-invasive pathology and known HER2+ status, specifically for controls: any distant recurrence during follow up period and specifically for cases: distant recurrence in years 0-10. The year 1998 was chosen as the upper bound of the date range to allow for identification of controls who were alive and distant disease free at 20 years. The lower bound of 1988 was chosen because the first two Oxford Overview meta-analyses concerning Tamoxifen and the survival advantage for early breast cancer were published in 1988 and 1990 and so by 1988 most women with ER+ breast cancer were receiving some adjuvant endocrine therapy[145,146]. Cases were defined as distant recurrence beyond ten years from diagnosis and controls were defined as alive and distant recurrence free at 20 years from diagnosis. Death from any cause prior to 10 years of follow up was excluded. Deaths without recurrence in years 10-20 due to non-breast cancer causes were excluded. Cases and controls were not matched.

The following data were collected from the EPR of all eligible patients: date of diagnosis, age at diagnosis, menopausal status (where this was not documented, a woman older than 51 at diagnosis was presumed to be post-menopausal), tumour size, histopathological tumour grade, lymph node burden, ER status, use of systemic therapy (chemotherapy and endocrine therapy), date of distant recurrence, date of last follow up, date and cause of death where appropriate.

6.3.2 Laboratory Methods

As described in chapter 1, the study of both gene expression (in the form of RNA) and DNA aberrations are fundamental to understanding the molecular make-up of a cancer. Both these elements were explored in this study.

6.3.2.1 Sample Preparation

Blocks of FFPE tissue from the patients' original surgery were obtained from the hospital archives. Where the patient had neo-adjuvant treatment, FFPE blocks of the diagnostic cores were requested. Tissue blocks were sectioned (3 x 10 microns) plus one 3 micron section which was stained for haematoxylin and eosin (H&E). The H&E stained slide was marked for invasive tissue by a senior laboratory scientist. If no invasive tissue is seen another block (if available) was requested.

6.3.2.2 DNA and RNA Extraction

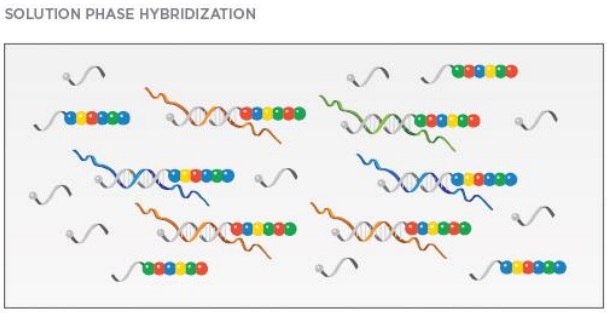
RNA and DNA were extracted manually using the Qiagen Allprep FFPE kit (Appendix 9.2). Sections of tissue were macrodissected to enrich for malignant cells guided by the marked up H&E. Macrodissection is essential for DNA analysis as mutations found in a cancer, particularly driver mutations, will not be found in normal breast tissue or in situ disease. Therefore to exclude those areas is vital as their inclusion will dilute the signal from mutant DNA within the tumour. Extracted RNA and DNA were quantified using the Qubit Analyzer. RNA samples containing more than 200ng RNA in a maximum of 10µl were taken forward to Nanostring analysis. This quantity was increased after the first batch was run to 250ng where possible which was felt necessary in these older samples. DNA samples containing more than 200ng DNA were taken forward to targeted sequencing. In some cases where there was a lower than expected yield, further sections were cut, and tissue dissected.

6.3.2.3 nCounter Analysis (RNA)

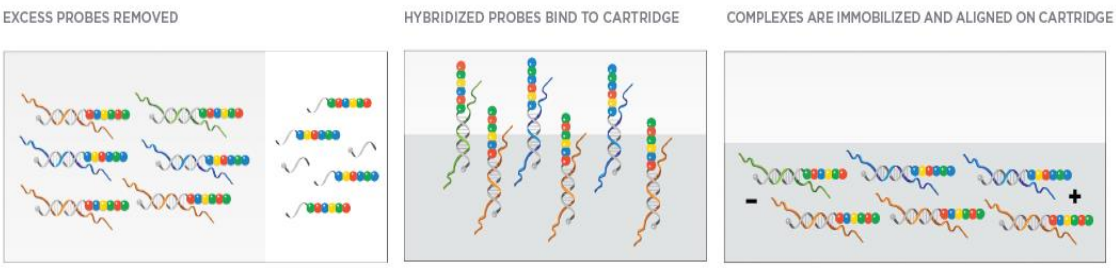
The nCounter Analysis System is a fully automated platform to analyse gene expression levels of up to 800 genes in a multiplex and has been validated for use on RNA extracted FFPE specimens and is employed for the widely used ProSigna prognostic test. The user selects or designs a panel of target genes to quantify based on the disease of interest.

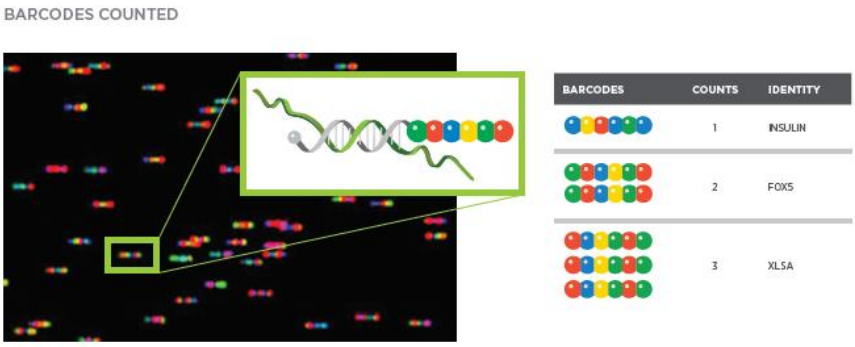
The nCounter process occurs over 2 days and comprises multiple steps (Figure 6.1). On day 1 each extracted RNA sample is mixed with the capture probe and colour barcoded reporter probes (codesets) which are specific to different RNA sequences. RNA sequences bind to the capture and reporter probes and the combined solution is placed on a thermocycler at 65°C for 23 hours. On day 2 excess probes are removed and the capture probe/RNA/reporter probe complexes are immobilised on the surface of the nCounter cartridge. An electric current is applied to the cartridge to align the complexes. The cartridge is then manually transferred to the digital analyser where the reporter probe carries the colour-coded signal from the labelled barcodes. This is detected by a fluorescent microscope. The analyser records the presence of a particular RNA sequence and how many times that sequence is read.

a.



b.





C.

Figure 6.1 **Credit:** https://www.nanostring.com/scientific-content/technology-overview/ncounter-technology
Process of nCounter analysis of RNA. A. incubation of target RNA (green), capture probe (grey) and reporter probe (coloured). B. removal of unbound probes and binding of hybridized probes onto the nanostring cartridge via the capture probe and application of an electrical current across the cartridge resulting in alignment of the probes. C. The fluorescent signal emitted by the bound probes is counted.

For this study, the Nanostring breast cancer 360 panel was chosen as the gene panel of interest. This panel contains 776 genes (Appendix 9.3) which are involved in multiple different pathways in breast cancer such as proliferation, invasion and also those linked with the tumour microenvironment and the immune response. Among the 776 genes are housekeeping genes, positive controls and negative controls. The positive controls inform on whether the assay has worked and the negative controls ensure that contamination of the samples can be ruled out. The housekeeping genes allow for adjustments to be made on the basis on background signal in the samples.

6.3.2.4 DNA Library Preparation and capture

The kits used for library preparation and DNA capture were the SureSelect XT Low Input Reagent Kit. This required 200ng of DNA per sample from FFPE material. In this process the DNA is broken into single stranded fragments of 150-200 base pairs. Specialised adaptors are ligated at both ends and the DNA is PCR amplified which enables the DNA being studied to bind to the capture microarray.

DNA capture occurs via a microarray onto which a capture library is bound. The capture library consists of oligonucleotides complementary to specific DNA sequences of interest. In the capture step of the process the prepared sample DNA libraries are hybridised and the adaptors bind to complementary oligonucleotides (also known as baits) on the microarray. Any unbound DNA is then washed away. These captured libraries are then PCR

amplified and the resulting library is quantified. The resulting target enriched DNA is then sequenced.

6.3.2.5 DNA Sequencing

For this study, DNA was sequenced according to a targeted exome panel spanning approximately 1.3 megabases, incorporating the exons of 87 genes (Appendix 9.4). These genes are those known to be affected by driver mutations and mutations associated with AI resistance[147]. The panel also captures a backbone of genetic material across the genome which spans 1 to 3 megabases and is used to detect chromosomal instability. This makes copy number alteration analysis more accurate. The backbone is comprised of randomly spaced DNA between 2 probes. Distance between the probes determines the coverage. As described above, sequencing occurs according to micro-array based capture where sample DNA adheres to oligonucleotides on the surface. In this case oligonucleotides are present for the backbone, the exons of interest (at double the quantity compared to the backbone) and specifically in this study *ESR1* and *TP53* exons (at 4 times that of the backbone). This enables better coverage of exons and even greater coverage of these specific genes to reduce the likelihood of sequencing error. Sequencing took place using the NovaSeq platform by Illumina in the laboratories of the tumour profiling unit within the Institute of Cancer Research.

Sequencing starts with the addition of the first sequencing primer. After this, fluorescently tagged nucleotides are added sequentially in a process called sequencing by synthesis. This occurs by the sequential addition of chemically modified nucleotides to their complementary bases within the hybridised DNA strand. Each nucleotide contains a fluorescent tag and a reversible terminator that blocks incorporation of the next base. When a nucleotide binds to the DNA strand, the fluorescent signal is released and an optical sensor scans the flow cell to capture colours that indicate addition of a base in each cluster. At the end of the cycle, the terminator is cleaved so that the next base can be added.

After reading the forward strand, the reads are washed away and the process repeats for the reverse strand known as paired end sequencing in which the forward and reverse

strands are paired up and both are sequenced. This enables double the amount of reads for the same amount of preparation. Each base contains a different colour fluorescent dye so the sequence can be interpreted by image analysis software.

Sequencing on a targeted panel of a limited number of genes allows study of specific genes known to be important for the disease or phenotype being studied. The benefits of targeted sequencing as opposed to whole exome or whole genome sequencing is that the former is much less costly, has a shorter turnaround time, enables deep sequencing with higher coverage and it focuses on genes of interest and therefore increases confidence that a positive finding is truly associated with the phenotype.

6.3.2.6 Copy number alteration analysis

For CNA detection, pooled blood diploid normal controls were used. These were comprised of 6 patients from the POETIC study[148] and 1 commercially sourced female control. Sequences of DNA within these samples are aligned to a human reference genome (version GRCh38) with BWA software (version 0.7.15) to detect genetic variants. Any duplicated reads were identified using Picard tools (version 2.8.2). The average DNA segment size was 20 megabases and 3 megabases for SNP array.

A senior bio-informatician set up the CNVkit pipeline for detection of CNA. This method combines reads from both the on and off target chromosome regions to improve accuracy in determining a true alteration. Although the sequencing here was targeted exome sequencing, a large proportion of off-target DNA remains within the library after hybridization. Using this off-target DNA improves the accuracy of copy number variation detection. The output of CNVkit is an estimation of the log2 ratio of study samples versus the normal pooled blood controls. If the log2 ratio for a segment is > 0.4 then this region has gained a copy in the tumour and if is < -0.4, then this region has lost a copy in the tumour. The bedtools program then identifies the largest overlapping segments across all tumours.

6.3.2.7 Mutation detection

Mutation detection was conducted using Mutect2 software from GATK (version 4.0.5.1). This software uses the same human reference genome for alignment as for copy number variation. Mutations are called and then germline mutations are removed using a panel of normal and a database of common alleles. Mutect2 has an orientation bias filter that helps reduce false positives of C>T caused by long term storage of FFPE tissue. Ensembl variant effect predictor (VEP) is then used to map the mutation to coding genes. Only somatic mutations with a moderate to high impact on the translated protein are included in the analysis.

As this study was undertaken using archived FFPE samples that had been stored for many years, analysis had to take into account DNA fragmentation which occurs over time. In this process cytosine bases in particular can be spontaneously deaminated into a thymidine base creating the potential for C > T false positive mutations.

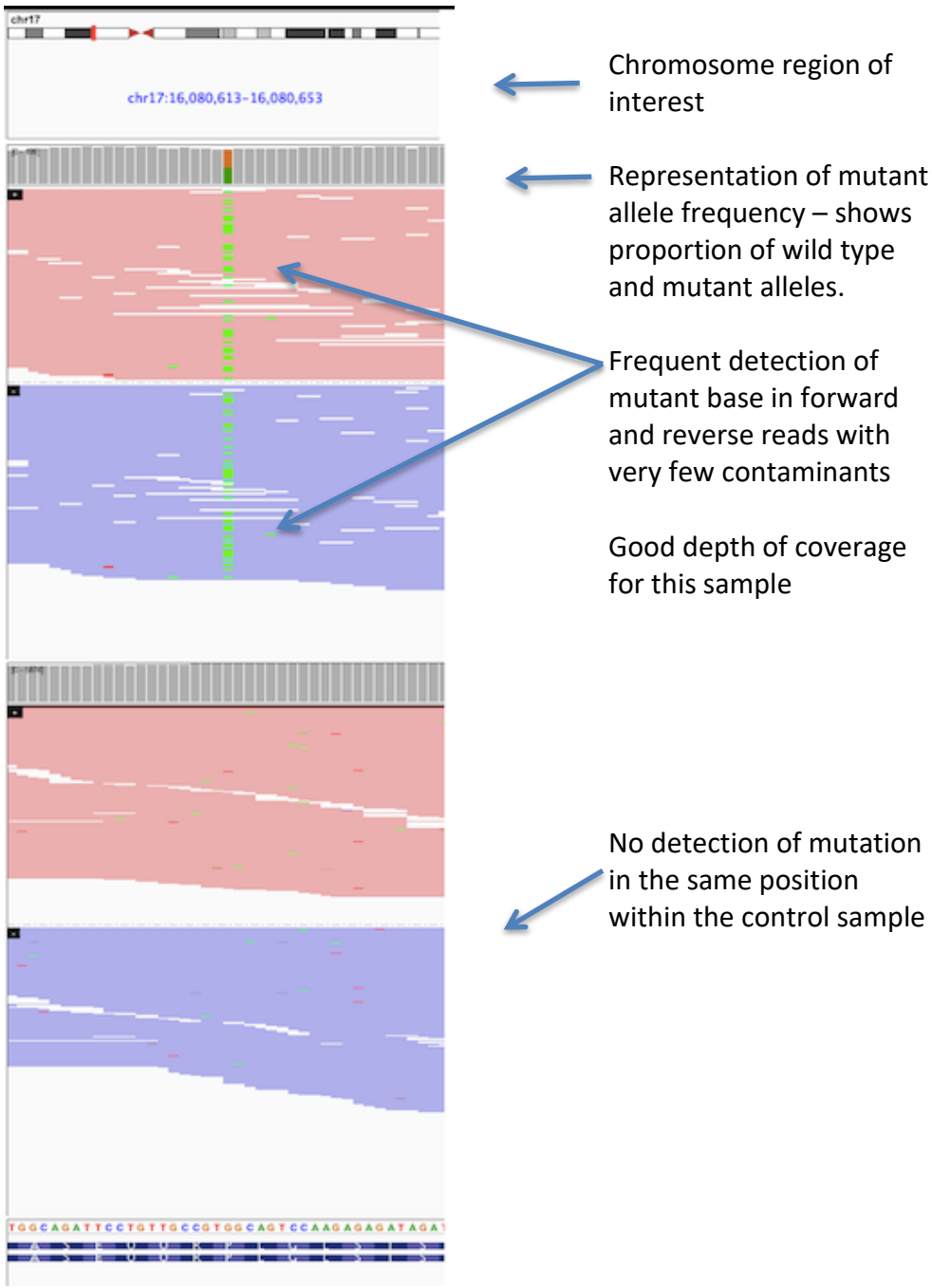
6.3.2.8 Manual mutation calling

Once mutations were detected by Mutect2, they required manual checking to determine whether they met a given threshold for being a true mutation. In this process, the software generates images of a specific chromosomal region within a single sample. This region corresponds to one of the target regions within the panel of interest. The degree of coverage for that region will be shown and the forward and reverse reads are demonstrated with different colours (Figure 6.2). An identical image is shown for the control. Mutations are shown for all reads in which it was detected. This can then be visually analysed and compared to the control specimen in order to conclude whether this is likely to be a true aberration. In order to meet the threshold for being a true aberration, the mutation had to be visible in a substantial number of both the forward and reverse reads. For samples where coverage for that region was low, a higher mutant allele frequency was required. Additionally, the mutation must not be present in the control sample.

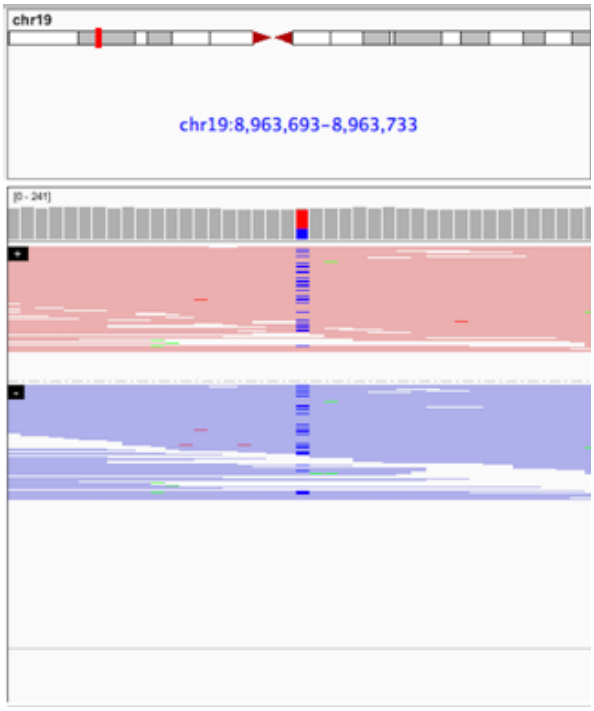
6.3.2.9 RNA expression data processing

Raw NanoString gene expression data was normalised with NanoStringNorm package (ref paper: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3356845/) using R software. Briefly, for each sample, positive control probes were normalised to the geometric mean followed by the subtraction of the negative probes mean from all expression values. RNA content was normalised using the housekeeping genes geometric mean. Values of <1 were set to 0 before log2 transformation of the data.

a.

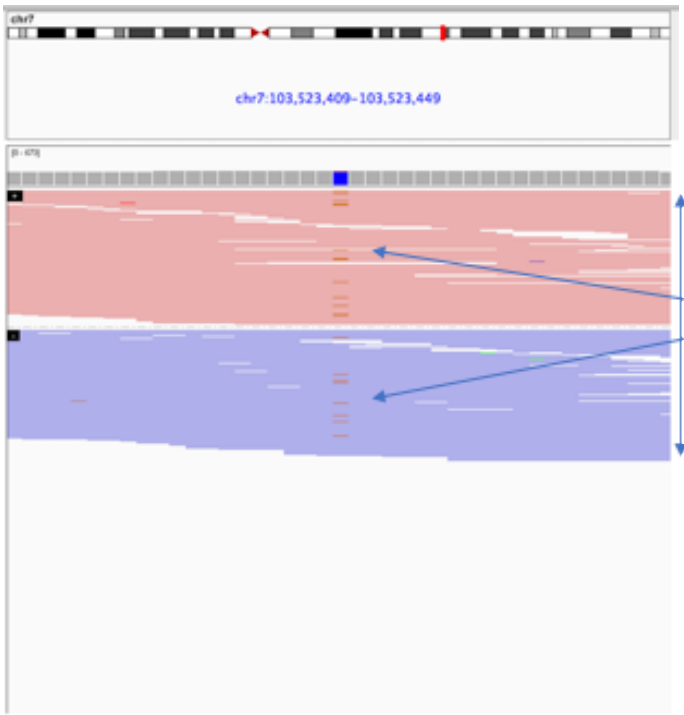


b.



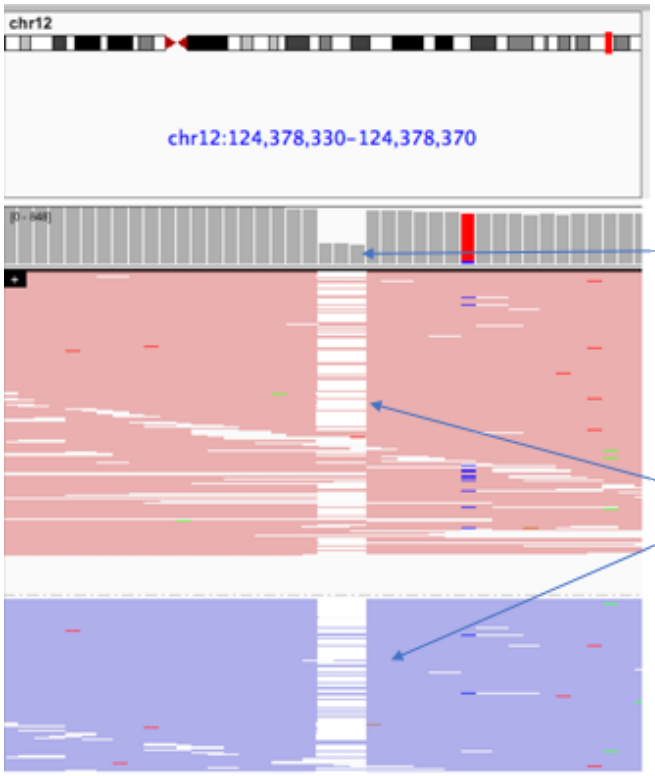
Reduced depth of coverage for this sample so a high frequency of mutations would be required to pass the threshold

c.



Mutation detected in both forward and reverse reads but not at such high frequency. Lower than ideal coverage would class this mutation as weak.

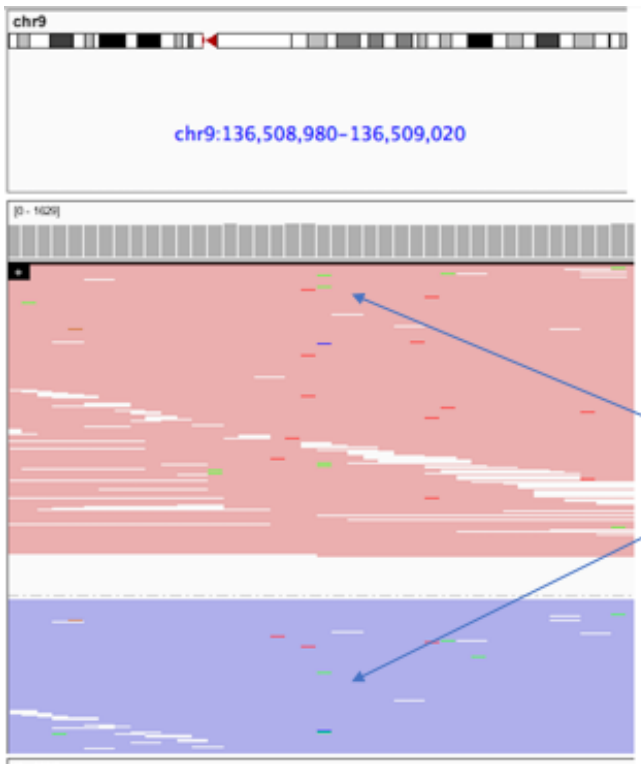
d.



Significant drop in proportion of DNA for this specific region due to deletion

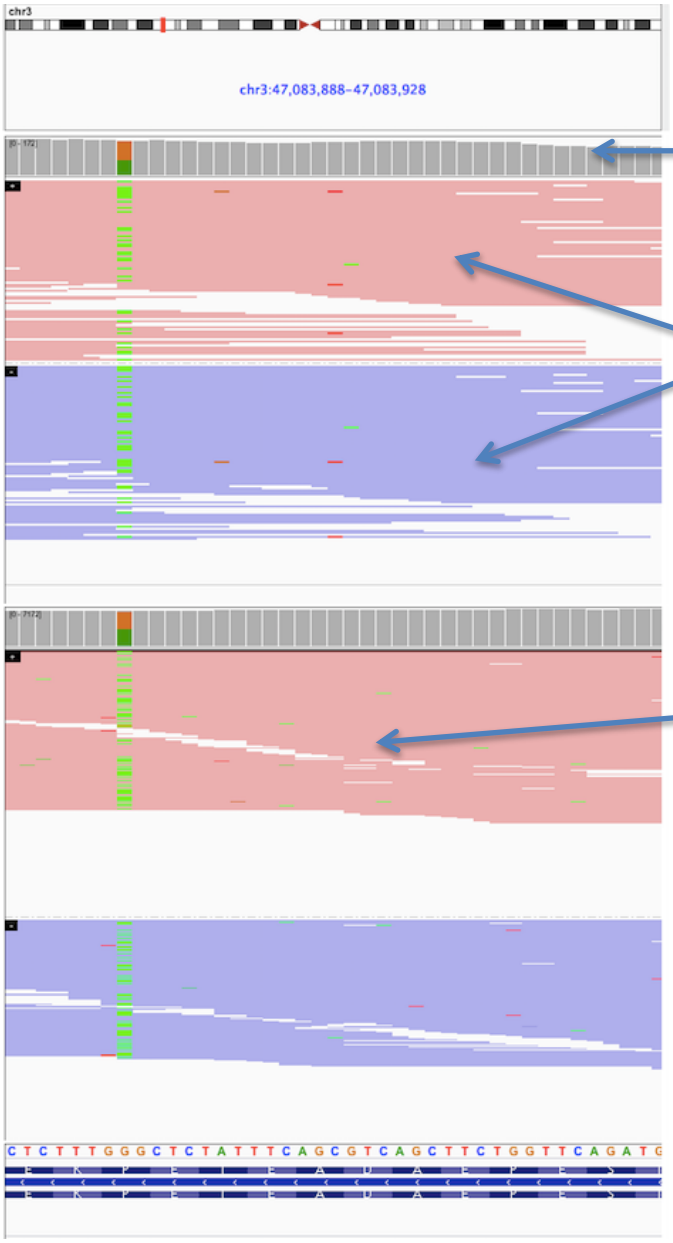
High frequency detection of deletion in both the forward and reverse reads with good depth of coverage

e.



