JOHN VAN GEEST
CANCER RESEARCH CENTRE
Nottingham Trent University

Improving the Diagnostic Yield of Prostate Cancer

Masood Ahmed Khan

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of 'Doctor of Philosophy by Published Work'.

October 2021

Copyright Statement

Signature: …………………………   **Date:** …01/10/2021…

Acknowledgements

I am very grateful to my wife, Tahsina, for her enduring support over the past twenty-five years. I am also very grateful to our daughter, Sahar, and son, Raei, who have both been a source of inspiration to me.

I would like to offer a special thanks to Professor Graham Pockley who encouraged me to submit this thesis and for his tremendous support and guidance in compiling the thesis.

Table of Contents

Abbreviations

ALL	Acute Lymphoblastic Leukaemia
ANN	Artificial Neural Network
AUC	Area Under the Curve
BPH	Benign Prostatic Hyperplasia
bPSA	benign Prostate Specific Antigen
DRE	Digital Rectal Examination
EPD	Extra-Prostatic Disease
FDA	Food and Drug Administration
FPR	False Positive Rate
fPSA	free Prostate Specific Antigen
IL-2	Interleukin 2
NK	Natural Killer
OCD	Organ Confined Disease
PBMC(s)	Peripheral Blood Mononuclear Cell(s)
PCa	Prostate Cancer
PHI	Prostate Health Index
PSA	Prostate Specific Antigen
PSAP	Prostate Specific Acid Phosphatase
tPSA	total Prostate Specific Antigen
ROC	Receiver Operator Characteristic
TCGA	The Cancer Genome Atlas
TPR	True Positive Rate
TPTPB	Transperineal Template Prostate Biopsy(ies)
TRUS	Transrectal Ultrasound

Improving the Diagnostic Yield of Prostate Cancer

Abstract

Although the introduction of PSA as a biomarker for detecting prostate cancer has enabled the disease to be diagnosed at an earlier stage, the primary drawback to PSA as a biomarker and basis for prostate cancer screening and diagnosis is its lack of specificity. To address this issue, PSA density, velocity and percent free PSA have all been assessed as potential approaches for improving the diagnosis of prostate cancer, albeit with variable and limited success.

These challenges and the desire to improve the management and treatment of patients with prostate cancer prompted studies that investigated the capacity of measuring pro-PSA levels to improve diagnostic yield over that of PSA alone. Our study confirmed that combining pro-PSA, total and percent free PSA improved the specificity of prostate cancer detection from that of 23% for PSA alone to 44%. To further improve the utility of pro-PSA, we subsequently determined that the ratio of pro-PSA and benign-PSA can identify the presence of prostate cancer with greater accuracy when the percent free PSA is below 15%.

Traditionally, prostate cancer has been definitively diagnosed by performing transrectal ultrasound (TRUS) guided prostate biopsies. However, such a biopsy technique has a cancer detection rate of less than 30% in a benign feeling prostate. In addition, when TRUS biopsies are repeated due to rising PSA, the cancer detection rate significantly reduces to below 10% for men undergoing a third set of such biopsies. We therefore undertook a study to assess the diagnostic strength of the transperineal template prostate biopsies (TPTPB) in this group of men and demonstrated that the cancer detection rate of the TPTPB is significantly better than that of the TRUS biopsy (52%-68%). We subsequently directly compared TRUS against TPTPB in biopsy naïve men and revealed that TPTPB significantly outperforms TRUS in prostate cancer detection (60% *versus* 32%).

Based on the established concept that there is a reciprocal relationship between cancer and the immune system, and the proposition that the presence of cancer will influence the phenotype of immune cells in the blood which can be detected by profiling peripheral

blood mononuclear cells, we subsequently investigated whether profiling the phenotype of immune cells in the blood could further improve our ability not only to detect prostate cancer, but clinically significant prostate cancer. For this, flow cytometric immune profiling of the peripheral blood from men with benign prostate disease and patients with confirmed prostate cancer identified phenotypic features 'fingerprints' within the lymphocyte populations which, when incorporated into machine learning based algorithms, can be used to distinguish between the presence of benign prostate disease and prostate cancer. Furthermore, we identified a panel of eight natural killer (NK) cell phenotypic features which can be used to very accurately differentiate between high risk and low/intermediate risk prostate cancer when incorporated into machine learning algorithms.

In summary, over the past two decades, my work has not only resulted in an improvement in the utility of PSA as a biomarker for the detection of prostate, but I also demonstrated that performing transperineal prostate biopsies significantly improves prostate cancer detection. Furthermore, we have recently revealed that our immune system can aid us to differentiate between indolent and clinically significant prostate cancer. Taken together, this programme of work has greatly impacted clinical practice.

