Master of Philosophy Thesis

Identifying biomarkers with predictive utility in the clinical management of prostate cancer

Mr William Green

BSc BMBS FRCS

Acknowledgments:

I would like to thank my supervisors Dr Des Powe and Prof Graham Ball for their invaluable support and guidance throughout my MPhil project and also Miss Catherine Johnson for all her help carrying out the immunohistochemistry.

Many thanks are due to the Nottingham University Hospitals Special Trustees for providing a grant award that funded the TMA production and contributing significantly to the IHC.

In addition I would like to thank the John van Geest Cancer Research Centre for funding my higher degree registration fees, it is much appreciated.

Publications based on work carried out during this MPhil thesis:

Green W**,** Ball G, Hulman G, Johnson C, Van Schalwyk G, Ratan H, Soria D, Garibaldi J, Parkinson R, Hulman J, Rees R, Powe D. *KI67 and DLX2 predict high risk of metastasis formation in prostate cancer – a targeted molecular approach.* British Journal of Cancer 2016 Jul 12;115(2):236-42. PMID: 27336609

Green WJF, Ball G, Powe D. *Does the molecular classification of breast cancer point the way for biomarker identification in prostate cancer?* World J Clin Urol 2016; 5(2): 80-89

Buczek E, Miles A, Green W, Johnson C, Boocock D, Pockley A, Rees R, Hulman G, Van Schalkwyk G, Parkinson R, Hulman J, Powe D and Regad T. *Cytoplasmic PML Promotes TGFβ-Associated Epithelial-Mesenchymal Transition and Invasion in Prostate Cancer.* Oncogene. 2016 Jun 30;35(26):3465-75. PMID: 26549027

Contents

Tables and Figures

Abbreviations

ADT	Androgen Deprivation Therapy
ANN	Artificial Neural Network
AR	Androgen Receptor
CaP	Prostate Cancer
CK	Cytokine
DRE	Digital Rectal Examination
DSS	Disease Specific Survival
DTC	Disseminated Tumour Cells
EAU	European Association of Urology
ESC	Embryonic Stem Cells
GEP	Gene Expression Profiling
IHC	Immunohistochemistry
ISUP	International Society for Urological Pathology
LRP	Laparoscopic Radical Prostatectomy
NICE	National Institute for Clinical Excellence
OS	Overall Survival
PSA	Prostate Specific Antigen
RARP	Robot Assisted Radical Prostatectomy
TMA	Tissue Micro Array
TRUS	Trans-rectal Ultrasound
TURP	Transurethral Resection of Prostate

1.0 Abstract

Background

There is significant variation in clinical outcome between patients diagnosed with prostate cancer (CaP). Although useful, statistical nomograms and risk stratification tools alone do not always accurately predict an individual's need for and response to treatment. As a result there remains a need to identify and validate biomarkers for predicting prostate cancer outcomes using robust and routinely available pathology techniques to recognize men at most risk of premature death due to prostate cancer.

The day-to-day treatment options available to clinicians are continually evolving, with newer technologies and a greater understanding of the tumour biology prompting innovative approaches. However, despite this, all techniques have considerable associated side effects and there is a great deal of disagreement regarding which patients need radical treatment and which can be safely monitored, therefore avoiding unnecessary morbidity.

If more accurate risk stratification can be achieved using newly developed biomarkers (probably in addition to conventional staging techniques) then the aim is to reduce unnecessary treatment and assist timely and appropriate surgical and oncological intervention.

Aims and Objectives

We aimed to develop biomarkers predictive of outcome in prostate cancer, in particular ones that could be used in a mainstream NHS laboratory to help clinicians and patients make informed decisions regarding the management of their disease. We

also intended to investigate whether bioinfomatic techniques such as artificial neural network analysis (ANN) could play a role in prostate cancer biomarker identification. We also felt it was important to validate our clinical data set and associated tissue micro array (TMA) by comparing its performance against previously identified biomarkers shown to have predictive utility in prostate cancer.

Methods

A tissue microarray (TMA) was constructed from transurethral resection of prostate (TURP) and transrectal ultrasound-guided (TRUS) prostate biopsy samples that were histologically proven to demonstrate prostate cancer. Patients had undergone these procedures either to deal with troublesome urinary symptoms or had presented with an elevated prostate specific antigen (PSA) blood test. A comprehensive clinical data set of parameters conventionally used to decide upon treatment and monitor clinical response was collected. ANN analysis was used to identify candidate markers conferring increased risk of death and metastasis interrogating a public cDNA array, alongside a conventional literature review identifying previously published biomarkers that could be used to validate the TMA and clinical dataset. Immunohistochemical analysis of the TMA was carried out and univariate and multivariate tests performed to explore the association of tumour protein levels of identified biomarkers with various clinical endpoints, particularly time to death and metastasis.

Results

We successfully demonstrated associations between various biomarkers, and in particular validated our TMA and clinical dataset against the previously published

marker Ki67, showing that Ki67 is predictive of CaP-specific survival and development of future metastases. In addition we were able to identify an entirely novel prostate cancer marker, DLX2, using artificial neural network analysis and demonstrate it has a statistically significant association with the development of prostate cancer metastases.

Conclusion

The Nottingham TMA has been shown to have utility in the investigation of candidate biomarkers in prostate cancer. We have also demonstrated that bioinfomatic techniques such as artificial neural network analysis can be employed to isolate candidate markers. During this work we have identified two cancer cell proliferation markers, Ki67 and DLX2, that may be able to inform clinical decision making when identifying patients for suitable for prompt active treatment versus active surveillance.

2.0 Introduction

Prostate cancer (CaP) is the most common male malignancy in the United Kingdom with an incidence of 176.4 per 100,000 in 2015[1]. It is the second most common male cancer worldwide[2] and confers significant morbidity and mortality. With rising incidence it is a tremendous health economic burden, with annual expenditure in the UK of £94.2 million and in the US of $11.5 billion in 2010 alone[3].

Diagnosis of CaP is based on clinical examination of the prostate, serum prostate specific antigen (PSA) levels and histological assessment. Assessment of tumour aggression is based on the Gleason grading system. This was developed by Dr Donald Gleason, a pathologist in the USA, in the 1970s and has remained the single most important diagnostic and prognostic component in CaP assessment and treatment. The Gleason score is based on a microscopic examination of prostate tissue. The most common architectural pattern is identified and graded on a scale of 1 to 5 (5 being the most aggressive) and then the second most common architectural pattern is similarly graded. The two scores are then added together giving a total out of 10. More information on the Gleason score can be found in section 2.3.

While technology has evolved quickly, particularly in the field of imaging used to identify and stage prostate cancer, the formal diagnosis still rests on histological analysis of prostate tissue samples. Since the 1980s the most common technique of acquiring tissue for targeted diagnosis has been using a trans-rectal needle biopsy. Initially the method involved taking 6 cores, 3 from each lateral lobe, but over the years evolving to employing a ‘sextant’ pattern of 12 cores, 6 from each lateral lobe. The procedure is usually guided by an ultrasound probe inserted into the rectum, and

covered with injected local anaesthetic. Despite this the procedure is still extremely uncomfortable for the patient, particularly in the early stages of inserting the ultrasound probe, as the local anaesthetic can only be inserted once this is done. The procedure also confers a risk of rectal bleeding and infection, which in approximately 1% of patients can necessitate admission to hospital with sepsis and the need for intravenous antibiotics. The benefits of the procedure are that it is relatively cheap, easy to learn and can be performed in an outpatient setting.

More recently attention has turned to trans-perineal prostate biopsies, which involve multiple needle biopsies taken through the perineum, again guided by a rectal ultrasound probe, but carried out under general or spinal anaesthesia. Trans-perineal biopsies have been repeatedly demonstrated to find higher-grade cancer in approximately 40% of patients originally diagnosis with TRUS biopsies[4]. This is thought to be due to the techniques ability to access the anterior aspect of the prostate and also through a simple increase in the volume of tissue taken (approximately 40-50 cores compared to a 'standard' 12 in TRUS biopsy). The trans-perineal approach has far less risk of infection, but does have a higher (approx. 10%) risk of triggering a period of urinary retention. It is also much more expensive and labour intensive, requiring a full theatre team, general anaesthetic and often an overnight stay in hospital.

In a resource limited NHS, many Trusts have taken the pragmatic approach to perform TRUS biopsy in the initial setting and if the patient is thought suitable for surveillance rather than radical treatment (for example if the TRUS biopsy shows Gleason 6 or low volume Gleason 3+4) then proceeding to trans-perineal biopsies to

ensure that the patient is not one of the 40% who are upgraded and therefore may indeed have to undergo radical treatment.

There are various ways to stage prostate cancer, and a combination of modalities are usually employed. Clinical staging is based on a digital rectal examination of the prostate and if an abnormality is felt can be described as:

Stage 2a	A tumour involving less than half of one lobe of the prostate
Stage 2b	A tumour involving more than one half of one lobe of the prostate
Stage 2c	Tumour involving any amount of both lobes of the prostate
Stage 3	Tumour likely to have grown beyond prostatic capsule
Stage 4	Tumour has involved other surrounding organs

By definition impalpable disease cannot be staged clinically, so stage 1 disease is based on histological analysis of tissue.

In addition, the staging can be further subdivided when histological and imaging information becomes available.

Prostate cancer is also diagnosed from tissue not taken with the primary goal of looking for cancer. A commonly performed procedure, transurethral resection of prostate (TURP) is carried out to improve the urinary flow from the bladder that was being prevented from emptying efficiently by prostatic overgrowth. Here, the prostatic tissue surrounding the urethra is ‘resected’ (usually cut with a heated loop of wire), leaving a larger channel for urine to flow through and the resected tissue is

routinely sent to the laboratory for analysis. Given that this operation is performed most frequently in older men, and prostate cancer is more common as men get older, it is not unlikely that a diagnosis of prostate cancer will be made. If this occurs they should be staged in the same way as biopsy-proven disease.

Correlation between the above staging factors and subsequent cancer outcomes has led to the development of well validated risk stratification tools[5] that broadly classify newly diagnosed patients into low (47%), intermediate (38%) and high (15%) risk groups[6,7]. These tools, and meticulously populated nomograms[8,9], continue to inform clinical decision making during the investigation, management and follow-up of prostate cancer[10].

However, it is increasingly apparent that these tools alone are not sufficient to determine an individual's likelihood of being affected by clinically significant disease, particularly in the large 'intermediate' risk group, which accounts for approximately 38% of patients. Some patients require radical treatment but in others their disease is likely to remain indolent, having no demonstrable effect on their quality of life, or indeed life expectancy.

Over the last decade there has been increasing recognition that most patients with low risk disease do not require radical treatment, as they are unlikely to die from their disease. Major studies published recently have confirmed this observation, demonstrating no survival benefit from radical treatment in any patient group with low risk disease[11]. As a result, both the National Institute for Clinical Excellence (NICE) and the European Association of Urology (EAU) guidelines suggest active

surveillance (close monitoring of patients' PSA and follow-up biopsies) as first line management in low risk disease[12,13].

The factors that determine this variation in cancer aggression between patients are not fully elucidated. In particular, cellular response to androgen ablation and subsequent paracrine/autocrine adaptation is poorly understood and despite best therapies, median survival in castrate resistant patients is only approximately 35 months[14].

Prostate cancer is an extremely heterogeneous disease process and further work is required to characterise its complex molecular biological mechanisms and genetic aberrations. This heterogeneity presents obstacles and opportunities for identifying and developing more accurate diagnostic and prognostic tests and new therapeutic avenues. There is now a realisation that CaP has an intricate relationship with its stromal microenvironment[15,16] and it may develop from different progenitor cells resulting in cancers with basal and luminal lineages, resulting in divergent disease pathways[17].

Increasingly there is awareness that many malignancies share causative, pathophysiological and genetic features. Prostate cancer and breast cancer share a number of these characteristics. It is now recognised that breast cancer has multiple genetic phenotypes reflecting the tumour cell of origin and cell signalling pathways involved in disease progression and are generally characterised by differing patient outcomes[18–20]. However the biggest impact on survivorship has arguably resulted from advances in targeted adjuvant therapy derived from the identification of individual cell surface protein receptors on primary tumours. In operable breast cancer,

surgical excision with adjuvant chemotherapy and radiotherapy is performed to remove and 'sterilise' residual proliferating cancer cells. In addition adjuvant hormonal therapy is given to inhibit growth in oestrogen receptor responsive cancers, or the humanised antibody, trastuzumab, to inhibit growth and metastasis in patients with cancers expressing HER2[21].

There are similarities in the treatment approach used for prostate cancer, but the repertoire of therapeutic options is more restricted. The mainstay curative treatments for localised prostate cancer are surgical prostatectomy alone or radiotherapy combined with a period of chemical castration involving hormonal-based androgen deprivation therapy (ADT) whose primary purpose is to prevent testosterone-responsive growth in any residual viable tumour cells[22]. However, 20-30% of patients treated for localised prostate cancer will fail therapy and require long term ADT[23]. Unfortunately, castrate-refractory disease is essentially inevitable at some point along the disease pathway, associated with poor prognosis due to metastasis formation. Therefore, there remains an unmet need to inhibit metastasis formation, possibly resulting from circulating tumour cells[24] or the activation of dormant disseminated tumour cells (DTCs) present at the time of diagnosis[25]. Crucially, it has been proposed that the biology of DTCs is fundamentally different to the primary tumour supporting the need for characterisation of DTCs so that appropriate therapeutic approaches can be designed to successfully neutralise the threat posed by DTCs[26] . Novel combinative strategies that target the primary and disseminated tumour cells may be required to achieve significant improvement in treatment success.

As described above, treatment for prostate cancer with curative intent is described as ‘radical’ treatment. In the vast majority of cases this is either approached from a surgical perspective or through radiotherapy. The decision of which route to take is based on a multi-disciplinary team meeting, involving Urologists, Oncologists, Radiologists, Pathologists and associated support specialities. The meeting is an obligatory part of the clinical decision making process and will issue a recommendation that is subsequently discussed with the patient. The patient can then choose to accept the recommendation (sometimes various options are presented to them) or with the help of their medical team discuss other options that may not necessarily be considered ‘first line therapy’. If the patient has capacity and understands the risks of not following the initial recommendation then every effort is made to assist them with their choice of therapy.

Historically, the mainstay of radical prostate surgery has been the open prostatectomy. Most frequently this involved a lower midline incision, dissection of the prostate from the bladder and the bulbar urethra and then re-anastamosis of the bladder to the urethral stump. This procedure is still occasionally carried out in the UK, but has largely been superseded by laparoscopic radical prostatectomy (LRP) and more recently robot-assisted radical prostatectomy (RARP), both of which aim to achieve the same result, but are more ‘minimally-invasive’. While each evolution of the procedure has improved recovery time (for example a patient would likely remain in inpatient for 7 days post open surgery, but for only 1 day following RARP) and lessened surgical complications there are still risks associated with the approach. Even with extremely experienced surgeons patients should expect at least a 50% chance of erectile dysfunction (permanent) post-procedure, up to 12 months of varying degrees

of urinary incontinence (5% with permanent problems) and a 0.5% risk of serious rectal injury necessitating an emergency bowel stoma. In addition there are the normal surgical risks including, but not limited to, bleeding, infection, deep vein thrombosis and pulmonary embolus. Clearly surgical intervention should not be taken lightly.

Radical radiotherapy can either be delivered by an external beam approach, or via insertion of radioactive seeds or temporary radioactive probes known as 'brachytherapy'. While an advantage of radiotherapy over surgery is its lack of immediate significant complications, it certainly has inevitable consequences. Erectile dysfunction is delayed, but has a similar incidence to that of surgery. Incontinence is less frequent, but radiotherapy can trigger lower urinary tract symptoms such as urinary frequency and urgency and can exacerbate pre-existing symptoms. In addition radiotherapy can cause rectal irritation and inflammation leading to diarrhoea and rectal bleeding. It can also lead to urethral/bladder neck strictures (although so can surgery in certain circumstances). A major issue of radiotherapy is that it can trigger secondary malignancies, usually 10 to 20 years after treatment, and confers a relative risk of bladder cancer of approximately 1.3. In addition radiotherapy, particularly external beam, is highly time consuming and a typical regimen would involve coming to hospital every weekday for 6-7 weeks, which in younger patients who are still working is an extremely challenging logistical issue.

There is no convincing evidence of surgery or radiotherapy being superior when assessed against cancer survival. Most studies have been hard to interpret as historically older, less fit patients have been offered radiotherapy rather than surgery so would inevitably not do so well when compared with younger, fitter patients

undergoing surgery. One commonly used argument, particularly when counselling younger patients who would be anticipated to live for many years is that if they undergo surgery primarily and the disease returns, then radiotherapy is an option at that stage. If a patient has radiotherapy up front and the disease returns then 'salvage' surgery is extremely difficult as the tissues are very damaged from radiation exposure and are therefore very hard to operate on, with a much higher complication rate. Many surgical units do not offer salvage surgery, as this is highly sub-specialised.

2.1 Prostate epithelial cell lineage

Anatomically the prostate has a lobular structure, with lateral and anterior lobes. However, seminal work by McNeil described a zonal architecture, with each zone demonstrating different characteristics and propensity to develop cancers. The three main cellular populations are the luminal columnar epithelial cells that line the approximately 30 prostatic ducts, the basal epithelial cells on which the luminal cells rest and small numbers of neuroendocrine cells within the basal layer[27,28]. In addition, anteriorly is a mixture of smooth muscle and fibrous tissue, the fibromuscular stroma. There is accumulating evidence that malignant potential, disease aggression and prognosis may be determined by the subset of cells from which the cancer is derived.

The majority of malignancies are thought to develop in epithelial cells located in the peripheral zone, whereas the majority of benign prostatic hyperplasia develops in the transitional zone.

There has been considerable debate as to the cell lineage pathways of prostate epithelium. It is becoming apparent that the basal compartment contains a pool of

multipotent stem cells[29–32] that are capable of differentiation into basal and secretory luminal epithelium. These different cellular subtypes can be identified through discrete expression patterns of certain cell surface proteins. For example luminal cells commonly express cytokines (CKs) 8 and 18, whereas basal cells express CK 5 and 14[28]. However, further work has demonstrated an 'intermediate' cell type that co-express these markers along with others such as CD24[33,35]. This intermediate population is believed to represent a transition or amplification stage in the progression from multipotent stem cells to more differentiated basal and luminal epithelium[33].

Further work has shown that human basal cells *in vivo* can be triggered to develop prostate cancer when exposed to common gene mutations[34]. This evidence fits with a hypothesis that stem cells are highly likely to be the origin of CaP as they have an inherent ability to self-renew, and their subsequent longevity provides sufficient time for repeated genetic mutations to finally trigger carcinogenesis.

However, the question at which point in the cellular differentiation pathway CaP is initiated remains. Evidence is growing that there are also populations of luminal cells that retain some stem-cell like qualities, perhaps because they are still 'early' in the differentiation phase[35], or because they derive from an entirely separate stem cell population[36,37]. The lineage of prostate epithelial cell development has certainly not yet been fully mapped, and as a result the exact cell, or cells, of origin of prostate cancer remain uncertain. Another unanswered question is whether the cell of origin determines tumour aggression, metastatic potential and likelihood of developing

castrate resistance as it has recently been proposed that selective clonal stem cell expansion is associated with prostate cancer aggressiveness[38].

2.2 Circulating tumour cells

Isolation of circulating prostate cancer cells has been shown to be prognostic in prostate cancer[24,39]. Also, it has been shown that the activation of dormant disseminated tumour cells (DTCs) and consequential metastasis involves a balance between three opposing processes: cellular dormancy (mitotic arrest); angiogenic dormancy (vascular-delivered nutrient restriction); and immune-mediated dormancy resulting from immune system cytotoxicity[40]. Procedures exist for the isolation of DTCs[41] and biomarkers have been proposed for assessment of their functional state.

Interestingly, given that part of this project has been to study the association between stem cell markers and cancer, some of proteins expressed in stem cell populations have been implicated in DTC function and control. For example, HER2 is commonly expressed in DTCs of various cancers, in particular those displaying a stem-cell like phenotype[42]. Further work has shown that prostate cancer DTCs have the ability to replace haematopoietic stem cells of the bone marrow stem cell 'niche'[43] and that this environment can promote cellular dormancy of the DTCs, rendering them less sensitive to taxotere chemotherapy and thus enable them to 'weather the storm' of treatment and emerge to seed metastases as some point after chemotherapy has finished. These 'dormancy enriched' DTCs have recently been found to express lower than usual levels of Ki67[44], another marker we have examined in this project. Also, evidence has emerged that DTCs are likely to differ from their primary tumour at both the genomic and gene expression level, leading to differences seen in the marker

profile between primary tumour and metastatic tissue and indeed between different sites of metastasis[45]

Circulating tumour cells may offer the possibility of a 'liquid' biopsy, which could potentially obviate the need for conventional trans-rectal or template tissue biopsies and avoid their attendant risks. Currently there are many studies evaluating the utility of using CTCs in a diagnostic role, but this technique is certainly not being used in clinical practice at present. CTCs are perhaps more likely to be used as part of a predictive/prognostic array to help stratify an individual patient's risk of progression of prostate cancer, such as developing metastases.