Weak detection of mutant allele along with high level of contamination mutations – this mutation would be rejected

f.



Representation of proportion of mutant bases

Detection of mutation in forward and reverse reads

BUT – detection of mutation in control samples therefore this mutation is rejected.

g.

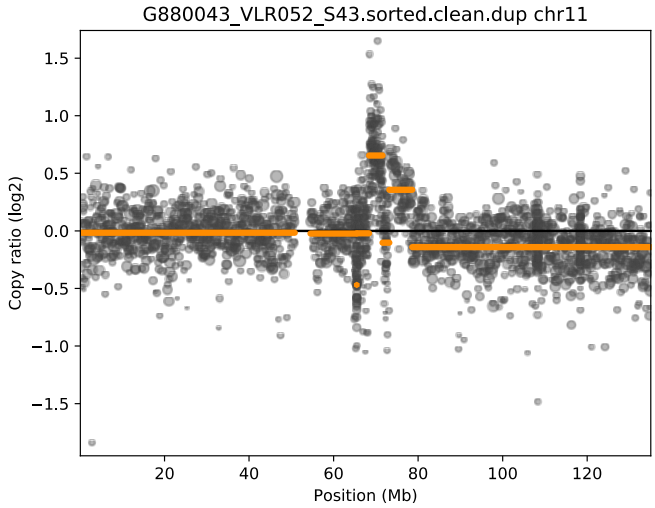


Figure 6.2 Examples of images for mutation and CNA detection– all are actual samples from the VLR study. Images a to f show columns for individual bases in the reference human genome and rows are the individual reads from the Illumina machine. Pink horizontal lines are reads aligning to forward strand and light blue the reverse strand. Base pairs that do not match the reference genome are marked within a read by the colour code adenine (A) - green, cytosine (C) - blue, guanine (G) - orange, and thymine (T) - red. Deletions within a individual read are marked by a black line and the location of insertions within an individual read are marked by a purple vertical line. Samples a, b and d are clear passes. Sample c is a weak pass due to reduced frequency of detection and lower coverage. Sample e would fail due to low frequency of mutation and high contamination. Sample f would fail as the mutation is detected in the control sample. 2g shows a scatter image of the CNVkit output with segments. This shows log2 ratio at nearly probe level (grey circles) and the segmentation output from CNVkit (orange line). Example shows gain of a region encoding CCND1 in chromosome 11.

6.3.3 Statistical Considerations

6.3.3.1 Endpoints

The primary endpoint of the study was proportion of overall CNA (measured by percentage of genome altered) in very late recurrence (cases) versus no recurrence (controls) groups for the whole population. Secondary endpoints can be divided into clinical and molecular. The clinical endpoints were: differences in tumour size, grade and lymph node burden as well as chemotherapy treatment and endocrine therapy treatment. Molecular secondary endpoints were proportion of overall copy number gains and losses (by percentage of genome altered), proportion of individual gene CNA, proportion of overall mutations (by percentage of genome altered), proportion of individual gene mutations, mean gene

expression for individual genes, gene expression for separate gene modules. All endpoints were compared between cases and controls.

6.3.3.2 Sample size

In a previous analysis of 2428 unselected RMH patients who were diagnosed between 2000 and 2007, who were alive and distant recurrence free at 5 years, there were 83 distant recurrences beyond 10 years giving a rate of 3.4%.

In this older cohort with a wider inclusion period for date of diagnosis, approximately 90 distant recurrences beyond 10 years are anticipated. Given this estimate, the maximum possible number of cases was identified with an aim to ultimately obtaining sufficient RNA and DNA for at least 65 cases and 65 controls thereby allowing for attrition due to missing tissue, inadequate tissue or inadequate RNA/DNA for analysis.

6.3.4 Statistical Analysis

Data generated for the primary, secondary endpoints and other patients' characteristics and treatment variables were summarised descriptively in the two matched cohorts using frequency and percentages for categorical variables and mean/median and standard deviation and interquartile ranges for continuous data as appropriate.

6.3.4.1 Primary Endpoints

Mann Whitney U test was used to compare overall CNA in cases and controls.

6.3.4.2 Secondary Endpoints

For clinical and pathological variables, Mann Whitney test was used to compare tumour size and CTS5 value. Comparison of categorical variables was assessed using χ^2 test and comparison of grade was made by χ^2 test for trend.

For CNA and mutations, Mann Whitney test was used to compare overall differences between cases and controls. Fisher's exact test was used to identify differences in individual gene CN or mutations between cases and controls.

For gene expression data, unpaired t-test was used to compare mean expression of each individual gene with corresponding p values. Clusters within hierarchical clustering were compared using Fisher's exact test. All analyses were 2 sided with an alpha level of 0.05. Owing to the large number of genes being tested, all analyses were subject to correction for multiple testing using the Benjamini Hochberg false discovery rate (FDR) method[149].

6.3.5 Gene set enrichment

For the gene expression data, gene set enrichment analyses were performed according to four sets of gene annotations produced by different organisations. For this, genes are grouped together according to biological processes or their associations with disease phenotypes. In this way, one can see if there are specific clusters of genes whose expression is different between cases and controls. The 4 different gene sets analysed were the Nanostring breast cancer 360 panel gene modules (analysed using Enrichr) the gene ontology biological processes annotations, KEGG annotations and Hallmarks of cancer. Each gene annotation or group is analysed for its degree of enrichment within cases and controls and p value is given. A ratio of genes within each gene set and the overall significant genes that were either up in cases or down in cases is given (gene ratio) as well as a ratio of the number of genes within that annotation and total number of genes in the breast cancer 360 panel (Bg Ratio). The KEGG gene set enrichment also produces visual maps of key pathways which show, through colour, the intensity of expression of genes within the pathway for cases or controls.

6.3.6 Comparisons with other datasets

In order to validate findings and provide better context for the results, comparisons in the gene expression and DNA aberration data were drawn between results from this study and 2 other studies. For copy number alteration, comparisons were made with samples from the POETIC study[148]. This was a phase 3 , window of opportunity study in post-menopausal, pre-surgical women with ER+ early breast cancer who were randomised to receive 2 weeks of neoadjuvant AI or no treatment. Data from 69 patients was used from this dataset as this was the number of samples that had undergone sequencing by this timepoint. Within

these 69, approximately 25% had suffered a recurrence in years 0-5. Further follow up data from this study is not yet mature. The second study used for comparison of copy number, mutations and gene expression was the METABRIC dataset[150]. This was a large bank of 2000 breast tumours from pre- and post-menopausal women with outcome data that was compiled to study the gene expression and DNA aberrations within these cancers. It should be noted that this included a proportion of women who did not receive endocrine therapy. Comparisons are drawn with POETIC for copy number data by chromosomal distribution. The METABRIC study was used to compare with overall copy number alterations between recurrence and non-recurrence in different time periods – 1450 patients in total. Additionally, patients with recurrences beyond 10 years (n=103) and patients with >20 years of no recurrence (n=111) were used to compare the mutation data. For the gene expression analysis, a sample of 1226 METABRIC patients were used to compare significantly altered levels of gene expression.

6.4 Results

6.4.1 Patients

An information search identified 1335 records of women who met the initial search criteria. The EPR was reviewed, and cases and controls were identified based on the inclusion criteria. Exclusions were made for loss of follow up, early deaths, early recurrences, ER- cancer and no available tissue. 50 cases and 67 controls had RNA and DNA extracted from the tumour. It was decided, owing to the high cost of DNA sequencing, that only 80 samples of DNA would be taken forward. After further exclusions due to insufficient material, 111 samples were sequenced for RNA and 80 samples for DNA (Figure 6.3).

Initial gene expression analyses highlighted a greater than expected proportion of samples that had a basal like gene expression pattern. The patient records of all the basal subtype samples were interrogated and it was found that 2 of them were ER- and had been included in error and the remaining 5 were ER unknown. ER unknown cancers were included in the study protocol as the majority of these would be ER+. For these 5 samples however, based on their ER unknown status coupled with their basal molecular subtype, it was decided that they, along with the 2 ER- should be excluded from further analyses.

In addition, a further subset of samples was found to have high expression of *HER2* and *GRB7* (genes associated with *HER2* amplification (Figure 6.4a). *HER2* copy number was analysed and plotted against *HER2* expression to confirm the HER2 status of the borderline cases identified by gene expression and sub-typing (Figure 6.4b). These samples had not been tested for *HER2* amplification at the time of diagnosis but owing to this finding and the criteria in this study to exclude known HER2+ cancers, these 6 samples were excluded from further analyses.

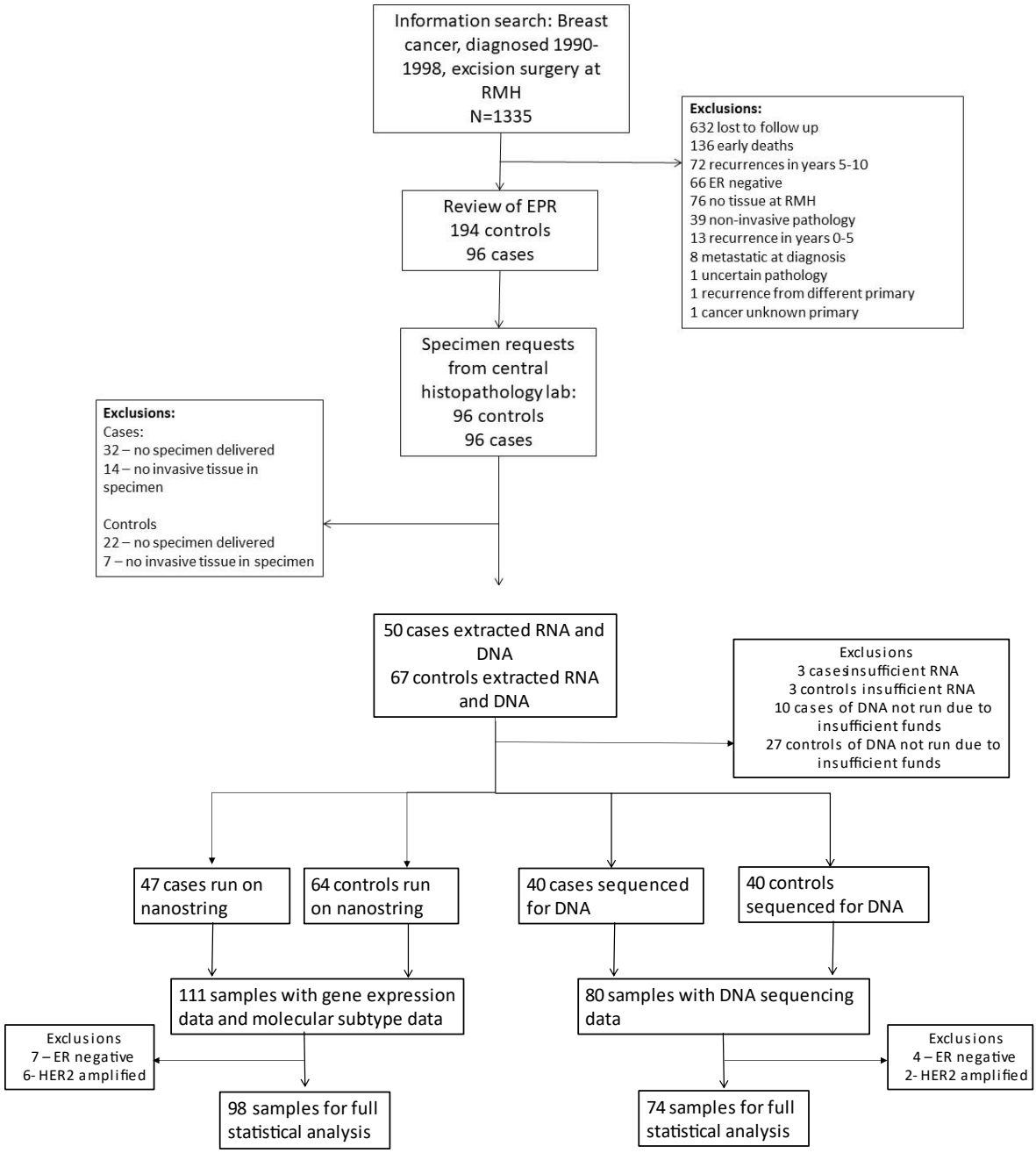
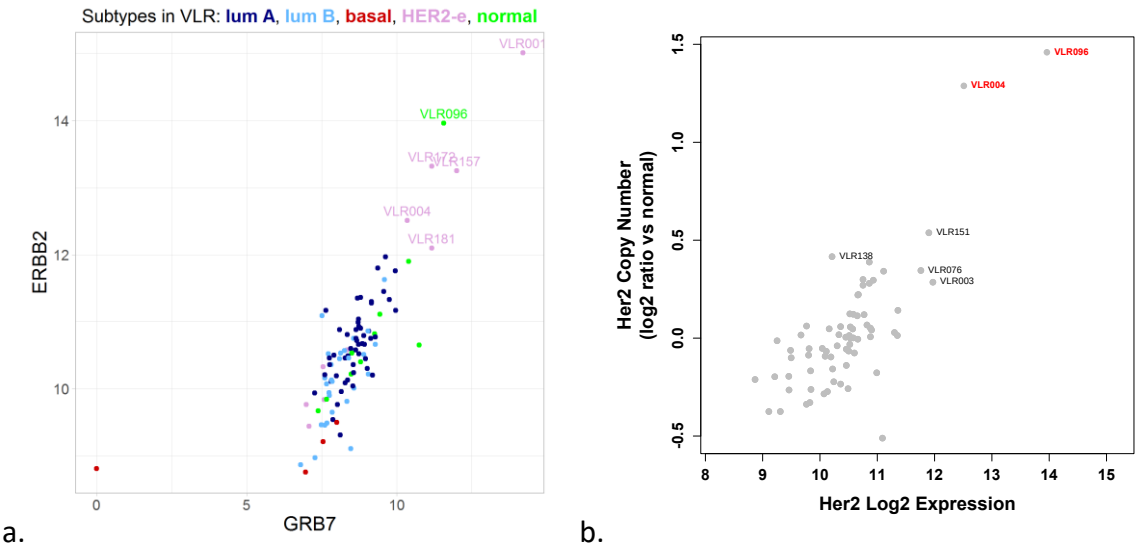


Figure 6.3 Consort diagram



a. b.

Figure 6.4 Identification of likely amplified samples. A. Using expression of ERBB2 and GRB7 and molecular subtyping – the 6 samples at the top right of the plot are those that were excluded. B. Using HER2 copy number against HER2 expression – aided identification of sample VLR 004 as a likely HER2 amplified sample.

Following exclusion of the 7 ER-/unknown with basal subtype and HER2+ samples 45 cases and 53 controls were taken forward for full gene expression analysis. 39 cases and 35 controls were analysed for mutation and copy number alteration data.

6.4.2 Clinical features

Cases and controls had a similar median age. Tumours in cases were significantly larger than controls (21mm v 16mm, p=0.01) and significantly more were node positive (p=0.0002). There was a larger proportion of grade 3 cancers among cases but this was not statistically significant (p=0.25). A larger proportion of cases were treated with chemotherapy compared to controls (76% vs 42% p=0.0007). Median time to distant recurrence for cases was 15.9 years and median follow up time for controls was 21.4 years (Table 6.1). Owing to the significant differences in clinico-pathological variables between cases and controls, CTS5 was calculated and was significantly higher in cases compared to controls 3.63 versus 2.91 (p=0.0003).

	Cases (45)	Controls (53)	All (98)	p value
Age (median)	48	51	50	0.1983
ER positive **ER unknown**	40 (89%) 5 (11%)	42 (79%) 11 (21%)	82 (84%) 16 (16%)	0.275
Tumour size, mm (median)	21	16	17.5	0.01
Grade **Grade 1** **Grade 2** **Grade 3** **Unknown**	 5 (11%) 23 (51%) 10 (22%) 7 (16%)	 4 (8%) 38 (72%) 4 (8%) 7 (13%)	 9 (9%) 61 (63%) 14 (14%) 14 (14%)	 0.25
Lymph node burden **Node negative** **Node positive** **Unknown**	 15 (31%) 25 (58%) 5 (9%)	 36 (68%) 12 (23%) 5 (9%)	 50 (51%) 38 (39%) 9 (95%)	 0.0002 (pos vs neg)
Chemotherapy treated	33 (75%)	23 (44%)	56 (57%)	0.0007
Endocrine therapy treated	43 (96%)	47 (89%)	90 (92%)	0.282
Duration endocrine therapy, months (median)	60	60	60	0.626
Time to distant recurrence, years (median)	15.9	N/A	N/A	
Follow up, years (median)	N/A	21.4	N/A	
CTS5 Value (median)	3.63 (35 values)	2.91 (46 values)	3.19	0.0003

Table 6.1 Baseline characteristics and outcomes of all patients whose tumours underwent molecular analysis

6.4.3 Copy Number Alteration

On average the sequencing produced 34 million reads of DNA per sample. Average coverage (i.e. the number of unique reads containing a specific nucleotide within a reconstructed DNA sequence) for the overall genetic material sequenced was 165. For the backbone this was to 132 fold, for the exons - 218 fold and for the exons of ESR1 and TP53 - 320 fold. The CNVkit analysis software broke down the data for individual patient samples

into a total of 2589 segments based on positions of copy number alterations. These segments were mapped to different chromosome regions and areas of overlap were compared between cases and controls.

6.4.3.1 Overall percentage of genome altered

After correction for multiple testing, there were no significant differences (in either direction) in percentage of genome with a CNA between cases and controls (Figure 6.5). This was the case for both gains and losses. In-silico analyses from the METABRIC study showed significant differences in copy number alteration between early recurrences and those who were recurrence free at 5 years, p=<0.0001 (Figure 6.6 box A&B). This difference was preserved when comparing patients who recurred in years 5-10 with those who had no recurrence by 10 years, p= 0.015 (box C&D), and when comparing those who recurred in years 10-20 with those who were 20 years recurrence free, p=0.011 (box E&F). The latter comparison is akin to the cases and controls within this study.

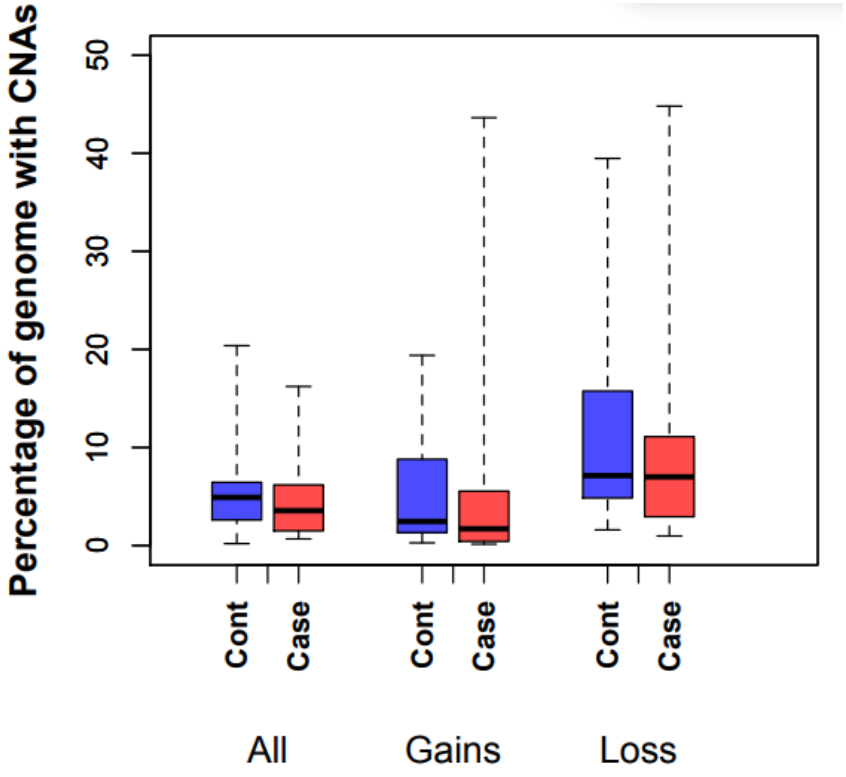


Figure 6.5 Boxplot showing overall CNA and gains and losses in cases and controls.

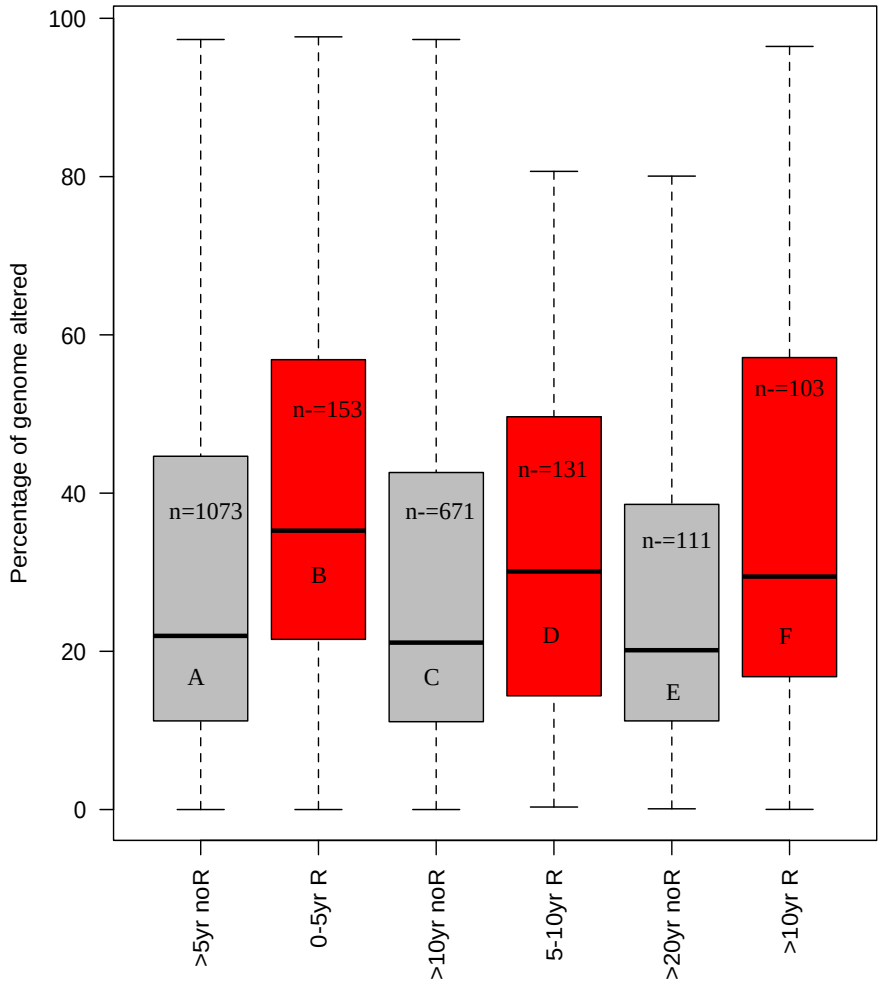


Figure 6.6 Comparisons in copy number alteration between recurrences and non-recurrences for different time periods all from the METABRIC dataset. Grey boxes A, C and E are patients who did not recur within the specified time period. Red boxes are patients who recurred within the specified time periods The two right-most box plots (E and F) are analogous to the cases and controls in this very late recurrence study.

6.4.3.2 Individual chromosome/gene copy number alterations

The full list of chromosome regions, genes and proportions of gains and losses in cases and controls is shown at http://tiny.cc/appendixtables . Significant differences in copy number gain or loss between cases and controls were found in certain genes (Appendix 9.5), however significance was lost on correction for multiple testing. Below a discussion of 'significant' genes or gene regions refers to those that were significant *before* correction for multiple testing. This discussion must therefore be interpreted with caution.

All significant copy number gains were of a higher proportion in controls compared to cases. *ESR1* was the gene most frequently included in the overlapping segments of DNA where gains were seen. Chromosome 6q25 was the most frequently gained chromosome region which is the locus of the *ESR1* gene. Gains in *CDKN1B* and *FOXP1* were also found to

be significant and chromosome 5q12-13, which incorporates the *PIK3R1* gene, had significantly higher copy number gains in controls compared to cases.

In terms of losses, the most notable finding was the frequent loss of segments within chromosome 17q12 (not including the *HER2* gene locus). The overlapping segments of this chromosome region often included the *AAFT* gene and loss of *CDC4a* was also significantly greater in controls. Additionally, losses in chromosome 14q32 were more prevalent in controls rather than cases and this is the location of many immunoglobulin heavy chain genes.

Notably, there was an excess of gains and losses in the controls compared to cases but these alterations occurred within a small number of chromosome regions.

When analysing the data based on chromosome, key chromosome regions with common CNA in breast cancer were seen in in both cases and controls. On chromosome 1, 5, 8 and 11 gains were noted. Chromosomes with noticeable losses were chromosomes 8, 11, 15, 16, 17, and 19. None of these chromosomal regions were altered significantly differently between cases and controls (Figure 6.7). When comparing CNA in this study with 69 samples from the of the POETIC study, the latter had globally increased level of CNA but similar distributions to the samples here.