Biography

After qualifying as a doctor from Cambridge University, I completed by basic surgical training on the Oxford rotation. I subsequently obtained my Doctor of Medicine (MD) research degree from the Royal Free Hospital, University of London, before commencing my higher Urology training in London. During my Urology training, I had the opportunity to expand my clinical and research knowledge by working as a Fellow for 2 years at the Brady Urological Institute at Johns Hopkins Hospital, Baltimore, USA (2002-2004). In my first year I undertook prostate cancer research and obtained my Master of Surgery (MS) degree, also from the University of London.

After completing my training in London, I was appointed as a Consultant Urological Surgeon at University Hospitals of Leicester in April 2006, for which I was Head-of-Service for Urology between August 2009 and April 2015.

My passion for prostate cancer research, and in particular, for improving the diagnosis of prostate cancer, started during my time at John Hopkins and continues today despite running a very busy clinical practice. My desire for advancing the field of prostate cancer research led to the start of a very fruitful collaborative relationship with the John van Geest Cancer Research Centre (JvGCRC) at Nottingham Trent University in 2012. I was initially appointed as a Visiting Professor to Nottingham Trent University in March 2012 and became an Honorary Professor in 2020.

I have a strong academic background that has resulted in more than 120 peer-reviewed publications and greater than 2000 citations.

My long-term collaboration with the JvGCRC has led to the awarding of a number of grants as well as many peer-reviewed publications.

Summary of Approach

There is no definitive structure for a thesis submitted for a 'Doctor of Philosophy by Published Work' at Nottingham Trent University, rather the content and structure can vary according to the content of the individual study / body of work. This submission therefore provides a list of the peer-reviewed publications that form the basis to the PhD followed by introduction to the field and the knowledge to which the work is contributing. Each of the published papers is then reviewed, following which is a copy of the published work to which reference is being made. Published works contain all the relevant literature review, rationale, methods, results and discussions that were peer-reviewed and published. The aim of this structured approach is to demonstrate the progression and significance of this programme of work and its coherency with regards to *Improving the Diagnostic Yield of Prostate Cancer.* Integrated into the critical review are copies of the published papers contained in the publications list.

The Research / Clinical Question

How do we improve the diagnosis of prostate cancer beyond the utilisation of PSA and digital rectal examination (DRE) alone? Unfortunately, measuring blood PSA levels lacks specificity and DRE lacks both sensitivity and specificity. Despite the shortfalls of both PSA as a biomarker for the detection of prostate cancer and our ability to suspect prostate cancer based on an abnormal DRE, we have traditionally relied on these two tools to determine whether a man has prostate cancer. Accordingly, men are subjected to invasive prostate biopsies to detect prostate cancer, which is present in only a minority of those that are biopsied. Prostate biopsies are not only associated with significant complications such as urosepsis, bleeding and urinary retention, PSA and DRE measurements do not necessarily differentiate between clinically significant prostate cancer, which requires treatment, and indolent cancer, for which the current recommendation is to 'watch and wait' or active surveillance. The challenge over the past two decades has therefore not only been to improve the diagnostic yield of prostate cancer, but also to develop new approaches for more specifically distinguishing between benign prostate disease and prostate cancer and, arguably more importantly, between low-risk disease which requires no treatment and clinically significant disease which requires treatment.

Summary Analysis of Included Publications and Candidate Contributions

PUBLICATION 1: Khan MA, Partin AW, Rittenhouse HG, Mikolajczyk SD, Sokoll LJ, Chan DW, Veltri RW. Evaluation of proprostate specific antigen for early detection of prostate cancer in men with a total prostate specific antigen range of 4.0 to 10.0 ng/ml. *Journal of Urology* 2003, 170:723-6. **doi:** 10.1097/01.ju.0000086940.10392.93

Citations: 115

Summary/Abstract: Total PSA consists of bound and free PSA. It has previously been demonstrated that percent free PSA (free PSA divided by total PSA) improves the specificity of prostate cancer diagnosis in men with an overall elevated total PSA. However, free PSA consists of several different components. We determined that measuring Sum-proPSA, total PSA and percent-free PSA, in combination, further improves the specificity of early prostate cancer detection in men with a total PSA of 4 to 10 ng/ml, compared with the results when individual molecular forms PSA are measured.

Contribution: As the 1ˢᵗ author, I conducted the study, analysed the data and wrote the paper.

PUBLICATION 2: Khan MA, Sokoll LJ, Chan DW, Mangold LA, Mohr P, Mikolajczyk SD, Linton HJ, Evans CL, Rittenhouse HG, Partin AW. Clinical utility of proPSA and "benign" PSA when percent free PSA is less than 15%.
Urology 2004, 64:1160-4. **doi:** 10.1016/j.urology.2004.06.033.