2.3 Molecular classification of breast cancer and the parallels to prostate cancer

Prostate cancer and breast cancer share a number of characteristics. For this reason it can be postulated that they might share similar approaches for investigating disease pathways and identifying biomarkers for clinical management. Both prostate and breast cancers are hormonally manipulated, the stromal microenvironment plays an integral role in each and they are more common in the presence of certain gene mutations such as BRCA1 and BRCA2[46,47]. Clearly, it is important to review the evidence for the possible existence of different molecular phenotypes in prostate cancer to assess if classification could lead to similar risk profiling and specific targeted therapies utilised in breast cancer.

It is now recognised that breast cancer has multiple genetic phenotypes that were initially identified by gene expression profiling (GEP)[48]. This technique was used to identify differentially expressed intrinsic genes in breast cancer and subsequent

hierarchical clustering models to define four molecular classes: normal breast, luminal (oestrogen receptor (ER) positive), basal-like and HER2 (epithelial growth factor receptor 2; ERBB2 gene/neu). Subsequent work demonstrated further subtypes such as luminal A and B and claudin-low[49].

An important aspect of this work was the association between molecular subtype and cancer specific survival, allowing the development of risk assessment and targeted therapy based on gene/protein expression profiling. These were the first steps towards personalised cancer treatment, which is now accepted as the gold standard in oncology.

2.3.1 Luminal-like prostate cancer

Similar to breast cancer, over the last decade many studies concluded that CaP derives mainly from terminally differentiated luminal cells, based on the observation that the majority of cancer specimens stained negative for basal cell markers and the cell surface protein p63[50,51]. However, other studies suggest that CaP is by no means a homogenous entity. Although the absence of P63 is used as an adjunct in the histological classification of CaP[52], it is occasionally expressed in prostate cancer tissue, with higher rates of expression in tissue with higher Gleason scores[51].

Androgen receptor (AR) signalling is critical in the development of normal prostate tissue and is analogous to the ER in breast tissue. Like the ER in breast cancer, the androgen receptor also plays a key role in mediating the various stages of prostate cancer and subsequent castrate resistance. AR ‘promiscuity’ is likely to contribute to

this process by triggering transcriptional activation in response to antiandrogens or other endogenous hormones[53].

Development of castrate resistance in advanced prostate cancer is associated with poor clinical outcome. Identifying which patients will succumb is currently a key research objective to aid clinical management and identify novel targets for therapy. The AR receptor and related genes are also implicated in the durability of ADT treatment. Fujimura *et al* proposed two panels of gene expression markers for determining clinical failure (defined by PSA recurrence) and cancer specific survival in treatment naïve prostate cancer patients with bone metastasis[54]. They found expression of *Sox2, Her2* and *CRP* in cancer cells to be predictive of clinical failure; panels comprising *Oct1, TRIM36, Sox2* and *c-Myc* AR, Klf4 and ERα were found to be prognostic of survival in cancer and stromal cells respectively.

2.3.2 Basal-like prostate cancer

In breast cancer, the basal phenotype has been shown to be associated with more aggressive disease, poor patient outcomes and as yet has no specific targeted treatment[55,56]. The basal phenotype is commonly defined by a lack of expression of oestrogen receptor, progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) and has been referred to in many papers as the 'triple negative' phenotype. However evidence now suggests that the basal and triple negative types may actually be two distinct groups, albeit with similar poor clinical outcomes[57,58], although the two terms are still used interchangeably in many recent studies. In prostate cancer, the steroid nuclear androgen receptor is expressed in luminal, basal and stromal cells, but importantly its regulatory function varies with each population.

It enhances cell survival in luminal cells, stimulates proliferation and metastases in stromal cells and suppresses proliferation and metastasis in basal cells respectively[59].

There is strong evidence demonstrating the basal phenotype as a cell of origin for some prostate cancers. Recent work has suggested that genetic signatures commonly associated with embryonic stem cells (ESCs) are up regulated in the tumours of patients with more poorly differentiated prostate cancers[60,61]. Interestingly, the same characteristic ESC signature has been found in high grade breast cancers, particularly the basal subtype[62]. ESC+ prostate cancers have been associated with higher Gleason scores and poorer prognosis than those not expressing this signature. The population of phenotypically positive prostate stem cells (PPSC) was higher in metastatic bone cancer compared to the primary prostate cancer[63]. Columbel et al suggest using the putative stem cell markers integrin alpha-2 or -6 in combination with c-met and a 5% cutoff threshold to predict reduced survival associated with bone metastasis. Interestingly, these markers appear to be confined to stem cells localised in the basal cell layer of normal and benign prostate hyperplasia tissue[64].

Contradicting the existence of a pure basal class of prostate cancer is the observation that basal prostate cancer cells tend to lose their basal-defining cell marker characteristics and transform into a more luminal phenotype. However, although appearing histologically homogenous, prostate cancers still maintain lineage-specific genetic signatures. In contrast to breast cancer, basal cell derived cancers appear to be a rarer event and may on occasion have a better prognosis than their luminal cell derived counterparts[17]. But, identifying the legacy of basal-transformed cells presents difficulties and limits its clinical usefulness.

2.3.3 HER2 prostate cancers

Given the similarities between prostate and breast cancer it is unsurprising that the HER2 oncogene has demonstrated an association with outcome in CaP. HER2 overexpression has been found in approximately 20% of localised, untreated prostate cancers, and this rises to over 60% in metastatic disease and those cancers treated with ADT, although there is significant variation between studies, based on definition of 'overexpression' and also the assay used [65]. Increased expression of HER2 in prostate cancer has been associated with higher Gleason grade, cancer stage and rate of proliferation (as demonstrated by the Ki67 index)[66] and also poorer outcome than those tumours that express lower levels[67]. However, anti HER2 antibodies such as trastuzumab (Herceptin) that have proven extremely effective in HER2-positive breast cancer have not shown any clinical efficacy in prostate cancer. Interestingly trastuzumab is most effective in breast cancers in which HER2 overexpression is mediated by gene amplification. In prostate cancer, while HER2 expression is upregulated, gene amplification is uncommon, and thus the target may not be as important in this disease[68].

In summary, stratification of prostate cancer based on similar principles to that used for the molecular classification of breast cancer may be conceptually possible for the luminal and basal classes, but they do not represent the full heterogeneity seen in prostate cancer disease and its progression. Based on current academic knowledge and the development of breast cancer therapy, future clinical management of prostate cancer is going to require an individualised approach built on assessment of cell signalling biomarkers that inform about cell functional activity.

2.4 Patient cohort and clinical context

The identification of molecular pathways and their significance in prostate cancer is made difficult due to the clinical course of the disease. CaP has a long latent phase where the patient is entirely asymptomatic and even patients who present late with metastases at diagnosis have an extended life expectancy (for example the median survival for patients with metastases in the recent Stampede study was 42 months[69]).

Patients diagnosed with organ confined disease can expect to be cured in the majority of cases, be it surgically or using radiotherapy, and those with disease recurrence after attempted curative treatment are likely to live many years and indeed may die of another condition before their prostate cancer becomes an issue. The indolent course of disease progression means that interventions may not demonstrate efficacy for many years after they are used as it often takes ten to fifteen years for survival curves to start to separate, as has been demonstrated in many seminal papers looking at cancer outcomes post radical treatment[70]. This then leads to further difficulties with interpretation of the clinical relevance of these results as often in the intervening years clinical practice has changed and they are no longer carried out in the same way.

The patient cohort examined in the study reported upon herein is unusual in that it has an extended period of clinical follow up – over 13 years – which makes it a powerful tool in the investigation of prostate cancer, particularly using high throughput immunohistochemical techniques that can interrogate multiple protein biomarkers in a relatively short period of time. The presence or absence of these markers can be

recorded, their levels of expression quantified and then associations with subsequent clinical outcome looked for.

Patients in our research cohort were diagnosed with CaP between 1999 and 2001 and were being managed according to local best practice at that time. Patients were asked if they would allow their tissue to be used for research purposes, and if in agreement were included in the study. These were consecutive, non-selected patients. Initial diagnosis was made from tissue taken from either prostate needle core biopsy or trans-urethral resection of prostate (TURP) specimens. More detail can be found in section 3.2.

The majority of patients in this cohort were diagnosed with prostate cancer from tissue taken during TURP, a surgical procedure that is done to relieve urinary obstruction by an enlarged prostate. This procedure is not intended to detect or treat cancer and does not reflect modern diagnostic practice. In the last 10 to 15 years the use of opportunistic PSA screening has dramatically increased the number of men investigated for prostate cancer and the standard of care was a trans-rectal biopsy of the prostate followed by staging investigations (such as MRI pelvis and bone scan) if cancer was detected. More recently the diagnostic pathway has begun to shift again, with increasing use of multi-parametric MRI and targeted trans-perineal biopsies of the prostate. This has led to a change in the proportion of patients diagnosed with various stages of prostate cancer. In the last decade the majority of prostate cancer was initially detected in its very early stages (i.e. organ confined).

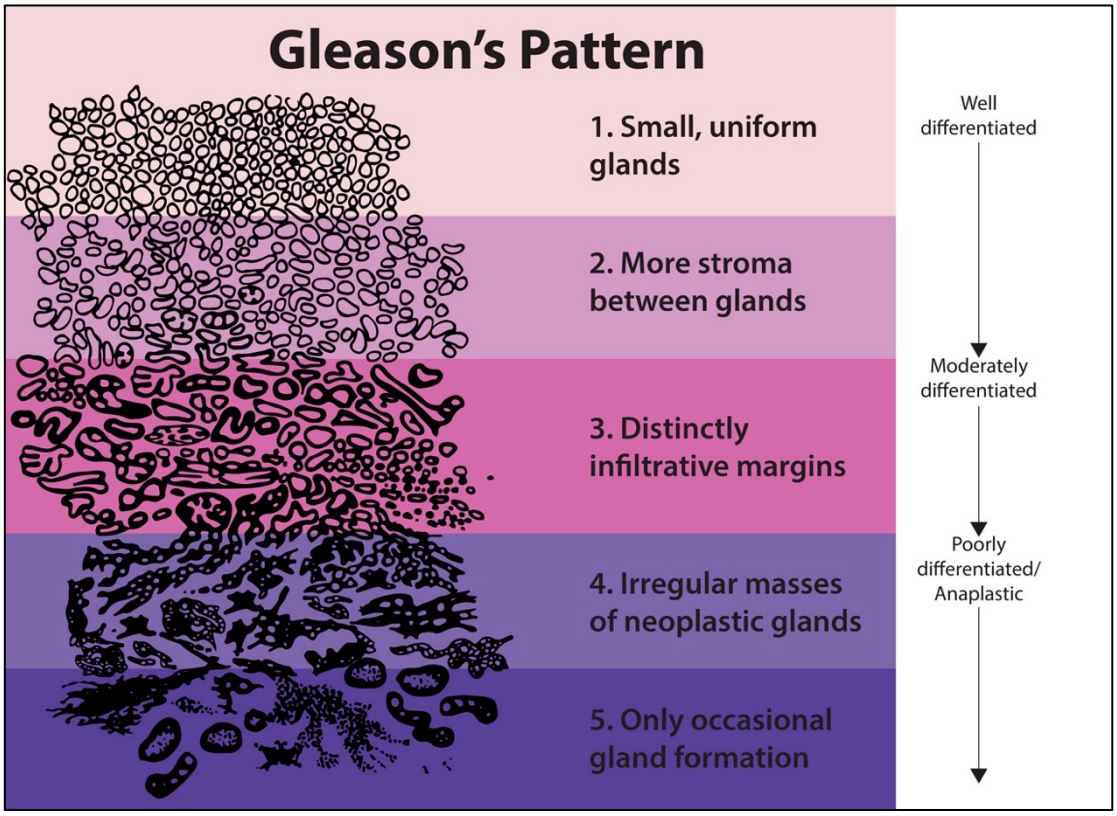
Multi-parametric MRI has been shown to be extremely accurate at predicting the presence of high risk and intermediate risk (Gleason primary pattern 4) disease, and increasingly is being used to direct prostate biopsies and in the recent PROMIS trial demonstrated a corresponding increase in the diagnostic accuracy of targeted biopsies[4]. While MRI certainly provides excellent staging information and can be useful in operative planning, and may have a role in preventing unnecessary biopsies it is not yet accurate enough to replace biopsies when a diagnosis is required and should therefore be viewed as an adjunct to diagnosis rather than as a stand-alone tool.

This shift in detection has led to criticisms of many studies carried out in the 'pre-screen detected' era on the basis that they are no longer reflective of current practice. While the proportion of patients now diagnosed with CaP by TURP is now small they do however represent a group of patients that need addressing. Current European Association of Urology guidelines suggest treating TURP-detected CaP in a similar fashion to biopsy-detected cancer and as most are asymptomatic and have a low PSA and Gleason grade they are likely to be managed by active surveillance – i.e. have no treatment unless the clinical situation changes. This means our cohort is still relevant as the majority of our patients were only treated with hormones when they became symptomatic.

As described earlier, a key factor in assessing the significance of a patient's prostate cancer is the Gleason score (Figure 1). The Gleason score is based on a histological examination of prostate tissue. The most common architectural pattern is identified and graded on a scale of 1 to 5 and then the second most common architectural pattern is similarly graded. The two scores are then added together giving a total out of 10.

While the fundamentals of this classification remain the same there have been some changes in the way cells are assigned to a particular grade category and this has implications for our cohort. Our patients' tissue was collected between 1999 and 2002 and was graded based on the Gleason score at the time. However, in 2005 the International Society of Urological Pathology (ISUP; Tables 1 and 2) made some changes that particularly affected low and intermediate grading[71]. For example, modifications included taking into account grading of variants of prostate carcinoma and unusual morphologies, such as ductal adenocarcinoma, considered as Gleason grade 4 (GS 8 if pure), pseudohyperplastic variant, graded as Gleason score 3+3=6 and mucinous fibroplasia which is subtracted and gland graded (mostly Grade 3).

Figure 1: An example of the histological patterns used to assign Gleason Score (Image taken from 'Commons.WikiMedia.org' with permission to share and reproduce)



Interestingly, very recently the ISUP have altered the system again with slight variations as to how cell architecture should be classified, but more importantly a significant difference in how the overall Gleason score is presented[72] (Table 3). For some years, Gleason score 6 has been the lowest total score that is considered a true prostate cancer and is therefore the least significant and is unlikely to cause patient morbidity and mortality. However, there was a gradual appreciation that patients were unduly worried by this system, as they perceived their score was 6/10 and was therefore a potentially dangerous cancer. The new system assigns the various Gleason

Score combinations into 5 grade categories from 1 to 5, with 1 the least significant and 5 the worst (Table 3).

Table 1: Changes in Gleason patterns proposed at the ISUP meeting 2005[73]

Gleason Pattern	**ISUP modified Gleason grading changes compared to conventional Gleason grading**
1	A Gleason score of 1+1=2 should not be diagnosed regardless of type of specimen. Extremely rare exceptions
2	A Gleason score of 2+2=4 should be diagnosed rarely, if ever. Glands should not infiltrate between non-neoplastic prostate acini
3	Typical pattern 3 consists of circumscribed, variably sized but often small individual glands that may infiltrate among non-neoplastic acini
4	3 main variants of pattern 4: • Most cribriform patterns should be pattern 4. Subtle features such as slight irregularity of the outer border of glands should be sufficient to move the glands from pattern 3 to pattern 4. • Fused Glands • Incomplete or poorly formed glands
5	Comedocarcinoma with central necrosis

Table 2: Changes in the reporting of Gleason grading proposed at the ISUP meeting 2005[74]

Specimen	**ISUP modified Gleason grading changes to conventional Gleason Grading**
Biopsies	The Gleason score of tumours on biopsy with tertiary higher grade should include the tertiary pattern and not be listed with primary and secondary patterns with a note relating to tertiary pattern
Biopsies	A Gleason score should be reported for each individual core or container
Biopsies	The highest grade would typically be the one selected by the clinician as the grade of the entire case. One also has the option to give an overall score at the end of the case, in addition to individual scores
Prostatectomy Specimen	A separate Gleason score should be assigned to each dominant tumour nodule
Prostatectomy Specimen	Tertiary patterns of higher grade should not be included in the Gleason score but rather be mentioned separately

Table 3: Changes in the reporting of Gleason grade groups proposed at the ISUP meeting 2014[72]

Grade Group 1 (3+3)	Only individual discrete well-formed glands
Grade Group 2 (3+4)	Predominantly well-formed glands with lesser component of poorly- formed/fused/cribriform glands
Grade Group 3 (4+3)	Predominantly poorly formed/fused/cribriform glands with lesser component of well-formed glands
Grade Group 4 (4+4 / 3+5 / 5+3)	Only poorly-formed/fused/cribriform glands or Predominantly mix of well-formed and lack of glands
Grade Group 5 (4+5 / 5+4 / 5+5)	Lack gland formation (or with necrosis) with or w/o poorly formed/fused/cribriform glands

As can be seen below in the results section, our main analysis was based on the pre-ISUP 2005 Gleason classification. After discussion with the editorial office at the British Journal of Cancer (where our data relating to the markers Ki67 and DLX2 have been published[75]) it was felt to be more clinically relevant to re-score our cohort using ISUP 2005, as this is the most widespread classification used in contemporary day-to-day clinical practice. We therefore reviewed our histological specimens and reassessed the relevant data. For comparison we include pre- and post-ISUP 2005 results for the markers Ki67 and DLX2.

2.5 Artificial Neural Networks

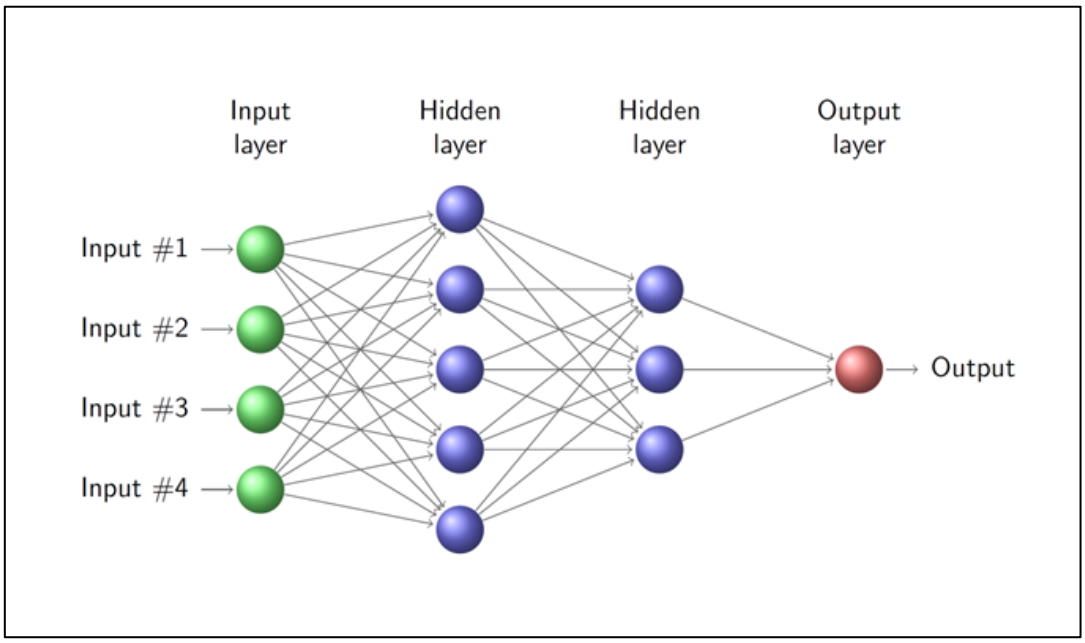
Artificial neural networks (ANNs) are a form of machine learning from the field of artificial intelligence with proven pattern recognition capabilities and have been utilized in many areas of bioinformatics[76]. ANNs have been widely employed in the identification and stratification of molecular biomarkers[77,78] and are inspired by the way organic systems learn and processes complex information. ANNs have the ability to handle complex (non-linear) features within data in order to generalise and predict outcomes accurately in future cases[76].

ANNs are essentially information processing tools that utilise methods observed in biological systems such as the human brain. The networks are made up of a large number of interconnected processing elements that work in parallel to produce an output, based on input data chosen by the system operator (Figure 2). These processing elements make decisions based on a simple set of rules, and the outputs contribute to the rest of the network elements which then make further decisions with the entire system acting in a gestalt way allowing extremely complex data analysis and pattern recognition.

In general, ANNs utilise 3 layers: an input layer, a series of 'hidden' interconnected nodes and an output layer. The difference between conventional computational algorithms and ANN systems is that ANNs have the ability to self-correct if the outcome is not sufficiently close to that required by the user, i.e. they can learn. This learning process actually takes place by altering the strength of connection between certain nodes involved in the computational pathway, so if a connection leads to a

correct result then the strength or likelihood of this connection being used again increases.