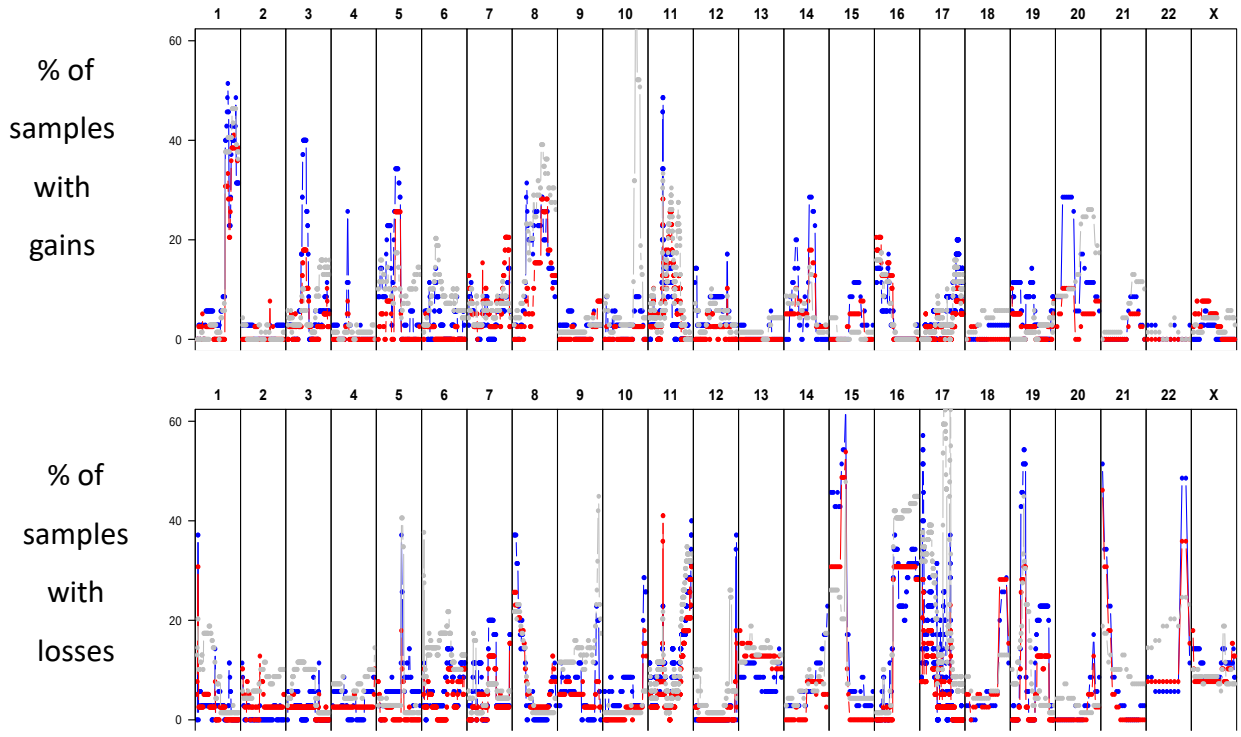


Figure 6.7 CNA in VLR cases (red) and VLR controls (blue) compared to POETIC study (grey), broken down by chromosome

6.4.4 Mutations

754 DNA sites were inspected for potential aberrations with a moderate to high impact on protein function. Based on manual checking, a cut-off of 5% allele frequency with a minimum of 10 mutant allele counts was selected for an aberration to be called a true mutation.

Overall, there was no difference in somatic mutational burden between cases and controls (Figure 6.8). The full list of mutations can be found in Appendix 9.6. There were more *MAP3K1* mutations in controls compared to cases (p=0.04) but this did not remain significant after correction for multiple testing. *CDH1* and *GATA3* mutations showed a trend for higher prevalence in controls versus cases (p=0.06 and 0.07 respectively, Figure 6.9a). *PIK3CA* was the most commonly detected mutation in both cases and controls. The pattern for *MAP3K1* and particularly for *GATA3* was concordant with data from METABRIC which showed increasing prevalence of mutations in these 2 genes in non-recurrent cancers the longer the individual was disease free (Figure 6.9b and c). Levels of *PIK3CA* and

TP53 mutations in this study were similar to those who were >10 years disease free in METABRIC. Other differences between cases and controls (although not statistically significant) were the higher proportion of mutations in cases of *SPEN, MK167, JAK1, BRCA2, SETD2 and FOXP1.*

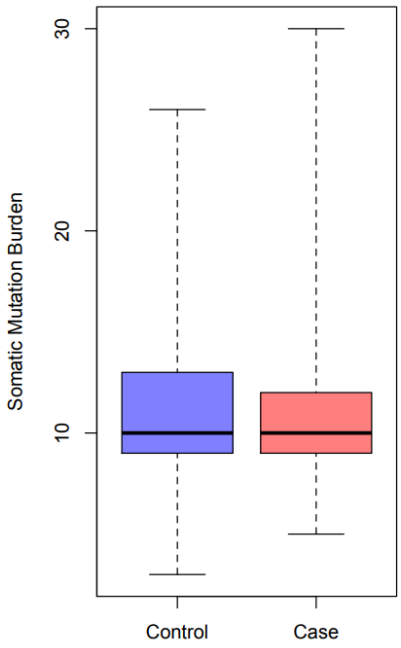
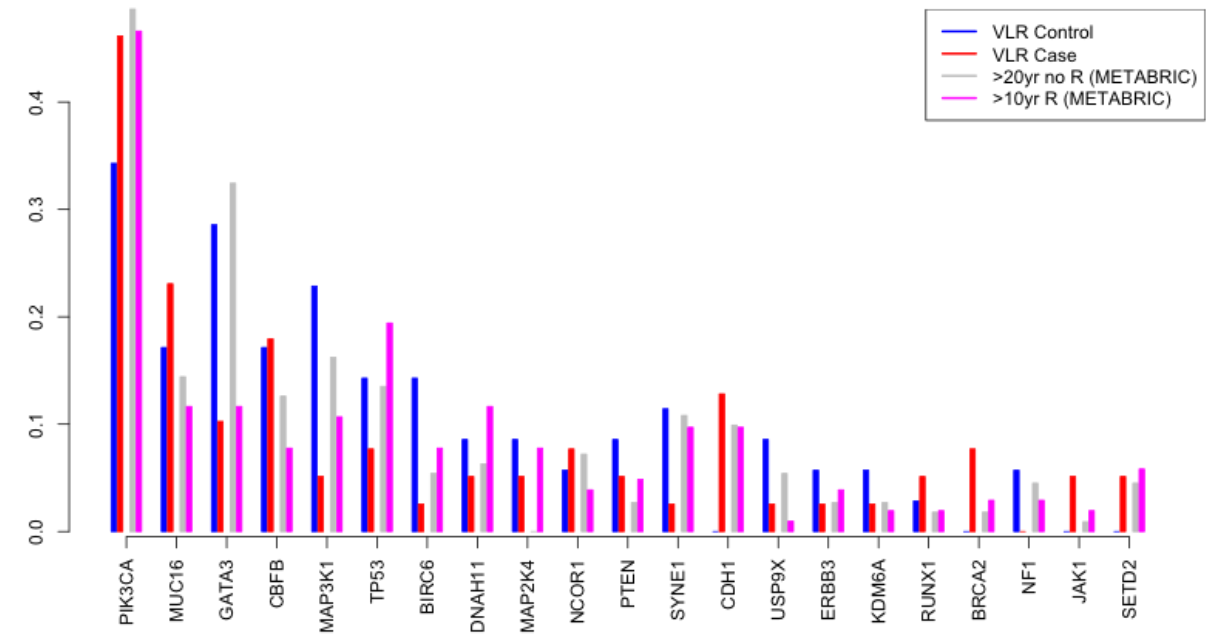


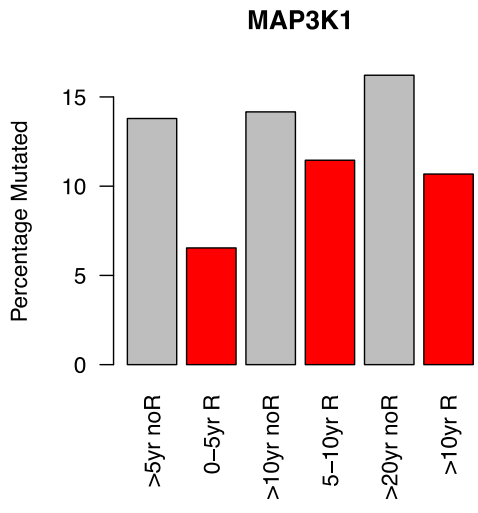
Figure 6.8 Boxplot showing no. of overall somatic mutations per sample in cases and controls (includes synonymous and non-synonymous mutations)

High mutation rates are seen in the *MUC16* and *SYNE1* genes (Figure 6.9), which are not considered to be clinically significant. These high rates of mutations are likely to be caused by C>T deamination described previously which long genes (such as these two) are particularly susceptible to.

A.



B.



C.

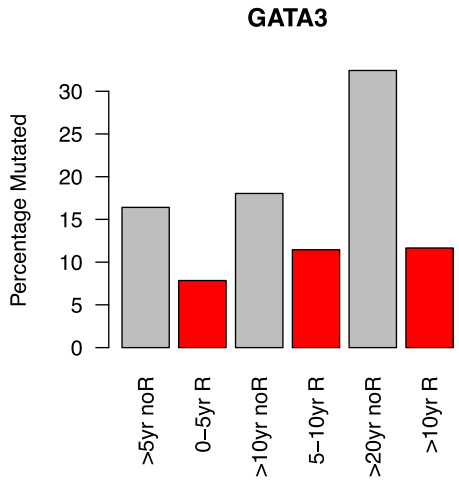


Figure 6.9 Percentage of samples of VLR cases (red) and VLR controls (blue) with mutations to individual genes. Also shown are cases from METABRIC with recurrences beyond 10 years (pink) and non-recurrences beyond 20 years (grey). B. Percentage of samples within METABRIC with MAP3K1 mutations for different time periods. C. Percentage of samples within METABRIC with GATA3 mutations for different time periods (R=recurrence, no R= no recurrence)

6.4.5 Gene Expression

Of the 758 genes of interest included in the breast cancer 360 panel, 5 were removed from the analysis owing to a low detection rate of <10%. In addition to expression data for individual genes and gene modules, intrinsic molecular subtyping was performed on each

sample. The full set of gene expression data for all cases and controls is shown in appendix 9.7.

6.4.6 Molecular subtyping

Distribution of subtypes was as expected for an ER+ population (Table 6.2). There was a similar proportion of cases and controls within each of the luminal A and luminal B subtypes.

	Basal	HER2-e	LumA	LumB	Normal	Total
Case	1	3	24	13	3	44
Control	3	2	27	16	6	54
Total	4	5	51	29	9	98

Table 6.2 Molecular subtype of the 98 samples sequenced for RNA expression

After excluding those with a low detection rate, 65 individual genes were significantly differentially expressed between cases and controls on univariate analysis (Figure 6.10 and Appendix 9.7). After correction for multiple testing, none remained significant.

The top 5 genes that were expressed at significantly (univariate) higher levels in cases compared to controls were *PLA2G4F* (p=0.0011), *HGF* (0.0028), *PAX* 8 (p=0.0111), *CCNE2* (p=0.0112) and *BAIAP3* (p=0.0122). The top 5 genes that were expressed at significantly higher levels in controls compared to cases were *CDKN2D* (p=0.0025), *SMAD5* (p=0.0042), *ITGB3* (p=0.0043), *SPRY4* (p=0.0056) and *CDC25B* (p=0.0065). Of the 65 significant genes within this analysis, 9 genes were also significantly differentially expressed in the same direction within the METABRIC dataset (Figure 6.11). Those increased in cases were *CCR2, PKMYT1, PTTG1, MMP 11* and *CCNE2*. Those reduced in cases were *MAP2K4* and *BNIP3*. Of note, within the METABRIC dataset, *ESR1* and *PgR* showed higher expression in non-recurrent tumours in years 0-5 but this was no longer seen beyond ten years in either METABRIC or VLR.

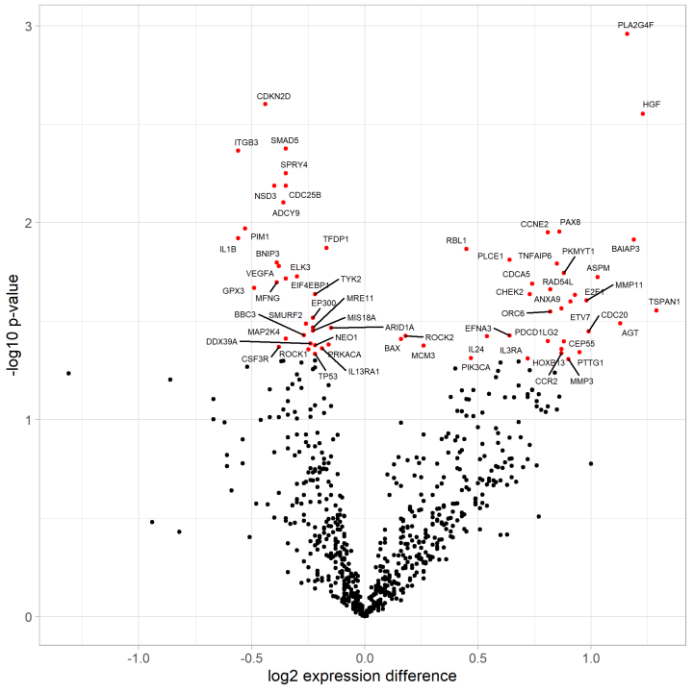


Figure 6.10 Volcano plot of gene expression. Dots represent genes (black - non significant, red - significant) Red dots to the right of the 0.0 point on the x axis represent genes expressed more highly in cases compared to controls. Red dots to the left of the 0.0 mark on the x axis represent genes expressed more highly in controls compared to cases. The further to the right or left of 0 indicates a greater magnitude of difference. The y axis shows increasing degree of significance

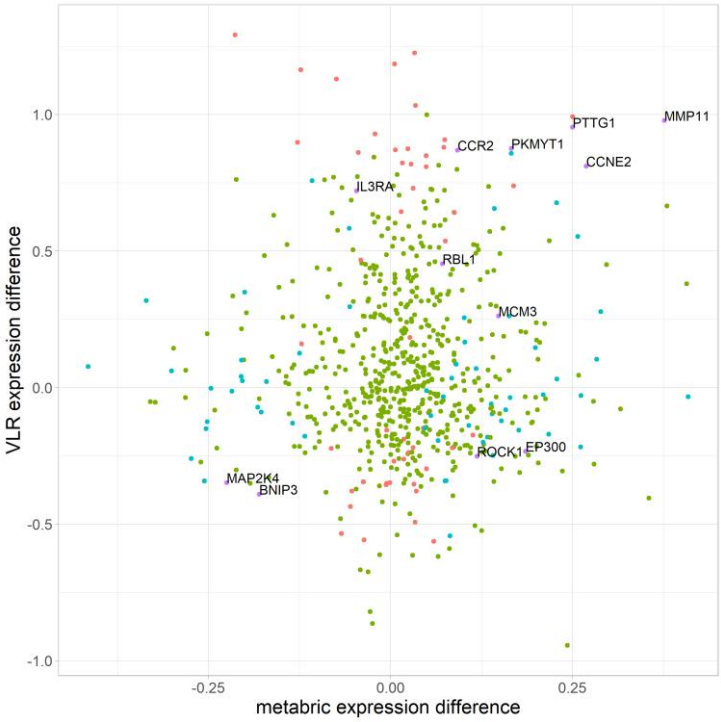


Figure 6.11 Scatter plot of expression level of genes within this study compared with the METABRIC dataset. Purple dots with labels are those genes expressed at significantly different levels in cases compared to controls in both this study and the METABRIC dataset. Red dots denote genes expressed significantly differently in this study and not in METABRIC, blue dots denote significant genes in METABRIC and not this study and green dots denote genes significant in neither.

Unsupervised hierarchical clustering of the samples according to expression of all analysed genes showed the presence of 2 distinct clusters which separated more according to their molecular subtype rather than their DR outcome status (Figure 6.12). Luminal B cancers dominated one cluster and these samples were enriched for genes involved in cell cycle progression such as *MYBL2, CDK1, CDC20* and *UBE2C* and DNA replication such as *ORC6, TYMS, SPC 25* (Figure 6.12, yellow box). The second cluster was dominated by luminal A cancers, and these were enriched for genes involved with immunity and inflammation such as *CXCL12, PLA2G4A, CCL21* and *SP1*, DNA repair genes such *as PRKDC, PARP 1, PARP4* and *RAD51C* and also cell cycle regulators such as *SMAD5, PLD1* (Figure 6.12, green box).

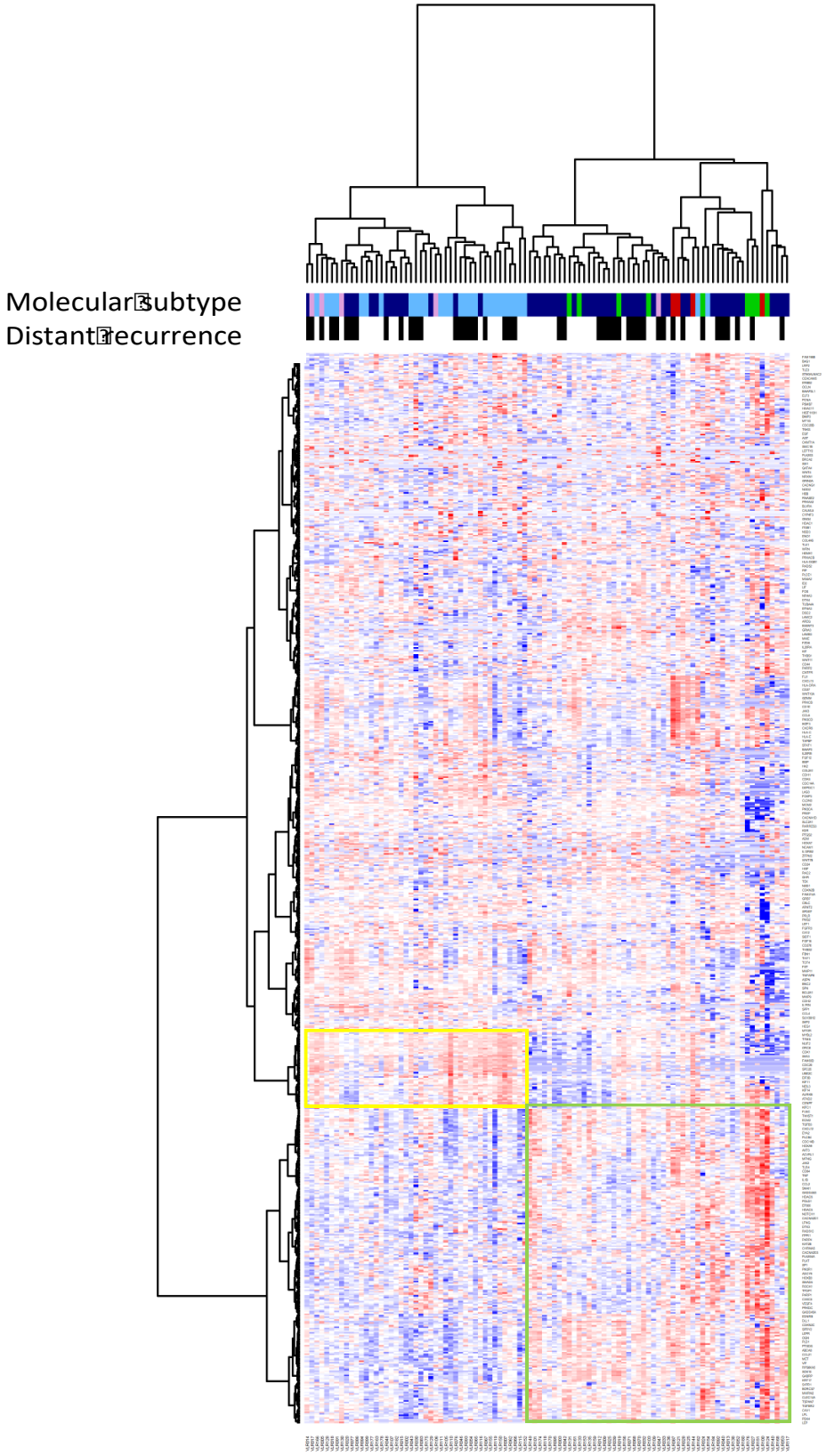


Figure 6.12 **Heatmap with unsupervised clustering of all samples analysed according to patterns of expression of significant genes. Molecular subtype shown by coloured bars (dark blue – luminal A, pale blue – luminal B, pink – HER2 enriched, red – basal, green – normal like. Recurrence is shown by black bar, non-recurrence shown by no bar. Yellow box shows cluster dominated by luminal B cancers. Green box shows cluster dominated by luminal A cancers.**

Unsupervised hierarchical clustering of the 65 significant genes highlighted three separate clusters (Figure 6.13). The most distinct of the 3 contains a subset of 24 samples. This cluster was enriched for luminal A and normal-like subtypes and contained more controls than cases. These samples were characterised by high expression of immune related genes and low expression of proliferation genes. The second cluster was again predominated by luminal A subtype but in contrast to the first cluster was enriched for recurrences. Prevalence of recurrence was statistically significant across the 3 clusters with a χ^2 of 9.74 (p=0.0077). Cluster 2 had a gene expression pattern that contrasted that of the first. The third cluster was dominated by luminal B subtype and had similar numbers of recurrences and non- recurrences. This cluster had a more random gene expression pattern but of note there was a high expression of cell cycle and DNA replication related genes (Figure 6.13, yellow box).

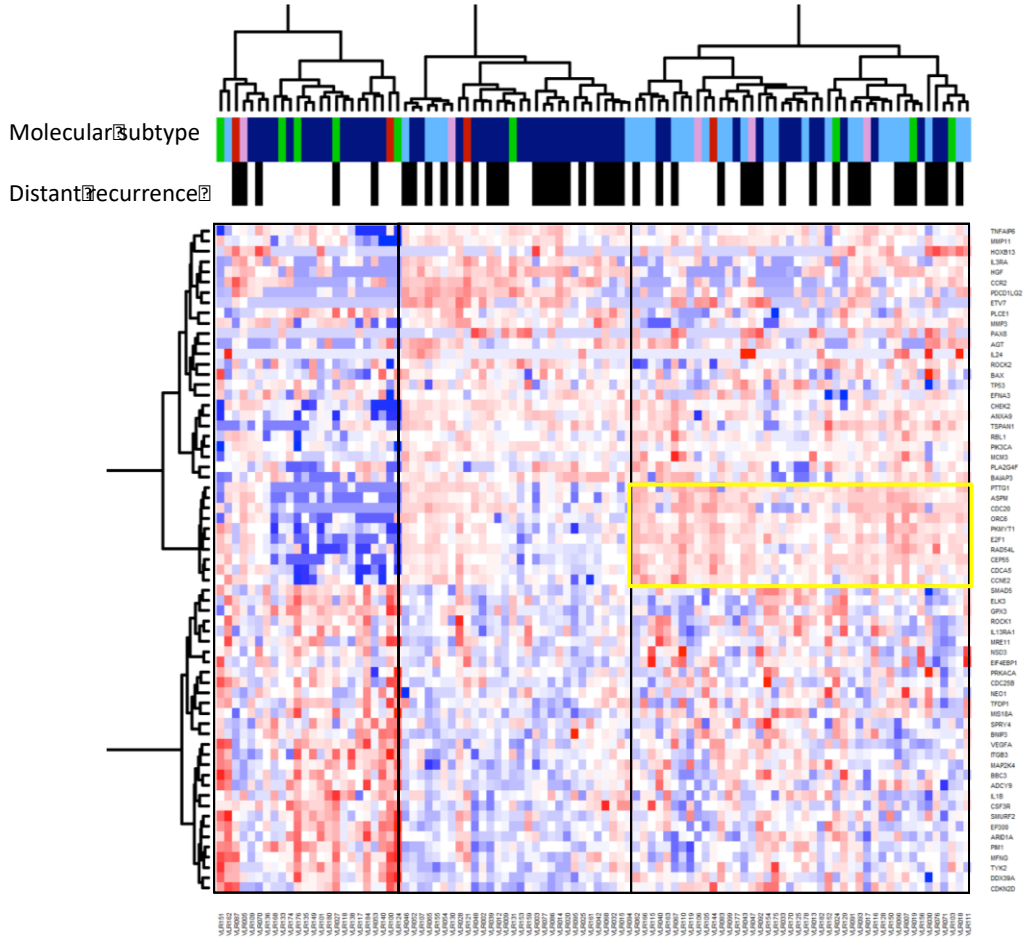


Figure 6.13 Heatmap with unsupervised clustering of all samples analysed according to patterns of expression of individual genes found to be significant. Molecular subtype shown by coloured bars (dark blue – luminal A, pale blue – luminal B, pink – HER2 enriched, red – basal, green – normal like. Recurrence is shown by black bar, non-recurrence shown by no bar. Yellow box showing the cluster dominated by luminal B subtype cancers.

6.4.7 Gene set enrichment

6.4.7.1 Reduced expression in cases compared to controls

From the 4 different gene set enrichment analyses, significantly reduced expression of gene sets involved with epigenetic regulation and cytokine and chemokine signalling were found in cases compared to controls from the breast cancer 360 module analysis (Appendix 9.8 and Figure 6.14). These were mainly incorporating genes associated with apoptosis and immune/inflammatory response such as *ARID1A* and *IL3RA1*. From the gene ontology analysis, gene sets involved in cell immune response and apoptosis such as *IL1B* and *TP53* were significantly reduced and from the KEGG gene set analysis, groups of apoptotic genes such as *BNIP3* and *BBC3* were significantly reduced. The Hallmarks gene set enrichment identified reduced expression of groups of genes associated with cell cycle inhibition and DNA damage response such as *CDKN2D* and immunity such as *IL1B* and *IL3RA1*. Following correction for multiple testing none of these gene sets remained significant.