Figure 2: Diagrammatic representation of the concepts of an ANN (image taken from www.texample.net with permission)



ANNs are now used widely in the field of medical bioinfomatics and have been used successfully in the field of prostate cancer research over the last 10 to 20 years, although not on a large scale. For example, ANN techniques were shown to accurately predict trans-rectal prostate cancer biopsy outcome using a 'feed-forward' network with the input variables 'PSA, DRE, age, and percentage of free PSA' as early as 2003[79]. However, over the last few years a growing realisation and awareness of the technique has led to increased interest, not least because of the vast complexity of data now available for analysis due to high throughput techniques such as gene expression arrays and the desire to find reliable biomarkers that can accurately stage

the disease, with prognostic utility, without necessarily having to carry out invasive procedures[80].

2.6 Aims and objectives

The aim of this research project was to identify a panel of prostate cancer associated protein biomarkers that can reliably predict patient outcome at the point of diagnosis. Outcome in this context is the likelihood of a patient subsequently developing metastatic disease or dying of prostate cancer. This will allow patients to be more accurately 'risk stratified' into groups, subsequently helping to inform and guide treatment decisions. This will improve outcomes in those that need radical intervention and minimise the number of patients undergoing unnecessary treatment, thereby avoiding inherent complications.

Candidate biomarkers were identified using two approaches. The first was the use of bioinfomatic techniques, in particular artificial neural network analysis of publically available CaP gene expression arrays, carried out by Professor Graham Ball (Nottingham Trent University). The second was based on a comprehensive review of the current literature examining the links between the molecular classification of prostate cancer and subsequent clinical outcomes.

A tissue microarray (TMA) has already been constructed from archival wax-embedded prostate cancer taken from 365 patients (fully consented and ethically approved) between 1999 and 2002. A major component of this project was populating a database of patient outcomes over the following 12 years for this cohort. This involved reviewing the medical records of every patient and recording pathological,

treatment and outcome data. This data allowed eventual comparison of biomarker panels with clinical risk groupings.

Statistical analysis was carried out to look for associations between biomarker expression and clinico-pathological variables and their ability to predict prostate specific survival, tumour recurrence, metastasis development, and treatment failure.

3.0 Materials and Methods

3.1 Overview

Candidate biomarkers associated with metastases and outcome in prostate cancer were selected using the techniques described in detail below. A tissue microarray (TMA) of patients diagnosed with CaP between 1999 and 2001 has recently been created by Dr Des Powe (DGP), Histopathology Department, NUH, comprising tumour and adjacent samples of tissue from each patient. These were then immunostained for protein expression of each marker, with markers being selected using a bioinformatic approach and additional literature search. Staining was quantified and independently scored by WG & DGP, with levels of agreement checked to ensure consistent results.

The North West 7 Research Ethics Committee approved use of the tissue samples for this study – Greater Manchester Central REC number 10/H1008/72.

Patient outcome data was collated in a spreadsheet (Microsoft Excel) with multiple parameters recorded, as described below. This data was used to determine the clinical progression of the disease and drive subsequent statistical analysis and risk stratification.

3.2 Patient cohort and data collection

365 Patients diagnosed with CaP between 1999 and 2001 were incorporated into the TMA. These were consecutive non-selected patients and all underwent contemporary 'best-practice' treatment at Nottingham City Hospital, UK. Initial histological cancer diagnosis was made using tissue obtained by prostate needle core biopsy, transurethral

resection of prostate (TURP) or radical prostatectomy (RP) specimens and cohort patient characteristics can be seen in Table 4 below.

Multiple data points were recorded for each patient (Table 5), after initial discussions between clinicians and research scientists to determine which factors were required to demonstrate variation in patient demographics, initial histological diagnosis, treatment modalities and clinical outcome.

Table 4: Clinical characteristics of prostate cancer patients incorporated in the TMA, including method of tissue extraction, PSA at diagnosis, Gleason score, D'Amico score, androgen deprivation treatment status, metastatic status and mortality.

Clinical Variable		**Number of patients (%)**
Surgical procedure	TURP	279 (76%)
	Prostatectomy	26 (7%)
	Biopsy	54 (15%)
	Not recorded	6 (2%)
PSA (ng/ml) at diagnosis	<4	34 (9%)
	>4	237 (65%)
	Not recorded	94 (26%)
Gleason Score	≤7	141 (39%)
	≥8	156 (43%)
	Not recorded	68 (19%)
D'Amico risk	Low	54 (15%)
	Intermediate	33 (9%)
	High	202 (55%)
	Unclear	76 (21%)
Antigen deprivation therapy	Yes	197 (54%)
	No	87 (24%)
	Unclear	81 (22%)
Castration resistant	Yes	127 (65% of those rendered castrate)
	No	70 (35% of those rendered castrate)
Metastasis at diagnosis	Yes	44 (12%)
	No	257 (70%)
	Not recorded	64 (18%)
Subsequent metastases (in those patients without metastases at diagnosis)	Yes	85 (23%)
	No	167 (46%)
	Not recorded	69 (19%)
Death due to prostate cancer	Yes	134 (37%)
	No	92 (25%)
	Unknown	87 (24%)
	Still alive	52 (14%)

Table 5: Patient and clinical factors recorded in the patient database

Racial origin	Date of Diagnosis	Initial PSA	Initial DRE	Use of 5aRI	Mode of tissue diagnosis	Gleason1	Gleason2	Overall Gleason score	Initial D'Amico risk
Alternate histology	Extra prostatic invasion	Perineural invasion	Vascular invasion	Nodes negative	Negative resection margins	% Cancer	Bone mets at diagnosis	Radical prostatectomy histology	Active surveillance
Radiotherapy (DXT)	Date of DXT	Recurrence after DXT	Subsequent mets	Date of mets	Location of mets	Months to mets from diagnosis	Chemo	1st Androgen deprivation therapy (ADT)	Date of 1st ADT
2nd ADT	Date of 2nd ADT	3rd ADT	Date of 3rd ADT	Death	Date of death	CaP related death	Months to all cause death	Months to castrate resistance	Months to CaP specific death
Other cause of death	B-Blocker	Date of initiation of B-blocker	Comments	PSA surveillance values	Dates of PSA				

Patient data were obtained from a combination of review of paper-based medical notes, and Nottingham University Hospitals' computer-based results system 'NOTIS'. All data were recorded in a spreadsheet (Microsoft Excel) and all patient identifiers were removed other than a derived 'patient analysis number' used to ensure that multiple TMA cores taken from a single patient could be matched to the corresponding clinical data.

3.3 Tissue microarray construction and layout

A TMA was constructed using archival wax-embedded TURP and radical prostatectomy samples sourced via the Nottingham Health Science BioBank. Histology sections were reviewed by a pathologist (Dr Geoffrey Hulman, GH) and 0.6mm diameter donor cores were sampled from at least two different tumour regions per patient using an automated TMA Grand Master instrument (3DHistech Ltd, Hungary) and placed in paraffin blocks. Each block accommodates 100 cores, with the majority of patients being represented by at least 2 cores, and some up to 6 cores. Each block was arranged in a grid pattern from A1 to J10, and the blocks were serially numbered. Each core was assigned a unique 'NPN' number (starting at '2000'

for core A1 in the first block and continuing in sequence to 2968 for core E10 in the final block) that allows subsequent biomarker scores to be linked to a particular tissue sample from a particular patient, without compromising patient personal data. All TMA tissue sections for IHC were cut at a thickness of 4μm on a microtome. The pathology of constituent cores in the finished TMA was confirmed by GH.

3.4 Biomarker selection

3.4.1 Artificial neural network analysis

Artificial neural analysis of publically available prostate cancer cDNA gene expression arrays (U133A – data source GSE8218, U133Plus2.0 – data source GSE17951 and U95Av2 – data source GSE1431)[81] was carried out by Professor Graham Ball (Nottingham Trent University), looking at markers that confer an increased risk of progression and metastases. The ANN model was reiterated 50 times with random sampling and the average mean square error of a test subset for each input variable was considered to determine the predictive capability for metastasis class.

The top 10 genes ranked for association with metastasis development are included in Table 6. Four of these had commercially available antibodies (AMACR (Racemase), DLX2, PAICS and MYO6) directed against proteins and were validated using immunohistochemical staining of the Nottingham prostate TMA.

3.4.2 Literature search

For comparison, a curated literature search was performed by Mr Will Green and Dr Des Powe to identify biomarkers previously proposed for predicting disease-specific

survival (DSS). DSS is clearly an outcome measure of extreme importance when investigating cancer biomarkers and we felt it was important to include this as one of the key aims of the study was to identify biomarkers with practical clinical utility. The search was performed using 'Web of Science' and the following terms:

1) 'Review', 'genomic studies', 'prostate cancer' – this search returned 167 results, of which 19 papers were appropriate for full review.
2) 'Prostate cancer', 'transcriptome', 'classification' – 2 results, both fully reviewed.
3) 'Prostate cancer', 'cluster analysis', 'stratification' – 8 results, of which 2 papers were appropriate for full review
4) 'Prostate cancer', 'cluster analysis' – 955 results, of which 22 papers were appropriate for full review.

The literature search highlighted 5 further candidate biomarkers that were reported to be associated with prostate cancer associated survival. These included the proliferation marker Ki67, the tumour suppressor gene p53, the basal cell marker CK5/6, the proposed stem cell-like markers C-MET and integrin alpha2 and the gene fusion product TMPRSS-ERG. Table 6 shows the antibodies used to stain the Nottingham prostate TMA.

Table 6: ANN ranked gene list showing association with prostate cancer metastasis

Rank	Gene accession number	Gene name	Gene Product (protein)
1	AK022765.1	AMACR	Alpha-methylacyl-CoA racemase
2	AI796120	AMACR	Alpha-methylacyl-CoA racemase
3	AF047020.1	AMACR	Alpha-methylacyl-CoA racemase
4	NM_004405.2	DLX2	Distal-less homeo box 2
5		PCA3	Prostate cancer antigen 3
6	NM_012485.1	HMMR	Hyaluronan-mediated motility receptor (RHAMM or CD168)
7	U90236.2	MYO6	Myosin VI
8	NM_017636.1	FLJ20041	Hypothetical protein FLJ20041 Alias: TRPM4B
9	BF511718	RHO7	GTP-binding protein Rho7 Alias: RND2
11	NM_006452.1	ADE2H1	Multifunctional polypeptide similar to SAICAR synthetase and AIR carboxylase Alias: PAICS
12	NM_002570.1	PACE4	Paired basic amino acid cleaving system
13	NM_004503.1	HOXC6	Homeo box C6

Table 7: Antibodies used for immunohistochemical staining of the Nottingham prostate TMA. (Microwave antigen retrieval - MAR)

Gene name	Gene Product (protein)	Manufacturer/clone/+ve Control	Dilution/MAR
AMACR	Alpha-methylacyl-CoA racemase	Dako M3616 Clone 13H4 Control – Human Breast Tissue	1:200/Citrate
DLX2	Distal-less homeo box 2	Abcam Ab18188 Polyclonal Control – Mouse Brain	1:1500/Citrate
MYO6	Myosin VI	Abcam ab170522 Polyclonal Control – Human Breast Tissue	1:300/Citrate
PAICS/ADE2H	Multifunctional polypeptide similar to SAICAR synthetase and AIR carboxylase Alias: PAICS	Abcam ab174685 Polyclonal Control – Human Breast Tissue	1:1000/Citrate
Ki67	Proliferation marker	Leica NCL-L-MM1 Clone MM1 Control – Colorectal Tissue	1:25/Citrate
P53	Tumour suppressor gene	Leica ICL-L-p53-DO Clone DO-7 Control – Human Breast Tissue	1:50/Citrate
CK5/6	Basal-like epithelial cell marker	Dako M7237 Clone D5/16 B4 Control – Tonsillar Tissue	1:50/Citrate
C-MET	Receptor for hepatocyte growth factor	Abcam ab51067 Clone EP1454Y Control – Human Breast Tissue	1:400/Citrate
Integrin alpha2	Cell adhesion molecule	Abcam Ab133557 Clone EPR5788 Control – Human Breast Tissue	1:300/Citrate

3.5 Immunohistochemistry (IHC)

Optimal antibody dilutions and antigen retrieval conditions were initially performed using positive and negative control tissues at dilutions suggested by the antibody suppliers. Positive and negative samples were then examined by Dr Des Powe (DP) and Mr Will Green (WG) and an assessment made of whether it was possible to clearly discriminate between the control samples at these dilutions. In samples that

were not clearly demarcated the dilutions were titrated as appropriate. 4μm sections from each TMA block were mounted on Dako REALTM Capillary Gap Microscope Slides, 75 μm (Grey), S2024 (Dako UK Ltd.). After initial optimisation each set of TMA sections representing the entire clinical cohort underwent IHC according to the protocol developed by Dr TMA Abdel-Fatah[82] (Clinical Oncology Department, Nottingham City Hospital), as described below.

The sections were dewaxed by being heated at 60°C and then sequentially immersed in xylene, IMS and dH20 using an autostainer (Leica Autostainer XL). The sections then underwent microwave antigen retrieval for 20 minutes in an antigen retrieval buffer (10mM solution of sodium citrate, pH6). Slides were then loaded into a sequenza reservoir and underwent exposure to serial reagents using a Novolink kit. First, they were exposed to a peroxidase block for 5 minutes, then a protein block for 5 minutes. They were then incubated for 1 hour at room temperature with the relevant antibody in each experiment (see Table 7 for antibodies and concentrations). After incubation, slides were then exposed to the post primary block, a polymer, DAB working solution and finally haematoxylin. Slides were washed with Tris-buffered saline (TBS) between each stage. Slides were then dehydrated in the autostainer using industrial methylated spirits IMS and xylene and mounted in DPX mounting medium.

Immunostained TMA sections were assessed to determine the appropriate scoring technique for quantifying protein expression levels. Sections were independently scored (WG, DGP) without knowledge of pathology grade. Staining thresholds used for dichotomous categorisation were chosen using the software program X-tile, or by those given in previously published studies.

H-scoring was used for Ki67 (>110 = positive), P53 nuclear (>90 = positive), AMACR (>30 = positive) and DLX2 (>10 = positive). Categorical scoring was used for PAICS and Integrin alpha (0=absent, 1=weak, 2=strong; 2 considered positive), CMET (0=absent, 1=weak, 2 strong; 1 and 2 considered positive), MYO6 (0=absent, 1=weak, 2=moderate, 3=strong; 2 and 3 considered positive). Presence/absence was used for P53 cytoplasmic, CMET membranous, Integrin alpha membranous, and CK 5/6.

We used REMARK (REporting recommendations for tumour MARKer prognostic studies) guidelines for reporting on prognostic biomarkers in the whole patient series. REMARK guidelines are an internationally recognised set of parameters developed to ensure a robust standard of study design, pre-planned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods[83].

The proportion of patients with scorable tissue cores was less than the total number of patients originally incorporated in the TMA due to detachment of cores during processing and insufficient cancer tissue. Missing data were assessed for randomness using a Little's test and Wilcoxon-Mann-Whitney test, both at 95% confidence level. We failed to reject the null hypothesis of data being missing completely at random ($p>0.05$). Table 8 demonstrates the number of patients that could be scored each individual marker. No marker had a score reflecting the whole patient cohort as some core samples were lost in processing and others contained no cancer cells in the microtome section analysed.

Table 8: The number of patients within the prostate cancer cohort that were dichotomously categorised for each biomarker

Biomarker	Number of cancer patients scored	Percentage positive	Percentage negative
Ki67	182	6.6	93.4
P53 Nuclear	223	10.3	89.7
PAICS	175	80.6	19.4
MYO 6	219	48.4	51.6
CMET cytoplasmic	193	80.8	19.2
INTa cytoplasmic	203	34.0	66.0
Racemase	214	72.9	27.1
CK 5/6	221	28.5	71.5
DLX2	209	72.7	27.3

3.6 Statistical analysis

Statistical analysis was performed using SPSS (Version 21; IBM, US) applied to verified cancer samples. Pearson Chi-square tests were performed to assess biomarker associations with clinicopathological variables including initial PSA, Gleason score and initial risk (D'Amico). Kaplan-Meier plots with log-rank tests were used to model biomarker associations with disease-specific survival (DSS), time to metastasis development and time to castrate resistance. Biomarkers that showed an association with DSS or metastasis were included in a multivariate Cox proportional hazards regression model to assess the additional prognostic value to the PSA and Gleason score. The significance level used was $P<0.05$.

4.0 Results – Initial analysis of all markers with pre ISUP 2005 histology

Univariate associations between candidate biomarkers and clinicopathological variables are shown in Tables 9-12. Multivariate Kaplan-Meier models showing the associations between biomarkers with DSS, time to metastasis and castrate resistance are shown in Table 13. Example biomarker staining patterns are shown in Figure 5.

Table 9 Association of Racemase, DLX2 and MYO6 with clinical pathology variables. Significant p-values are shown in bold.

		Racemase negative	**Racemase positive**		**DLX2 negative**	**DLX2 positive**		**MYO6 negative**	**MYO6 positive**	
Clinical Variable		Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)
PSA (ng/ml) at diagnosis	<4	3 (27.3%)	8 (72.7%)	0.007 (p=0.93)	2 (20%)	8 (80%)	0.458 (p=0.499)	8 (66.7%)	4 (33.3%)	1.649 (p=0.969)
	>4	40 (26.1%)	113 (73%)		132 (30.1%)	187 (69.9%)		74 (47.4%)	82 (52.6%)	
Gleason Score	≤7	22 (23.2%)	73 (76.8%)	1.409 (0.235)	22 (25.3%)	65 (74.7%)	0.483 (p=0.487)	43 (43%)	57 (57%)	**6.808 (p=009)**
	≥8	31 (30.7%)	70 (69.3%)		31 (29.8%)	73 (70.2%)		62 (61.4%)	39 (38.6)	
D'Amico risk	Low	8 (33.3%)	16 (66.7%)	**6.114 (p=0.04)**	3 (15%)	17 (85%)	2.102 (p=0.350)	16 (66.7%)	8 (33%)	**14.667 (p=0.001)**
	Intermediate	1 (5%)	19 (95%)		4 (23.5%)	13 (76.5%)		2 (11.1%)	16 (88.9%)	
	High	40 (31%)	89 (69%)		41 (29.9%)	96 (70.1%)		75 (55.1%)	61 (44.9%)	
Antigen deprivation therapy (ADT)	Yes	39 (28.3%)	99 (71.7%)	0.56 (p=0.813)	41 (29.1%)	100 (70.9%)	1.165 (p=0.280)	73 (51.8%)	68 (48.2%)	0.007 (p=0.935)
	No	10 (26.3%)	28 (73.7%)		7 (20%)	28 (80%)		21 (52.5%)	19 (47.5%)	
Castration resistant	Yes	24 (27%)	65 (73%)	0.145 (p=0.703)	29 (31.2%)	64 (68.8%)	1.891 (p=0.169)	49 (54.4%)	41 (45.6%)	0.206 (p=0.650)
	No	26 (29.5%)	62 (70.5%)		18 (22%)	64 (78%)		47 (51.1%)	45 (48.9%)	

Table 10 Association of PAICS, Ki67 and P53 with clinical pathology variables. Significant p-values are shown in bold.

		PAICS negative	**PAICS positive**		**Ki67 negative**	**Ki67 positive**		**P53 negative**	**P53 positive**	
Clinical Variable		Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)
PSA (ng/ml) at diagnosis	<4	3 (33.3%)	6 (66.6%)	2.620 (p=0.106)	9 (100%)	0 (0%)	0.645 (p=0.422)	10 (83.3%)	2 (16.7%)	0.289 (p=0.591)
	>4	17 (13.5%)	109 (86.5%)		125 (93.3%)	9 (6.7%)		139 (88.5%)	18 (11.5%)	
Gleason Score	≤7	17 (21.8%)	61 (78.2%)	0.857 (p=0.355)	84 (96.6%)	3 (3.4%)	2.690 (p=0.101)	95 (92.2%)	8 (7.8%)	2.040 (p=0.153)
	≥8	13 (16.0%)	68 (84.0%)		75 (90.4%)	8 (9.6%)		86 (86.0%)	14 (14.0%)	
D'Amico risk	Low	4 (23.5%)	13 (76.5%)	0.460 (p=0.795)	20 (95.2%)	1 (4.8%)	1.493 (p=0.474)	25 (92.6%)	2 (7.4%)	0.514 (p=0.773)
	Intermediate	3 (18.8%)	13 (81.2%)		18 (100%)	0 (0 %)		18 (90.0%)	2 (10.0%)	
	High	18 (16.8%)	89 (83.2%)		103 (92.8%)	8 (7.2%)		117 (88.0%)	16 (12.0%)	
Antigen deprivation therapy (ADT)	Yes	20 (17.7%)	9 (82.3%)	0.185 (p=0.667)	108 (92.3%)	9 (7.7%)	2.781 (p=0.095)	122 (87.1%)	18 (12.9%)	2.276 (p=0.131)
	No	4 (14.3%)	24 (85.7%)		34 (100%)	0 (0%)		41 (95.3%)	2 (4.7%)	
Castration resistant	Yes	13 (18.6%)	57 (81.4%)	0.201 (p=0.654)	69 (90.8%)	7 (9.2%)	1.768 (p=0.184)	79 (86.8%)	12 (13.2%)	0.998 (p=0.318)
	No	11 (15.7%)	59 (84.3%)		74 (96.1%)	3 (3.9%)		85 (91.4%)	8 (8.6%)	

Table 11 Association of CK5/6, C-MET and Integrin alpha2 with clinical pathology variables. Significant p-values are shown in bold.