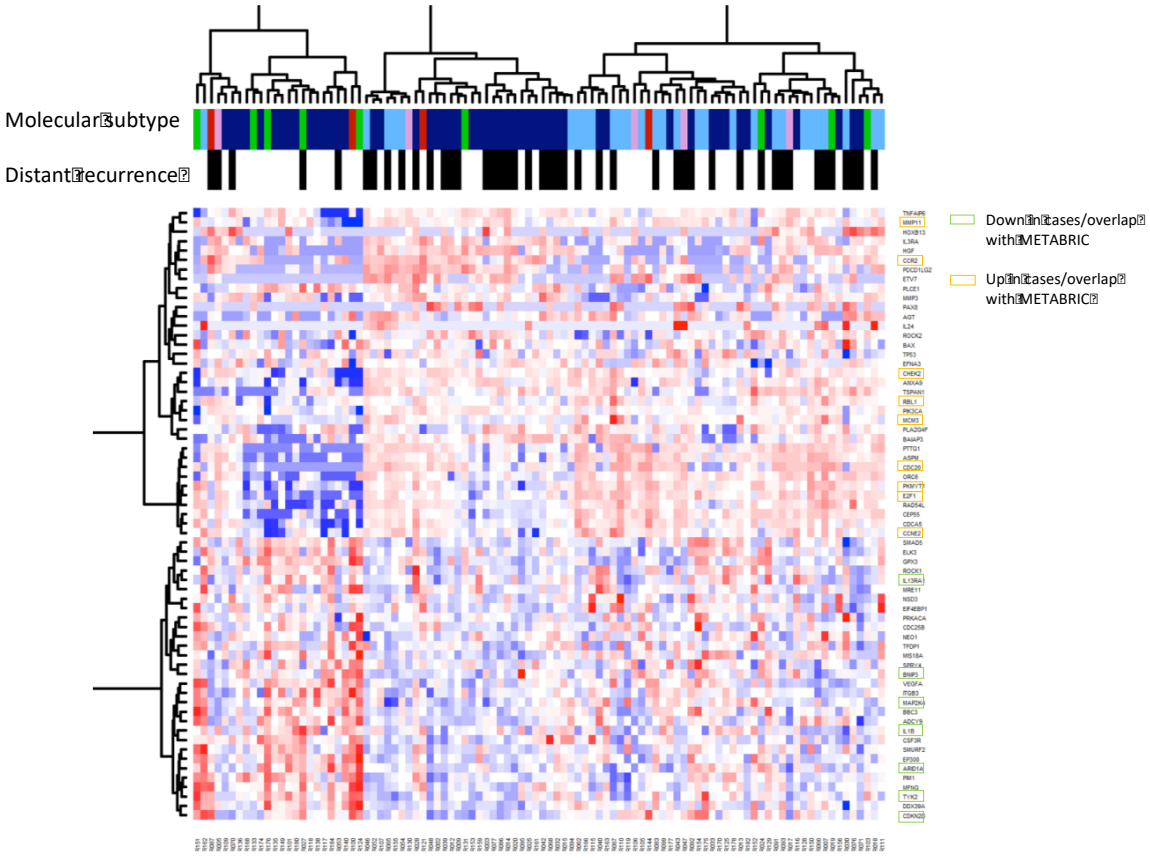


Figure 6.14 Heatmap with unsupervised clustering with key genes highlighted that feature in the gene annotations most significantly altered in cases. Highlighted genes are those that feature several times within the different annotations, overlap with those that are significant in the METABRIC study or are well described in the literature

6.4.7.2 Increased expression in cases compared to controls

None of the breast cancer 360 modules were significantly increased in cases. From the gene ontology analysis, gene sets associated with DNA replication and cell cycle progression such as *CCNE2, CDC20, CDCA5* and *CHEK2* (which also has anti-apoptotic properties) were found to be significantly increased in expression in cases compared to controls. A similar picture was seen on analysis of the KEGG (Figure 6.15 and Appendix 9.8) and Hallmark gene sets (Appendix 9.8). Many of the genes expressed at high levels in cases are those involved in promoting cell cycle pathways whilst those that were reduced included several genes involved in cell cycle inhibitory processes. None of these gene sets remained significant after correction.

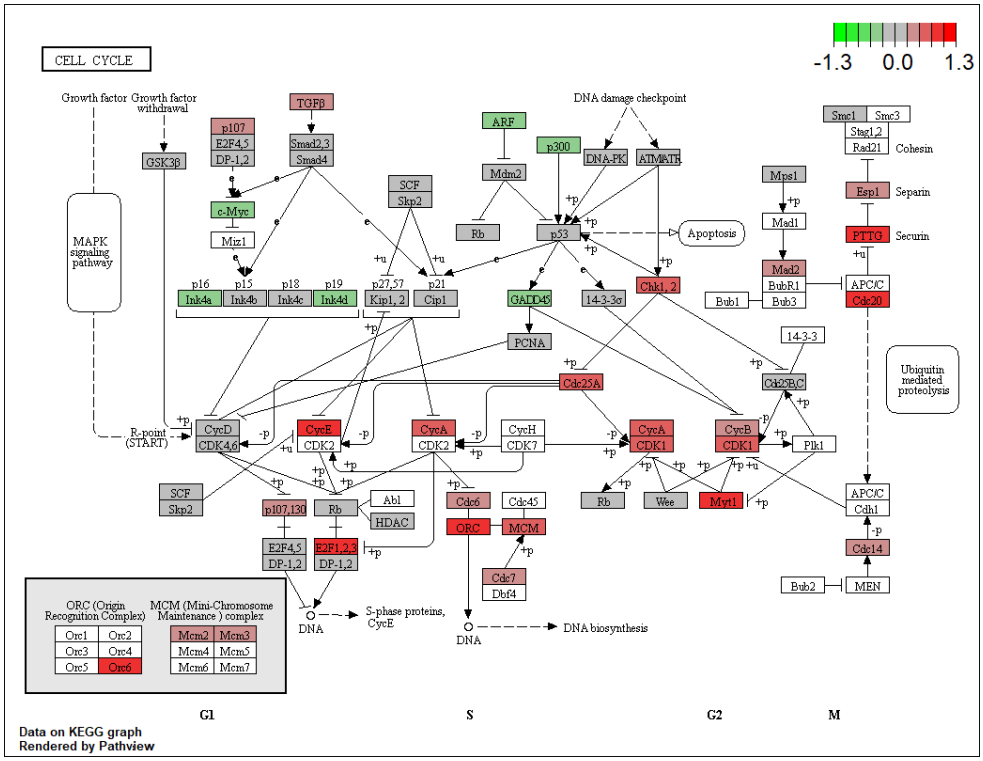


Figure 6.15 KEGG cell cycle gene set pathway showing genes expressed high in cases (red) and low in cases (green)

When comparing the breast cancer 360 and Illumina targeted sequencing panels, there were only 35 genes that overlapped. Of these genes, there was no genes significant for altered gene expression and also copy number gain or loss between cases and controls.

6.5 Discussion

6.5.1 Clinical features

In this study cases had significantly larger cancers and a higher proportion of node positive disease than controls. When incorporating the variables into the CTS5 model, the values for cases were significantly higher than for controls. It was initially anticipated that with a median time to recurrence of just under 16 years, clinico-pathological features would no longer have prognostic value for recurrence. However, clear differences in clinico-pathological stage at diagnosis between cases and controls are evident in this study. This confirms the strength of these prognostic biomarkers and the CTS5 model as a whole and how useful it could be as a prognostic tool for late recurrence even beyond 10 years.

6.5.2 Copy number alteration

In contrast to clinicopathological features, this study did not show a difference in overall percentage of genome altered in cases compared to controls or for overall gains and losses. This was compared to the METABRIC data which showed a significant difference between recurrence beyond 10 years and no recurrence beyond 20 years. The magnitude of this difference however was smaller than that seen between recurrences and non recurrences in years 0-5 suggesting that as time progresses, even those women destined to recur have greater genomic stability in comparison to those who recur early. Similarly, when comparing this case-control data with the POETIC study, patterns of CNA were similar but at a higher frequency in the latter. This is to be expected due to the 25% early recurrence rate among the POETIC samples used for comparison. Therefore, percentage of genome altered could be useful in risk stratifying patients for recurrence at time of diagnosis and possibly at 5 years but not later.

6.5.2.1 Individual copy number gains

In univariate analysis, significantly more copy number gains were seen in *CDKN1B* in controls compared to cases. *CDKN1B* encodes the p27 protein and inhibits the formation of cyclin and cyclin dependent kinase complexes thereby blocking G1 to S phase

progression of the cell cycle. SNPs in several genes including *CDKN1b* have shown associations with breast cancer risk[151] and reduced levels of p27 protein have been associated with high proliferation and poor outcomes[152]. Deletions in ER+ metastatic breast cancer samples[153]. Therefore, gains of *CDKN1B* would be consistent with the theory that stronger regulation of the cell cycle would reduce proliferation in residual micrometastatic disease and therefore prevent propagation into clinically overt metastases.

The *ESR1* gene which encodes the ER and is a key oncogenic driver of ER+ breast cancer was seen to be frequently implicated in gains in the overlapping regions of chromosome 6q25 and were more prevalent in controls compared to cases in this study. There is evidence to suggest that *ESR1* amplification correlates with ER protein levels which confers good prognosis with regards to disease free survival in women treated with endocrine therapy[154].

Gains in the *FOXP1* gene were seen in a significantly higher proportion of controls. This gene interacts with the ER protein and regulates transcription of downstream effector genes. Existing data on copy number on this gene is lacking, however expression correlates with improved relapse free survival in patients with ER+ breast cancer[155] and has shown correlation with favourable 5 years survival in tamoxifen treated patients[156].

The *PIK3R1* gene, involved in the PI3K/AKT/MTOR pathway and closely related to ER signalling pathways, showed both increased gains and deletions in controls suggesting that variation in this gene can confer good prognosis in ER+ breast cancer. Existing literature has demonstrated poor recurrence outcomes and overall survival[157,158] with low expression or knockdown of this gene.

Altogether this data suggests that gains in key genes in breast cancer may be important in long-term recurrence risk. The importance of cell cycle regulators, such as *CDKN1B*, in protecting from late recurrence is inferred as is estrogen related signalling as a key driver of ER+ breast cancer recurrence even after long periods of dormancy.

6.5.2.2 Individual copy number losses

The chromosome region encoding *AKT1* saw a significantly higher number of losses in controls. AKT1 is a downstream effector of PI3K which, in its activated form regulates cyclin and cyclin-dependent kinase control of the cell cycle. High expression of *AKT1* has been associated with poor recurrence outcomes and overall survival in breast cancer patients[159]. The PI3K/AKT pathway is frequently implicated in endocrine therapy resistance in metastatic disease and AKT inhibitor capivasertib has shown progression free survival benefit in combination with fulvestrant in advanced breast cancer that has progressed on an AI[160].

Genes involved in cell cycle progression such as *CDC4a, CCL5, PIMREG, PITPNM3* and *RIOK1* are all incorporated in areas of copy number loss that were significantly higher in controls rather than cases. This apparent protective advantage from long-term recurrence is consistent with pre-existing data. Such data reports the proliferative effect of high expression of these genes or their corresponding proteins such as *CCL5*[161], *PIMREG*[162] and *RIOK*[163]and also how low expression or knockdown of genes such as *CDC4a*[164] and *PITPNM3*[165,166] reduce cell cycle progression and inhibit formation of metastases.

The chromosome region of 14q32, which showed more frequent losses in controls, incorporates the *MTA1* oncogene which drives breast tumourigenesis and metastatic invasion. This has shown an association with tamoxifen resistance in breast cancer patients and higher levels have been associated with poor prognosis and increased recurrence following tamoxifen treatment[167].

17q is frequently altered in breast cancer however this is largely thought to be due to the presence of *HER2/ERBB2, BRCA1* and *NME2*. The overlapping regions of 17q that were found to be deleted in this study did not incorporate these genes perhaps suggesting that genes within this region other than *HER2* are important in breast cancer recurrence. Gains in 17q12-22 are associated with luminal A breast cancer[168,169] and gains in 17q23-25 are associated with non-recurrences when compared to recurrences. Whilst this finding was not reflected in this very late recurrence case control study, it would seem that this

chromosome region is relevant in the pathogenesis and progression of ER+ breast cancer and this is supported by the literature[170].

Altogether, data from the copy number losses seen in this study show the potential protective effect of loss of cell cycle drivers such as *CCL5, CDC4* and *PITPNM3* and that their loss may confer a good prognosis in terms of late recurrence. However, it is also clear that overall genomic stability is probably not a useful prognostic marker in this setting.

6.5.3 Mutations

Overall somatic mutational burden did not differ between cases and controls and this could be due to similar reasons for the lack of difference in percentage of genome altered. Several mutations showed significance or a trend for higher prevalence in cases over controls and vice versa. These, whilst losing significance on multiple testing, are worthwhile considering as potential candidate biomarkers owing to their consistency with published data.

In particular, significantly higher proportions of *MAP3K1* mutations in controls versus cases (reflected in the METABRIC analysis) supports existing data that in ER+ breast cancer, this mutation confers a good prognosis in terms of its association with lower stage clinico-pathological features[171] and association with improved survival outcomes[172]. Additionally, there was a trend, although not significant, for a higher proportion of *GATA-3* mutations in controls which was also reflected in the METABRIC analyses and is consistent with reports that *GATA-3* mutations are associated with lower pathological stage and improved survival outcomes[172,173]. *CDH1* mutations seen here at higher proportions in cases, are observed in ER+ breast cancer and are more frequently associated with lobular breast cancer [172,174]. The CDH1 gene functions as a tumour suppressor and mutations are usually inactivating. Association with recurrence outcome is not fully understood in the literature.

Altogether the DNA aberration data is consistent with published findings in the literature and reflect biologic theory that long term non-recurrent breast cancer has a low level of genomic instability. This might be useful to stratify risk between early and late recurrence but not for patients who have all reached 10 years without recurrence. In terms of

mutations, whilst none of the differences between cases and controls remained significant after correction, this data supports existing reports that *MAP3K1* and *GATA-3* mutations are likely to confer a good prognosis even after long follow up. Other mutations and copy number alterations identified here could serve as potential candidate biomarkers such as those involved in cell cycle regulation and warrant investigation going forward.

6.5.4 Gene expression

On an individual gene level, following correction for multiple testing, no single genes were found to be significantly associated with cases or controls. However, univariate analysis found several genes differentially expressed between cases and controls two of which (*CDC25B* and *CDKN2D*) are known regulators of cell cycle progression. In addition, several of the significant genes which overlap with published data from METABRIC are involved in cell division and cell cycling (*CCNE2, PKMYT*) and also are important in ER related cell signalling (*CCR2*). Observations from the hierarchical clustering suggest that while intrinsic subtyping lost prognostic importance beyond 10 years, 2 groups of largely luminal A tumours were revealed, one of which was strongly associated with increased risk of late recurrence. Cell cycle and proliferation genes were upregulated in this subgroup while immune-related genes and cell cycle inhibitors (eg *CDKN2D*) were down-regulated.

6.5.4.1 High gene expression in cases

Several of the genes implicated in the hierarchical clustering also feature in the gene set annotations examined here. The most notable finding was the increased expression in cases of genes involved in the cell cycle and proliferation. The KEGG pathway map (Figure 6.15) highlights how many proliferative genes were raised in cases along with reductions in expression of cell cycle inhibitors such as those within the *INK* family. High expression of *CCNE2* seen in this study and METABRIC (Figure 6.14) which encodes cyclin E, a key component in the cell cycle driving transition from G2 to S phase has been shown to be associated with shorter distant metastasis free survival[175] and with endocrine therapy resistance[176]. In a neo-adjuvant study of palbociclib, responders had a greater decrease in *CCNE2* levels compared with non-responders[177]. Additionally, higher levels of cyclin E confer a poor prognosis with regards to OS and BCSS in a meta-analysis[178].

In addition I observed higher expression of *E2F1, CDC20* and *MCM3* (Figure 6.14) which are key oncogenes and are also implicated in cell cycle progression. Their high expression has previously shown association with reduced relapse free survival in breast cancer[179,180] and associations with poor clinicopathological features as well as poor overall survival and response to endocrine therapy[181 182].

6.5.4.2 Low gene expression in cases

Many of the genes with lower expression in cases are associated with immune or inflammatory processes. High expression of *IL1B* has been associated with prevention of differentiation of metastasis initiating cells in cell lines and also with distant metastasis-free survival in breast cancer tissue[183]. It has been reported that high expression of genes involved in cytokine synthesis is associated with improved early and late recurrence free survival[184] suggesting the importance of immune infiltration in protecting from late recurrence . However, these findings are in contrast to research reporting high expression of these genes and immune cell infiltrates confer resistance to AIs perhaps suggesting a temporal change in the effect of these genes[47]. This switching of effect with time in breast cancer is seen with other prognostic markers such as ER gene levels[74] as discussed previously.

6.5.4.3 Comparison between copy number and gene expression data

Despite minimal overlap in genes featuring on both the DNA and RNA panels, trends in expression and DNA aberration in genes from the same family showed overlap in terms of likely importance in very late versus non-recurrence. Increased copy number gain of the chromosome region encoding *CDKN1B* in controls is consistent with the increased expression of *CDKN2D* RNA in controls supporting the protective effect of cell cycle inhibitors. Additionally, the finding of a greater number of deletions in 14q32 around the *CDCA4* locus in the controls is consistent with the higher expression of *CDCA5* RNA in cases. Both of these are oncogenes within the family of genes that drive cell cycle progression.

One possible explanation for the findings in this study is that the emergence of metastatic disease is dependent on both the physical tumour load (as determined by the size, grade etc) and the molecular features that drive behaviour of the cancer cells. These two factors,

along with a myriad of other determinants as discussed in chapter 1, result either in cure or recurrence after a varied duration of dormancy. The tumour load is not temporal and, following surgery, its ongoing effect on long term behaviour of residual cancer cells is pre-determined and does not change. Molecular features within any remaining cells however will continue to exert an effect over time with the most aggressive features bringing about early recurrence. In this study, which lacked an early recurrence comparator group, both cases and controls had all experienced long-term dormancy and therefore their tumours would be enriched for prognostically favourable molecular features making it difficult to find statistically significant differences in this regard.

6.5.5 Strengths of the study

A unique feature of this study is its long follow up time for both cases and controls within a single institution. For ER+ breast cancer, where recurrence can occur so many years after the initial primary, this is a key factor in studies of recurrence prediction and one that is not often practised or possible. The long follow up times for the patients here means that one can be confident that the controls were likely to be true non-recurrences and not merely women with dormant disease with a strong potential to recur. Additionally, comprehensive analyses and bio-informatics has enabled comparisons with existing datasets which give greater credence and clinical context to the findings presented here.

6.5.6 Limitations of the study

This exploratory study was hampered by the relatively small number of cases partly due to high attrition rates from the original number of cases and controls identified at the start. Even though RMH retains the majority of specimens indefinitely, there was an appreciable proportion of samples that could not be located or that had insufficient invasive tumour for analysis. The smaller number of samples than originally planned impaired the power of this study and contributed to the difficulty in identifying significant differences between cases and controls. In particular, with the low sample number, correction following multiple testing meant that true significant differences would be hard to detect despite the

use of a conservative correction method such as Benjamini Hochberg. The discussion above, therefore, needs to be viewed in a cautious fashion.

Additionally, this study did not include a third group of early recurrences with which to compare late recurrences and non-recurrent controls. This was intentional owing to the specific research question being addressed and due to limits in budget and resources but would be an interesting element to include in future.

6.5.7 Clinical Application

The most notable application of this research is the likely ongoing relevance of CTS5 as a tool for recurrence risk estimation beyond 10 years. This supports data showing that that women at all stages but particularly those with higher stage features remain at risk for 20 years or more [25]. This could be borne in mind now when planning appropriate follow up and management of risk amongst these patients.

6.5.8 Research application

The findings pertaining to CTS5 presented and discussed herein suggest that validating CTS5 in a population with 20+ years follow up is warranted. As stated in the introduction of this chapter, there remains a need to further risk stratify women after 5 additional years of endocrine therapy. Given its significant association with recurrences in this study validation should be sought to address this particular question of its applicability for very late recurrence.

DNA and RNA sequencing findings from this study merit further investigation in the context of larger translational studies. Candidate biomarkers such as GATA3 and MAP3K1 mutations should be explored further. The effect of expression of proliferation genes and immune related genes and their related gene pathways should be investigated in larger sample numbers from archival clinical trial tissue with long follow up times to truly delineate their impact on late and very late recurrence which is lacking in the literature. The finding of the association of immune-related genes which may impede early

recurrence with increased risk of late recurrence suggests that the use of check-point inhibitors might have a role in preventing late recurrence and could be explored.

Additionally, further exploration of already developed mutational and copy number signatures[104] could reveal trends in DNA aberrations or key subtypes which increase the risk of late or very late recurrence. It would also be sensible to include comparisons with molecular features of early recurrence as this may reveal patterns that were undetectable in this study owing to the likely indolent nature and low genomic burden of the micro-metastatic disease of the late recurrence cases that were analysed.

7. Overall conclusions and fulfilment of aims

7.1 Conclusions from CTS5 validation in RMH

The validation of CTS5 conducted here was the first performed on the re-calibrated CTS5 from the combined ATAC and BIG 1-98 data. This study concluded that CTS5 was prognostic for both pre- and post-menopausal women. For post-menopausal women, there was a strong prognostic association and crucially the model separated the cohort into three distinct risk categories which is useful for guiding decisions regarding extended endocrine

therapy. The data supported the ongoing use of CTS5 in post-menopausal women and the omission of extended endocrine therapy in women deemed to be low-risk by CTS5. Risk stratification was less distinct for pre-menopausal women and as it stands the data does not support the omission of extended endocrine therapy in the low-risk group.

7.1.1 Fulfilment of aims of this study

The main aim of this study was to test the validity of CTS5 within an unselected non-clinical trial population of RMH patients. Specific aims were to establish whether or not CTS5 had a similar prognostic performance in the RMH dataset compared to the development set, to establish whether CTS5 sufficiently accurately predicted late recurrence in pre- and post-menopausal women and to establish whether the model adequately divided the cohort into clinically useful risk groups. These three aims are all key determinants of validation.

The study conducted here was of robust methodology, with good sample size and had an appropriate analysis plan with which to test validity. The study was able to draw firm conclusions regarding the post-menopausal cohort but not so for the pre-menopausal cohort. Therefore, whilst all three aims were addressed and appropriately tested, this study alone was not able to conclusively establish or refute the validity of CTS5 in pre-menopausal women. It is therefore proposed that further tests of validation on pre-menopausal women would be useful as well as potentially re-training of the model if it performed poorly in other cohorts.

7.2 Conclusions from clinical impact study

This study examined the impact that knowledge of the CTS5 value for individual patients had on MDT recommendations regarding extended endocrine therapy. The study concluded that the lack of change in recommendations arose because the variables within the CTS5 algorithm are well known –particularly in this tertiary centre, and therefore were the basis on which clinicians were already making decisions. It was further concluded that although, in this centre, CTS5 did not make a large impact, there were few disadvantages to its use and therefore should continue to be used as a decision aid. Additionally, owing

to the striking concordance in recommendations for low and high-risk patients, it was proposed that, for those women, discussion in the MDT was probably not necessary.

It was further suggested that tests of impact of CTS5 should be conducted in a variety of settings including non-specialist centres. Generating further data for predictive value of CTS5 would also add to the evidence base for utility.

7.2.1 Fulfilment of aims of this study

The aim of this study was to test clinical utility of CTS5 for changing practise with regards to extended endocrine therapy recommendations using a decision impact study. Through this appropriately designed and conducted study, this aim was achieved. However, this study concluded that to have conducted it in a multicentre setting involving less experienced oncologists would probably have been more informative given the specialist knowledge available to clinicians in a tertiary referral centre. Although not an aim of the study, a potential pathway for managing intermediate risk women emerged which could have a positive clinical impact.

7.3 Conclusions from CTS5 validation in high stage women

The test of CTS5 in the TACT trial dataset assessed whether the model could extrapolate to higher stage women and pre-menopausal women. This study concluded that CTS5 was valid for use in high stage, post-menopausal women and data was consistent with existing evidence. These findings broaden the application of CTS5 to higher stage post-menopausal women in whom, a low-risk group who can avoid extended endocrine therapy were identified. For pre-menopausal women in this population, it was concluded that CTS5 lacks prognostic accuracy and did not adequately divide the cohort in a clinically useful way. It concluded that CTS5 in its current form cannot be applied to high stage pre-menopausal women. It was recommended that CTS5 be tested in a cohort of lower risk pre-menopausal women to further test its validity and in addition that a separate CTS5 for pre-menopausal women be trained.

7.3.1 Fulfilment of aims of this study

The aims of this study were the same as for the RMH validation but applied to this higher stage population with a higher proportion of younger and pre-menopausal women. The test of validation in post-menopausal women conducted here was in an appropriate population in order to address the question of higher stage women. This study robustly addressed the issue of higher stage women and therefore the central aim of this study was achieved. The appropriateness of testing validation of CTS5 for pre-menopausal women rather than building a new model from scratch was discussed in chapter 4 and the suitability of having this aim at all could be questioned. The study was unable to show validity of CTS5 for high stage, pre-menopausal women and it was concluded that stark differences in this population compared with the development set meant that extrapolating for this particular group was not achievable.

7.4 Conclusions from IHC biomarker study

This study tested the prognostic value of IHC biomarkers (including HER2) for late recurrence in the same high-stage TACT population in which CTS5 was investigated. No association was found between the IHC-measured variables and late recurrence and their presence in a prognostic model alongside CTS5 did not improve its prognostic value. Conclusions regarding the impact of HER2 status could not be drawn owing to the historic nature of this cohort and differences in management of these women compared to the present day. The quality of the data within this study was robust and the automated scoring system is worthwhile pursuing for increasing efficiency and consistency in IHC based research. This study concluded that the lack of prognostic association between biomarkers and late recurrence was likely due to the solely high clinical stage of the population studied and the fact that all patients within the study received chemotherapy. This is in stark contrast to most datasets used in previous translational studies of this type and therefore it is not unexpected that previously observed associations did not extrapolate. It was recommended that these biomarkers be explored further in a lower-risk and more generally representative population of early breast cancer sufferers. Additionally re-training CTS5 to incorporate data on HER2+ women (who have received HER2 targeted

treatment) is an important area of currently unmet need for this sub-group of breast cancer sufferers.

7.4.1 Fulfilment of aims of this study

The aims of this study were: to test the association of molecular biomarkers with late recurrence in pre- and post-menopausal women from the TACT population, develop a prognostic model using these biomarkers and test it for its additional prognostic value when combined with CTS5. It also aimed to explore the possibility for re-calibrating CTS5 for HER2+ women and to evaluate automated visual analysis as a biomarker scoring method.

Aim one was achieved in the study which used robust laboratory methods and an appropriate study design. The choice of population meant that the more generic question of biomarker association with late recurrence has not been fully answered as this study only addresses this issue within the high stage TACT population. The second and third aims were only partially investigated owing to the lack of prognostic association on univariate level of these biomarkers for late recurrence in this population. Complex modelling was not attempted and so this aim was not achieved. The fourth aim pertaining to HER2+ women was also not achieved as it was decided that a CTS5 model for HER2+ women (trained on this data) would not reflect the treatment effect of HER2 targeted therapy that is routine nowadays. The final aim however, of evaluating an automated method of scoring of IHC stained tissue was investigated and conclusions drawn. Further evaluative testing and optimisation would have been ideal but were not achievable in this study and should be further explored.

7.5 Conclusions from very late recurrence tissue study

This exploratory study analysed clinical and molecular differences in primary tumours of cancers that recurred very late versus non-recurrent tumours. The main conclusions of this study were that, clinico-pathological parameters significantly differ in late recurrences compared to non-recurrences. This was reflected in the differences in CTS5 between cases

and controls. It was suggested that CTS5 should be validated for very late recurrence in a larger sample size for its clinical validity but could be used as a guide immediately for extended endocrine therapy decision making at 10 years. In addition the study found that after long periods of dormancy, genomic stability is a less discerning biomarker for recurrence. Cell cycle regulators, estrogen signalling related genes and apoptosis related genes present potential candidate biomarker for future research into very late recurrence.

7.5.1 Fulfilment of aims of this study

The aim of this study was to explore clinical and molecular differences in primary cancers from women who recurred beyond 10 years compared to those who were recurrence free for at least 20 years from diagnosis. The rationale behind this aim was the need to identify biomarkers to risk stratify women who are 10 years or more from diagnosis given their ongoing risk of recurrence. This study comprehensively explored these differences using samples with very long follow up and drew conclusions which show consistency with published data. Conclusions were drawn regarding clinico-pathological differences and potential molecular biomarkers that can be taken forward for further translational investigation were identified. Therefore the aim of this study was realised although the limited number of samples available likely reduced its conclusiveness.

The long follow up, consistency with published work and comprehensive analysis means that this study contributes a unique set of both clinical and molecular data (investigating RNA and DNA) using robust methodology. This is rare in this field and will certainly have its place in the literature.

7.6 Overall aims

The first overall aim of this research was to test and develop existing prognostic models in a variety of independent patient cohorts in order to improve and broaden the scope of late recurrence risk estimation. The studies conducted here have robustly tested the CTS5 model and shown evidence of its validity in an independent dataset and consistent results

for post-menopausal women across two different datasets. The CTS5 model has been developed in the sense that the scope of women in whom risk prediction is possible has been broadened and therefore this aim has been achieved.

The aim to *develop* existing prognostic models was also tested in the IHC biomarker study, which, whilst successfully exploring prognostic associations, did not achieve its intended aim of developing existing prognostic models owing to an absence of prognostic association found in that study.

The second aim was to explore molecular differences in tumours that recur very late compared to those that do not recur at all. This aim was realised in the very late recurrence tissue study which has proposed potential biomarkers for future larger translational studies.

The practice of medical oncology is inextricably linked with research. Therefore, experience and appreciation of related research is essential and will undoubtedly improve clinical practice. For this reason, personal aims for this MD(Res) were identified.

One of these was to undertake an element of research with direct clinical relevance. This has been achieved in the work I have done with CTS5 – both in the tests of validation and the study of clinical utility. These studies were all clinically relevant whose outcomes several of which can be implemented immediately. The other aim was to develop translational laboratory-based skills and an overall understanding of techniques in this field and how they relate to medical oncology. The laboratory based studies enabled acquisition of a multitude of translational laboratory based skills, along with complex computer based analyses. The clinical studies involved data trawling and manual data extraction taking many weeks of work. All studies provided experience in literature reviews, proposal writing and ethics approval – all of which was undertaken independently as well as statistical analysis in which a confident knowledge base was required and collaborating with colleagues of many different disciplines.

7.7 Overarching Hypothesis

The overarching hypothesis for this research was: Predicting risk of late recurrence can be improved by development of existing clinical and molecular biomarkers and prognostic models.

This research has been partially focused on clinically validating and assessing utility of an existing prognostic model for late recurrence in ER+ breast cancer. One can strongly argue that the evidence presented in this thesis proves the hypothesis that risk prediction can be improved. This is with regards to the fact that there is now additional confidence of the accuracy of CTS5 in post-menopausal women including those with high stage disease and those not selected within a clinical trial. Additionally, evidence presented here demonstrates that clinicians are prepared to take management decisions which are concordant with those anticipated by the developers of the model. Altogether, data presented in this thesis has enhanced the evidence base for CTS5 and developed its scope and provides resources for clinicians to ensure its appropriate use.

Alternatively, it could be argued that the hypothesis set out in this thesis should be rejected on certain bases. The evidence reported in this thesis failed to conclusively validate or invalidate CTS5 in pre-menopausal women (although it conclusively invalidated it for high stage premenopausal women). Importantly, issues highlighted with study designs and population selection has meant that risk prediction has not been developed for HER2+ women and has not been shown to improve using IHC molecular biomarkers. Additionally, utility, highlighted as a fundamental part of validation, has not been demonstrated and the findings of the very late recurrence tissue study are too exploratory at this stage to show association with risk of late recurrence. However, it is likely that molecular biomarkers (either molecular of IHC based) will be able to demonstrate prognostic significance for late recurrence and will be instrumental in risk stratifying women at key junctures of their survivorship and this research has contributed towards this investigative process.

7.8 Concluding remarks

Risk prediction in ER+ breast cancer has developed into a large sphere of research over the past 15 years. With irrefutable evidence that distant recurrence is as likely beyond 5 years as before, it will and should remain a focus for clinicians and researchers. This thesis has explored a small corner of this field of research in an attempt to refine and develop existing models and biomarkers so that they may be applied clinically and investigated further in translational studies in the future. For survivors of breast cancer, risk stratification at various key time-points from their diagnosis will provide them with knowledge and armoury to manage their risk of recurrence and in turn enable their clinicians to tailor their long term management appropriately.

8. References

1. Boyd S. On Oophorectomy in the Treatment of cancer. British Medical Journal 1897;2:890-896.
2. Allen ED, E. An ovarian hormone:preliminary report on its localization, extraction and partial purification and action in test animals. Journal of the American Medical Association 1923;81:819-821.
3. Lacassagne A. Hormonal pathogenesis of adenocarcinoma of the breast. . American Journal of Cancer 1936;27:217-225.
4. Toft DGJ. A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization. Proc Natl Acad Sci U S A 1966;55:1574-1581.
5. Toft DS, G Gorski, J. . A receptor moleculefor estrogens: studies using a cell-free system. Proc Natl Acad Sci U S A 1967;57:1740-1743.
6. Jordan V. The development of tamoxifen for breast cancer therapy: a tribute to the late Arthur L. Walpole. Breast Cancer Res Treat 1988;11:197-209.
7. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. N Engl J Med 2006;354(3):270-82.
8. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747-52.
9. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98(19):10869-74.

10. Cancer Research UK. Breast cancer incidence (invasive). (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero). 18/11/19.
11. Li CI, Daling JR, Malone KE. Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. J Clin Oncol 2003;21(1):28-34.
12. Dodson A, Parry S, Ibrahim M, et al. Breast cancer biomarkers in clinical testing: analysis of a UK national external quality assessment scheme for immunocytochemistry and in situ hybridisation database containing results from 199 300 patients. J Pathol Clin Res 2018;4(4):262-273.
13. Cancer Research UK. Breast cancer incidence by age. (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One). 18/11/19.
14. Cancer Research UK. Projections of incidence for breast cancer. (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Five). 18/11/19.
15. Cancer Research UK. Breast Cancer Risk. (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/risk-factors). 18/11/19.
16. Cancer Research UK. Breast Cancer incidence by stage at diagnosis. (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Three). 18/11/19.
17. Cancer Research UK. Routes to diagnosis of breast cancer. (https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/diagnosis-and-treatment#heading-Zero). 18/11/19.
18. Davies C, Godwin J, Gray R, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378(9793):771-84.
19. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. Early Breast Cancer Trialists' Collaborative Group. Lancet 1992;339(8785):71-85.
20. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet 1998;351(9114):1451-67.
21. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet 2015;386(10001):1341-1352.
22. Dowsett M, Cuzick J, Ingle J, et al. Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen. J Clin Oncol 2010;28(3):509-18.
23. Cancer Research UK. Breast Cancer Statistics. (http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero).May 2018
24. Schairer C, Mink PJ, Carroll L, Devesa SS. Probabilities of death from breast cancer and other causes among female breast cancer patients. J Natl Cancer Inst 2004;96(17):1311-21.

25. Pan H, Gray R, Braybrooke J, et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. N Engl J Med 2017;377(19):1836-1846.
26. Esserman LJ, Moore DH, Tsing PJ, et al. Biologic markers determine both the risk and the timing of recurrence in breast cancer. Breast Cancer Res Treat 2011;129(2):607-16.
27. Rosenberg PS, Barker KA, Anderson WF. Estrogen Receptor Status and the Future Burden of Invasive and In Situ Breast Cancers in the United States. J Natl Cancer Inst 2015;107(9).
28. Fisher B, Dignam J, Bryant J, Wolmark N. Five versus more than five years of tamoxifen for lymph node-negative breast cancer: updated findings from the National Surgical Adjuvant Breast and Bowel Project B-14 randomized trial. J Natl Cancer Inst 2001;93(9):684-90.
29. Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381(9869):805-16.
30. Gray RR, D. Handley, K. . aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. Journal of Clinical Oncology 2013;31 (18_suppl)(18):5.
31. Goss PE, Ingle JN, Martino S, 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):1262-71.
32. Mamounas EP, Jeong JH, Wickerham DL, et al. Benefit from exemestane as extended adjuvant therapy after 5 years of adjuvant tamoxifen: intention-to-treat analysis of the National Surgical Adjuvant Breast And Bowel Project B-33 trial. J Clin Oncol 2008;26(12):1965-71.
33. Goss PE, Ingle JN, Pritchard KI, et al. Extending Aromatase-Inhibitor Adjuvant Therapy to 10 Years. N Engl J Med 2016;375(3):209-19.
34. Tjan-Heijnen VCG, van Hellemond IEG, Peer PGM, et al. Extended adjuvant aromatase inhibition after sequential endocrine therapy (DATA): a randomised, phase 3 trial. Lancet Oncol 2017;18(11):1502-1511.
35. Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg E, et al. Optimal Duration of Extended Adjuvant Endocrine Therapy for Early Breast Cancer; Results of the IDEAL Trial (BOOG 2006-05). J Natl Cancer Inst 2018;110(1).
36. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nat Rev Cancer 2007;7(11):834-46.
37. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. Nat Rev Cancer 2014;14(9):611-22.
38. Yeh AC, Ramaswamy S. Mechanisms of Cancer Cell Dormancy--Another Hallmark of Cancer? Cancer Res 2015;75(23):5014-22.
39. Ogba N, Manning NG, Bliesner BS, et al. Luminal breast cancer metastases and tumor arousal from dormancy are promoted by direct actions of estradiol and progesterone on the malignant cells. Breast Cancer Res 2014;16(6):489.
40. Fluegen G, Avivar-Valderas A, Wang Y, et al. Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. Nat Cell Biol 2017;19(2):120-132.

41. Johnson RW, Finger EC, Olcina MM, et al. Induction of LIFR confers a dormancy phenotype in breast cancer cells disseminated to the bone marrow. Nat Cell Biol 2016;18(10):1078-1089.
42. Gao H, Chakraborty G, Lee-Lim AP, et al. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. Cell 2012;150(4):764-79.
43. Ghajar CM, Peinado H, Mori H, et al. The perivascular niche regulates breast tumour dormancy. Nat Cell Biol 2013;15(7):807-17.
44. Lu X, Mu E, Wei Y, et al. VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging alpha4beta1-positive osteoclast progenitors. Cancer Cell 2011;20(6):701-14.
45. Lawson DA, Bhakta NR, Kessenbrock K, et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature 2015;526(7571):131-5.
46. Naumov GN, Bender E, Zurakowski D, et al. A model of human tumor dormancy: an angiogenic switch from the nonangiogenic phenotype. J Natl Cancer Inst 2006;98(5):316-25.
47. Dunbier AK, Ghazoui Z, Anderson H, et al. Molecular profiling of aromatase inhibitor-treated postmenopausal breast tumors identifies immune-related correlates of resistance. Clin Cancer Res 2013;19(10):2775-86.
48. Denkert C, Loibl S, Noske A, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol 2010;28(1):105-13.
49. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, Yuan Y. Relevance of Spatial Heterogeneity of Immune Infiltration for Predicting Risk of Recurrence After Endocrine Therapy of ER+ Breast Cancer. J Natl Cancer Inst 2018;110(2).
50. Kim RS, Avivar-Valderas A, Estrada Y, et al. Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. PLoS One 2012;7(4):e35569.
51. Gawrzak S, Rinaldi L, Gregorio S, et al. MSK1 regulates luminal cell differentiation and metastatic dormancy in ER(+) breast cancer. Nat Cell Biol 2018;20(2):211-221.
52. Walter D, Lier A, Geiselhart A, et al. Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. Nature 2015;520(7548):549-52.
53. Richman J, Dowsett M. Beyond 5 years: enduring risk of recurrence in oestrogen receptor-positive breast cancer. Nat Rev Clin Oncol 2019;16(5):296-311.
54. Lee ES, Han W, Kim MK, et al. Factors associated with late recurrence after completion of 5-year adjuvant tamoxifen in estrogen receptor positive breast cancer. BMC Cancer 2016;16:430.
55. Sestak I, Dowsett M, Zabaglo L, et al. Factors predicting late recurrence for estrogen receptor-positive breast cancer. J Natl Cancer Inst 2013;105(19):1504-11.
56. Yamashita H, Ogiya A, Shien T, et al. Clinicopathological factors predicting early and late distant recurrence in estrogen receptor-positive, HER2-negative breast cancer. Breast Cancer 2016;23(6):830-843.
57. Baek SY, Kwon JY, Lee YJ, et al. Prediction of Late Breast Cancer-Specific Mortality in Recurrence-Free Breast Cancer Survivors Treated for Five Years with Tamoxifen. J Breast Cancer 2019;22(3):387-398.

58. Dowsett M, Sestak I, Regan MM, et al. Integration of Clinical Variables for the Prediction of Late Distant Recurrence in Patients With Estrogen Receptor-Positive Breast Cancer Treated With 5 Years of Endocrine Therapy: CTS5. J Clin Oncol 2018;36(19):1941-1948.
59. Candido Dos Reis FJ, Wishart GC, Dicks EM, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 2017;19(1):58.
60. Dowsett M, Sestak I, Regan MM, et al. Integration of Clinical Variables for the Prediction of Late Distant Recurrence in Patients With Estrogen Receptor-Positive Breast Cancer Treated With 5 Years of Endocrine Therapy: CTS5. J Clin Oncol 2018:Jco2017764258.
61. Cuzick J, Dowsett M, Pineda S, 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):4273-8.
62. Howell A, Cuzick J, Baum M, 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):60-2.
63. Sgroi D, Zhang Y, Schnabel CA. Comparative performance of Breast Cancer Index (BCIN+) and CTS5 in hormone receptor-positive (HR+) lymph node-positive (N+) breast cancer patients with one to three positive nodes (N1). Journal of Clinical Oncology 2019;37(15_suppl):555-555.
64. Kennecke HF, Olivotto IA, Speers C, 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):45-51.
65. Nordenskjold A, Fohlin H, Fornander T, Lofdahl B, Skoog L, Stal O. Progesterone receptor positivity is a predictor of long-term benefit from adjuvant tamoxifen treatment of estrogen receptor positive breast cancer. Breast Cancer Res Treat 2016;160(2):313-322.
66. Noordhoek I, de Groot AF, Cohen D, Liefers GJ, Portielje JEA, Kroep JR. Higher ER load is not associated with better outcome in stage 1-3 breast cancer: a descriptive overview of quantitative HR analysis in operable breast cancer. Breast Cancer Res Treat 2019;176(1):27-36.
67. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. Lancet Oncol 2010;11(2):174-83.
68. Stuart-Harris R, Caldas C, Pinder SE, Pharoah P. 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):323-34.
69. Song F, Zhang J, Li S, et al. ER-positive breast cancer patients with more than three positive nodes or grade 3 tumors are at high risk of late recurrence after 5-year adjuvant endocrine therapy. Onco Targets Ther 2017;10:4859-4867.
70. Conforti F, Pala L, Pagan E, et al. Endocrine-responsive lobular carcinoma of the breast: features associated with risk of late distant recurrence. Breast Cancer Res 2019;21(1):153.
71. Hosoda M, Yamamoto M, Nakano K, et al. Differential expression of progesterone receptor, FOXA1, GATA3, and p53 between pre- and postmenopausal women with estrogen receptor-positive breast cancer. Breast Cancer Res Treat 2014;144(2):249-61.

72. Dunbier AK, Anderson H, Ghazoui Z, et al. Relationship between plasma estradiol levels and estrogen-responsive gene expression in estrogen receptor-positive breast cancer in postmenopausal women. J Clin Oncol 2010;28(7):1161-7.
73. Bianchini G, Pusztai L, Karn T, et al. Proliferation and estrogen signaling can distinguish patients at risk for early versus late relapse among estrogen receptor positive breast cancers. Breast Cancer Res 2013;15(5):R86.
74. Dowsett M, Sestak I, Buus R, et al. Estrogen Receptor Expression in 21-Gene Recurrence Score Predicts Increased Late Recurrence for Estrogen-Positive/HER2-Negative Breast Cancer. Clin Cancer Res 2015;21(12):2763-70.
75. Lin YC, Lee YC, Li LH, Cheng CJ, Yang RB. Tumor suppressor SCUBE2 inhibits breast-cancer cell migration and invasion through the reversal of epithelial-mesenchymal transition. J Cell Sci 2014;127(Pt 1):85-100.
76. Cheng CJ, Lin YC, Tsai MT, et al. SCUBE2 suppresses breast tumor cell proliferation and confers a favorable prognosis in invasive breast cancer. Cancer Res 2009;69(8):3634-41.
77. Wilson TR, Yu J, Lu X, et al. The molecular landscape of high-risk early breast cancer: comprehensive biomarker analysis of a phase III adjuvant population. NPJ Breast Cancer 2016;2:16022.
78. Dubsky P, Brase JC, Jakesz R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. Br J Cancer 2013;109(12):2959-64.
79. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res 2011;17(18):6012-20.
80. Abubakar M, Figueroa J, Ali HR, et al. Combined quantitative measures of ER, PR, HER2, and KI67 provide more prognostic information than categorical combinations in luminal breast cancer. Mod Pathol 2019;32(9):1244-1256.
81. Sgroi DC, Sestak I, Cuzick J, 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):1067-76.
82. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27(8):1160-7.
83. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol 2013;31(22):2783-90.
84. Sestak I, Buus R, Cuzick J, et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol 2018.
85. Sestak I, Cuzick J, Dowsett M, et al. Prediction of late distant recurrence after 5 years of endocrine treatment: a combined analysis of patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. J Clin Oncol 2015;33(8):916-22.

86. Buus R, Sestak I, Kronenwett R, et al. Comparison of EndoPredict and EPclin With Oncotype DX Recurrence Score for Prediction of Risk of Distant Recurrence After Endocrine Therapy. J Natl Cancer Inst 2016;108(11).
87. Jerevall PL, Ma XJ, Li H, 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):1762-9.
88. Zhang Y, Schnabel CA, Schroeder BE, et al. Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence. Clin Cancer Res 2013;19(15):4196-205.
89. Schroeder B, Zhang Y, Stal O, et al. Risk stratification with Breast Cancer Index for late distant recurrence in patients with clinically low-risk (T1N0) estrogen receptor-positive breast cancer. NPJ Breast Cancer 2017;3:28.
90. Sgroi DC, Carney E, Zarrella E, et al. Prediction of late disease recurrence and extended adjuvant letrozole benefit by the HOXB13/IL17BR biomarker. J Natl Cancer Inst 2013;105(14):1036-42.
91. Sanft T, Aktas B, Schroeder B, et al. Prospective assessment of the decision-making impact of the Breast Cancer Index in recommending extended adjuvant endocrine therapy for patients with early-stage ER-positive breast cancer. Breast Cancer Res Treat 2015;154(3):533-41.
92. Foldi J, O'Meara T, Marczyk M, Sanft T, Silber A, Pusztai L. Defining Risk of Late Recurrence in Early-Stage Estrogen Receptor-Positive Breast Cancer: Clinical Versus Molecular Tools. J Clin Oncol 2019;37(16):1365-1369.
93. Cheng Q, Chang JT, Gwin WR, et al. A signature of epithelial-mesenchymal plasticity and stromal activation in primary tumor modulates late recurrence in breast cancer independent of disease subtype. Breast Cancer Res 2014;16(4):407.
94. Buus R, Cheang, M., Brentnall, A., et al. B14: Novel 18-gene signature predicts early and late relapse in ER+/HER2- breast cancer patients. NCRI Cancer Conference. Liverpool2015.
95. Yates LR, Knappskog S, Wedge D, et al. Genomic Evolution of Breast Cancer Metastasis and Relapse. Cancer Cell 2017;32(2):169-184.e7.
96. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature 2012;490(7418):61-70.
97. Jeselsohn R, Yelensky R, Buchwalter G, et al. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. Clin Cancer Res 2014;20(7):1757-1767.
98. Fribbens C, O'Leary B, Kilburn L, et al. Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer. J Clin Oncol 2016;34(25):2961-8.
99. Schiavon G, Hrebien S, Garcia-Murillas I, et al. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Sci Transl Med 2015;7(313):313ra182.
100. Allouchery V, Beaussire L, Perdrix A, et al. Circulating ESR1 mutations at the end of aromatase inhibitor adjuvant treatment and after relapse in breast cancer patients. Breast Cancer Res 2018;20(1):40.
101. López-Knowles E, Wilkerson PM, Ribas R, et al. Integrative analyses identify modulators of response to neoadjuvant aromatase inhibitors in patients with early breast cancer. Breast Cancer Res 2015;17(1):35.

102. Zhang Y, Martens JW, Yu JX, et al. Copy number alterations that predict metastatic capability of human breast cancer. Cancer Res 2009;69(9):3795-801.
103. Lopez-Knowles E, Wilkerson PM, Ribas R, et al. Integrative analyses identify modulators of response to neoadjuvant aromatase inhibitors in patients with early breast cancer. Breast Cancer Res 2015;17:35.
104. Rueda OM, Sammut SJ, Seoane JA, et al. Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. Nature 2019;567(7748):399-404.
105. O'Leary B, Hrebien S, Morden JP, et al. Early circulating tumor DNA dynamics and clonal selection with palbociclib and fulvestrant for breast cancer. Nat Commun 2018;9(1):896.
106. Cristofanilli M, Budd GT, Ellis MJ, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med 2004;351(8):781-91.
107. Sparano J, O'Neill A, Alpaugh K, et al. Association of Circulating Tumor Cells With Late Recurrence of Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol 2018.
108. Garcia-Murillas I, Schiavon G, Weigelt B, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med 2015;7(302):302ra133.
109. Garcia-Murillas I, Chopra N, Comino-Méndez I, et al. Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer. JAMA Oncol 2019;5(10):1473-1478.
110. Wu NC, Wong W, Ho KE, et al. Comparison of central laboratory assessments of ER, PR, HER2, and Ki67 by IHC/FISH and the corresponding mRNAs (ESR1, PGR, ERBB2, and MKi67) by RT-qPCR on an automated, broadly deployed diagnostic platform. Breast Cancer Res Treat 2018;172(2):327-338.
111. Thurlimann B, Keshaviah A, Coates AS, et al. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med 2005;353(26):2747-57.
112. (https://cts5-calculator.com). 26/11/19.
113. Altman DGV, Y. Royston, P. . Prognosis and prognostic research: validating a prognostic model. British Medical Journal 2009;338.
114. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. Bmj 2009;338:b604.
115. Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol 2008;26(25):4063-71.
116. Ibraheem A, Olopade OI, Huo D. Propensity score analysis of the prognostic value of genomic assays for breast cancer in diverse populations using the National Cancer Data Base. Cancer 2020;126(17):4013-4022.
117. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. New England Journal of Medicine 2018;379(2):111-121.
118. Sestak I, Crager M, Cuzick J, et al. Abstract GS4-03: Validation of the clinical treatment score post 5 years (CTS5) in women with hormone receptor positive, HER2-negative, node-negative disease from the TAILORx study. Cancer Research 2020;80(4 Supplement):GS4-03-GS4-03.