		CK5/6 negative	**CK5/6 positive**		**C-MET negative**	**C-MET positive**		**Integrin α2 negative**	**Integrin α2 positive**	
Clinical Variable		Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)	Number (%)	Number (%)	Chi-square (p-value)
PSA (ng/ml) at diagnosis	<4	4 (36.4%)	7 (63.6%)	**6.859 (p=0.009)**	3 (27.3%)	8 (72.7%)	0.316 (p=.574)	6 (60.0%)	4 (40.0%)	0.243 (p=0.622)
	>4	116 (73.4%)	42 (26.6%)		28 (20.1%)	111 (79.9%)		100 (67.6%)	48 (32.4%)	
Gleason Score	≤7	66 (65.3%)	35 (34.7%)	3.324 (p=.068)	17 (19.1%)	72 (80.9%)	0.352 (p=0.553)	53 (59.6%)	36 (40.4%)	2.598 (p=0.107)
	≥8	77 (77.0%)	23 (23.0%)		20 (22.7%)	68 (77.3%)		68 (70.8%)	28 (29.2%)	
D'Amico risk	Low	15 (60.0%)	10 (40.0%)	**6.179 (p=0.046)**	7 (31.8%)	15 (68.2%)	**6.513 (p=0.039)**	9 (45.0%)	11 (55.0%)	5.034 (p=0.081)
	Intermediate	10 (52.6%)	9 (47.4%)		0 (0.0%)	18 (100.0%)		13 (65.0%)	7 (35.0%)	
	High	101 (75.9%)	32 (24.1%)		28 (24.1%)	88 (75.9%)		88 (70.4%)	37 (29.6%)	
Antigen deprivation therapy (ADT)	Yes	102 (73.4%)	37 (36.6%)	1.303 (p=0.254)	27 (22.3%)	93 (76.9%)	0.320 (p=0.852)	90 (67.7%)	43 (32.3%)	0.727 (p=0.394)
	No	27 (64.3%)	15 (35.7%)		8 (21.6%)	29 (78.4%)		21 (60.0%)	14 (40.0%)	
Castration resistant	Yes	67 (74.4%)	23 (25.6%)	0.794 (p=0.373)	20 (26.0%)	56 (72.7%)	2.822 (p=0.244)	60 (69.0%)	27 (31.0%)	0.357 (p=0.550)
	No	63 (68.5%)	29 (31.5%)		14 (17.5%)	66 (82.5%)		53 (64.6%)	29 (35.4%)	

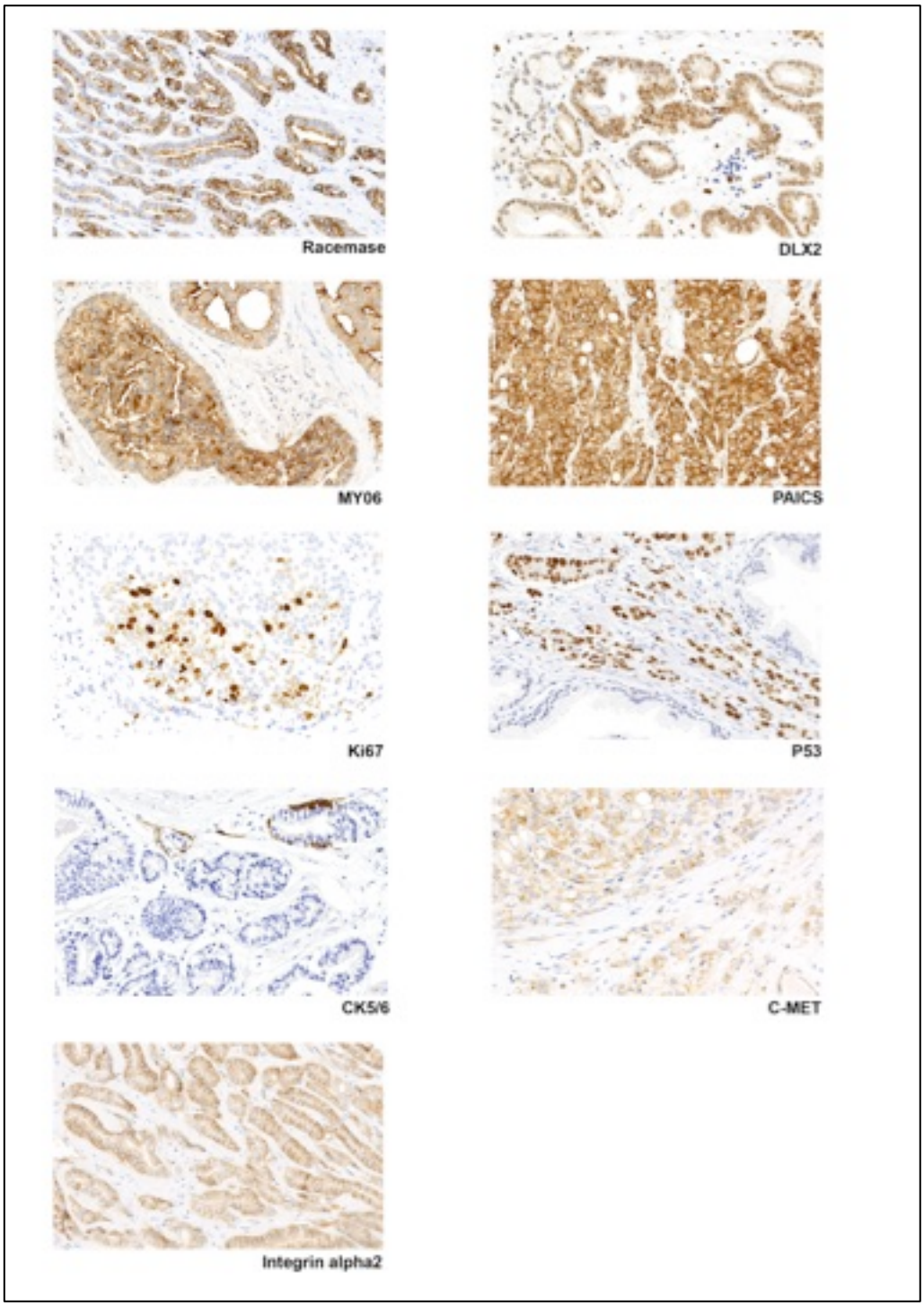
Table 12 Association of ERG2 with clinical pathology variables. Significant p-values are shown in bold.

		ERG2 Negative	**ERG2 positive**	
Clinical Variable		Number %	Number %	Chi-square (p-value)
PSA (ng/ml) at diagnosis	<4	6 (85.7%)	1 (14.3%)	1.270 (p=0.260)
	>4	76 (65.0%)	41 (35.0%)	
Gleason Score	≤7	45 (67.2%)	22 (32.8%)	0.046 (p=0.831)
	≥8	53 (68.8%)	24 (31.2%)	
D'Amico risk	Low	11 68.8%	5 31.2%	0.012 (p=0.994)
	Intermediate	11 68.8%	5 31.2%	
	High	65 67.7%	31 32.3%	
Antigen deprivation therapy (ADT)	Yes	67 65.7%	35 34.3%	0.871 (p=0.351)
	No	21 75.0%	7 25.0%	
Castration resistant	Yes	44 (67.7%)	21 (32.3%)	0.016 (p=0.901)
	No	44 (66.7%)	22 (33.3%)	

Table 13: Kaplan-Meier modelling was used to assess the association of biomarkers in predicting disease specific mortality, time to metastasis (months) and castrate resistance (months). Significant associations are shown in bold.

Biomarker	CaP specific mortality		Time to metastasis		Time to castrate resistance	
	$\chi 2$	**p-value**	**$\chi 2$**	**p-value**	**$\chi 2$**	**p-value**
Ki67	5.069	**0.024**	4.822	**0.028**	4.838	**0.028**
p53 nuclear	3.897	**0.048**	0.012	0.914	0.194	0.659
PAICS	0.002	0.966	0.383	0.536	1.11	0.292
MYO6	3.026	0.082	0.423	0.515	1.314	0.252
CMET cytoplasmic	2.945	0.086	16.208	**0.00005**	2.023	0.155
INT-a cytoplasmic	0.45	0.502	1.362	0.243	0.399	0.528
INT-a membranous	0.044	0.833	0.13	0.718	2.262	0.133
Racemase	0.705	0.401	2.974	0.085	0.213	0.644
CK5/6	0.344	0.558	0.091	0.762	2.161	0.142
DLX2	3.536	0.06	12.12	**0.0005**	1.182	0.277

Figure 3: Examples of IHC staining



4.1 Single biomarkers

Racemase (AMACR)

- Increased cytoplasmic racemase expression showed positive association with initial D'Amico risk at diagnosis (χ^2=6.114, p=0.047).
- No association was seen between racemase expression and initial serum PSA, Gleason score, castrate resistance, metastasis development, CaP-specific survival (CaPSS) or overall survival (OS).

DLX2

- Increased nuclear DLX2 expression showed a positive association with metastasis development (χ^2=12.12, p=0.0005), independently of PSA concentration and Gleason score using multivariate Cox regression analysis (HR=3.311, p=0.0002, 95%CI=1.756-6.241).
- No association was seen between DLX2 expression and initial D'Amico risk, initial serum PSA, Gleason score, castrate resistance, CaPSS or OS.

MYO6

- Increased cell membrane MYO6 expression showed a positive association with initial D'Amico risk at diagnosis (χ^2=14.667, p=0.001) and initial Gleason score at diagnosis (χ^2=6.808, p=0.009).
- No association was seen between MYO6 expression and initial serum PSA, castrate resistance, metastasis development, CaPSS or OS.

PAICS

- Cytoplasmic PAICS protein expression showed no association with PAICS expression, initial D'Amico risk, initial serum PSA, Gleason score, castrate resistance, metastasis development, CaPSS or OS.

Ki67

- Increased nuclear Ki67 expression showed a negative association with OS (χ 2=8.481, p=0.004) and CaPSS (χ 2=5.069, p=0.024) and a positive association with metastatic disease (χ 2=4.822, p=0.028) and castrate resistance (χ 2=4.838, p=0.028).
- Subsequent multivariate Cox regression analysis demonstrated that Ki67 contributed additional predictive ability to PSA concentration and Gleason score for CaPSS (HR=2.190, p=0.05, 95%CI=1.001-4.792) and metastasis risk (HR=2.746, p=0.046, 95%CI=1.020-7.390).
- No association was seen between Ki67 expression and initial serum PSA, D'Amico risk or Gleason Score.

P53

- P53 protein was localised to the cytoplasm and nucleus of malignant prostate tissue. Increased P53 nuclear expression was negatively associated with CaPSS (χ 2=3.897, p=0.048).
- No association was seen between P53 (nuclear) or P53 (cytoplasmic) and initial serum PSA, D'Amico risk, Gleason Score, metastases development,

castrate resistance and OS. P53 (cytoplasmic) showed no association with CaPSS.

CK5/6

- Increased cytoplasmic CK5/6 expression showed positive association with initial D'Amico risk category (χ 2=6.179, p=0.046).
- No association was seen between CK5/6 expression and initial serum PSA, Gleason score, castrate resistance, metastasis development, CaPSS or OS.

CMET

- CMET expression was localised to the cytoplasm and cell membrane of CaP.
- Increased cytoplasmic CMET expression showed positive association with metastasis development (χ 2=16.208, p=0.00005) and initial D'Amico risk (χ 2=6.513, p=0.039).
- Multivariate Cox regression analysis demonstrated that this association was not independent of Gleason score and PSA concentration at diagnosis (HR=0.311, p=0.002, 95%CI=0.148-0.651).
- No association was seen between cytoplasmic CMET and initial serum PSA, Gleason score, castrate resistance, CaPSS or OS. Increased membranous CMET expression showed positive association with initial serum PSA at diagnosis (χ 2=8.336, p=0.015).
- No association was seen between membranous CMET and initial D'Amico risk, Gleason score, castrate resistance, metastasis development, CaPSS or OS.

Integrin alpha2

- Integrin alpha 2 expression was localised to the cytoplasm and cell membrane of CaP.
- No association was seen between Integrin alpha 2 and initial D'Amico risk, initial serum PSA, Gleason score, castrate resistance, metastasis development, CaPSS or OS.

TMPRSS-ERG

- Increased TMPRSS-ERG expression was negatively associated with CaPSS (x 2=6.926, p=0.008).
- No association was seen between TMPRSS-ERG expression and initial serum PSA, Gleason score, castrate resistance, metastasis development or OS.

4.2 Biomarker combinations

Individually PAICS and RACEMASE do not predict CaPSS. However, when both were present and scored positive they were highly predictive of CaPSS ($\chi 2$=13.65, p=0.0002) and metastasis formation ($\chi 2$=6.775, p=0.009).

4.3 Initial results summary and rationale for further analysis

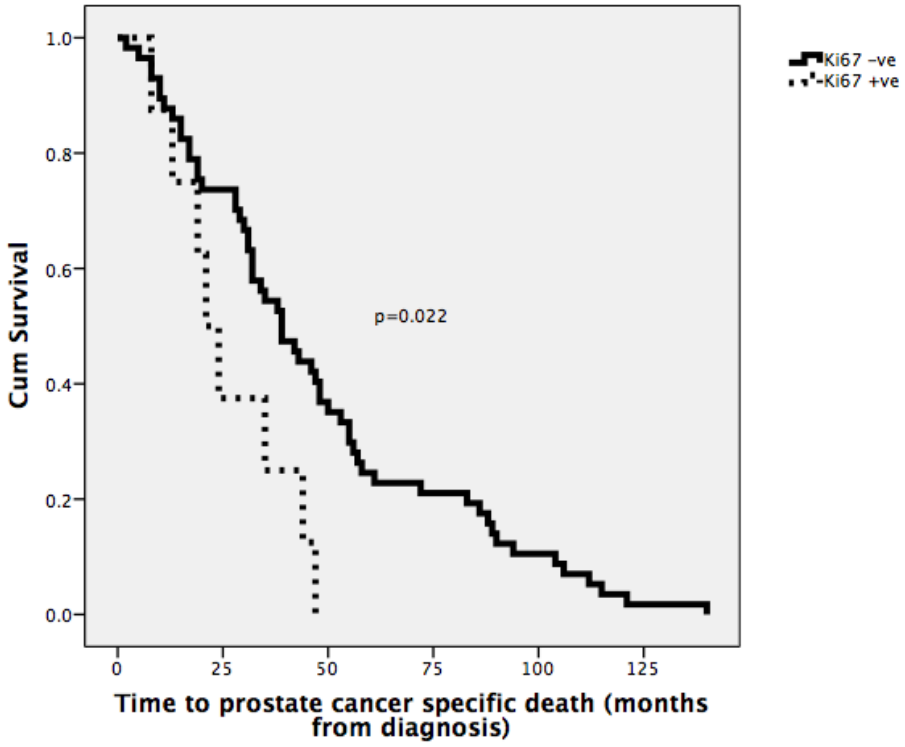
As can be seen from the above results we have demonstrated that biomarkers selected using bioinformatics and targeted literature search are predictive of clinical outcome in prostate cancer. As is the focus of all translational research the aim is to contribute information that will confer direct clinical benefit to a patient group. We therefore identified prostate cancer specific death and the development of metastatic disease as the two most significant clinical outcomes that can occur during the disease pathway. In addition, we felt that as DLX2 is an entirely novel marker that it was important to

disseminate our findings in the form of publication. To do that we also needed to validate our cohort by demonstrating similar biomarker associations to those seen by other groups and we therefore elected to include our Ki67 data. After discussion with the editorial team at the British Journal of Cancer regarding the issue of Gleason grading pre- and post- the ISUP 2005 consensus (as described in section 3.7), we asked Geoffrey Hulman (Consultant Histopathologist, Nottingham University Hospitals NHS Trust), to re-score the cancer cohort to post-ISUP 2005 standards. Once this had taken place we repeated our statistical analysis, focussing on DLX2 and Ki67, the results of which can be found below in section 4.4.

4.4 Ki67 and DLX2 with ISUP 2005 histological review

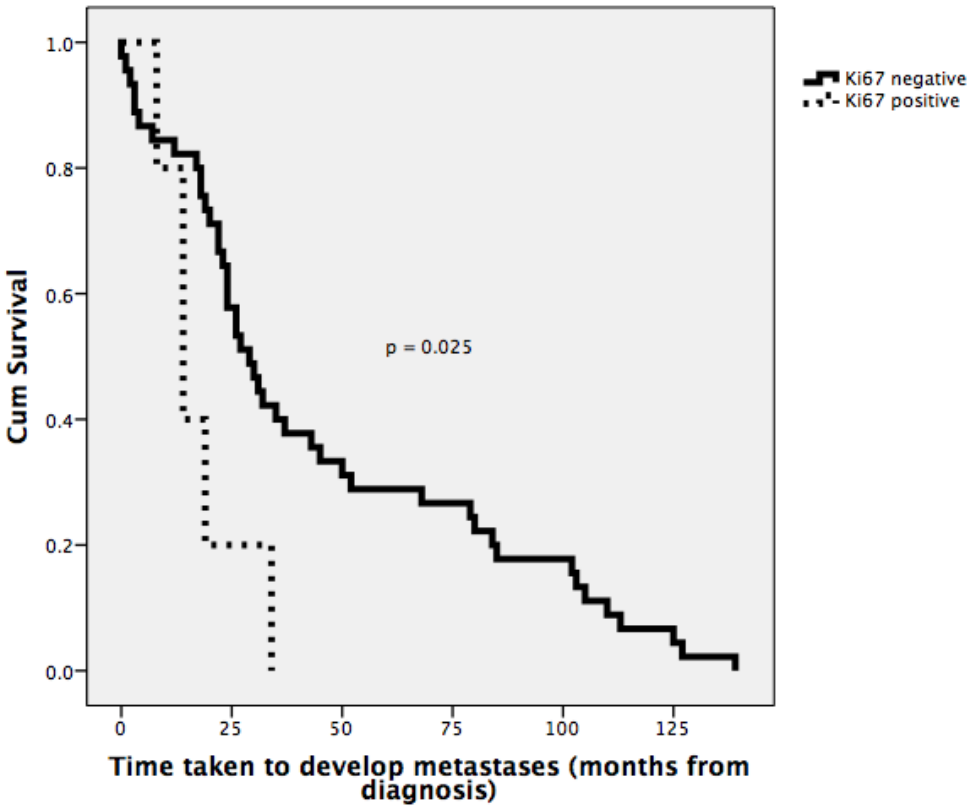
4.4.1 Ki67 and time to CaP-specific death

Figure 4: Kaplan Meier Chart demonstrating the association between Ki67 and CaP-specific death



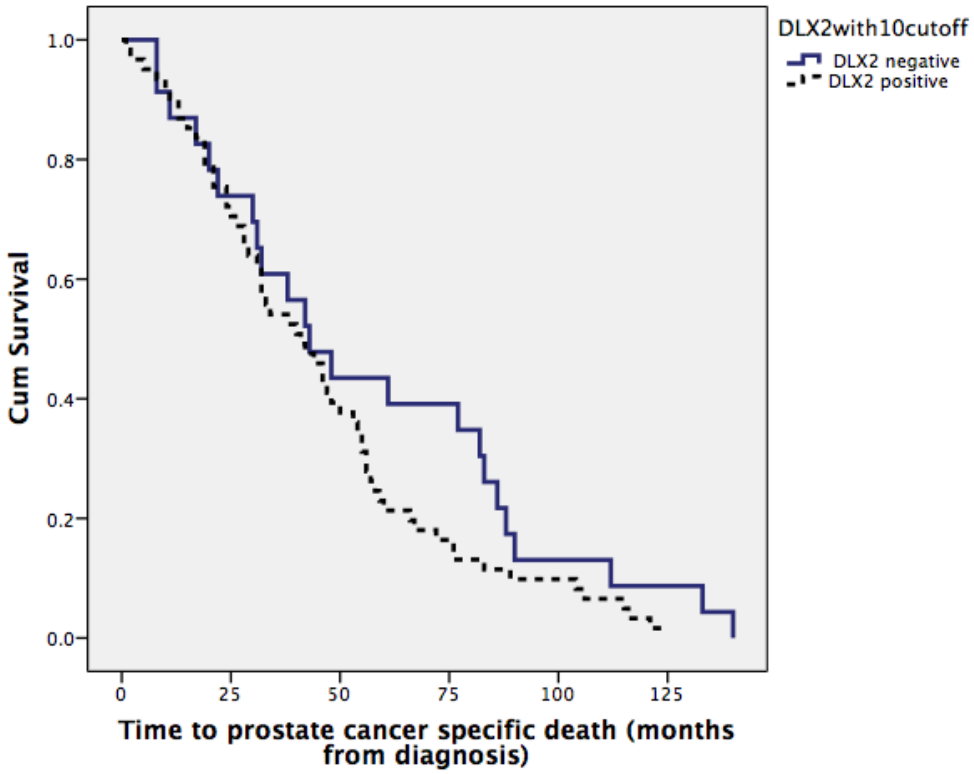
4.4.2 Ki67 and time to metastases

Figure 5. Kaplan Meier Chart demonstrating the association between Ki67 and time to development of metastases.



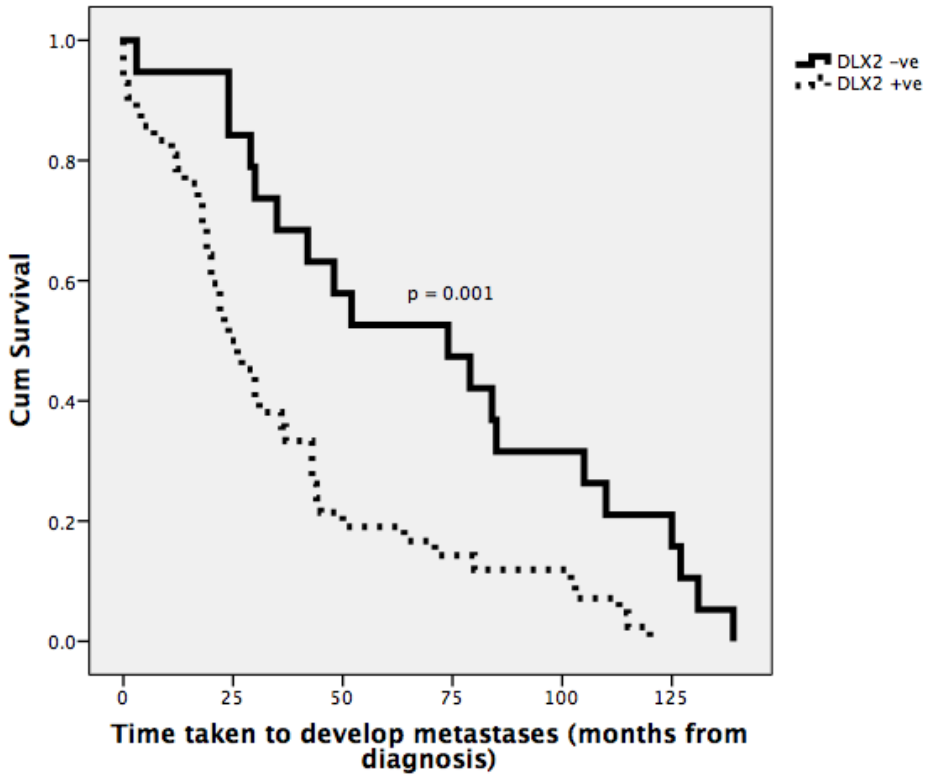
4.4.3 DLX2 and time to CaP-specific death

Figure 6: Kaplan Meier Chart demonstrating the association between DLX2 and CaP-specific death



4.4.4 DLX2 and time to metastases

Figure 7 Kaplan Meier Chart demonstrating the association between DLX2 and time to development of metastases



5.0 Discussion

5.1 Overview

This study examines biomarkers of prostate cancer that can be practically used to further enhance the current prognostic staging of this important disease. We employed highly sophisticated statistical techniques such as artificial neural network analysis to identify potential new markers. We also developed a large, clinically comprehensive tissue micro-array to allow quantification of relevant protein expression against multiple recorded clinical outcomes.

We demonstrated several biomarkers with the potential to predict disease progression and clinical outcome in an unselected group of prostate cancer patients over an extended period of clinical follow up. In particular, we showed that the markers Ki67 and DLX2 have a statistically significant relationship with key study outcomes prostate cancer mortality (Ki67 and DLX2) and prostate cancer metastases (DLX2). We also summarised the potential biological pathways where these markers may be involved in the pathogenesis of prostate cancer.

Given that these markers are relatively easy to assay, and that tissue taken from diagnostic prostate biopsies is sufficient in volume to allow for this assay without compromising the ability to diagnose the disease our intention going forward is to set up a prospective trial examining whether the expression of KI67 and DLX2 at diagnosis could practically predict outcome of a patient's prostate cancer, particularly in the 'intermediate' risk group where the decision to undergo radical treatment or active surveillance is a difficult one, with potentially extremely significant repercussions if the incorrect decision is made.

The techniques employed in this study have the capacity to be applied to other urological malignancies. For example, urothelial cancer is extremely prevalent, but with a wide-ranging clinical outcome, from fairly indolent disease to rapid development of metastases and death. In addition, in comparison to prostate cancer it has had only a fraction of the research effort applied to it, probably due to the lack of equivalent funding sources - prostate cancer has multiple successful charities that have provided financial backing over the last decade. It would therefore seem a logical next step to develop an appropriate tissue bank and clinical database to allow practical assessment of new biomarkers identified with similar methods to those we have used in this study.

5.2 Background

The development and progression of prostate cancer is an extremely complex process and a number of important factors are involved, including genetic abnormalities, oxidative stress and cellular inflammation, altered epithelial – stromal interaction and androgen receptor signalling. The diagnosis of prostate cancer is based on a combination of clinical signs, examination findings, molecular markers and histological examination of biopsy samples. While improvements are always being sought to make diagnosis a quicker, more reliable process with as few risks to the patient as possible it is essentially straightforward – a patient either has prostate cancer or does not. The complex part is predicting which patients are going to progress to life-threatening disease without treatment, and those that will not develop clinically-significant disease in their lifetime. Current risk stratification models are more accurate in predicting those with very low and very high risk, but are less able to

predict the clinical outcome for those in the intermediate group. In addition, once patients have undergone treatment the molecular biology of biochemical relapse, progression and the development of metastatic spread are still not fully understood.

At present, single biomarkers are currently used for diagnostic and predictive assessment of prostate cancer. The most widely used and evidence-based is prostate-specific antigen (PSA), a 34kD serine protease encoded by a gene on chromosome 19 and uniquely produced by prostate epithelial cells. A raised PSA level can indicate an increased risk of prostate cancer, although presence of other factors such as urinary tract infection and significant lower urinary tract symptoms can cause similar rises. There is no absolute value above which prostate cancer is present, but studies have shown that a PSA of 4ng/ml confers approximately a 25% risk of cancer[84]. PSA is particularly useful in monitoring response to treatment in CaP. For example, a rise in PSA in a patient who has undergone radical treatment is an early indicator of disease recurrence. Another marker, less widely used and mainly as a diagnostic adjunct, is prostate cancer antigen-3 (PCA3). It is a non-coding segment of mRNA produced by prostate epithelial cells approximately 60 to 100 times more in prostate cancer than benign tissue. The most common assay is marketed as Progensa[85] and samples for analysis are collected in the urine after prostatic massage and a ratio of PCA3 to PSA mRNA is calculated and a prostate cancer risk is determined.

This study identifies several tumour biomarkers, selected using a bioinformatic ANN approach and literature survey, which provide clinically relevant information concerning prostate cancer progression and survival.

Development of metastasis was associated with increased expression of the tumour markers DLX2, Ki67 and CMET. DLX2 is entirely novel and this is the first study to identify it as a potential marker of disease progression in prostate cancer. The markers Ki67, P53 and TMPRSS-ERG are predictors of disease specific survival in CaP and we have validated this finding in the present study. Individually, Ki67 contributed predictive information additional to the PSA and histological Gleason score. Interestingly in some instances combining biomarkers strengthened predictive accuracy. For example, while neither PAICS nor Racemase demonstrated statistically significant association with clinical outcome in isolation, when combined and both scored positive they were shown to be highly predictive of disease specific survival and metastasis.

The biomarkers tested here are functionally associated with cell replication, apoptosis, cell migration or tumour cell of origin and their relationships and our findings are discussed in detail below.

5.3 Cellular proliferation markers and their application in prostate cancer

We have demonstrated a clear association between the proliferation marker Ki67 and prostate cancer specific survival and time to metastases. It should be noted that when analysing the data, patients were only included if the appropriate endpoint (eg. metastases or death) was recorded in the clinical follow-up, and that there was an IHC result available for that particular marker (as with all TMA assessment some samples are lost during processing, and due to tissue heterogeneity multiple samples taken from the same core will not always show cancer). This means that for any particular marker the number of patients included was less than the entire cohort in the TMA –

individual figures can be seen in Table 8. This relationship was independent of whether the prostate tissue was scored on the pre- or post-ISUP 2005 histological Gleason grades. In addition, we demonstrated that Ki67 provides additional prognostic utility (HR:2.19) to the PSA and Gleason score, validating the study by Fisher et al.[86] who recently reported that Ki67 independently adds significant predictive information (HR:2.78) in prostate biopsies. Only 6.6% of our cancer cohort had raised Ki67 and we propose that such patients could be counselled regarding an increased risk of death, metastasis and castrate resistance, particularly if they are being considered for an active surveillance programme. These findings fit with a hypothesis that tumour proliferation rates are a surrogate for tumour aggression.

Assessment of cell proliferation in cancer has received much attention because proliferation is a key requirement for tumour growth and its progression. Unregulated cell turnover occurs as a result of genetic abnormalities at all stages of tumour development and can broadly be grouped by where in this pathway they occur. For example, alterations to genes such as ER-1B [87,88] and NKX3-1[89,90] have been linked to dysregulated cell proliferation. Loss or mutation of genes such as HPC1 (that codes for the tumour suppressor protein RNaseL) are thought to lead to altered apoptotic processes in response to cellular stress[91,92].

Ki67 is a nuclear proliferation marker that represents cell turnover and has been investigated extensively as a clinical marker, particularly in the field of breast cancer research where it has been used as an adjunct to assessing disease prognosis, predicting response to treatment and in disease surveillance. Results of studies, and its subsequent clinical application have been limited, however, by inconsistency in

immunohistochemical techniques and scoring boundaries[93]. This has been clearly demonstrated in a number of recent papers showing that factors such as tumour region selection while scoring, counting method and subjective assessment of staining intensity all affect outcome[93–95]. Even if these factors are accounted for (which is one of the goals of the Ki67 international reproducibility study[93]) then scoring boundaries need standardisation as in a recent review cutoffs were found to range from 0 to 28%.

The biological mechanism of Ki67 is via its action as a key controller of the cell cycle, and is expressed in all active phases of the cell cycle. During interphase, Ki67 can be exclusively detected within the nucleus, whereas in mitosis most of the protein is relocated to the surface of the chromosomes. The fact that the Ki-67 protein is present during all active phases of the cell cycle (G1, S, G2, and mitosis), but is absent from resting cells (G0), makes it an excellent marker for determining the so-called growth fraction of a given cell population[96].

Like in breast cancer, there has been a great deal of interest in the use of Ki67 as a prognostic marker in prostate cancer, and in various trials it has shown both univariate and multivariate clinical significance[86,87,97–100]. For example, using a cutoff of >5% cancer nuclei positively stained, Ki67 is prognostic of cancer specific death in tissues derived by trans-urethral resection of prostate (TURP)[87]. Similar findings were obtained using a 10% nuclei cutoff in diagnostic biopsies[86]. A recent investigation found high (>6.2%) levels of histologically detected Ki67 were prognostic of disease specific death, metastasis and biochemical failure (rising PSA) in low to intermediate (PSA<20ng/ml) patients treated with a combination of short term ADT and

radiotherapy[101]. Ki67 is also a component of the Cell Cycle Progression signature proposed by the Transatlantic Prostate Group[102].

While the assessment of Ki67 expression with immunohistochemistry has the potential for providing a cost-effective and robust laboratory technique applicable to routinely processed pathology samples for similar reasons to breast cancer (as described above) it has so far failed to make the conversion from research/investigative tool into daily clinical use. It is likely, however, that with efforts to reach consensus in laboratory IHC techniques and with determining an agreed value for what level of expression should be deemed 'positive' that this marker will play an increasingly important role in prostate cancer risk stratification strategies.

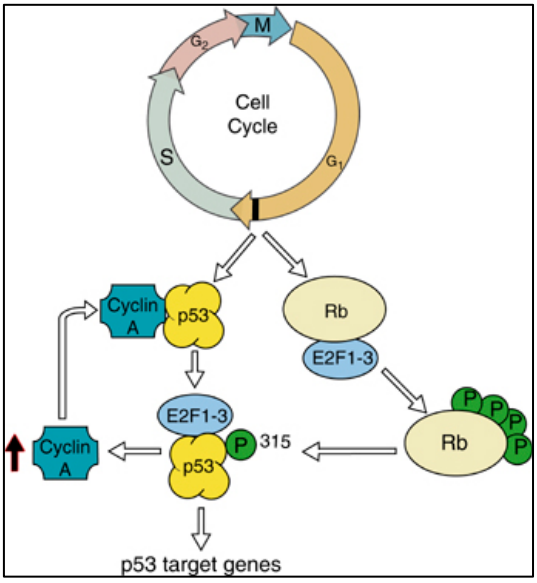
5.4 Cell cycle regulation

Cell cycle regulation is a key arbiter of normal cellular growth, development and death. The markers P53 and DLX2 are both important proteins in this mechanism. We have demonstrated that there is an association between increased nuclear P53 expression and reduced CaPSS. In addition, increased nuclear DLX2 expression showed a positive association with metastasis development independently of PSA concentration and Gleason score using multivariate Cox regression analysis.

P53 is an extremely important tumour suppressor gene, with mutations seen in over 50% of all human malignancies and is coded for on the short arm of chromosome 17. P53 exerts its effect at the G1/S cell cycle checkpoint where it can pause the cell cycle if it recognises DNA damage to allow repair and subsequent cycle progression, or if the damage is significant it can trigger apoptosis. When activated, p53 increases

expression of many other genes, most commonly through E2F protein family mediation. In particular it up regulates p21 that in turn complexes with the cyclin-dependent kinases, which normally facilitate transition through the G1/S checkpoint, thus inhibiting their action.

FIGURE 8**:** Cyclic control of p53 cell cycle checkpoint function by E2F family members. p53 becomes phosphorylated at ser315 by cell cycle-associated kinases. This provokes E2F1–3 to bind p53 displacing cyclin A interaction and inducing p53 to become transcriptionally competent. As cyclin A and other kinase levels increase, E2F1–3 binding will be disrupted thereby reducing p53 activity[103]



In addition to the relationship between p53 dysregulation and CaP further important associations have been demonstrated between PTEN (phosphatase and tensin homologue), which acts as a cell cycle regulator, and the development of prostate cancer. PTEN deletions have been demonstrated in 5% of localised CaP but over 30% of metastatic CaP[104], suggesting it may be an important target in the molecular transition between organ confined and widespread disease.

Distal-less homeobox 2 (DLX2) is a transcription factor involved in cell cycle regulation and is one of a family of six known genes that are involved in embryonic development, tissue homeostasis, lymphocyte development, cell cycle and apoptosis. While only recently becoming a target of interest to cancer research scientists the evidence base linking the distal-less homeobox family to oncogenic processes is expanding rapidly, both in urological and non-urological cancers.

The interaction between DLX2 and various cellular pathways is being investigated currently with one area of research examining its role in regulation of Transforming Growth Factor beta (TGF-β). TGF-β is a protein involved in proliferation and cellular differentiation in many cells and its dysregulation plays an integral part in the development and propagation of prostate cancer. The distal-less homeobox gene family has been implicated in triggering this dysregulation. A recent study[105] showed that DLX2 is involved in shifting TGFβ from a tumour suppressor to a tumour promoting function by repressing TGFβRII and the cell cycle inhibitor p21CIP1, and simultaneously increasing the mitogenic transcription factor c-Myc and epidermal growth factor (EGF). The impact of this has been shown to increase tumour growth and metastasis formation in melanoma and lung cancer

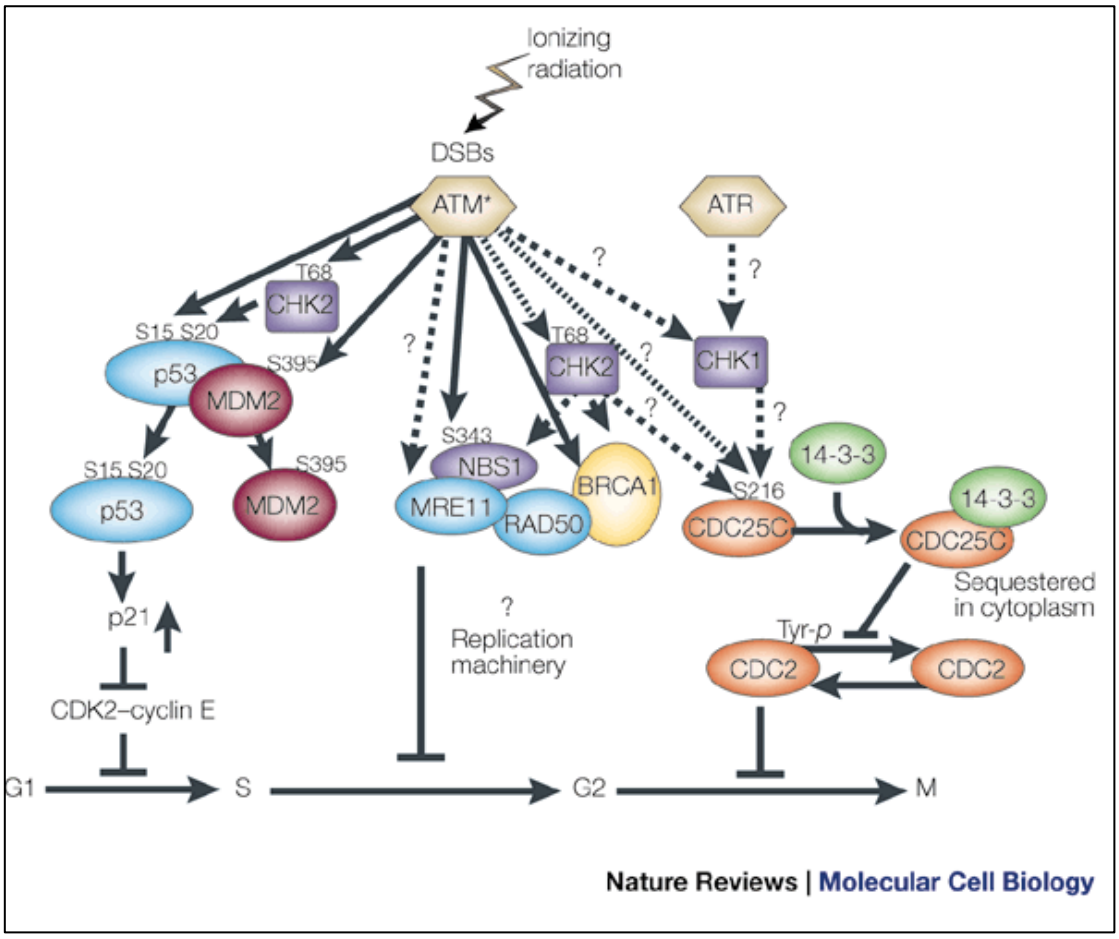
Another very interesting field of study is the relationship between DLX2 and cell senescence. Cellular senescence describes the process by which cells enter a period of growth arrest that essentially determines their life span. When senescence markers are triggered in a regulated fashion cells stop replicating and will succumb to apoptosis. However, if the process of senescence becomes dysregulated cells can continue to

replicate, despite the presence of cellular stressors or genetic aberration. It appears that DLX2 expression can switch cells into a 'senescence-bypass' phenotype, probably though mediation of the P53 and p21 cell cycle proteins.

The p53 tumor suppressor pathway plays a pivotal role in the initiation of senescence cell cycle arrest. Triggers of function include DNA damage and telomere shortening[106]. DNA abnormalities activate a pivotal serine/threonine protein kinase called Ataxia Telangectasia Mutated (ATM) which is a key molecule in the regulation of cell cycle and DNA repair[107]. ATM activates the p53 tumor suppressor through phosphorylation, which subsequently activates a cascade of interactions that result in the upregulation of p21, which suppresses inhibitory phosphorylation of the Retinoblastoma protein RB to arrest the cell cycle[108] (Figure 5).

DLX2 has been shown to down-regulate the p53-p21 cell cycle control pathway, both in ageing cells where shortened telomeres trigger senescence and in younger cells with laboratory-induced ionizing radiation DNA damage. This appears to be mediated by a reduction in expression of the ATM protein. Interestingly the same group has shown that in breast cancer (a malignancy with many similarities to prostate cancer) DLX2 and P53 mutations are generally mutually exclusive suggesting that overexpressed DLX2 may negate the need for p53 mutations in cancer cells.