119. Arvold ND, Taghian AG, Niemierko A, et al. Age, Breast Cancer Subtype Approximation, and Local Recurrence After Breast-Conserving Therapy. Journal of Clinical Oncology 2011;29(29):3885-3891.
120. Filipits M, Nielsen TO, Rudas M, et al. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. Clin Cancer Res 2014;20(5):1298-305.
121. NICE Guidance. Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer. December 2018. 15/07/19.
122. Holt S, Bertelli G, Humphreys I, et al. A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the U.K. Br J Cancer 2013;108(11):2250-8.
123. McVeigh TP, Hughes LM, Miller N, et al. The impact of Oncotype DX testing on breast cancer management and chemotherapy prescribing patterns in a tertiary referral centre. Eur J Cancer 2014;50(16):2763-70.
124. Augustovski F, Soto N, Caporale J, Gonzalez L, Gibbons L, Ciapponi A. Decision-making impact on adjuvant chemotherapy allocation in early node-negative breast cancer with a 21-gene assay: systematic review and meta-analysis. Breast Cancer Res Treat 2015;152(3):611-25.
125. Martin M, Gonzalez-Rivera M, Morales S, et al. Prospective study of the impact of the Prosigna assay on adjuvant clinical decision-making in unselected patients with estrogen receptor positive, human epidermal growth factor receptor negative, node negative early-stage breast cancer. Curr Med Res Opin 2015;31(6):1129-37.
126. Muller BM, Keil E, Lehmann A, et al. The EndoPredict Gene-Expression Assay in Clinical Practice - Performance and Impact on Clinical Decisions. PLoS One 2013;8(6):e68252.
127. Jakesz R, Greil R, Gnant M, 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):1845-53.
128. Ellis P, Barrett-Lee P, Johnson L, et al. Sequential docetaxel as adjuvant chemotherapy for early breast cancer (TACT): an open-label, phase III, randomised controlled trial. Lancet 2009;373(9676):1681-92.
129. Willson ML, Burke L, Ferguson T, Ghersi D, Nowak AK, Wilcken N. Taxanes for adjuvant treatment of early breast cancer. Cochrane Database Syst Rev 2019;9(9):Cd004421.
130. Qin YY, Li H, Guo XJ, et al. Adjuvant chemotherapy, with or without taxanes, in early or operable breast cancer: a meta-analysis of 19 randomized trials with 30698 patients. PLoS One 2011;6(11):e26946.
131. Howlader N, Altekruse SF, Li CI, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst 2014;106(5).
132. Kaufman B, Mackey JR, Clemens MR, et al. Trastuzumab plus anastrozole versus anastrozole alone for the treatment of postmenopausal women with human epidermal growth factor receptor 2-positive, hormone receptor-positive metastatic breast cancer: results from the randomized phase III TAnDEM study. J Clin Oncol 2009;27(33):5529-37.

133. Strasser-Weippl K, Horick N, Smith IE, et al. Long-term hazard of recurrence in HER2+ breast cancer patients untreated with anti-HER2 therapy. Breast Cancer Res 2015;17:56.
134. Cameron D, Piccart-Gebhart MJ, Gelber RD, et al. 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. Lancet 2017;389(10075):1195-1205.
135. Carey LB, WT. Pitcher B. Gene expression signatures in pre- and post-therapy (Rx) specimens from CALGB 40601 (Alliance), a neoadjuvant phase III trial of weekly paclitaxel and trastuzumab with or without lapatinib for HER2-positive breast cancer (BrCa). Journal of Clinical Oncology 2014;32:5s(suppl; abstr 506).
136. Chumsri S, Li Z, Serie DJ, et al. Incidence of Late Relapses in Patients With HER2-Positive Breast Cancer Receiving Adjuvant Trastuzumab: Combined Analysis of NCCTG N9831 (Alliance) and NRG Oncology/NSABP B-31. J Clin Oncol 2019;37(35):3425-3435.
137. Goss PE, Ingle JN, Martino S, 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):1793-802.
138. Regan MM, Pagani O, Fleming GF, et al. Adjuvant treatment of premenopausal women with endocrine-responsive early breast cancer: design of the TEXT and SOFT trials. Breast 2013;22(6):1094-100.
139. Garcia Fernandez A, Gimenez N, Fraile M, et al. Survival and clinicopathological characteristics of breast cancer patient according to different tumour subtypes as determined by hormone receptor and Her2 immunohistochemistry. a single institution survey spanning 1998 to 2010. Breast 2012;21(3):366-73.
140. Buus R, Yeo B, Brentnall AR, et al. Novel 18-gene signature for predicting relapse in ER-positive, HER2-negative breast cancer. Breast Cancer Res 2018;20(1):103.
141. Bliss JM, J. Evans A et al. Abstract P2-05-01: Clinico-pathological relationships with Ki67 in POETIC (CRUK/07/015) - Critical lessons for assessing Ki67 for prognosis and as a pharmacodynamic marker. Cancer Research 2017;77(4(Suppl)).
142. Sun J, Chen C, Wei W, et al. Associations and indications of Ki67 expression with clinicopathological parameters and molecular subtypes in invasive breast cancer: A population-based study. Oncol Lett 2015;10(3):1741-1748.
143. Alataki AZ, L. Tovey, H et al. Abstract PO03: Comparison of manual and digital histopathology scoring for Ki67 in breast cancer core biopsies. UK International Breast Cancer Symposium. Birmingham, UK 2020.
144. Dowling RJO, Kalinsky K, Hayes DF, et al. Toronto Workshop on Late Recurrence in Estrogen Receptor–Positive Breast Cancer: Part 1: Late Recurrence: Current Understanding, Clinical Considerations. JNCI Cancer Spectrum 2019;3(4).
145. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. N Engl J Med 1988;319(26):1681-92.
146. EBCTCG. Treatment of Early Breast Cancer, Volume 1: worldwide evidence 1985-1990. Oxford: Oxford University Press, 1990.

147. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. Nat Rev Cancer 2015;15(5):261-75.
148. Smith I, Robertson J, Kilburn L, et al. Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): an open-label, multicentre, parallel-group, randomised, phase 3 trial. Lancet Oncol 2020;21(11):1443-1454.
149. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995;57(1):289-300.
150. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486(7403):346-52.
151. Driver KE, Song H, Lesueur F, et al. Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population. Carcinogenesis 2008;29(2):333-41.
152. Alkarain A, Slingerland J. Deregulation of p27 by oncogenic signaling and its prognostic significance in breast cancer. Breast Cancer Res 2004;6(1):13-21.
153. Basudan A, Priedigkeit N, Hartmaier RJ, et al. Frequent ESR1 and CDK Pathway Copy-Number Alterations in Metastatic Breast Cancer. Mol Cancer Res 2019;17(2):457-468.
154. Tomita S, Zhang Z, Nakano M, et al. Estrogen receptor alpha gene ESR1 amplification may predict endocrine therapy responsiveness in breast cancer patients. Cancer Sci 2009;100(6):1012-7.
155. Fox SB, Brown P, Han C, et al. Expression of the forkhead transcription factor FOXP1 is associated with estrogen receptor alpha and improved survival in primary human breast carcinomas. Clin Cancer Res 2004;10(10):3521-7.
156. Shigekawa T, Ijichi N, Ikeda K, et al. FOXP1, an estrogen-inducible transcription factor, modulates cell proliferation in breast cancer cells and 5-year recurrence-free survival of patients with tamoxifen-treated breast cancer. Horm Cancer 2011;2(5):286-97.
157. Cizkova M, Vacher S, Meseure D, et al. PIK3R1 underexpression is an independent prognostic marker in breast cancer. BMC Cancer 2013;13:545.
158. Qi L, Zhou B, Chen J, et al. Significant prognostic values of differentially expressed-aberrantly methylated hub genes in breast cancer. J Cancer 2019;10(26):6618-6634.
159. Bonin S, Pracella D, Barbazza R, Dotti I, Boffo S, Stanta G. PI3K/AKT Signaling in Breast Cancer Molecular Subtyping and Lymph Node Involvement. Dis Markers 2019;2019:7832376.
160. Jones RH, Casbard A, Carucci M, et al. Fulvestrant plus capivasertib versus placebo after relapse or progression on an aromatase inhibitor in metastatic, oestrogen receptor-positive breast cancer (FAKTION): a multicentre, randomised, controlled, phase 2 trial. Lancet Oncol 2020;21(3):345-357.
161. Svensson S, Abrahamsson A, Rodriguez GV, et al. CCL2 and CCL5 Are Novel Therapeutic Targets for Estrogen-Dependent Breast Cancer. Clin Cancer Res 2015;21(16):3794-805.
162. Yao Z, Zheng X, Lu S, et al. Knockdown of FAM64A suppresses proliferation and migration of breast cancer cells. Breast Cancer 2019;26(6):835-845.

163. Weinberg F, Reischmann N, Fauth L, et al. The Atypical Kinase RIOK1 Promotes Tumor Growth and Invasive Behavior. EBioMedicine 2017;20:79-97.
164. Xu Y, Wu X, Li F, Huang D, Zhu W. CDCA4, a downstream gene of the Nrf2 signaling pathway, regulates cell proliferation and apoptosis in the MCF-7/ADM human breast cancer cell line. Mol Med Rep 2018;17(1):1507-1512.
165. Chen J, Yao Y, Gong C, et al. CCL18 from tumor-associated macrophages promotes breast cancer metastasis via PITPNM3. Cancer Cell 2011;19(4):541-55.
166. Liu Y, Zheng H, Li Q, et al. Discovery of CCL18 antagonist blocking breast cancer metastasis. Clin Exp Metastasis 2019;36(3):243-255.
167. Lee MH, Koh D, Na H, et al. MTA1 is a novel regulator of autophagy that induces tamoxifen resistance in breast cancer cells. Autophagy 2018;14(5):812-824.
168. Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006;10(6):529-41.
169. Chin SF, Wang Y, Thorne NP, et al. Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. Oncogene 2007;26(13):1959-70.
170. Han W, Han MR, Kang JJ, et al. Genomic alterations identified by array comparative genomic hybridization as prognostic markers in tamoxifen-treated estrogen receptor-positive breast cancer. BMC Cancer 2006;6:92.
171. Carene D, Tran-Dien A, Lemonnier J, et al. Association between FGFR1 copy numbers, MAP3K1 mutations, and survival in axillary node-positive, hormone receptor-positive, and HER2-negative early breast cancer in the PACS04 and METABRIC studies. Breast Cancer Res Treat 2020;179(2):387-401.
172. Pereira B, Chin SF, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun 2016;7:11479.
173. Jiang YZ, Yu KD, Zuo WJ, Peng WT, Shao ZM. GATA3 mutations define a unique subtype of luminal-like breast cancer with improved survival. Cancer 2014;120(9):1329-37.
174. Ciriello G, Gatza ML, Beck AH, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell 2015;163(2):506-19.
175. Sieuwerts AM, Look MP, Meijer-van Gelder ME, et al. Which cyclin E prevails as prognostic marker for breast cancer? Results from a retrospective study involving 635 lymph node-negative breast cancer patients. Clin Cancer Res 2006;12(11 Pt 1):3319-28.
176. Caldon CE, Sergio CM, Kang J, et al. Cyclin E2 overexpression is associated with endocrine resistance but not insensitivity to CDK2 inhibition in human breast cancer cells. Mol Cancer Ther 2012;11(7):1488-99.
177. Arnedos M, Bayar MA, Cheaib B, et al. Modulation of Rb phosphorylation and antiproliferative response to palbociclib: the preoperative-palbociclib (POP) randomized clinical trial. Ann Oncol 2018;29(8):1755-1762.
178. Wang L, Shao ZM. Cyclin e expression and prognosis in breast cancer patients: a meta-analysis of published studies. Cancer Invest 2006;24(6):581-7.
179. Sun CC, Li SJ, Hu W, et al. Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer. Mol Ther 2019;27(6):1153-1165.

180. Liu ZL, Bi XW, Liu PP, et al. Expressions and prognostic values of the E2F transcription factors in human breast carcinoma. Cancer Manag Res 2018;10:3521-3532.
181. Alfarsi LH, Ansari RE, Craze ML, et al. CDC20 expression in oestrogen receptor positive breast cancer predicts poor prognosis and lack of response to endocrine therapy. Breast Cancer Res Treat 2019;178(3):535-544.
182. Kwok HF, Zhang SD, McCrudden CM, et al. Prognostic significance of minichromosome maintenance proteins in breast cancer. Am J Cancer Res 2015;5(1):52-71.
183. Castano Z, San Juan BP, Spiegel A, et al. IL-1beta inflammatory response driven by primary breast cancer prevents metastasis-initiating cell colonization. Nat Cell Biol 2018;20(9):1084-1097.
184. Takeshita T, Yan L, Asaoka M, Rashid O, Takabe K. Late recurrence of breast cancer is associated with pro-cancerous immune microenvironment in the primary tumor. Sci Rep 2019;9(1):16942.

9. Appendix

9.1 IHC staining protocol/SOP for ER, PgR, Ki67, CK7, CK19 and p63

9.1.1 Working Solutions (according to manufacturer's instructions)

- Dako DAB Away DAB Chomogen Removal System (S1967) reagents to be used as described to clean the Autostainer when prompted.
- Envision™ FLEX Target Retrieval Solution low pH (50X) (Dako Agilent, DM829) for ER, PgR and Ki67.
- Envision™ FLEX Target Retrieval Solution high pH (50X) (Dako Agilent, DM829) for CK7/19 and p63
- Envision™ FLEX Wash Buffer (Dako, DM831 or equivalent).
- Envision™ FLEX Peroxidase blocking reagent SM801.
- Envision™ FLEX Linker (Mouse) SM804.
- Envision™ FLEX /HRP SM802.
- Envision™ FLEX Substrate buffer SM803.
- Envision™ FLEX DAB+Chromogen DM827 (1 drop of DAB/ml Substrate)
- Primary antibody:
  Oestrogen Receptor: NCL-L-ER6F11 (Novocastra, Leica).
  Progesterone Receptor: L-PgR 312 (Novocastra/Leica)
  Ki67: MIB1, M7240 (DAKO)
  CK7: OV-TL 12/30, M7018 (DAKO)
  CK19: RCK 108, M0888 (DAKO)

P63： DAK p63, M7317 (DAKO)

- Antibody diluent (Dako, S2022).
- Mayer's Haematoxylin (TCS Biosciences Ltd., HS315-1L).
- Mounting medium (Leica Surgipath Ref 3801731)
- IMS (Genta, CAS No 64-17-5).
- Xylene (Genta, CAS No 1330-20-7).

9.1.2 Staining procedure

1. De-wax the rack of sections in two changes of Xylene for approximately 5 minutes each change. (Coated/dipped slides will take approximately 30 minutes to de-wax. Do not let the sections dry out during any stage of the immunohistochemical staining process.

2. Rehydrate sections and bring to water by immersing for 1-2 minutes in each of decreasing grades of Industrial methylated spirits (IMS) 100%, 90% and 70%. Transfer the rack containing the sections into running tap water for a few minutes and then load into autostainer racks into long troughs containing wash buffer whilst waiting for the PT link to reach 65

3. Switch on the Dako Autostainer

4. Programme Autostainer

5. When the menu appears select "Program". Go to "Slides" option and enter total number of slides in the analysis, including the tests and the controls.

6. If the position of the negative control is not already pre-programmed, decide where to position the negative control slide.

7. Change "Program" back to AUTO and click on cell you want to alter.

8. To change default information on grid select EDIT SLIDE from scroll down list and change as required then click OK. Repeat this for the number of slides being tested.

9. Alternatively open a previously programmed run, save as the new autostainer run title, enter the number of slides for the current run and modify to current requirements using edit, delete and save, as instructed on screen. Program grid should now be ready (Table 1).

10. Press NEXT (slide positions shown)

11. Press NEXT – Run Time window will show Probe Washes, Litres of Buffer and Distilled Water required in the labelled aspirators under the machine bench. Then the Autostainer screen will show the volume of each reagent that is required (Figure 2).

12. Dilute primary antibody with DAKO Antibody diluent according to the following dilutions:
    ER antibody 1:400
    PgR antibody: 1:300
    Ki67 antibody 1:1200
    CK7 antibody 1:100
    CK19 antibody 1:100
    P63 antibody 1:100

13. For negative diluent controls without primary antibody use the diluent neat.

14. Decant reagents/antibodies into labelled (REA/001) vials and place in the appropriate slots as indicated by the Autostainer reagent layout in Figure 2 which shows the co- ordinate position of each vial of reagent to go in the white rack if using the FLEX Kit.

15. Check the non-hazardous waste receptacle is empty. In the hazardous waste receptacle starting off from empty cover the bottom with diluted Milton stock solution.

16. Dilute Dako Target Retrieval Solution low pH (for ER, PgR and Ki67) or high pH (for CK7/19 and p63), into 1.5 litres of de-ionised water in the long trough/s of the PT Link. Only use the antigen retrieval solution 3 times then discard and make up fresh.

17. Tap menu to start heating to 65°C (re 10.4) and when the temperature reaches 65°C place the slide rack/s into the chamber/s containing hot low pH antigen retrieval solution, close and secure lid. Press preheat button again to regain 65°C. Press ‘start’ and listen for a click sound. The temperature will rise to 97°C and target retrieval will occur automatically for 20 minutes and after this the temperature cools down to 65°C.

18. The PT Link apparatus will alarm indicating the process has been completed. Immediately remove the rack/s of slides and place into two washes of five minutes each of Dako/FLEX wash buffer.

19. Place the slide rack/s into the slots in the pre-programmed Dako Autostainer. Check the slides are positioned with the labelled end first.

20. Ensure the sections are totally covered with wash buffer and check they are in the correct grid positions.
21. Click on NEXT twice

22. Before starting run check pathway clear then start to run the program, press “yes” as appropriate. Continue as prompted.

23. When the programme has ended, remove slides one by one into a vessel containing water ensuring the sections do not dry out.

24. Gently rinse the slides for approximately 5 minutes in running tap water.

25. Sign off from the main menu and shut down the computer.

26. Counter-stain with Mayers Haematoxylin for approx. 1-2 minutes.

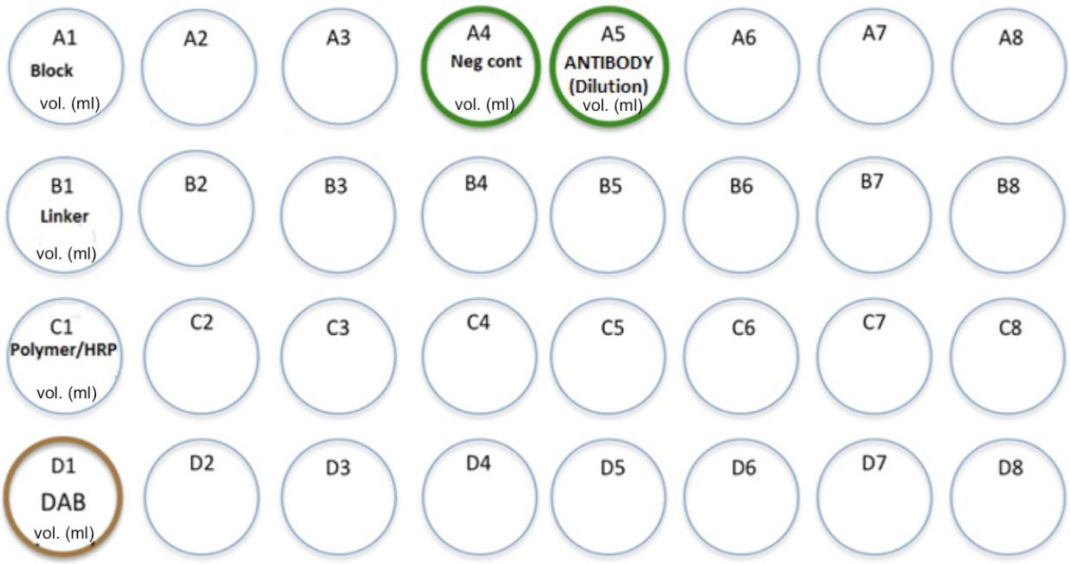
27. Place the slides under running tap water for at least 5 minutes.

28. Dehydrate, clear and mount.

Table 1 Grid run sequence of Autostainer (eg FLEX Kit with ER antibody clone 6F11).

Rinse wash buffer	Endo-genous Block (H202)	Rinse water	Rinse wash buffer	Primary antibody	Rinse wash buffer	Mouse Linker	Rinse wash buffer	Polymer HRP	Rinse wash buffer	Rinse Wash buffer	switch	DAB	Rinse water
	5 Minutes			20 minutes		20 minutes		30 minutes				10 minutes	

Figure 2. Example of the Autostainer reagent grid positions. Always use the grid produced by the Autostainer for the specific run.



Block *= Peroxidase block;* ***Linker*** *= Mouse linker;* ***Polymer/HRP*** *= Polymer Horseradish peroxidase;* ***Neg cont*** *= Diluent;* ***DAB*** *= Diaminobenzidine.*

9.2 FFPE Allprep Extraction Protocol

For each new kit, make up reagents that need reconstituting (this can be done on the day before extraction if necessary)

9.2.1 Make DNAse 1 stock solution

- Pipette out 550µl RNAse free water into a small syringe with needle already attached (pull out plunger, pipette it in and then replace the plunger)
- Push needle into the DNAse 1 vial and inject the water
- Mix gently by inverting the vial – do not vortex
- Store in required amounts.

9.2.2 Prepare FFPE sections

1. Take slides out of -80 and prepare tubes and reagents
2. Deparaffinise: xylene, ethanols…
    a. 2 mins in 2 x xylene and 2 x ethanol

3. Using a scalpel, dip blade into tube with 150µl Buffer PKD. Using microscope, scrape off the invasive tissue as marked on the H&E
4. Add 10µl proteinase K, and vortex.
5. Incubate at 56°C for 15 min. (always ensure heat block is at temperature before placing on it)
6. Incubate on ice for 3 min.
7. Centrifuge for 15 min at 20,000 x *g*.- in the big centrifuge
8. Pipette out the supernatant (RNA) into a 2ml tube. Keep the pellet (DNA) in the original tube
9. Move on to RNA step 1 and then DNA step 1 then RNA steps 2-16

9.2.3 RNA extraction

1. Incubate supernatant at 80°C for 15 min and centrifuge at 8000 for 10 seconds to remove droplets from lid
2. Add 320µl Buffer RLT + 1120µl ethanol (96–100%), and mix by vortexing.
3. Transfer 700µl of the sample including any precipitate into RNeasy MinEliute spin column (pink one) and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm). Discard the flow through.
4. Repeat step 3 until the whole sample has passed through the spin column
5. Add 350µl Buffer FRN (needs to be re-constituted before) and centrifuge for 15 s at ≥8000 x *g*. Discard the flow through.
6. Add 10µl DNase I stock solution to 70µl Buffer RDD in a separate tube and mix gently and centrifuge at 8000 for 10 seconds to get all liquid at bottom.
7. Add the whole 80µl DNAse1+buffer mix to RNeasy minelute spin column. MUST PIPETTE DIRECTLY ONTO THE MEMBRANE WITHOUT TOUCHING THE MEMBRANE and place at RT for 15 min.
8. Add 500µl Buffer FRN and centrifuge for 15 s at ≥8000. **Save the flow-through** as it contains RNA including small RNAs.
9. Place the RNeasy MinElute spin column in a **new 2 ml collection tube** (supplied).
10. Pipette out the flow through from step 8 and apply to the RNeasy MinElute spin column and centrifuge for 15 s at ≥8000 x *g.* Discard the flow through. Re-use the collection tube in next step

11. Add 500µl Buffer RPE (needs to be reconstituted before use) to the RNeasy MinElute spin column and centrifuge for 15 s at ≥8000 x *g.* Discard the flow through and the collection tube.
12. Place the RNeasy MinElute spin column in a **new 2 ml collection tube** (supplied). Add 500µl Buffer RPE and centrifuge for 15 s at ≥8000 x *g.* Discard the flow through and the collection tube.
13. Place the RNeasy MinElute spin column in a **new 2 ml collection tube** (supplied). **Open the lid**, and centrifuge at full speed for 5 min. Ensure the lid is opposite the direction of the rotor so as to not affect the centrifugation.
14. Place the RNeasy MinElute spin column in a **new 1.5 ml collection tube** (supplied). Add 30 µl RNase-free water directly to the spin column membrane and incubate for 1 min at RT. Centrifuge at full speed for 1 min
15. Reelute in 20ul and repeat procedure (optional)
16. Discard the spin column and store the collecting tube in -80 freezer

9.2.4 DNA extraction

1. Resuspend the pellet in 180µl Buffer ATL, add 40µl proteinase K, and mix by vortexing. Leave overnight 56°C.
2. Incubate at 90°C for 2 h without agitation.
3. Centrifuge at 8000 for 10 seconds to get all the liquid down.
4. Allow the sample to cool to room temperature and then add 4µl RNase A (100 mg/ml). Incubate for 2 min at room temperature
5. Add 200µl Buffer AL + 200µl ethanol (96–100%) and vortex. (Can make a master mix of this and add 400µl to each tube – make a little more than what is needed)
6. Transfer the entire sample to a QIAamp MinElute spin column (white) and centrifuge for 1 min at ≥8000 x *g.* Discard the flow through and the collection tube. If some liquid remains in the spin column centrifuge again until the column is empty.
7. Place QIAamp MinElute spin column in a **new 2 ml collection tube** (supplied). Add 700µl Buffer AW1 (needs to be reconstituted before use) and centrifuge for 15 s at ≥8000 x *g.* Discard the flow through but can re-use the collection tube.

8. Add 700µl Buffer AW2 (needs to be re-constituted before use) and centrifuge for 15 s at ≥8000 x *g.* Discard the flow through but can re-use the collection tube.
9. Add 700µl ethanol (96–100%). Centrifuge for 15 s at ≥8000 x *g.* Discard the flow through and the collection tube.
10. Place QIAamp MinElute spin column in a **new 2 ml collection tube**. Open the lid and centrifuge at full speed for 5 min. Orientate lids so that they are opposite the direction of the rotor.
11. Place QIAamp MinElute spin column in a **new 1.5 ml collection tube** (supplied). Add 30µl Buffer ATE directly to the spin column membrane and incubate for 1 min at RT and centrifuge at full speed for 1 min to elute the DNA.
12. Reelute in 20ul and repeat procedure (optional)
13. Put in freezer at -80.

9.3 Nanostring 360 panel – individual genes

Gene	Full name	Gene	Full name
ABCA8	ATP binding cassette subfamily A member 8	IL3RA	interleukin 3 receptor subunit alpha
ACTR3B	ARP3 actin related protein 3 homolog B	IL4R	interleukin 4 receptor
ACVR1B	activin A receptor type 1B	IL6	interleukin 6
ACVR1C	activin A receptor type 1C	IL6R	interleukin 6 receptor
ACVRL1	activin A receptor like type 1	IL7R	interleukin 7 receptor
ADAM12	ADAM metallopeptidase domain 12	INHBA	inhibin beta A subunit
ADCY9	adenylate cyclase 9	INHBB	inhibin beta B subunit
ADD1	adducin 1	IRF6	interferon regulatory factor 6
ADM	adrenomedullin	IRX1	iroquois homeobox 1
AGR2	anterior gradient 2, protein disulphide isomerase family member	ISG15	ISG15 ubiquitin-like modifier
AGT	angiotensinogen	ISM1	isthmin 1
AGTR1	angiotensin II receptor type 1	ITGA6	integrin subunit alpha 6
AKT3	AKT serine/threonine kinase 3	ITGAV	integrin subunit alpha V
ALDH1A1	aldehyde dehydrogenase 1 family member A1	ITGB1	integrin subunit beta 1
ALDOA	aldolase, fructose-bisphosphate A	ITGB3	integrin subunit beta 3
ANGPT1	angiopoietin 1	ITGB6	integrin subunit beta 6
ANLN	anillin actin binding protein	ITPR1	inositol 1,4,5-trisphosphate receptor type 1
ANXA9	annexin A9	JAG1	jagged 1
APH1B	aph-1 homolog B, gamma-secretase subunit	JAG2	jagged 2
APOD	apolipoprotein D	JAK1	Janus kinase 1

APOE	apolipoprotein E	JAK2	Janus kinase 2
AR	androgen receptor	JAK3	Janus kinase 3
AREG	amphiregulin	JAM2	junctional adhesion molecule 2
ARID1A	AT-rich interaction domain 1A	JCAD	junctional cadherin 5 associated
ARNT2	aryl hydrocarbon receptor nuclear translocator 2	JUN	Jun proto-oncogene, AP-1 transcription factor subunit
ASPM	abnormal spindle microtubule assembly	KAT2B	lysine acetyltransferase 2B
ASPN	asporin	KCNB1	potassium voltage-gated channel subfamily B member 1
ATAD2	ATPase family, AAA domain containing 2	KDR	kinase insert domain receptor
ATM	ATM serine/threonine kinase	KIAA0040	KIAA0040
ATP10B	ATPase phospholipid transporting 10B (putative)	KIF11	kinesin family member 11
AURKA	aurora kinase A	KIF14	kinesin family member 14
AURKB	aurora kinase B	KIF23	kinesin family member 23
AXIN1	axin 1	KIF2C	kinesin family member 2C
AXIN2	axin 2	KIFC1	kinesin family member C1
B3GNT3	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3	KIT	KIT proto-oncogene receptor tyrosine kinase
BAD	BCL2 associated agonist of cell death	KLRK1	killer cell lectin like receptor K1
BAG1	BCL2 associated athanogene 1	KRT14	keratin 14
BAIAP2L1	BAI1 associated protein 2 like 1	KRT17	keratin 17
BAIAP3	BAI1 associated protein 3	KRT5	keratin 5
BAMBI	BMP and activin membrane bound inhibitor	KRT6B	keratin 6B
BAX	BCL2 associated X, apoptosis regulator	KRT7	keratin 7
BBC3	BCL2 binding component 3	LAD1	ladinin 1
BBOX1	gamma-butyrobetaine hydroxylase 1	LAG3	lymphocyte activating 3
BCAS1	breast carcinoma amplified sequence 1	LAMA3	laminin subunit alpha 3
BCL11A	B-cell CLL/lymphoma 11A	LAMB3	laminin subunit beta 3
BCL2	BCL2, apoptosis regulator	LAMC2	laminin subunit gamma 2
BCL2A1	BCL2 related protein A1	LEF1	lymphoid enhancer binding factor 1
BCL2L1	BCL2 like 1	LEFTY2	left-right determination factor 2
BCL6B	B-cell CLL/lymphoma 6B	LEMD1	LEM domain containing 1
BDNF	brain derived neurotrophic factor	LEP	leptin
BIRC5	baculoviral IAP repeat containing 5	LEPR	leptin receptor
BLM	Bloom syndrome RecQ like helicase	LFNG	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase
BLVRA	biliverdin reductase A	LIF	LIF, interleukin 6 family cytokine
BMP2	bone morphogenetic protein 2	LIFR	LIF receptor alpha
BMP4	bone morphogenetic protein 4	LINC02381	long intergenic non-protein coding RNA 2381
BMP5	bone morphogenetic protein 5	LPL	lipoprotein lipase
BMP6	bone morphogenetic protein 6	LRP2	LDL receptor related protein 2
BMP7	bone morphogenetic protein 7	LRRC32	leucine rich repeat containing 32
BMP8A	bone morphogenetic protein 8a	LTB	lymphotoxin beta
BMPR1A	bone morphogenetic protein receptor type 1A	LTBP1	latent transforming growth factor beta binding protein 1
BMPR1B	bone morphogenetic protein receptor type 1B	MAD2L1	mitotic arrest deficient 2 like 1

BMPR2	bone morphogenetic protein receptor type 2	MAF	MAF bZIP transcription factor
BNC2	basonuclin 2	MAML2	mastermind like transcriptional coactivator 2
BNIP3	BCL2 interacting protein 3	MAP2K4	mitogen-activated protein kinase kinase 4
BORCS7	BLOC-1 related complex subunit 7	MAP3K12	mitogen-activated protein kinase kinase kinase 12
BRCA1	BRCA1, DNA repair associated	MAPK1	mitogen-activated protein kinase 1
BRCA2	BRCA2, DNA repair associated	MAPK10	mitogen-activated protein kinase 10
BTG2	BTG anti-proliferation factor 2	MAPK3	mitogen-activated protein kinase 3
C5orf38	chromosome 5 open reading frame 38	MAPK8IP2	mitogen-activated protein kinase 8 interacting protein 2
CA12	carbonic anhydrase 12	MAPT	microtubule associated protein tau
CACNA1D	calcium voltage-gated channel subunit alpha1 D	MARCO	macrophage receptor with collagenous structure
CACNA1H	calcium voltage-gated channel subunit alpha1 H	MCM2	minichromosome maintenance complex component 2
CACNA2D1	calcium voltage-gated channel auxiliary subunit alpha2delta 1	MCM3	minichromosome maintenance complex component 3
CACNA2D3	calcium voltage-gated channel auxiliary subunit alpha2delta 3	MDM2	MDM2 proto-oncogene
CACNG1	calcium voltage-gated channel auxiliary subunit gamma 1	MED1	mediator complex subunit 1
CACNG4	calcium voltage-gated channel auxiliary subunit gamma 4	MELK	maternal embryonic leucine zipper kinase
CACNG6	calcium voltage-gated channel auxiliary subunit gamma 6	MEOX2	mesenchyme homeobox 2
CALML5	calmodulin like 5	MET	MET proto-oncogene, receptor tyrosine kinase
CAMK2B	calcium/calmodulin dependent protein kinase II beta	MFNG	MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase
CAV1	caveolin 1	MIA	melanoma inhibitory activity
CBLC	Cbl proto-oncogene C	MIS18A	MIS18 kinetochore protein A
CCL2	C-C motif chemokine ligand 2	MKI67	marker of proliferation Ki-67
CCL21	C-C motif chemokine ligand 21	MLH1	mutL homolog 1
CCL3L1	C-C motif chemokine ligand 3 like 1	MLLT3	MLLT3, super elongation complex subunit
CCL4	C-C motif chemokine ligand 4	MLPH	melanophilin
CCL5	C-C motif chemokine ligand 5	MME	membrane metalloendopeptidase
CCL7	C-C motif chemokine ligand 7	MMP11	matrix metallopeptidase 11
CCL8	C-C motif chemokine ligand 8	MMP14	matrix metallopeptidase 14
CCNA1	cyclin A1	MMP3	matrix metallopeptidase 3
CCNA2	cyclin A2	MMP7	matrix metallopeptidase 7
CCNB1	cyclin B1	MMP9	matrix metallopeptidase 9
CCND1	cyclin D1	MMRN2	multimerin 2
CCND2	cyclin D2	MRE11	MRE11 homolog, double strand break repair nuclease
CCNE1	cyclin E1	MS4A2	membrane spanning 4-domains A2
CCNE2	cyclin E2	MSR1	macrophage scavenger receptor 1
CCR1	C-C motif chemokine receptor 1	MT1G	metallothionein 1G
CCR2	C-C motif chemokine receptor 2	MTOR	mechanistic target of rapamycin kinase
CCR5	C-C motif chemokine receptor 5 (gene/pseudogene)	MUC1	mucin 1, cell surface associated

CD163	CD163 molecule	MUS81	MUS81 structure-specific endonuclease subunit
CD19	CD19 molecule	MYBL2	MYB proto-oncogene like 2
CD1E	CD1e molecule	MYC	MYC proto-oncogene, bHLH transcription factor
CD24	CD24 molecule	MYCN	MYCN proto-oncogene, bHLH transcription factor
CD27	CD27 molecule	MYCT1	MYC target 1
CD274	CD274 molecule	NASP	nuclear autoantigenic sperm protein
CD276	CD276 molecule	NAT1	N-acetyltransferase 1
CD34	CD34 molecule	NCAM1	neural cell adhesion molecule 1
CD36	CD36 molecule	NCAPH2	non-SMC condensin II complex subunit H2
CD44	CD44 molecule (Indian blood group)	NDC80	NDC80, kinetochore complex component
CD68	CD68 molecule	NDP	NDP, norrin cystine knot growth factor
CD84	CD84 molecule	NEIL1	nei like DNA glycosylase 1
CD8A	CD8a molecule	NEIL2	nei like DNA glycosylase 2
CD8B	CD8b molecule	NEIL3	nei like DNA glycosylase 3
CDC14A	cell division cycle 14A	NEO1	neogenin 1
CDC14B	cell division cycle 14B	NETO2	neuropilin and tolloid like 2
CDC20	cell division cycle 20	NFATC1	nuclear factor of activated T-cells 1
CDC25A	cell division cycle 25A	NFKBIZ	NFKB inhibitor zeta
CDC25B	cell division cycle 25B	NGFR	nerve growth factor receptor
CDC25C	cell division cycle 25C	NKG7	natural killer cell granule protein 7
CDC6	cell division cycle 6	NOD2	nucleotide binding oligomerization domain containing 2
CDC7	cell division cycle 7	NOTCH1	notch 1
CDCA5	cell division cycle associated 5	NOTCH2	notch 2
CDCA7L	cell division cycle associated 7 like	NOTCH3	notch 3
CDCA8	cell division cycle associated 8	NPEPPS	aminopeptidase puromycin sensitive
CDH1	cadherin 1	NPR1	natriuretic peptide receptor 1
CDH2	cadherin 2	NR4A1	nuclear receptor subfamily 4 group A member 1
CDH3	cadherin 3	NR4A3	nuclear receptor subfamily 4 group A member 3
CDH5	cadherin 5	NRCAM	neuronal cell adhesion molecule
CDK1	cyclin dependent kinase 1	NRXN1	neurexin 1
CDK4	cyclin dependent kinase 4	NRXN3	neurexin 3
CDK6	cyclin dependent kinase 6	NSD1	nuclear receptor binding SET domain protein 1
CDKN1A	cyclin dependent kinase inhibitor 1A	NSD3	nuclear receptor binding SET domain protein 3
CDKN1B	cyclin dependent kinase inhibitor 1B	NTRK2	neurotrophic receptor tyrosine kinase 2
CDKN1C	cyclin dependent kinase inhibitor 1C	NUDT1	nudix hydrolase 1
CDKN2A	cyclin dependent kinase inhibitor 2A	NUF2	NUF2, NDC80 kinetochore complex component
CDKN2B	cyclin dependent kinase inhibitor 2B	NUMBL	NUMB like, endocytic adaptor protein
CDKN2C	cyclin dependent kinase inhibitor 2C	NUPR1	nuclear protein 1, transcriptional regulator
CDKN2D	cyclin dependent kinase inhibitor 2D	OAS3	2'-5'-oligoadenylate synthetase 3
CDKN3	cyclin dependent kinase inhibitor 3	OCLN	occludin

CEACAM5	carcinoembryonic antigen related cell adhesion molecule 5	OGN	osteoglycin
CEACAM6	carcinoembryonic antigen related cell adhesion molecule 6	OLFML2B	olfactomedin like 2B
CENPF	centromere protein F	ORC6	origin recognition complex subunit 6
CEP55	centrosomal protein 55	PALB2	partner and localizer of BRCA2
CETN2	centrin 2	PALMD	palmdelphin
CFD	complement factor D	PARP1	poly(ADP-ribose) polymerase 1
CHAD	chondroadherin	PARP2	poly(ADP-ribose) polymerase 2
CHEK2	checkpoint kinase 2	PARP4	poly(ADP-ribose) polymerase family member 4
CHI3L1	chitinase 3 like 1	PAX5	paired box 5
CHIT1	chitinase 1	PAX8	paired box 8
CHRNA5	cholinergic receptor nicotinic alpha 5 subunit	PBX3	PBX homeobox 3
CKB	creatine kinase B	PCK1	phosphoenolpyruvate carboxykinase 1
CKMT1A	creatine kinase, mitochondrial 1A	PCNA	proliferating cell nuclear antigen
CKS1B	CDC28 protein kinase regulatory subunit 1B	PDCD1	programmed cell death 1
CLDN1	claudin 1	PDCD1LG2	programmed cell death 1 ligand 2
CLDN3	claudin 3	PDE9A	phosphodiesterase 9A
CLDN4	claudin 4	PDGFB	platelet derived growth factor subunit B
CLDN7	claudin 7	PDGFRA	platelet derived growth factor receptor alpha
CLEC14A	C-type lectin domain containing 14A	PDGFRB	platelet derived growth factor receptor beta
CLEC5A	C-type lectin domain containing 5A	PDK4	pyruvate dehydrogenase kinase 4
CMKLR1	chemerin chemokine-like receptor 1	PECAM1	platelet and endothelial cell adhesion molecule 1
CNTFR	ciliary neurotrophic factor receptor	PFDN2	prefoldin subunit 2
COL11A1	collagen type XI alpha 1 chain	PGK1	phosphoglycerate kinase 1
COL27A1	collagen type XXVII alpha 1 chain	PGR	progesterone receptor
COL2A1	collagen type II alpha 1 chain	PHGDH	phosphoglycerate dehydrogenase
COL4A6	collagen type IV alpha 6 chain	PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
COL6A3	collagen type VI alpha 3 chain	PIK3CD	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta
COL7A1	collagen type VII alpha 1 chain	PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma
COL9A3	collagen type IX alpha 3 chain	PIK3R1	phosphoinositide-3-kinase regulatory subunit 1
COLEC12	collectin subfamily member 12	PIK3R2	phosphoinositide-3-kinase regulatory subunit 2
COMP	cartilage oligomeric matrix protein	PIK3R3	phosphoinositide-3-kinase regulatory subunit 3
CPA3	carboxypeptidase A3	PIK3R5	phosphoinositide-3-kinase regulatory subunit 5
CREBBP	CREB binding protein	PIM1	Pim-1 proto-oncogene, serine/threonine kinase
CRYAB	crystallin alpha B	PIP	prolactin induced protein
CSF3R	colony stimulating factor 3 receptor	PKMYT1	protein kinase, membrane associated tyrosine/threonine 1
CTSW	cathepsin W	PLA1A	phospholipase A1 member A
CXADR	CXADR, Ig-like cell adhesion molecule	PLA2G2A	phospholipase A2 group IIA

CXCL10	C-X-C motif chemokine ligand 10	PLA2G3	phospholipase A2 group III
CXCL12	C-X-C motif chemokine ligand 12	PLA2G4A	phospholipase A2 group IVA
CXCL13	C-X-C motif chemokine ligand 13	PLA2G4F	phospholipase A2 group IVF
CXCL5	C-X-C motif chemokine ligand 5	PLAT	plasminogen activator, tissue type
CXCL8	C-X-C motif chemokine ligand 8	PLCB1	phospholipase C beta 1
CXCL9	C-X-C motif chemokine ligand 9	PLCB4	phospholipase C beta 4
CXCR6	C-X-C motif chemokine receptor 6	PLCE1	phospholipase C epsilon 1
CXorf36	chromosome X open reading frame 36	PLD1	phospholipase D1
CXXC5	CXXC finger protein 5	PMS2	PMS1 homolog 2, mismatch repair system component
CYBB	cytochrome b-245 beta chain	POLD1	DNA polymerase delta 1, catalytic subunit
CYP4F3	cytochrome P450 family 4 subfamily F member 3	POLQ	DNA polymerase theta
DCN	decorin	POPDC3	popeye domain containing 3
DDB2	damage specific DNA binding protein 2	PPARG	peroxisome proliferator activated receptor gamma
DDR2	discoidin domain receptor tyrosine kinase 2	PPARGC1A	PPARG coactivator 1 alpha
DDX39A	DExD-box helicase 39A	PPP2CB	protein phosphatase 2 catalytic subunit beta
DEPDC1	DEP domain containing 1	PPP2R1A	protein phosphatase 2 scaffold subunit Aalpha
DHRS2	dehydrogenase/reductase 2	PPP2R2C	protein phosphatase 2 regulatory subunit Bgamma
DKK1	dickkopf WNT signaling pathway inhibitor 1	PRC1	protein regulator of cytokinesis 1
DKK2	dickkopf WNT signaling pathway inhibitor 2	PREP	prolyl endopeptidase
DLGAP5	DLG associated protein 5	PRF1	perforin 1
DLL1	delta like canonical Notch ligand 1	PRKAA2	protein kinase AMP-activated catalytic subunit alpha 2
DLL3	delta like canonical Notch ligand 3	PRKACA	protein kinase cAMP-activated catalytic subunit alpha
DLL4	delta like canonical Notch ligand 4	PRKACB	protein kinase cAMP-activated catalytic subunit beta
DNAJC12	DnaJ heat shock protein family (Hsp40) member C12	PRKCA	protein kinase C alpha
DPT	dermatopontin	PRKCB	protein kinase C beta
DSC2	desmocollin 2	PRKDC	protein kinase, DNA-activated, catalytic polypeptide
DTX1	deltex E3 ubiquitin ligase 1	PRKX	protein kinase, X-linked
DTX3	deltex E3 ubiquitin ligase 3	PRLR	prolactin receptor
DTX4	deltex E3 ubiquitin ligase 4	PROM1	prominin 1
DUSP4	dual specificity phosphatase 4	PSAT1	phosphoserine aminotransferase 1
DUSP6	dual specificity phosphatase 6	PSMB10	proteasome subunit beta 10
E2F1	E2F transcription factor 1	PSMB7	proteasome subunit beta 7
E2F5	E2F transcription factor 5	PSMB9	proteasome subunit beta 9
ECM2	extracellular matrix protein 2	PTCH1	patched 1
EDN1	endothelin 1	PTEN	phosphatase and tensin homolog
EDNRB	endothelin receptor type B	PTGDS	prostaglandin D2 synthase
EFNA3	ephrin A3	PTGER3	prostaglandin E receptor 3
EFNA5	ephrin A5	PTGS2	prostaglandin-endoperoxide synthase 2
EGF	epidermal growth factor	PTTG1	pituitary tumor-transforming 1

EGFR	epidermal growth factor receptor	PYCARD	PYD and CARD domain containing
EGLN2	egl-9 family hypoxia inducible factor 2	RAC2	Rac family small GTPase 2
EGLN3	egl-9 family hypoxia inducible factor 3	RAC3	Rac family small GTPase 3
EIF2AK3	eukaryotic translation initiation factor 2 alpha kinase 3	RAD51	RAD51 recombinase
EIF3B	eukaryotic translation initiation factor 3 subunit B	RAD51C	RAD51 paralog C
EIF4E2	eukaryotic translation initiation factor 4E family member 2	RAD52	RAD52 homolog, DNA repair protein
EIF4EBP1	eukaryotic translation initiation factor 4E binding protein 1	RAD54L	RAD54 like
ELF3	E74 like ETS transcription factor 3	RARRES3	retinoic acid receptor responder 3
ELK3	ELK3, ETS transcription factor	RASAL1	RAS protein activator like 1
ELOVL2	ELOVL fatty acid elongase 2	RASGRF1	Ras protein specific guanine nucleotide releasing factor 1
EMCN	endomucin	RASGRF2	Ras protein specific guanine nucleotide releasing factor 2
ENO1	enolase 1	RASGRP1	RAS guanyl releasing protein 1
ENPP2	ectonucleotide pyrophosphatase/phosphodiesterase 2	RB1	RB transcriptional corepressor 1
EP300	E1A binding protein p300	RBL1	RB transcriptional corepressor like 1
EPAS1	endothelial PAS domain protein 1	RBL2	RB transcriptional corepressor like 2
ERBB2	erb-b2 receptor tyrosine kinase 2	RBX1	ring-box 1
ERBB4	erb-b2 receptor tyrosine kinase 4	RELN	reelin
ERCC1	ERCC excision repair 1, endonuclease non-catalytic subunit	RFC4	replication factor C subunit 4
EREG	epiregulin	RNASE2	ribonuclease A family member 2
ESPL1	extra spindle pole bodies like 1, separase	RNF103	ring finger protein 103
ESR1	estrogen receptor 1	ROBO4	roundabout guidance receptor 4
ETV4	ETS variant 4	ROCK1	Rho associated coiled-coil containing protein kinase 1
ETV7	ETS variant 7	ROCK2	Rho associated coiled-coil containing protein kinase 2
EXO1	exonuclease 1	RORA	RAR related orphan receptor A
EYA1	EYA transcriptional coactivator and phosphatase 1	RORB	RAR related orphan receptor B
EYA2	EYA transcriptional coactivator and phosphatase 2	RPS6KA5	ribosomal protein S6 kinase A5
EYA4	EYA transcriptional coactivator and phosphatase 4	RPS6KB1	ribosomal protein S6 kinase B1
F3	coagulation factor III, tissue factor	RPS6KB2	ribosomal protein S6 kinase B2
FAM124B	family with sequence similarity 124 member B	RRM2	ribonucleotide reductase regulatory subunit M2
FAM198B	family with sequence similarity 198 member B	RUNX3	runt related transcription factor 3
FAM214A	family with sequence similarity 214 member A	S100A14	S100 calcium binding protein A14
FAM83D	family with sequence similarity 83 member D	S100A7	S100 calcium binding protein A7
FANCF	Fanconi anemia complementation group F	S1PR1	sphingosine-1-phosphate receptor 1
FAP	fibroblast activation protein alpha	SCARA5	scavenger receptor class A member 5
FBN1	fibrillin 1	SCUBE2	signal peptide, CUB domain and EGF like domain containing 2