Figure 8: The ATM molecular cascade: ATM and other molecular signals controlling cell-cycle checkpoints prompted by DNA damage (reused with permission from Nature Publishing Group, Licence number 4226361274939)

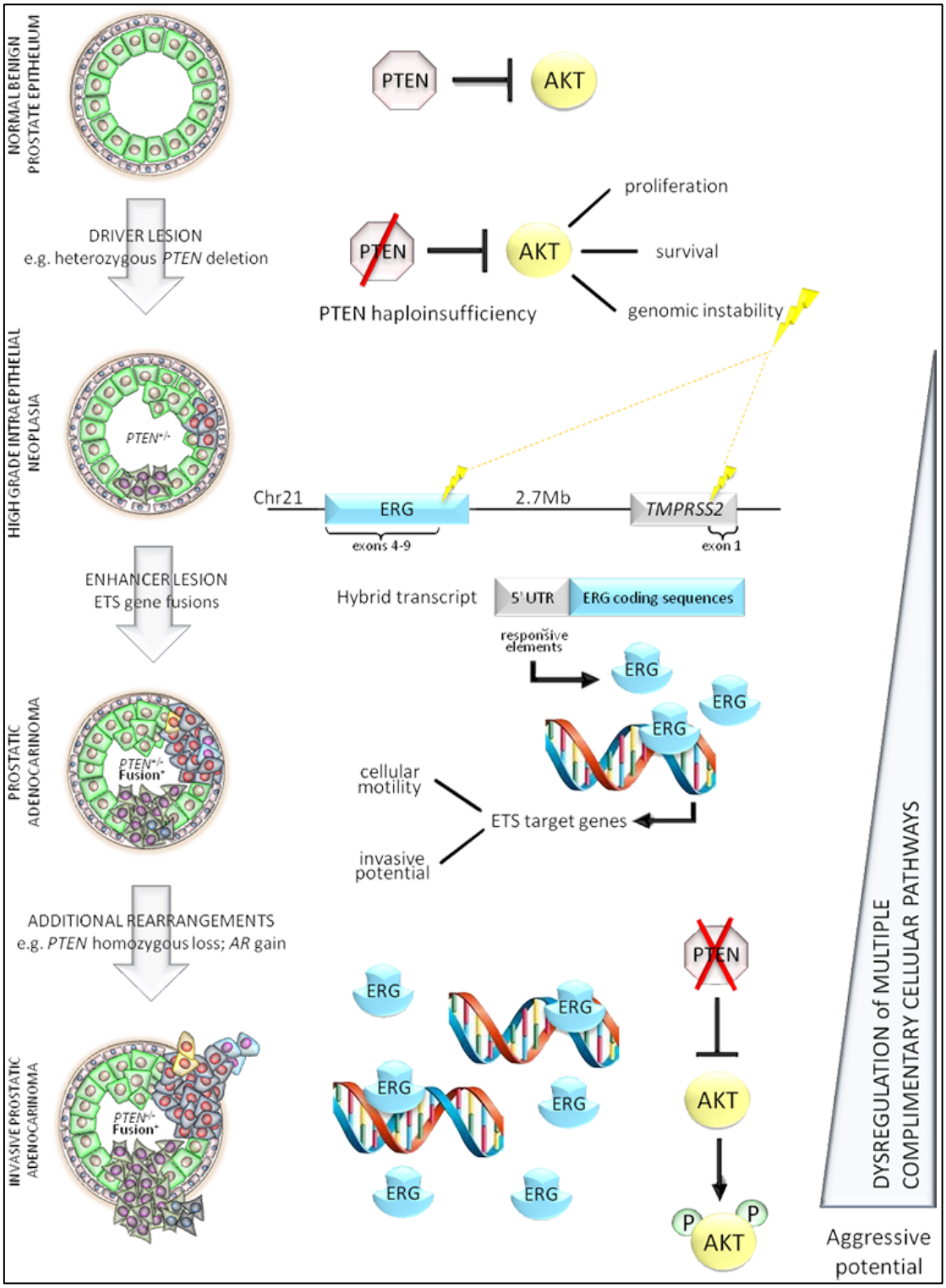


Another key genetic abnormality in prostate cancer is TMPRSS2-ERG fusion, which is found to occur in at least 50% of CaP patients[109]. Our study has demonstrated that an increased TMPRSS-ERG expression is negatively associated with CaPSS ($\chi 2=6.926$, $p=0.008$). TMPRSS-ERG has been shown to promote cancer invasion and metastasis and some groups have linked TMPRSS2-ERG fusion to poorer overall prognosis, particularly in those patients in a 'watchful waiting' cohort[110].

However, in early studies those patients undergoing surgery for their CaP demonstrated no clear difference in cancer specific survival caused by this gene fusion[111]. The general consensus until recently was that TMPRSS2-ERG expression for predicting outcomes in prostate cancer was controversial, but it certainly became an established adjunct in the diagnosis of prostate cancer in men wishing to avoid a biopsy if possible. Studies have shown that raised urinary TMPRSS2-ERG levels (and in particular if combined with urinary PCA3 levels) have improved clinical decision curve analysis characteristics than PSA value alone in predicting the presence of significant and high grade prostate cancer[112].

However, more recently TMPRSS2-ERG gene fusion as a prognostic marker in prostate cancer has generated a greater evidence base[113] and often seems to be associated with concomitant PTEN gene aberration in many situations, particularly those patients with high grade prostate cancer[114]. One theory that has been examined is that PTEN loss increases the susceptibility of prostatic epithelial cells to switch to a prostatic intraepithelial neoplasia (PIN) phenotype (Figure 6), which is a known precursor of prostate cancer. Subsequent TMPRSS2-ERG gene fusion promotes the final step of CaP development and increased activation of the androgen receptor, increasing cell turnover rate and likelihood of metastases[115].

Figure 10: Model of prostate cancer progression showing ETS gene fusions as an enhancer lesion (reproduced with permission from the Atlas of Genetics and Cytogenetics in Oncology and Haematology, www.atlasgeneticsoncology.org) [116].



Cooperation of unregulated pathways downstream of PTEN with effectors of ERG overexpression is likely a crucial event in the progression of an invasive and aggressive prostatic adenocarcinoma. Heterozygous genomic deletion of *PTEN* in benign prostatic precursors may represent an early event, and act as a driver lesion leading to proliferation, survival and genomic instability-all initial requisites of cancer. As a consequence of such heightened genomic instability, PTEN haploinsufficiency may facilitate the selective formation of the fusion gene with consequent acquisition of additional invasive properties. The presence of both rearrangements within a lesion is associated with accelerated disease progression and poor prognosis, indicating that synergistic molecular interactions exist between their complementary pathways. Continuing instability generates genotypic heterogeneity and diversity, such that subclones bearing *PTEN* homozygous deletions and amplified *AR* loci have further selective advantage for aggressive tumour progression, androgen escape and metastases.

5.5 Tumour Lineage

As described above there is evidence that genetic signatures commonly associated with embryonic stem cells (ESCs) are up regulated in the tumours of patients with more poorly differentiated prostate cancers[60,61]. Those prostate cancers expressing an ESC signature are more likely to be of higher Gleason grade and therefore have a poorer outcome than those cancers that do not.

Eaton et al. have demonstrated an interesting relationship between the proportion of phenotypically positive prostate stem cells (PPSC) in the primary tumour and its derived metastases. Their work showed that PPSC was significantly higher in bone metastases compared to the primary prostate cancer[63]. Further work by Columbel et al described an association between the stem cell markers integrin alpha-2 and -6 in

combination with c-met, with higher levels of expression predicting poorer CaP outcome[64].

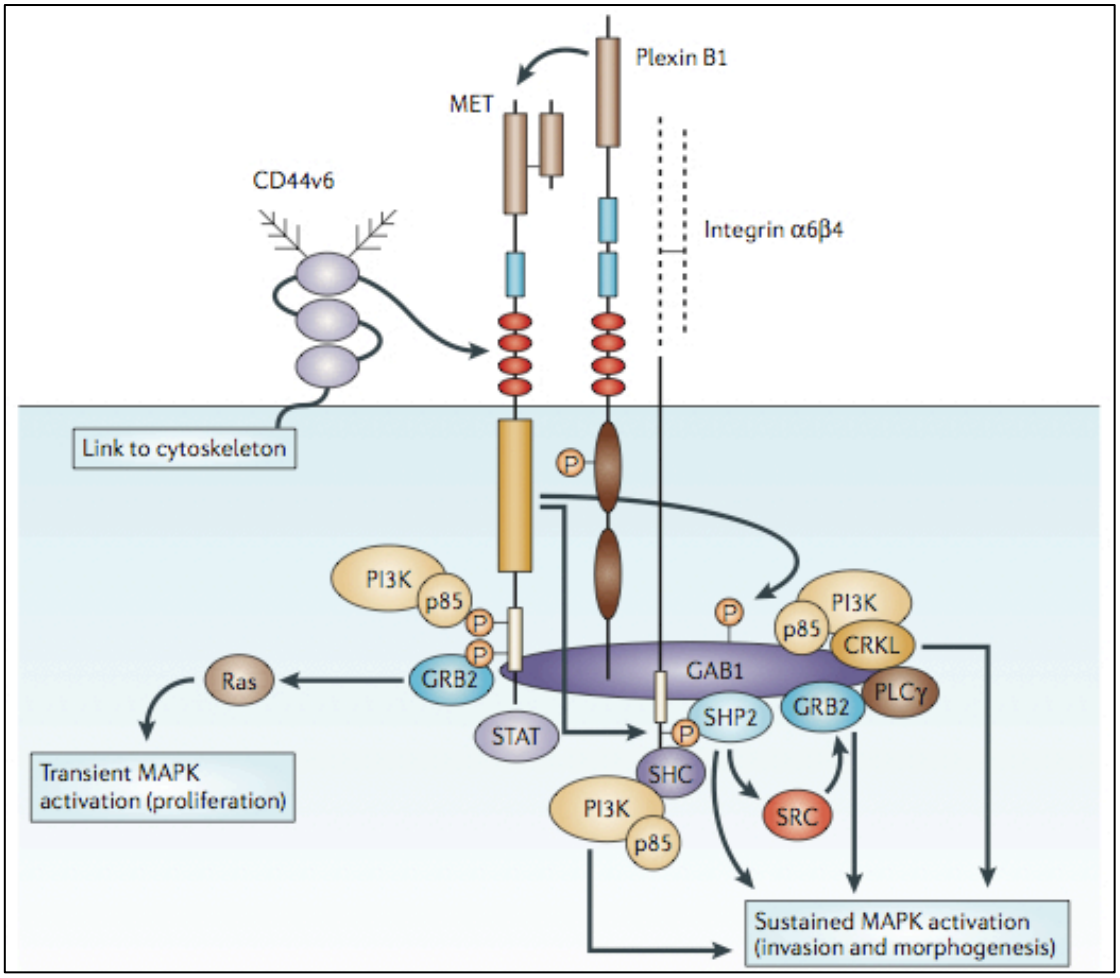
We have confirmed that there is a relationship between the stem cell marker CMET and the development of CaP metastases, however we were unable to demonstrate a further association with integrin-alpha.

It is likely that CaP may develop from progenitor cells resulting in cancers with basal and luminal lineages, and possibly conferring divergent disease pathways[17]. The basal epithelial compartment contains a pool of multipotent stem cells[29,31,32] capable of differentiation into basal and secretory luminal epithelium, and an 'intermediate' amplification cell type[33] and interestingly PTEN loss has been associated with basal cell derived prostate cancers[117]. These rare tumours have been considered fairly indolent in past studies, however increasingly there is evidence that this may not be the case and that in fact up to 44% of patients with basal cell derived prostate cancers in recent series have high risk disease[117].

Human prostate basal cells can be triggered to develop prostate cancer *in vivo*[34] and in addition populations of luminal cells that retain some stem-cell like qualities have been identified. These may still be 'early' in the differentiation phase[35], or derive from an entirely separate stem cell population[17,37]. Colombel et al.[64] examined the expression of three putative stem cell markers: alpha2 and alpha6 integrin and CMET, in men with high risk CaP. They concluded that the proportion of stem cell-like cancer cells is predictive of metastatic bone progression, and that the accuracy was increased when either of the integrin markers was combined with CMET. Here, we

report that increased CMET expression is highly predictive of metastatic disease ($\chi 2=16.208$, $p=0.00005$), however did not find increased significance when combined with integrin alpha, possibly because very few patients concurrently expressed CMET and integrin alpha. CMET (or hepatocyte growth factor receptor) is coded for on chromosome 7 and acts as a tyrosine kinase that is appropriately active during embryonic development but if subject to genetic aberration is associated with tumour growth, angiogenesis and metastases. It triggers a series of downstream protein complexes, particularly the RAS, PI3K and STAT signal pathways.

Figure 11: CMET-mediated downstream cell signalling pathways (reused with permission from Nature Publishing Group, License number 4226380237905)[118]



5.6 Genetic risk stratification

Molecular classification techniques have identified candidate genes that might influence hormone/receptor biological pathways. Bioscience companies have been quick to recognise that array-based gene expression tools could potentially be used to guide treatment and predict outcomes.

In breast cancer, early stratification studies of patients using gene expression profiling revealed an association between tumour biology genotype, tumour behaviour and

response to targeted therapy[48]. This approach has been refined and it is the case that whilst a single biomarker can inform about likely response to targeted therapy (theranostics eg. ER status and tamoxifen), panels of biomarkers are needed to inform about individualised risk of disease progression and survival. Risk assessment can be used to assist chemotherapy decision-making. For example, a 21 multi-gene PCR-based assay was developed for predicting tumour recurrence in tamoxifen treated, node negative, ER expressing breast cancer[119]. This assay has been commercialised and marketed as the Oncotype DX Breast Cancer Test (Genomic Health) for predicting the likelihood of chemotherapy benefit in women with early stage breast cancer. Further assays such as Mammaprint and Prosigna have also been developed in breast cancer, and differ in their gene signatures and target population[120].

Similar techniques have been applied in prostate cancer and the resulting Oncotype DX Prostate Cancer test is a multigene PCR-based assay that assesses risk of disease progression in patients with apparent low risk disease. This 17 gene profile assesses 4 distinct biological targets: the androgen pathway, cellular organisation, proliferation and stromal response[121]. This assay gives a 'genomic prostate score' (GPS) that predicts the likelihood of high grade or high stage disease at the time of diagnosis[122]. This array, like its competitors 'Prolaris' and 'Decipher' is not yet widely used and a recent systematic review concluded that they have yet to clearly demonstrate any significant advantage over more established predictive nomograms as a general clinical application[39].

Multi-gene/multi-protein biomarker panels are likely to be the mainstay of risk stratification and prognostic tools in the foreseeable future. However biomarker

assays have been significantly hindered by the known prostate cancer tumour multi-focality and indeed intra-tumoral heterogeneity[123,124]. These two factors can cause difficulty when quantifying biomarkers as when limited tissue is being stained and scored it is possible to 'miss' other tumor foci of varying significance and in the presence of multiple clones of tumour cells with different characteristics, a single sample from one small tumour region might not be optimal for predicting a tumour's aggressiveness[125]. There is an increasing awareness that multiple samples from a particular patient may need to be analysed to generate a 'representative' biomarker score.

Further work employing hierarchical clustering techniques has identified expression of the gene product of Hey2 as being an independent predictor of biochemical failure, local recurrence and distant metastasis in prostate cancer[126]. The same group has demonstrated another gene (CYP4Z1) as an independent predictor of indolent disease.

5.7 Limitations

As with many studies there were limitations in this project that need to be acknowledged to allow accurate, informed opinion to be developed when assessing the work's validity.

The majority of tissue incorporated in the TMA, and thus exposed to IHC was from patients who underwent TURP rather than diagnostic biopsy. While we have described above that current guidelines state that patient's diagnosed with CaP from TURP should be treated in the same way as biopsy-proven cancer, in practice patients undergoing TURP are older than men undergoing raised PSA-driven biopsy and

therefore are more likely to have higher volume and more aggressive disease. This is reflected in our high proportion of Gleason 8 and above cancer (and therefore high-risk disease), which is not reflective of current trends. As a result, our data may be less applicable to those newly diagnosed intermediate risk patients who need further risk stratification to guide the choice between radical treatment and surveillance. An important priority for our group is to examine our markers again in a more up to date clinical cohort to ascertain whether the same results will be found.

Our study used a TMA to score for immuno-reactivity, which is an established and validated method employed to process large numbers of patient's tissue in an efficient manner. However staining can sometimes be unreliable or misleading as due to tumour heterogeneity samples can be missed when only staining a very thin tissue sample removed from the TMA with a microtome (as described above).

6.0 Conclusion

There is a need for more accurate markers of disease outcome in prostate cancer. Currently many patients undergo highly invasive and expensive treatments that carry significant side effects that may have been unnecessary, as their disease would never have become clinically apparent. Others will be stratified as low or intermediate risk but will subsequently develop highly aggressive disease with its attendant morbidity and mortality.

This study has demonstrated several biomarkers with the potential to predict disease progression and clinical outcome in an unselected group of prostate cancer patients

over an extended period of clinical follow up. Several of the markers have a known biological function and disease mechanism including Ki67, P53 and DLX2.

We have successfully added to the scientific literature by confirming that Ki67 is indeed associated with clinical outcome in CaP and we are the first group to identify the novel marker DLX2 as having prognostic utility in this disease, particularly showing great promise in predicting metastatic disease.

Understanding the cell lineage of prostate cancer and applying the highly successful techniques used in breast cancer research has led to the development of gene signature arrays that reveal different molecular classifications in CaP. Emerging evidence suggests that molecular phenotyping is possible in prostate cancer and identification of distinct subtypes may allow personalised risk stratification way beyond that currently available.

While initial results are promising, further work is required to define a robust panel of predictive markers in prostate cancer; this may involve selection of predictive/prognostic biomarkers that inform about the potential biological behaviour of circulating and disseminated tumour cells in addition to those detected in the primary organ (prostate).

The use of gene expression arrays coupled with bioinformatic techniques has led to the identification of clinically useful multigene PCR assays and protein-based biomarkers. The former are generally more complex and require specialised tissue processing. Protein based assays are mostly applied to routinely processed

histological based samples or liquid samples. Currently, all would be used in-conjunction with nomograms and risk algorithms currently employed by clinicians managing patients with prostate cancer.

As we have demonstrated that DLX2 and Ki67 have prognostic function in CaP when applied retrospectively to a TMA based population it is important that we now investigate their application in a prospective clinical trial. The techniques required to quantify the protein expression of each biomarker (immunohistochemistry) are used widely in most mainstream NHS diagnostic laboratories and as prostate cancer is diagnosed through multi-core biopsy or transurethral resection of prostate (TURP) there is likely to be no logistical problem in acquiring sufficient tissue to assess biomarker expression. A major advantage of using these biomarkers is that their analysis and quantification are likely to be logistically and technically possible in a 'real world' NHS setting, which unlike other first-world health systems such as the US is constrained financially and politically, making the uptake of expensive multi-gene assays unlikely in the short to medium term.

7.0 References

1. Cancer Statistics Registrations, England (Series MB1). Office for National Statistics. http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations--england--series-mb1-/no--42--2011/index.html. Published June 26, 2013. Accessed May 3, 2015.
2. Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z,Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2011, National Cancer Institute. Bethesda, MD. 2014.
3. Roehrborn CG, Black LK. The economic burden of prostate cancer. *BJU Int*. 2011;108(6):806-813. doi:10.1111/j.1464-410X.2011.10365.x
4. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet Lond Engl*. 2017;389(10071):815-822. doi:10.1016/S0140-6736(16)32401-1
5. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA J Am Med Assoc*. 1998;280(11):969-974.
6. Korets R, Motamedinia P, Yeshchina O, Desai M, McKiernan JM. Accuracy of the Kattan nomogram across prostate cancer risk-groups. *BJU Int*. 2011;108(1):56-60. doi:10.1111/j.1464-410X.2010.09838.x
7. Cooperberg MR, Lubeck DP, Mehta SS, Carroll PR, CaPSURE. Time trends in clinical risk stratification for prostate cancer: implications for outcomes (data from CaPSURE). *J Urol*. 2003;170(6 Pt 2):S21-S25; discussion S26-S27. doi:10.1097/01.ju.0000095025.03331.c6
8. Eifler JB, Feng Z, Lin BM, et al. An updated prostate cancer staging nomogram (Partin tables) based on cases from 2006 to 2011. *BJU Int*. 2013;111(1):22-29. doi:10.1111/j.1464-410X.2012.11324.x
9. van Vugt HA, Roobol MJ, Kranse R, et al. Prediction of prostate cancer in unscreened men: external validation of a risk calculator. *Eur J Cancer Oxf Engl 1990*. 2011;47(6):903-909. doi:10.1016/j.ejca.2010.11.012
10. Hernandez DJ, Nielsen ME, Han M, Partin AW. Contemporary evaluation of the D'amico risk classification of prostate cancer. *Urology*. 2007;70(5):931-935. doi:10.1016/j.urology.2007.08.055
11. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. *N Engl J Med*. 2012;367(3):203-213. doi:10.1056/NEJMoa1113162
12. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol*. 2017;71(4):618-629. doi:10.1016/j.eururo.2016.08.003
13. Prostate cancer: diagnosis and management | Guidance and guidelines | NICE. https://www.nice.org.uk/guidance/cg175/chapter/1-Recommendations#localised-and-locally-advanced-prostate-cancer-2. Accessed September 25, 2017.
14. Nakabayashi M, Hayes J, Taplin M-E, et al. Clinical predictors of survival in men with castration-resistant prostate cancer: evidence that Gleason score 6 cancer can evolve to lethal disease. *Cancer*. 2013;119(16):2990-2998. doi:10.1002/cncr.28102

15. Cunha GR, Hayward SW, Wang YZ, Ricke WA. Role of the stromal microenvironment in carcinogenesis of the prostate. *Int J Cancer J Int Cancer*. 2003;107(1):1-10. doi:10.1002/ijc.11335
16. Gravina GL, Mancini A, Ranieri G, et al. Phenotypic characterization of human prostatic stromal cells in primary cultures derived from human tissue samples. *Int J Oncol*. 2013;42(6):2116-2122. doi:10.3892/ijo.2013.1892
17. Wang ZA, Mitrofanova A, Bergren SK, et al. Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. *Nat Cell Biol*. 2013;15(3):274-283. doi:10.1038/ncb2697
18. Perreard L, Fan C, Quackenbush JF, et al. Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res BCR*. 2006;8(2):R23. doi:10.1186/bcr1399
19. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*. 2005;353(16):1673-1684. doi:10.1056/NEJMoa052122
20. 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):805-816. doi:10.1016/S0140-6736(12)61963-1
21. Dent S, Oyan B, Honig A, Mano M, Howell S. HER2-targeted therapy in breast cancer: a systematic review of neoadjuvant trials. *Cancer Treat Rev*. 2013;39(6):622-631. doi:10.1016/j.ctrv.2013.01.002
22. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol*. 2014;65(1):124-137. doi:10.1016/j.eururo.2013.09.046
23. Shore ND, Karsh L, Gomella LG, Keane TE, Concepcion RS, Crawford ED. Avoiding obsolescence in advanced prostate cancer management: a guide for urologists. *BJU Int*. 2015;115(2):188-197.
24. Doyen J, Alix-Panabières C, Hofman P, et al. Circulating tumor cells in prostate cancer: a potential surrogate marker of survival. *Crit Rev Oncol Hematol*. 2012;81(3):241-256. doi:10.1016/j.critrevonc.2011.05.004
25. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. *Nat Rev Cancer*. 2014;14(9):611-622. doi:10.1038/nrc3793
26. Klein CA. Selection and adaptation during metastatic cancer progression. *Nature*. 2013;501(7467):365-372. doi:10.1038/nature12628
27. Ousset M, Van Keymeulen A, Bouvencourt G, et al. Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nat Cell Biol*. 2012;14(11):1131-1138. doi:10.1038/ncb2600
28. Shen MM, Abate-Shen C. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev*. 2010;24(18):1967-2000. doi:10.1101/gad.1965810
29. Robinson EJ, Neal DE, Collins AT. Basal cells are progenitors of luminal cells in primary cultures of differentiating human prostatic epithelium. *The Prostate*. 1998;37(3):149-160.
30. English HF, Santen RJ, Isaacs JT. Response of glandular versus basal rat ventral prostatic epithelial cells to androgen withdrawal and replacement. *The Prostate*. 1987;11(3):229-242.