FGF1	fibroblast growth factor 1	SELE	selectin E
FGF10	fibroblast growth factor 10	SERBP1	SERPINE1 mRNA binding protein 1
FGF12	fibroblast growth factor 12	SERPINB5	serpin family B member 5
FGF13	fibroblast growth factor 13	SERPINH1	serpin family H member 1
FGF18	fibroblast growth factor 18	SFN	stratifin
FGF2	fibroblast growth factor 2	SFRP1	secreted frizzled related protein 1
FGF7	fibroblast growth factor 7	SFRP2	secreted frizzled related protein 2
FGF9	fibroblast growth factor 9	SFRP4	secreted frizzled related protein 4
FGFR2	fibroblast growth factor receptor 2	SHC2	SHC adaptor protein 2
FGFR3	fibroblast growth factor receptor 3	SHC4	SHC adaptor protein 4
FGFR4	fibroblast growth factor receptor 4	SHE	Src homology 2 domain containing E
FGL2	fibrinogen like 2	SHMT2	serine hydroxymethyltransferase 2
FHL1	four and a half LIM domains 1	SIDT1	SID1 transmembrane family member 1
FLI1	Fli-1 proto-oncogene, ETS transcription factor	SIGIRR	single Ig and TIR domain containing
FLNC	filamin C	SIX1	SIX homeobox 1
FLRT3	fibronectin leucine rich transmembrane protein 3	SKA3	spindle and kinetochore associated complex subunit 3
FLT3	fms related tyrosine kinase 3	SKP1	S-phase kinase associated protein 1
FNBP1	formin binding protein 1	SKP2	S-phase kinase associated protein 2
FOS	Fos proto-oncogene, AP-1 transcription factor subunit	SLC2A1	solute carrier family 2 member 1
FOSL1	FOS like 1, AP-1 transcription factor subunit	SLC39A6	solute carrier family 39 member 6
FOXA1	forkhead box A1	SLC44A4	solute carrier family 44 member 4
FOXC1	forkhead box C1	SLPI	secretory leukocyte peptidase inhibitor
FOXC2	forkhead box C2	SMAD1	SMAD family member 1
FOXM1	forkhead box M1	SMAD3	SMAD family member 3
FOXP3	forkhead box P3	SMAD4	SMAD family member 4
FREM2	FRAS1 related extracellular matrix protein 2	SMAD5	SMAD family member 5
FST	follistatin	SMC1B	structural maintenance of chromosomes 1B
FSTL1	follistatin like 1	SMO	smoothened, frizzled class receptor
FSTL3	follistatin like 3	SMURF2	SMAD specific E3 ubiquitin protein ligase 2
FUT3	fucosyltransferase 3 (Lewis blood group)	SNAI1	snail family transcriptional repressor 1
FXYD3	FXYD domain containing ion transport regulator 3	SNAI2	snail family transcriptional repressor 2
FZD10	frizzled class receptor 10	SOCS1	suppressor of cytokine signaling 1
FZD7	frizzled class receptor 7	SOCS2	suppressor of cytokine signaling 2
FZD8	frizzled class receptor 8	SOCS3	suppressor of cytokine signaling 3
FZD9	frizzled class receptor 9	SOX10	SRY-box 10
GABRP	gamma-aminobutyric acid type A receptor pi subunit	SOX17	SRY-box 17
GADD45A	growth arrest and DNA damage inducible alpha	SOX2	SRY-box 2
GADD45B	growth arrest and DNA damage inducible beta	SOX9	SRY-box 9
GADD45G	growth arrest and DNA damage inducible gamma	SP1	Sp1 transcription factor
GAS1	growth arrest specific 1	SPC25	SPC25, NDC80 kinetochore complex component

GATA3	GATA binding protein 3	SPDEF	SAM pointed domain containing ETS transcription factor
GATA4	GATA binding protein 4	SPN	sialophorin
GDF15	growth differentiation factor 15	SPP1	secreted phosphoprotein 1
GDF5	growth differentiation factor 5	SPRY1	sprouty RTK signaling antagonist 1
GGH	gamma-glutamyl hydrolase	SPRY2	sprouty RTK signaling antagonist 2
GHR	growth hormone receptor	SPRY4	sprouty RTK signaling antagonist 4
GJB2	gap junction protein beta 2	SRPX	sushi repeat containing protein, X-linked
GLI3	GLI family zinc finger 3	ST6GALNAC2	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
GNG4	G protein subunit gamma 4	STAT1	signal transducer and activator of transcription 1
GNLY	granulysin	STC1	stanniocalcin 1
GPC4	glypican 4	SUV39H2	suppressor of variegation 3-9 homolog 2
GPR160	G protein-coupled receptor 160	SYTL4	synaptotagmin like 4
GPX3	glutathione peroxidase 3	TAP1	transporter 1, ATP binding cassette subfamily B member
GRB2	growth factor receptor bound protein 2	TAP2	transporter 2, ATP binding cassette subfamily B member
GRB7	growth factor receptor bound protein 7	TAPBP	TAP binding protein
GREM1	gremlin 1, DAN family BMP antagonist	TBC1D9	TBC1 domain family member 9
GRIA3	glutamate ionotropic receptor AMPA type subunit 3	TBX1	T-box 1
GRIN1	glutamate ionotropic receptor NMDA type subunit 1	TCEAL1	transcription elongation factor A like 1
GRIN2A	glutamate ionotropic receptor NMDA type subunit 2A	TCF4	transcription factor 4
GSK3B	glycogen synthase kinase 3 beta	TCF7L1	transcription factor 7 like 1
GTF2H2	general transcription factor IIH subunit 2	TEK	TEK receptor tyrosine kinase
GZMA	granzyme A	TFDP1	transcription factor Dp-1
GZMB	granzyme B	TFF1	trefoil factor 1
GZMH	granzyme H	TFF3	trefoil factor 3
GZMM	granzyme M	TGFB1	transforming growth factor beta 1
HAPLN1	hyaluronan and proteoglycan link protein 1	TGFB2	transforming growth factor beta 2
HAS1	hyaluronan synthase 1	TGFB3	transforming growth factor beta 3
HBB	hemoglobin subunit beta	TGFBR2	transforming growth factor beta receptor 2
HDAC1	histone deacetylase 1	THBS1	thrombospondin 1
HDAC10	histone deacetylase 10	THBS2	thrombospondin 2
HDAC11	histone deacetylase 11	THBS4	thrombospondin 4
HDAC2	histone deacetylase 2	THY1	Thy-1 cell surface antigen
HDAC5	histone deacetylase 5	TIE1	tyrosine kinase with immunoglobulin like and EGF like domains 1
HDAC6	histone deacetylase 6	TIGIT	T-cell immunoreceptor with Ig and ITIM domains
HDC	histidine decarboxylase	TIMP4	TIMP metallopeptidase inhibitor 4
HEG1	heart development protein with EGF like domains 1	TLE3	transducin like enhancer of split 3
HELLS	helicase, lymphoid specific	TLR4	toll like receptor 4
HEMK1	HemK methyltransferase family member 1	TLX1	T-cell leukemia homeobox 1
HES1	hes family bHLH transcription factor 1	TMEM45B	transmembrane protein 45B

HGF	hepatocyte growth factor	TMPRSS2	transmembrane protease, serine 2
HIF1A	hypoxia inducible factor 1 alpha subunit	TMPRSS4	transmembrane protease, serine 4
HIST1H1C	histone cluster 1 H1 family member c	TNF	tumor necrosis factor
HIST1H2BH	histone cluster 1 H2B family member h	TNFAIP6	TNF alpha induced protein 6
HIST1H3H	histone cluster 1 H3 family member h	TNFSF10	TNF superfamily member 10
HIST3H2BB	histone cluster 3 H2B family member b	TNKS	tankyrase
HK2	hexokinase 2	TNKS2	tankyrase 2
HLA-A	H1 histone family member 0	TNN	tenascin N
HLA-B	major histocompatibility complex, class I, B	TOP2A	DNA topoisomerase II alpha
HLA-C	major histocompatibility complex, class I, C	TP53	tumor protein p53
HLA-DMA	major histocompatibility complex, class II, DM alpha	TPSAB1	tryptase alpha/beta 1
HLA-DMB	major histocompatibility complex, class II, DM beta	TRIP13	thyroid hormone receptor interactor 13
HLA-DOB	major histocompatibility complex, class II, DO beta	TSPAN1	tetraspanin 1
HLA-DPA1	major histocompatibility complex, class II, DP alpha 1	TSPAN7	tetraspanin 7
HLA-DPB1	major histocompatibility complex, class II, DP beta 1	TTK	TTK protein kinase
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	TTYH1	tweety family member 1
HLA-DQB1	major histocompatibility complex, class II, DQ beta 1	TUBA4A	tubulin alpha 4a
HLA-DRA	major histocompatibility complex, class II, DR alpha	TWIST1	twist family bHLH transcription factor 1
HLA-DRB1	major histocompatibility complex, class II, DR beta 1	TWIST2	twist family bHLH transcription factor 2
HLA-E	major histocompatibility complex, class I, E	TYK2	tyrosine kinase 2
HMGA1	high mobility group AT-hook 1	TYMP	thymidine phosphorylase
HNF1A	HNF1 homeobox A	TYMS	thymidylate synthetase
HOXA5	homeobox A5	UBE2C	ubiquitin conjugating enzyme E2 C
HOXA7	homeobox A7	UBE2T	ubiquitin conjugating enzyme E2 T
HOXA9	homeobox A9	VCAN	versican
HOXB13	homeobox B13	VEGFA	vascular endothelial growth factor A
HOXB3	homeobox B3	VEGFD	vascular endothelial growth factor D
HSPA2	heat shock protein family A (Hsp70) member 2	VIM	vimentin
IBSP	integrin binding sialoprotein	VIT	vitrin
ICAM1	intercellular adhesion molecule 1	WDR77	WD repeat domain 77
ID1	inhibitor of DNA binding 1, HLH protein	WEE1	WEE1 G2 checkpoint kinase
ID2	inhibitor of DNA binding 2	WIF1	WNT inhibitory factor 1
ID4	inhibitor of DNA binding 4, HLH protein	WNT10A	Wnt family member 10A
IDO1	indoleamine 2,3-dioxygenase 1	WNT11	Wnt family member 11
IFT140	intraflagellar transport 140	WNT2	Wnt family member 2
IGF1	insulin like growth factor 1	WNT4	Wnt family member 4
IGF1R	insulin like growth factor 1 receptor	WNT5A	Wnt family member 5A
IKZF3	IKAROS family zinc finger 3	WNT5B	Wnt family member 5B
IL10RA	interleukin 10 receptor subunit alpha	WNT6	Wnt family member 6
IL11RA	interleukin 11 receptor subunit alpha	WNT7B	Wnt family member 7B

IL12RB2	interleukin 12 receptor subunit beta 2	WRN	Werner syndrome RecQ like helicase
IL13RA1	interleukin 13 receptor subunit alpha 1	WT1	Wilms tumor 1
IL1B	interleukin 1 beta	XRCC2	X-ray repair cross complementing 2
IL1R2	interleukin 1 receptor type 2	XRCC3	X-ray repair cross complementing 3
IL1RN	interleukin 1 receptor antagonist	ZBTB16	zinc finger and BTB domain containing 16
IL20RA	interleukin 20 receptor subunit alpha	ZEB1	zinc finger E-box binding homeobox 1
IL20RB	interleukin 20 receptor subunit beta	ZEB2	zinc finger E-box binding homeobox 2
IL22RA2	interleukin 22 receptor subunit alpha 2	ZFPM2	zinc finger protein, FOG family member 2
IL24	interleukin 24	ZFYVE9	zinc finger FYVE-type containing 9
IL2RA	interleukin 2 receptor subunit alpha	ZIC2	Zic family member 2
IL2RB	interleukin 2 receptor subunit beta	ZNF205	zinc finger protein 205

9.4 Illumina DNA sequencing panel of genes

Full table showing genes by chromosome region is shown here:

http://tiny.cc/appendixtables sheet 1.

MTOR	CHD1	STMN2	RB1	RUNX1
SPEN	CDC25C	MYC	FOXA1	XBP1
LDLRAP1	CSF1R	CDKN2A	ZFP36L1	MYH9
ARID1A	DDR1	NOTCH1	AKT1	USP9X
JAK1	NOTCH4	GATA3	CYP19A1	KDM6A
NOTCH2	EPHA7	NRG3	CREBBP	MED12
INSRR	FOXO3	PTEN	CBFB	AGTR2
DNMT3A	MED23	MKI67	CTCF	
BIRC6	ESR1	MALAT1	CDH1	
SF3B1	SYNE1	CCND1	TP53	
SETD2	ARID1B	ATM	MAP2K4	
FOXP1	MAP3K4	KMT2A	NCOR1	
ATR	MLLT4	CDKN1B	NF1	
TBL1XR1	DNAH11	KRAS	ERBB2	
PIK3CA	EGFR	KMT2D	BRCA1	

PDGFRA	RELN	ERBB3	MUC16	
KIT	MET	MDM2	CCNE1	
FBXW7	BRAF	TBX3	AKT2	
MAP3K1	KMT2C	NCOR2	CDC25B	
PIK3R1	PRKDC	BRCA2	NCOA3	

9.5 Full copy number alteration results

Full copy number alteration results by patient available at http://tiny.cc/appendixtables sheet 2.

9.5.1 Significant chromosome region copy number losses

Chromosome	Start	End	Gene(s)
17p13	5740893	6309642	WSCD1
17q12	36556990	36557490	GGNBP2
17q12	36557490	36609611	GGNBP2, DHRS11, MRM1
17q12	36911328	36975929	LHX1, AATF
17q12	36976567	36977067	AATF
17q12	36977067	37463487	AATF, C17orf78, ACACA, TADA2A,
17q12	37635690	37636190	DDX52
17q12	37636190	37889955	DDX52, HNF1B
17q12	37890455	38557211	TBC1D3K, TBC1D3L, TBC1D3D, TBC1D3C, TBC1D3E, TBC1D3, MRPL45, GPR179, SOCS7, ARHGAP23, SRCIN1
14q32	104769515	106883218	IGH genes, AKT1, ZBTB42; CEP170B, AHNAK2, PLD4, CDCA4, CLBA1, GPR132, JAG2, NUDT14, CRIP2, CRIP1, TEDC1, BRF1, BTBD6, PACS2, TMEM121, ELK2BP, ADAM6, TEX22, MTA1
17q12	35856404	35909156	HEATR9, TAF15, CCL5
17p11	21332624	21616501	KCNJ12
17p13	6417265	6469214	AIPL1, PIMREG, PITPNM3
1q23	156840338	156858668	NTRK1, INSRR
5q13	68297626	68298126	PIK3R1

5q14-21	89934881	106305770	CHD1, FAM174A, ST8SIA4, SLCO4C1, SLCO6A1, PAM, LNPEP, ERAP2, GIN1, LIX1, RIOK2, CETN3, POLR3G, MBLAC2, LYSMD3 ADGRV1, ARRDC3, PPIP5K2, C5orf30, NUDT12, NR2F1, POU5F2 FAM172A, KIAA0825, MCTP1, SLF1, TTC37, ARSK, GPR150, RFESD, SPATA9, RHOBTB3, GLRX, ERAP1, PCSK1, ELL2
14q32	104769015	104769515	AKT1
17q12	37463487	37463987	TADA2A
17q12	37463987	37635690	DUSP14, SYNRG, DDX52, TADA2A
17q12	38557211	38609025	SRCIN1

9.5.2 Significant chromosome region copy number gains

Chromosome	Start	End	Gene(s)
5q12-13	65587954	68226162	ERBIN, SGTB, NLN, SHLD3, TRAPPC13, TRIM23, SREK1, MAST4 CD180, PIK3R1
6q25	151807394	151809333	ESR1
12p13	12122384	12271114	BCL2L14, LRP6
12p13	12271114	12717306	CREBL2, GPR19, CDKN1B, MANSC1, BORCS5, DUSP16
12p13	12717306	12721755	CDKN1B
8q11	45927765	47927950	SPIDR, CEBPD, PRKDC
8q11	47927950	47957482	PRKDC
8p12	34977915	37290924	UNC5D, KCNU1,
3p13	70966092	70969821	FOXP1
3p13	70969821	71593722	FOXP1

5q12	61765406	65587454	HTR1A, RNF180, RGS7BP, SHISAL2B, SREK1IP1, CWC27, ADAMTS6, CENPK, PPWD1, LRRC70, IPO11, DIMT, KIF2A
5q12	65587454	65587954	PPWD1
6q25	151379994	151380494	ZBTB2
6q25	151380494	151473607	ZBTB2, RMND1, ARMT1
6q25	151474107	151807394	CCDC170, ESR1
6q25	151809333	151821881	ESR1

9.6 Mutation results by gene

Gene	Mutations Controls	No mutations Controls	Mutations Cases	No mutations cases	P value (fisher)	FDR
PIK3CA	12	23	18	21	0.348	1
MUC16	6	29	9	30	0.575	1
GATA3	10	25	4	35	0.073	1
CBFB	6	29	7	32	1.000	1
MAP3K1	8	27	2	37	0.040	1
TP53	5	30	3	36	0.464	1
NOTCH2	1	34	6	33	0.111	1
BIRC6	5	30	1	38	0.095	1
DNAH11	3	32	2	37	0.662	1
MAP2K4	3	32	2	37	0.662	1
NCOR1	2	33	3	36	1.000	1
PTEN	3	32	2	37	0.662	1
SYNE1	4	31	1	38	0.183	1
CDH1	0	35	5	34	0.056	1
KMT2D	2	33	2	37	1.000	1
SPEN	1	34	3	36	0.617	1
USP9X	3	32	1	38	0.339	1
KMT2C	0	35	4	35	0.117	1
CREBBP	2	33	1	38	0.600	1
ERBB3	2	33	1	38	0.600	1
KDM6A	2	33	1	38	0.600	1
MKI67	1	34	2	37	1.000	1
PRKDC	2	33	1	38	0.600	1
RUNX1	1	34	2	37	1.000	1
BRCA2	0	35	3	36	0.242	1
ARID1A	1	34	1	38	1.000	1
FOXA1	2	33	0	39	0.220	1
FOXO3	1	34	1	38	1.000	1
KMT2A	1	34	1	38	1.000	1
MYH9	1	34	1	38	1.000	1
NF1	2	33	0	39	0.220	1
PIK3R1	1	34	1	38	1.000	1
RB1	1	34	1	38	1.000	1
RELN	1	34	1	38	1.000	1
TBX3	1	34	1	38	1.000	1
JAK1	0	35	2	37	0.495	1

MED12	0	35	2	37	0.495	1
SETD2	0	35	2	37	0.495	1
AKT1	1	34	0	39	0.473	1
ATM	1	34	0	39	0.473	1
BRCA1	1	34	0	39	0.473	1
EPHA7	1	34	0	39	0.473	1
TBL1XR1	1	34	0	39	0.473	1
XBP1	1	34	0	39	0.473	1
CCND1	0	35	1	38	1.000	1
CDC25B	0	35	1	38	1.000	1
CHD1	0	35	1	38	1.000	1
CSF1R	0	35	1	38	1.000	1
CYP19A1	0	35	1	38	1.000	1
EGFR	0	35	1	38	1.000	1
ERBB2	0	35	1	38	1.000	1
INSRR	0	35	1	38	1.000	1
MED23	0	35	1	38	1.000	1
MTOR	0	35	1	38	1.000	1
NCOA3	0	35	1	38	1.000	1
NCOR2	0	35	1	38	1.000	1
NOTCH1	0	35	1	38	1.000	1
SF3B1	0	35	1	38	1.000	1

9.7 Gene expression results

Full set of gene expression data available at http://tiny.cc/appendixtables sheet 3.

9.7.1 65 significant genes.

Gene	Case expression mean	Control expression mean	Expression differential	p.value	FDR
PLA2G4F	5.02	3.86	1.16	0.0011	0.5451
CDKN2D	6.81	7.25	-0.44	0.0025	0.5451
HGF	3.57	2.34	1.23	0.0028	0.5451
SMAD5	9.27	9.62	-0.35	0.0042	0.5451
ITGB3	6.08	6.63	-0.56	0.0043	0.5451
SPRY4	7.61	7.96	-0.35	0.0056	0.5451
CDC25B	8.38	8.73	-0.35	0.0065	0.5451
NSD3	9.63	10.02	-0.4	0.0065	0.5451
ADCY9	8.33	8.69	-0.36	0.0079	0.5451
PIM1	8.96	9.5	-0.53	0.0107	0.5451
PAX8	1.5	0.64	0.86	0.0111	0.5451
CCNE2	5.77	4.96	0.81	0.0112	0.5451
IL1B	5.22	5.79	-0.56	0.012	0.5451
BAIAP3	5.52	4.33	1.19	0.0122	0.5451

TFDP1	8.59	8.76	-0.17	0.0134	0.5451
RBL1	6.78	6.32	0.45	0.0136	0.5451
PLCE1	6.41	5.77	0.64	0.0154	0.5451
BNIP3	7.39	7.78	-0.39	0.0159	0.5451
TNFAIP6	6.47	5.62	0.85	0.0161	0.5451
VEGFA	10.02	10.4	-0.38	0.0166	0.5451
PKMYT1	6.24	5.36	0.88	0.018	0.5451
ELK3	8.81	9.11	-0.3	0.0187	0.5451
ASPM	5.1	4.06	1.03	0.0189	0.5451
EIF4EBP1	8.71	9.05	-0.35	0.0192	0.5451
MFNG	7.12	7.51	-0.39	0.0201	0.5451
CDCA5	5.6	4.87	0.74	0.0204	0.5451
GPX3	8.26	8.76	-0.49	0.0214	0.5451
RAD54L	5.42	4.6	0.82	0.0218	0.5451
CHEK2	5.71	4.98	0.73	0.023	0.5451
TYK2	8.68	8.9	-0.22	0.023	0.5451
E2F1	5.22	4.29	0.93	0.0232	0.5451
MMP11	10.58	9.6	0.98	0.0248	0.5451
ANXA9	6.67	5.76	0.91	0.0251	0.5451
ETV7	2.07	1.2	0.87	0.0272	0.5451
TSPAN1	6.45	5.16	1.29	0.0279	0.5451
ORC6	4.9	4.08	0.82	0.0282	0.5451
EP300	10.36	10.6	-0.23	0.0304	0.5451
AGT	4.41	3.28	1.13	0.0324	0.5451
SMURF2	8.7	8.97	-0.26	0.0325	0.5451
MRE11	8.07	8.3	-0.23	0.034	0.5451
ARID1A	9.74	9.89	-0.15	0.0342	0.5451
MIS18A	7.28	7.51	-0.23	0.0353	0.5451
CDC20	4.05	3.06	0.99	0.0356	0.5451
BBC3	8.49	8.76	-0.27	0.0372	0.5451
EFNA3	6.67	6.02	0.64	0.0374	0.5451
ROCK2	7.68	7.5	0.18	0.0375	0.5451
IL24	0.75	0.21	0.54	0.0377	0.5451
MAP2K4	9.62	9.97	-0.35	0.0388	0.5451
BAX	9.22	9.06	0.16	0.0389	0.5451
PDCD1LG2	2.76	1.95	0.81	0.0399	0.5451
CEP55	5.75	4.87	0.88	0.04	0.5451
DDX39A	8.46	8.7	-0.24	0.0411	0.5451
PRKACA	9.77	9.92	-0.16	0.0416	0.5451
NEO1	8.69	8.91	-0.22	0.0418	0.5451
MCM3	7.91	7.64	0.26	0.0421	0.5451
CSF3R	8.64	9.02	-0.38	0.0428	0.5451
IL13RA1	10.05	10.24	-0.19	0.0435	0.5451

HOXB13	1.88	1	0.87	0.0437	0.5451
ROCK1	8.5	8.75	-0.25	0.044	0.5451
PTTG1	4.63	3.68	0.95	0.0455	0.5451
CCR2	3.07	2.2	0.87	0.046	0.5451
TP53	8.59	8.81	-0.22	0.0463	0.5451
PIK3CA	6.61	6.15	0.47	0.0487	0.5451
IL3RA	3.11	2.39	0.72	0.0488	0.5451
MMP3	5.76	4.86	0.9	0.0493	0.5451

9.8 Gene Set Enrichment Results

Gene ratio describes the ratio of genes within each gene set to total significant genes either up or down in cases.

9.8.1 Down in cases

Gene set analysis	Gene set description	Gene Ratio	Bg Ratio	p value	p-adjust (FDR)	Gene name
BC360	Epigenetic Regulation	3/31	18/743	0.04	0.34	ARID1A/BNIP3/NSD3
	Cytokine and Chemokine Signalling	5/31	48/743	0.04	0.34	CSF3R/IL13RA1/IL1B/TYK2/ VEGFA
GO	Nuclear export	4/31	8/736	0.0002	0.26	DDX39A/IL1B/PRKACA/TP53
	Positive regulation of nuclear export	3/31	6/736	0.001	0.45	IL1B/PRKACA/TP53
	mitochondrial outer membrane permeability	4/31	14/736	0.002	0.44	BBC3/BNIP3/TFDP1/TP53
KEGG	Huntington disease	4/27	16/587	0.004	0.31	BBC3/EP300/GPX3/TP53
	Human cytomegalovirus infection	8/27	69/587	0.009	0.31	ADCY9/EIF4EBP1/IL1B/ITGB3/ PRKACA/ROCK1/TP53/VEGFA
HALLMARK	IL6/JAK/STAT signalling	6/22	31/451	0.002	0.07	CSF3R/IL13RA1/IL1B/ITGB3/ PIM1/TYK2
	Reactive oxygen species	2/22	3/451	0.007	0.11	CDKN2D/GPX3
	Unfolded protein response	2/22	6/451	0.03	0.32	EIF4EBP1/VEGFA

9.8.2 Up in cases

lysis	Gene set description	Gene Ratio	Bg Ratio	p value	p-adjust (FDR)	Gene name
GO	DNA replication initiation	3/34	6/736	0.002	0.70	CCNE2/MCM3/ORC6
	meiotic cell cycle	6/34	39/736	0.007	0.70	ASPM/CCNE2/CDC20/PKMYT1/ PTTG/RAD54L
	DNA dependent DNA replication	4/34	18/736	0.007	0.70	CCNE2/CHEK2/MCM3/ORC6
	Cell cycle	16/34	198/736	0.008	0.70	ASPM1/BAX/CCNE2/CDC20/ CDCA5/CEP55/CHEK2/E2F1/HGF/ MCM3/ORC6/PKMYT1/PTTG/ RAD54L/RBL1/ROCK2
KEGG	Cell cycle	9/27	65/587	0.002	0.20	CCNE2/CDC20/CHEK2/E2F1/ MCM3/ORC6/ PKMYT1/PTTG/RBL1
	Inositol phosphate metabolism	2/27	7/587	0.04	0.76	PIK3CA/PLCE1
	Human T-cell leukemia virus 1 infection	7/27	73/587	0.04	0.76	BAX/CCNE2/CDC20/CHEK2/E2F/ PIK3CA/PTTG
HALLMARK	G2M checkpoint	7/20	52/451	0.004	0.09	CDC20/E2F1/MCM3/ORC6/ PTTG1/RAD54L/RBL1