31. Zhou J, Feigenbaum L, Yee C, Song H, Yates C. Mouse prostate epithelial luminal cells lineage originate in the basal layer where the primitive stem/early progenitor cells reside: implications for identifying prostate cancer stem cells. *BioMed Res Int*. 2013;2013:913179. doi:10.1155/2013/913179
32. Blackwood JK, Williamson SC, Greaves LC, et al. In situ lineage tracking of human prostatic epithelial stem cell fate reveals a common clonal origin for basal and luminal cells. *J Pathol*. 2011;225(2):181-188. doi:10.1002/path.2965
33. Petkova N, Hennenlotter J, Sobiesiak M, et al. Surface CD24 distinguishes between low differentiated and transit-amplifying cells in the basal layer of human prostate. *The Prostate*. 2013;73(14):1576-1590. doi:10.1002/pros.22708
34. Mulholland DJ, Xin L, Morim A, Lawson D, Witte O, Wu H. Lin-Sca-1+CD49fhigh stem/progenitors are tumor-initiating cells in the Pten-null prostate cancer model. *Cancer Res*. 2009;69(22):8555-8562. doi:10.1158/0008-5472.CAN-08-4673
35. Tsujimura A, Koikawa Y, Salm S, et al. Proximal location of mouse prostate epithelial stem cells: a model of prostatic homeostasis. *J Cell Biol*. 2002;157(7):1257-1265. doi:10.1083/jcb.200202067
36. Wang X, Kruithof-de Julio M, Economides KD, et al. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature*. 2009;461(7263):495-500. doi:10.1038/nature08361
37. Liu J, Pascal LE, Isharwal S, et al. Regenerated luminal epithelial cells are derived from preexisting luminal epithelial cells in adult mouse prostate. *Mol Endocrinol Baltim Md*. 2011;25(11):1849-1857. doi:10.1210/me.2011-1081
38. Ju X, Casimiro MC, Gormley M, et al. Identification of a cyclin D1 network in prostate cancer that antagonizes epithelial-mesenchymal restraint. *Cancer Res*. 2014;74(2):508-519. doi:10.1158/0008-5472.CAN-13-1313
39. Boström PJ, Bjartell AS, Catto JWF, et al. Genomic Predictors of Outcome in Prostate Cancer. *Eur Urol*. April 2015. doi:10.1016/j.eururo.2015.04.008
40. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer*. 2007;7(11):834-846. doi:10.1038/nrc2256
41. Stott SL, Lee RJ, Nagrath S, et al. Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. *Sci Transl Med*. 2010;2(25):25ra23. doi:10.1126/scitranslmed.3000403
42. Pantel K, Alix-Panabières C. Bone marrow as a reservoir for disseminated tumor cells: a special source for liquid biopsy in cancer patients. *BoneKEy Rep*. 2014;3:584. doi:10.1038/bonekey.2014.79
43. Shiozawa Y, Pedersen EA, Havens AM, et al. Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *J Clin Invest*. 2011;121(4):1298-1312. doi:10.1172/JCI43414
44. Kim JK, Jung Y, Wang J, et al. TBK1 regulates prostate cancer dormancy through mTOR inhibition. *Neoplasia N Y N*. 2013;15(9):1064-1074.
45. Lam H-M, Vessella RL, Morrissey C. The Role of the Microenvironment – Dormant Prostate Disseminated Tumor Cells in the Bone Marrow. *Drug Discov Today Technol*. 2014;11:41-47. doi:10.1016/j.ddtec.2014.02.002
46. Sigurdsson S, Thorlacius S, Tomasson J, et al. BRCA2 mutation in Icelandic prostate cancer patients. *J Mol Med Berl Ger*. 1997;75(10):758-761.
47. Castro E, Eeles R. The role of BRCA1 and BRCA2 in prostate cancer. *Asian J Androl*. 2012;14(3):409-414. doi:10.1038/aja.2011.150
48. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747-752. doi:10.1038/35021093

49. Eroles P, Bosch A, Pérez-Fidalgo JA, Lluch A. Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer Treat Rev*. 2012;38(6):698-707. doi:10.1016/j.ctrv.2011.11.005
50. Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. *Science*. 2010;329(5991):568-571. doi:10.1126/science.1189992
51. Parsons JK, Gage WR, Nelson WG, De Marzo AM. p63 protein expression is rare in prostate adenocarcinoma: implications for cancer diagnosis and carcinogenesis. *Urology*. 2001;58(4):619-624.
52. Dabir PD, Ottosen P, Høyer S, Hamilton-Dutoit S. Comparative analysis of three- and two-antibody cocktails to AMACR and basal cell markers for the immunohistochemical diagnosis of prostate carcinoma. *Diagn Pathol*. 2012;7:81. doi:10.1186/1746-1596-7-81
53. Heinlein CA, Chang C. Androgen receptor in prostate cancer. *Endocr Rev*. 2004;25(2):276-308. doi:10.1210/er.2002-0032
54. Fujimura T, Takahashi S, Urano T, et al. Expression of androgen and estrogen signaling components and stem cell markers to predict cancer progression and cancer-specific survival in patients with metastatic prostate cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2014;20(17):4625-4635. doi:10.1158/1078-0432.CCR-13-1105
55. Park S, Koo JS, Kim MS, et al. Characteristics and outcomes according to molecular subtypes of breast cancer as classified by a panel of four biomarkers using immunohistochemistry. *Breast Edinb Scotl*. 2012;21(1):50-57. doi:10.1016/j.breast.2011.07.008
56. Banerjee S, Reis-Filho JS, Ashley S, et al. Basal-like breast carcinomas: clinical outcome and response to chemotherapy. *J Clin Pathol*. 2006;59(7):729-735. doi:10.1136/jcp.2005.033043
57. Kreike B, van Kouwenhove M, Horlings H, et al. Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Res BCR*. 2007;9(5):R65. doi:10.1186/bcr1771
58. Rao C, Shetty J, Prasad KH. Immunohistochemical profile and morphology in triple - negative breast cancers. *J Clin Diagn Res JCDR*. 2013;7(7):1361-1365. doi:10.7860/JCDR/2013/5823.3129
59. Niu Y, Chang T-M, Yeh S, Ma W-L, Wang YZ, Chang C. Differential androgen receptor signals in different cells explain why androgen-deprivation therapy of prostate cancer fails. *Oncogene*. 2010;29(25):3593-3604. doi:10.1038/onc.2010.121
60. Markert EK, Mizuno H, Vazquez A, Levine AJ. Molecular classification of prostate cancer using curated expression signatures. *Proc Natl Acad Sci U S A*. 2011;108(52):21276-21281. doi:10.1073/pnas.1117029108
61. Rye MB, Bertilsson H, Drabl?s F, Angelsen A, Bathen TF, Tessem M-B. Gene signatures ESC, MYC and ERG-fusion are early markers of a potentially dangerous subtype of prostate cancer. *BMC Med Genomics*. 2014;7:50. doi:10.1186/1755-8794-7-50
62. Ben-Porath I, Thomson MW, Carey VJ, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet*. 2008;40(5):499-507. doi:10.1038/ng.127
63. Eaton CL, Colombel M, van der Pluijm G, et al. Evaluation of the frequency of putative prostate cancer stem cells in primary and metastatic prostate cancer. *The Prostate*. 2010;70(8):875-882. doi:10.1002/pros.21121

64. Colombel M, Eaton CL, Hamdy F, et al. Increased expression of putative cancer stem cell markers in primary prostate cancer is associated with progression of bone metastases. *The Prostate*. 2012;72(7):713-720. doi:10.1002/pros.21473
65. Scher HI. HER2 in prostate cancer--a viable target or innocent bystander? *J Natl Cancer Inst*. 2000;92(23):1866-1868.
66. Minner S, Jessen B, Stiedenroth L, et al. Low level HER2 overexpression is associated with rapid tumor cell proliferation and poor prognosis in prostate cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2010;16(5):1553-1560. doi:10.1158/1078-0432.CCR-09-2546
67. Ahmad I, Patel R, Singh LB, et al. HER2 overcomes PTEN (loss)-induced senescence to cause aggressive prostate cancer. *Proc Natl Acad Sci U S A*. 2011;108(39):16392-16397. doi:10.1073/pnas.1101263108
68. Solit DB, Rosen N. Targeting HER2 in prostate cancer: where to next? *J Clin Oncol Off J Am Soc Clin Oncol*. 2007;25(3):241-243. doi:10.1200/JCO.2006.08.8187
69. James ND, Sydes MR, Clarke NW, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet Lond Engl*. 2016;387(10024):1163-1177. doi:10.1016/S0140-6736(15)01037-5
70. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med*. 2014;370(10):932-942. doi:10.1056/NEJMoa1311593
71. Epstein JI, Allsbrook WC, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005;29(9):1228-1242.
72. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol*. 2016;40(2):244-252. doi:10.1097/PAS.0000000000000530
73. Epstein JI, Allsbrook WC, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005;29(9):1228-1242.
74. Epstein JI, Allsbrook WC, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005;29(9):1228-1242.
75. Green WJ, Ball G, Hulman G, et al. KI67 and DLX2 predict increased risk of metastasis formation in prostate cancer-a targeted molecular approach. *Br J Cancer*. 2016;115(2):236-242. doi:10.1038/bjc.2016.169
76. Lancashire LJ, Lemetre C, Ball GR. An introduction to artificial neural networks in bioinformatics—application to complex microarray and mass spectrometry datasets in cancer studies. *Brief Bioinform*. 2009;10(3):315-329. doi:10.1093/bib/bbp012
77. Tong DL, Boocock DJ, Dhondalay GKR, Lemetre C, Ball GR. Artificial neural network inference (ANNI): a study on gene-gene interaction for biomarkers in childhood sarcomas. *PloS One*. 2014;9(7):e102483. doi:10.1371/journal.pone.0102483

78. Lancashire LJ, Rees RC, Ball GR. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. *Artif Intell Med*. 2008;43(2):99-111. doi:10.1016/j.artmed.2008.03.001
79. Combining Artificial Neural Networks and Transrectal Ultrasound in the Diagnosis of Prostate Cancer | Cancer Network. http://www.cancernetwork.com/review-article/combining-artificial-neural-networks-and-transrectal-ultrasound-diagnosis-prostate-cancer-1. Published October 1, 2003. Accessed November 6, 2016.
80. Hu X, Cammann H, Meyer H-A, Miller K, Jung K, Stephan C. Artificial neural networks and prostate cancer—tools for diagnosis and management. *Nat Rev Urol*. 2013;10(3):174-182. doi:10.1038/nrurol.2013.9
81. Wang Y, Xia X-Q, Jia Z, et al. In silico estimates of tissue components in surgical samples based on expression profiling data. *Cancer Res*. 2010;70(16):6448-6455. doi:10.1158/0008-5472.CAN-10-0021
82. Abdel-Fatah TM, Powe DG, Agboola J, et al. The biological, clinical and prognostic implications of p53 transcriptional pathways in breast cancers. *J Pathol*. 2010;220(4):419-434. doi:10.1002/path.2663
83. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). *Breast Cancer Res Treat*. 2006;100(2):229-235. doi:10.1007/s10549-006-9242-8
84. Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med*. 2004;350(22):2239-2246. doi:10.1056/NEJMoa031918
85. de la Taille A. Progensa PCA3 test for prostate cancer detection. *Expert Rev Mol Diagn*. 2007;7(5):491-497. doi:10.1586/14737159.7.5.491
86. Fisher G, Yang ZH, Kudahetti S, et al. Prognostic value of Ki-67 for prostate cancer death in a conservatively managed cohort. *Br J Cancer*. 2013;108(2):271-277. doi:10.1038/bjc.2012.598
87. Berney DM, Gopalan A, Kudahetti S, et al. Ki-67 and outcome in clinically localised prostate cancer: analysis of conservatively treated prostate cancer patients from the Trans-Atlantic Prostate Group study. *Br J Cancer*. 2009;100(6):888-893. doi:10.1038/sj.bjc.6604951
88. Dey P, Jonsson P, Hartman J, Williams C, Ström A, Gustafsson J-Å. Estrogen receptors β1 and β2 have opposing roles in regulating proliferation and bone metastasis genes in the prostate cancer cell line PC3. *Mol Endocrinol Baltim Md*. 2012;26(12):1991-2003. doi:10.1210/me.2012.1227
89. Ren S, Kang MR, Wang J, et al. Targeted induction of endogenous NKX3-1 by small activating RNA inhibits prostate tumor growth. *The Prostate*. 2013;73(14):1591-1601. doi:10.1002/pros.22709
90. Wang P, Ma Q, Luo J, et al. Nkx3.1 and p27(KIP1) cooperate in proliferation inhibition and apoptosis induction in human androgen-independent prostate cancer cells. *Cancer Invest*. 2009;27(4):369-375. doi:10.1080/07357900802232749
91. Shand RL, Gelmann EP. Molecular biology of prostate-cancer pathogenesis. *Curr Opin Urol*. 2006;16(3):123-131. doi:10.1097/01.mou.0000193384.39351.64
92. Malathi K, Paranjape JM, Ganapathi R, Silverman RH. HPC1/RNASEL mediates apoptosis of prostate cancer cells treated with 2',5'-oligoadenylates, topoisomerase I inhibitors, and tumor necrosis factor-related apoptosis-inducing ligand. *Cancer Res*. 2004;64(24):9144-9151. doi:10.1158/0008-5472.CAN-04-2226

93. Polley M-YC, Leung SCY, McShane LM, et al. An International Ki67 Reproducibility Study. *JNCI J Natl Cancer Inst*. 2013;105(24):1897-1906. doi:10.1093/jnci/djt306
94. Gudlaugsson E, Skaland I, Janssen EAM, et al. Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. *Histopathology*. 2012;61(6):1134-1144. doi:10.1111/j.1365-2559.2012.04329.x
95. Polley M-YC, Leung SCY, Gao D, et al. An international study to increase concordance in Ki67 scoring. *Mod Pathol Off J U S Can Acad Pathol Inc*. 2015;28(6):778-786. doi:10.1038/modpathol.2015.38
96. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol*. 2000;182(3):311-322. doi:10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9
97. Antonarakis ES, Keizman D, Zhang Z, et al. An immunohistochemical signature comprising PTEN, MYC, and Ki67 predicts progression in prostate cancer patients receiving adjuvant docetaxel after prostatectomy. *Cancer*. 2012;118(24):6063-6071. doi:10.1002/cncr.27689
98. Kachroo N, Gnanapragasam VJ. The role of treatment modality on the utility of predictive tissue biomarkers in clinical prostate cancer: a systematic review. *J Cancer Res Clin Oncol*. 2013;139(1):1-24. doi:10.1007/s00432-012-1351-7
99. Khor L-Y, Bae K, Paulus R, et al. MDM2 and Ki-67 predict for distant metastasis and mortality in men treated with radiotherapy and androgen deprivation for prostate cancer: RTOG 92-02. *J Clin Oncol Off J Am Soc Clin Oncol*. 2009;27(19):3177-3184. doi:10.1200/JCO.2008.19.8267
100. Pollack A, DeSilvio M, Khor L-Y, et al. Ki-67 staining is a strong predictor of distant metastasis and mortality for men with prostate cancer treated with radiotherapy plus androgen deprivation: Radiation Therapy Oncology Group Trial 92-02. *J Clin Oncol Off J Am Soc Clin Oncol*. 2004;22(11):2133-2140. doi:10.1200/JCO.2004.09.150
101. Verhoven B, Yan Y, Ritter M, et al. Ki-67 is an independent predictor of metastasis and cause-specific mortality for prostate cancer patients treated on Radiation Therapy Oncology Group (RTOG) 94-08. *Int J Radiat Oncol Biol Phys*. 2013;86(2):317-323. doi:10.1016/j.ijrobp.2013.01.016
102. Cuzick J, Berney DM, Fisher G, et al. Prognostic value of a cell cycle progression signature for prostate cancer death in a conservatively managed needle biopsy cohort. *Br J Cancer*. 2012;106(6):1095-1099. doi:10.1038/bjc.2012.39
103. Braithwaite AW, Del Sal G, Lu X. Some p53-binding proteins that can function as arbiters of life and death. *Cell Death Differ*. 2006;13(6):984-993. doi:10.1038/sj.cdd.4401924
104. Karayi MK, Markham AF. Molecular biology of prostate cancer. *Prostate Cancer Prostatic Dis*. 2004;7(1):6-20. doi:10.1038/sj.pcan.4500697
105. Yilmaz M, Maass D, Tiwari N, et al. Transcription factor Dlx2 protects from TGFβ-induced cell-cycle arrest and apoptosis. *EMBO J*. 2011;30(21):4489-4499. doi:10.1038/emboj.2011.319
106. Bodnar AG, Ouellette M, Frolkis M, et al. Extension of life-span by introduction of telomerase into normal human cells. *Science*. 1998;279(5349):349-352.
107. Kastan MB, Lim DS. The many substrates and functions of ATM. *Nat Rev Mol Cell Biol*. 2000;1(3):179-186. doi:10.1038/35043058

108. Wang Y, Xu Q, Sack L, Kang C, Elledge SJ. A gain-of-function senescence bypass screen identifies the homeobox transcription factor DLX2 as a regulator of ATM-p53 signaling. *Genes Dev*. 2016;30(3):293-306. doi:10.1101/gad.271445.115
109. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 2005;310(5748):644-648. doi:10.1126/science.1117679
110. Demichelis F, Fall K, Perner S, et al. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene*. 2007;26(31):4596-4599. doi:10.1038/sj.onc.1210237
111. Rubio-Briones J, Fernández-Serra A, Calatrava A, et al. Clinical implications of TMPRSS2-ERG gene fusion expression in patients with prostate cancer treated with radical prostatectomy. *J Urol*. 2010;183(5):2054-2061. doi:10.1016/j.juro.2009.12.096
112. Tomlins SA, Day JR, Lonigro RJ, et al. Urine TMPRSS2:ERG Plus PCA3 for Individualized Prostate Cancer Risk Assessment. *Eur Urol*. 2016;70(1):45-53. doi:10.1016/j.eururo.2015.04.039
113. Lahdensuo K, Erickson A, Saarinen I, et al. Loss of PTEN expression in ERG-negative prostate cancer predicts secondary therapies and leads to shorter disease-specific survival time after radical prostatectomy. *Mod Pathol Off J U S Can Acad Pathol Inc*. August 2016. doi:10.1038/modpathol.2016.154
114. Squire JA. TMPRSS2-ERG and PTEN loss in prostate cancer. *Nat Genet*. 2009;41(5):509-510. doi:10.1038/ng0509-509
115. Bismar TA, Yoshimoto M, Vollmer RT, et al. PTEN genomic deletion is an early event associated with ERG gene rearrangements in prostate cancer: ERG REARRANGEMENTS AND PTEN DELETIONS IN PROSTATE CANCER. *BJU Int*. 2011;107(3):477-485. doi:10.1111/j.1464-410X.2010.09470.x
116. Williams, JL. TMPRSS2:ETS gene fusions in prostate cancer. *Atlas Genet Cytogenet Oncol Haematol*. 2011;15(8):699-716.
117. Simper NB, Jones CL, MacLennan GT, et al. Basal cell carcinoma of the prostate is an aggressive tumor with frequent loss of PTEN expression and overexpression of EGFR. *Hum Pathol*. 2015;46(6):805-812. doi:10.1016/j.humpath.2015.02.004
118. Boccaccio C, Comoglio PM. Invasive growth: a MET-driven genetic programme for cancer and stem cells. *Nat Rev Cancer*. 2006;6(8):637-645. doi:10.1038/nrc1912
119. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351(27):2817-2826. doi:10.1056/NEJMoa041588
120. Zanotti L, Bottini A, Rossi C, Generali D, Cappelletti MR. Diagnostic tests based on gene expression profile in breast cancer: from background to clinical use. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. July 2014. doi:10.1007/s13277-014-2366-2
121. Knezevic D, Goddard AD, Natraj N, et al. Analytical validation of the Oncotype DX prostate cancer assay - a clinical RT-PCR assay optimized for prostate needle biopsies. *BMC Genomics*. 2013;14:690. doi:10.1186/1471-2164-14-690
122. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene Assay to Predict Prostate Cancer Aggressiveness in the Context of Gleason Grade Heterogeneity, Tumor Multifocality, and Biopsy Undersampling. *Eur Urol*. May 2014. doi:10.1016/j.eururo.2014.05.004

123. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495-501. doi:10.1038/nature12912
124. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013
125. Cyll K, Ersvær E, Vlatkovic L, et al. Tumour heterogeneity poses a significant challenge to cancer biomarker research. *Br J Cancer*. 2017;117(3):367-375. doi:10.1038/bjc.2017.171
126. Tradonsky A, Rubin T, Beck R, Ring B, Seitz R, Mair S. A search for reliable molecular markers of prognosis in prostate cancer: a study of 240 cases. *Am J Clin Pathol*. 2012;137(6):918-930. doi:10.1309/AJCPF3QWIG8FWXIH

8.0 Appendix

SPSS data output for Ki67 and DLX2 examining associations between biomarker expression and CaP metastases and death

Ki67 and time (months) to CaP-specific death

Case Processing Summary

Ki67with110positive	Total N	N of Events	Censored	
			N	Percent
.00	57	57	0	0.0%
1.00	8	8	0	0.0%
Overall	65	65	0	0.0%

Means and Medians for Survival Time

Ki67with110positive	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
.00	47.456	4.429	38.774	56.138	39.000	4.847	29.500	48.500
1.00	26.375	5.028	16.520	36.230	21.000	3.536	14.070	27.930
Overall	44.862	4.018	36.986	52.737	38.000	4.031	30.100	45.900

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	5.222	1	.022
Breslow (Generalized Wilcoxon)	3.396	1	.065

Test of equality of survival distributions for the different levels of Ki67with110positive.

Survival Table

Ki67with110positive		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
.00	1	2.000	1.00	.982	.017	1	56
	2	5.000	1.00	.965	.024	2	55
	3	8.000	1.00	.	.	3	54
	4	8.000	1.00	.930	.034	4	53
	5	10.000	1.00	.	.	5	52
	6	10.000	1.00	.895	.041	6	51
	7	11.000	1.00	.877	.043	7	50
	8	13.000	1.00	.860	.046	8	49
	9	15.000	1.00	.	.	9	48
	10	15.000	1.00	.825	.050	10	47
	11	17.000	1.00	.	.	11	46
	12	17.000	1.00	.789	.054	12	45
	13	19.000	1.00	.	.	13	44
	14	19.000	1.00	.754	.057	14	43
	15	20.000	1.00	.737	.058	15	42
	16	28.000	1.00	.	.	16	41
	17	28.000	1.00	.702	.061	17	40
	18	29.000	1.00	.684	.062	18	39
	19	30.000	1.00	.667	.062	19	38
	20	31.000	1.00	.	.	20	37
	21	31.000	1.00	.632	.064	21	36
	22	32.000	1.00	.	.	22	35
	23	32.000	1.00	.	.	23	34
	24	32.000	1.00	.579	.065	24	33

	25	34.000	1.00	.561	.066	25	32
	26	35.000	1.00	.544	.066	26	31
	27	38.000	1.00	.526	.066	27	30
	28	39.000	1.00	.	.	28	29
	29	39.000	1.00	.	.	29	28
	30	39.000	1.00	.474	.066	30	27
	31	42.000	1.00	.456	.066	31	26
	32	43.000	1.00	.439	.066	32	25
	33	46.000	1.00	.421	.065	33	24
	34	47.000	1.00	.404	.065	34	23
	35	48.000	1.00	.	.	35	22
	36	48.000	1.00	.368	.064	36	21
	37	50.000	1.00	.351	.063	37	20
	38	53.000	1.00	.333	.062	38	19
	39	55.000	1.00	.	.	39	18
	40	55.000	1.00	.298	.061	40	17
	41	56.000	1.00	.281	.060	41	16
	42	57.000	1.00	.263	.058	42	15
	43	58.000	1.00	.246	.057	43	14
	44	61.000	1.00	.228	.056	44	13
	45	72.000	1.00	.211	.054	45	12
	46	83.000	1.00	.193	.052	46	11
	47	86.000	1.00	.175	.050	47	10
	48	88.000	1.00	.158	.048	48	9
	49	89.000	1.00	.140	.046	49	8
	50	90.000	1.00	.123	.043	50	7
	51	94.000	1.00	.105	.041	51	6

	52	104.000	1.00	.088	.037	52	5
	53	106.000	1.00	.070	.034	53	4
	54	112.000	1.00	.053	.030	54	3
	55	115.000	1.00	.035	.024	55	2
	56	121.000	1.00	.018	.017	56	1
	57	140.000	1.00	.000	.000	57	0
1.00	1	8.000	1.00	.875	.117	1	7
	2	13.000	1.00	.750	.153	2	6
	3	19.000	1.00	.625	.171	3	5
	4	21.000	1.00	.500	.177	4	4
	5	24.000	1.00	.375	.171	5	3
	6	35.000	1.00	.250	.153	6	2
	7	44.000	1.00	.125	.117	7	1
	8	47.000	1.00	.000	.000	8	0

Ki67 and time (months) to metastases

Case Processing Summary

Ki67with110positive	Total N	N of Events	Censored	
			N	Percent
.00	45	45	0	0.0%
1.00	5	5	0	0.0%
Overall	50	50	0	0.0%

Survival Table

Ki67with110positive		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
.00	1	.000	1.00	.978	.022	1	44
	2	1.000	1.00	.956	.031	2	43
	3	2.000	1.00	.933	.037	3	42
	4	3.000	1.00	.	.	4	41
	5	3.000	1.00	.889	.047	5	40
	6	4.000	1.00	.867	.051	6	39
	7	7.000	1.00	.844	.054	7	38
	8	12.000	1.00	.822	.057	8	37
	9	17.000	1.00	.800	.060	9	36
	10	18.000	1.00	.	.	10	35
	11	18.000	1.00	.756	.064	11	34
	12	19.000	1.00	.733	.066	12	33
	13	20.000	1.00	.711	.068	13	32
	14	22.000	1.00	.	.	14	31
	15	22.000	1.00	.667	.070	15	30
	16	23.000	1.00	.644	.071	16	29
	17	24.000	1.00	.	.	17	28
	18	24.000	1.00	.	.	18	27
	19	24.000	1.00	.578	.074	19	26
	20	26.000	1.00	.	.	20	25
	21	26.000	1.00	.533	.074	21	24
	22	27.000	1.00	.511	.075	22	23
	23	29.000	1.00	.489	.075	23	22
	24	30.000	1.00	.467	.074	24	21

	25	31.000	1.00	.444	.074	25	20
	26	32.000	1.00	.422	.074	26	19
	27	35.000	1.00	.400	.073	27	18
	28	37.000	1.00	.378	.072	28	17
	29	43.000	1.00	.356	.071	29	16
	30	45.000	1.00	.333	.070	30	15
	31	50.000	1.00	.311	.069	31	14
	32	52.000	1.00	.289	.068	32	13
	33	68.000	1.00	.267	.066	33	12
	34	79.000	1.00	.244	.064	34	11
	35	80.000	1.00	.222	.062	35	10
	36	84.000	1.00	.200	.060	36	9
	37	85.000	1.00	.178	.057	37	8
	38	102.000	1.00	.156	.054	38	7
	39	103.000	1.00	.133	.051	39	6
	40	105.000	1.00	.111	.047	40	5
	41	110.000	1.00	.089	.042	41	4
	42	113.000	1.00	.067	.037	42	3
	43	125.000	1.00	.044	.031	43	2
	44	127.000	1.00	.022	.022	44	1
	45	139.000	1.00	.000	.000	45	0
1.00	1	8.000	1.00	.800	.179	1	4
	2	14.000	1.00	.	.	2	3
	3	14.000	1.00	.400	.219	3	2
	4	19.000	1.00	.200	.179	4	1
	5	34.000	1.00	.000	.000	5	0

Means and Medians for Survival Time

Ki67with110positive	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
.00	45.467	5.909	33.885	57.049	29.000	3.912	21.332	36.668
1.00	17.800	4.409	9.158	26.442	14.000	3.286	7.559	20.441
Overall	42.700	5.457	32.003	53.397	26.000	3.094	19.937	32.063

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	5.058	1	.025
Breslow (Generalized Wilcoxon)	3.825	1	.050

Test of equality of survival distributions for the different levels of Ki67with110positive.

DLX2 and time (months) to CaP-specific death

Case Processing Summary

DLX2with10cutoff	Total N	N of Events	Censored	
			N	Percent
.00	23	23	0	0.0%
1.00	61	61	0	0.0%
Overall	84	84	0	0.0%

Survival Table

DLX2with10cutoff		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
.00	1	8.000	1.00	.	.	1	22
	2	8.000	1.00	.913	.059	2	21
	3	11.000	1.00	.870	.070	3	20
	4	17.000	1.00	.826	.079	4	19
	5	20.000	1.00	.783	.086	5	18
	6	22.000	1.00	.739	.092	6	17
	7	30.000	1.00	.696	.096	7	16
	8	31.000	1.00	.652	.099	8	15
	9	32.000	1.00	.609	.102	9	14
	10	38.000	1.00	.565	.103	10	13
	11	42.000	1.00	.522	.104	11	12
	12	43.000	1.00	.478	.104	12	11
	13	48.000	1.00	.435	.103	13	10
	14	61.000	1.00	.391	.102	14	9
	15	77.000	1.00	.348	.099	15	8
	16	82.000	1.00	.304	.096	16	7
	17	83.000	1.00	.261	.092	17	6
	18	86.000	1.00	.217	.086	18	5
	19	88.000	1.00	.174	.079	19	4
	20	90.000	1.00	.130	.070	20	3
	21	112.000	1.00	.087	.059	21	2
	22	133.000	1.00	.043	.043	22	1
	23	140.000	1.00	.000	.000	23	0
1.00	1	1.000	1.00	.984	.016	1	60

2	2.000	1.00	.967	.023	2	59
3	5.000	1.00	.951	.028	3	58
4	8.000	1.00	.934	.032	4	57
5	10.000	1.00	.918	.035	5	56
6	11.000	1.00	.902	.038	6	55
7	13.000	1.00	.	.	7	54
8	13.000	1.00	.869	.043	8	53
9	15.000	1.00	.852	.045	9	52
10	17.000	1.00	.836	.047	10	51
11	19.000	1.00	.	.	11	50
12	19.000	1.00	.	.	12	49
13	19.000	1.00	.787	.052	13	48
14	21.000	1.00	.	.	14	47
15	21.000	1.00	.754	.055	15	46
16	24.000	1.00	.	.	16	45
17	24.000	1.00	.721	.057	17	44
18	25.000	1.00	.705	.058	18	43
19	26.000	1.00	.689	.059	19	42
20	28.000	1.00	.	.	20	41
21	28.000	1.00	.656	.061	21	40
22	29.000	1.00	.639	.061	22	39
23	31.000	1.00	.623	.062	23	38
24	32.000	1.00	.	.	24	37
25	32.000	1.00	.	.	25	36
26	32.000	1.00	.574	.063	26	35
27	33.000	1.00	.557	.064	27	34
28	34.000	1.00	.541	.064	28	33

29	38.000	1.00	.525	.064	29	32
30	40.000	1.00	.508	.064	30	31
31	41.000	1.00	.492	.064	31	30
32	42.000	1.00	.475	.064	32	29
33	44.000	1.00	.459	.064	33	28
34	46.000	1.00	.	.	34	27
35	46.000	1.00	.426	.063	35	26
36	47.000	1.00	.410	.063	36	25
37	48.000	1.00	.393	.063	37	24
38	50.000	1.00	.377	.062	38	23
39	53.000	1.00	.361	.061	39	22
40	54.000	1.00	.344	.061	40	21
41	55.000	1.00	.	.	41	20
42	55.000	1.00	.311	.059	42	19
43	56.000	1.00	.	.	43	18
44	56.000	1.00	.279	.057	44	17
45	57.000	1.00	.262	.056	45	16
46	58.000	1.00	.246	.055	46	15
47	59.000	1.00	.230	.054	47	14
48	60.000	1.00	.213	.052	48	13
49	66.000	1.00	.197	.051	49	12
50	67.000	1.00	.180	.049	50	11
51	72.000	1.00	.164	.047	51	10
52	76.000	1.00	.	.	52	9
53	76.000	1.00	.131	.043	53	8
54	83.000	1.00	.115	.041	54	7
55	89.000	1.00	.098	.038	55	6

56	104.000	1.00	.082	.035	56	5
57	106.000	1.00	.066	.032	57	4
58	115.000	1.00	.049	.028	58	3
59	116.000	1.00	.033	.023	59	2
60	121.000	1.00	.016	.016	60	1
61	123.000	1.00	.000	.000	61	0

Means and Medians for Survival Time

DLX2with10cutoff	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
.00	56.609	8.233	40.472	72.746	43.000	7.985	27.348	58.652
1.00	45.754	3.942	38.028	53.481	41.000	6.345	28.564	53.436
Overall	48.726	3.655	41.562	55.890	42.000	5.498	31.225	52.775

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	2.282	1	.131
Breslow (Generalized Wilcoxon)	.892	1	.345

Test of equality of survival distributions for the different levels of DLX2with10cutoff.

DLX2 and time (months) to metastases

Case Processing Summary

DLX2with10cutoff	Total N	N of Events	Censored	
			N	Percent
.00	19	19	0	0.0%
1.00	42	42	0	0.0%
Overall	61	61	0	0.0%

Survival Table

DLX2with10cutoff		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
.00	1	3.000	1.00	.947	.051	1	18
	2	24.000	1.00	.	.	2	17
	3	24.000	1.00	.842	.084	3	16
	4	29.000	1.00	.789	.094	4	15
	5	30.000	1.00	.737	.101	5	14
	6	35.000	1.00	.684	.107	6	13
	7	42.000	1.00	.632	.111	7	12
	8	48.000	1.00	.579	.113	8	11
	9	52.000	1.00	.526	.115	9	10
	10	74.000	1.00	.474	.115	10	9
	11	79.000	1.00	.421	.113	11	8
	12	84.000	1.00	.368	.111	12	7
	13	85.000	1.00	.316	.107	13	6
	14	105.000	1.00	.263	.101	14	5
	15	110.000	1.00	.211	.094	15	4

	16	125.000	1.00	.158	.084	16	3
	17	127.000	1.00	.105	.070	17	2
	18	131.000	1.00	.053	.051	18	1
	19	139.000	1.00	.000	.000	19	0
1.00	1	.000	1.00	.	.	1	41
	2	.000	1.00	.	.	2	40
	3	.000	1.00	.929	.040	3	39
	4	1.000	1.00	.905	.045	4	38
	5	3.000	1.00	.881	.050	5	37
	6	4.000	1.00	.857	.054	6	36
	7	7.000	1.00	.833	.058	7	35
	8	11.000	1.00	.810	.061	8	34
	9	12.000	1.00	.786	.063	9	33
	10	14.000	1.00	.762	.066	10	32
	11	17.000	1.00	.738	.068	11	31
	12	18.000	1.00	.	.	12	30
	13	18.000	1.00	.690	.071	13	29
	14	19.000	1.00	.	.	14	28
	15	19.000	1.00	.643	.074	15	27
	16	20.000	1.00	.	.	16	26
	17	20.000	1.00	.595	.076	17	25
	18	22.000	1.00	.	.	18	24
	19	22.000	1.00	.548	.077	19	23
	20	23.000	1.00	.524	.077	20	22
	21	24.000	1.00	.500	.077	21	21
	22	26.000	1.00	.476	.077	22	20
	23	27.000	1.00	.452	.077	23	19

24	30.000	1.00	.	.	24	18
25	30.000	1.00	.405	.076	25	17
26	31.000	1.00	.381	.075	26	16
27	36.000	1.00	.357	.074	27	15
28	37.000	1.00	.333	.073	28	14
29	43.000	1.00	.	.	29	13
30	43.000	1.00	.	.	30	12
31	43.000	1.00	.262	.068	31	11
32	44.000	1.00	.238	.066	32	10
33	45.000	1.00	.214	.063	33	9
34	50.000	1.00	.190	.061	34	8
35	64.000	1.00	.167	.058	35	7
36	71.000	1.00	.143	.054	36	6
37	80.000	1.00	.119	.050	37	5
38	102.000	1.00	.095	.045	38	4
39	103.000	1.00	.071	.040	39	3
40	113.000	1.00	.048	.033	40	2
41	115.000	1.00	.024	.024	41	1
42	120.000	1.00	.000	.000	42	0

Means and Medians for Survival Time

DLX2with10cutoff	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
.00	70.842	9.809	51.617	90.067	74.000	22.490	29.920	118.080
1.00	36.357	5.131	26.301	46.414	24.000	4.050	16.061	31.939
Overall	47.098	5.062	37.176	57.021	31.000	4.462	22.254	39.746

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	10.207	1	.001
Breslow (Generalized Wilcoxon)	9.475	1	.002

Test of equality of survival distributions for the different levels of DLX2with10cutoff.

4.4.5 Multivariate Analysis

Cox Regression

Omnibus Tests of Model Coefficients[a]

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
381.945	5.619	4	.229	4.629	4	.327	4.629	4	.327

a. Beginning Block Number 1. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Ki67with110positive	.890	.408	4.769	1	.029	2.436	1.096	5.416
PSAnew	.000	.000	1.121	1	.290	1.000	1.000	1.000
Gleasoncategorised	.018	.222	.006	1	.936	1.018	.659	1.572
Riskcategorised	-.085	.497	.029	1	.864	.919	.347	2.433

Covariate Means

	Mean
Ki67with110positive	.131
PSAnew	321.787
Gleasoncategorised	2.639
Riskcategorised	2.918

Cox Regression

Omnibus Tests of Model Coefficients[a]

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
754.571	20.937	4	.000	20.840	4	.000	20.840	4	.000

a. Beginning Block Number 1. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Ki67with110positive	.957	.369	6.729	1	.009	2.605	1.264	5.368
PSAnew	.000	.000	.989	1	.320	1.000	1.000	1.000
Gleasoncategorised	.133	.166	.646	1	.421	1.143	.825	1.582
Riskcategorised	.771	.344	5.023	1	.025	2.161	1.102	4.240

Covariate Means

	Mean
Ki67with110positive	.086
PSAnew	220.286
Gleasoncategorised	2.467
Riskcategorised	2.857

Cox Regression

Omnibus Tests of Model Coefficients[a]

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
266.620	9.628	4	.047	7.507	4	.111	7.507	4	.111

a. Beginning Block Number 1. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Ki67with110positive	1.193	.523	5.191	1	.023	3.296	1.181	9.196
PSAnew	.001	.001	2.282	1	.131	1.001	1.000	1.002
Gleasoncategorised	.385	.265	2.109	1	.146	1.469	.874	2.469
Riskcategorised	-.512	.519	.970	1	.325	.600	.217	1.660

Covariate Means

	Mean
Ki67with110positive	.106
PSAnew	136.149
Gleasoncategorised	2.596
Riskcategorised	2.894

Cox Regression

Omnibus Tests of Model Coefficients[a]

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
307.007	16.199	4	.003	14.375	4	.006	14.375	4	.006

a. Beginning Block Number 1. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
PSAnew	.001	.001	3.207	1	.073	1.001	1.000	1.002
Gleasoncategorised	.081	.257	.098	1	.755	1.084	.654	1.795
Riskcategorised	-.147	.525	.078	1	.780	.864	.309	2.414
DLX2with10cutoff	1.013	.337	9.038	1	.003	2.754	1.423	5.332

Covariate Means

	Mean
PSAnew	113.415
Gleasoncategorised	2.604
Riskcategorised	2.887
DLX2with10cutoff	.642