Citations: 69

Summary/Abstract: Although a level of percent-free PSA (%fPSA) below 15% is thought to be more likely associated with the presence of prostate cancer, this is not necessarily the case. In order to improve our ability to detect prostate cancer in men with a %fPSA below 15% we determined the clinical utility of the various subforms of fPSA, namely proPSA and "benign" PSA (BPSA). The results of our study revealed that the ratio of proPSA and BPSA can distinguish cancer with greater accuracy in this group of men, and may therefore provide better clinical utility in this lower range of percent free PSA.

Contribution: I performed the study, analysed the data and wrote the paper.

PUBLICATION 3: Pal RP, Elmussareh M, Chanawani M, **Khan MA**. The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies. *BJU International* 2012, 109:367-71. **doi:** 10.1111/j.1464-410X.2011.10355.x.

Citations: 55

Summary/Abstract: Traditionally we have relied on transrectal ultrasound (TRUS) guided prostate biopsies to determine the presence of prostate cancer. However, TRUS guided prostate biopsies can only interrogate the *posterior* aspect of the prostate. As such, if prostate cancer is present in the *anterior* area of the prostate, then such cancers are missed and the prostate is falsely diagnosed as negative for prostate cancer. These men are then are at a risk of representing at a future date with either more advanced or metastatic prostate cancer. In addition, as TRUS guided prostate biopsies are performed through the rectum, there is a risk of experiencing urosepsis. In view of the limitations and risks associated with TRUS guided prostate biopsies, we determined whether performing prostate biopsies using the transperineal route, which can interrogate the entire prostate, improves the detection of prostate cancer in men who have had two previous sets of negative TRUS guided prostate biopsies due to a rising PSA. Our study revealed that the transperineal approach does indeed significantly improve our ability to detect prostate cancer.

Contribution: As the senior author, I co-ordinated the study, performed all the biopsies and supervised the writing of the paper.

PUBLICATION 4: Nafie S, Pal RP, Dormer JP, **Khan MA**. Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies. *World Journal of Urology* 2014, 32:971-5. **doi:** 10.1007/s00345-013-1225-x.

Citations: 20

Summary/Abstract: Having shown that, when directly compared, transperineal biopsies are significantly better at detecting prostate cancer than TRUS guided biopsies in previously biopsy naïve men, we then determined whether this was still the case in the difficult group of men who have had two previous sets of negative TRUS guided prostate biopsies. We again performed both transperineal and TRUS guided prostate biopsies on the same occasion in order to rule out bias. Our unique study revealed that the

transperineal approach continues to be of benefit in detecting prostate cancer in men who have previously undergone two sets of negative TRUS guided prostate biopsies.

Contribution: As the senior author, I devised the study, performed all of the biopsies and supervised the writing of the paper. I was also involved in recruiting the patients.

PUBLICATION 5: Nafie S, Mellon JK, Dormer JP, **Khan MA**. The role of transperineal template prostate biopsies in prostate cancer diagnosis in biopsy naïve men with PSA less than 20 ng.ml^{-1}.

Prostate Cancer Prostatic Disease 2014, 17:170-3. **doi:** 10.1038/pcan.2014.4.

Citations: 28

Summary/Abstract: Having previously demonstrated the superiority of the transperineal approach in performing prostate biopsies over TRUS guided prostate biopsies in men who have previously had two negative sets of TRUS guided prostate biopsies, we determined if this benefit holds true in biopsy naïve men. Accordingly, ***we were the first group in the world to directly compare transperineal and TRUS guided prostate biopsies*** by performing both procedures at the same occasion. Our study revealed that transperineal prostate biopsies are significantly better at detecting prostate cancer than TRUS guided prostate biopsies.

Contribution: As the senior author, I devised the study, performed all of the biopsies and supervised the writing of the paper. I was also involved in recruiting the patients.

PUBLICATION 6: Nafie S, Wanis M, **Khan M.** The efficacy of Transrectal Ultrasound Guided Biopsy *versus* Transperineal Template Biopsy of the prostate in diagnosing prostate cancer in men with previous negative Transrectal Ultrasound Guided Biopsy.

Urology Journal 2017, 14:3008-12. **doi:** 10.22037/uj.v14i2.3702

Citations: 13

Summary/Abstract: We further determined whether the transperineal approach was of benefit and better than TRUS guided prostate biopsies in detecting prostate cancer in men who have a rising PSA despite one previous set of negative TRUS guided prostate biopsies. We again performed both of the procedures on the same occasion and determined that the transperineal approach outperformed TRUS guided prostate biopsies. ***These series of studies have had a paradigm shift in the way that we now biopsy men with an elevated PSA***. We now only perform transperineal prostate biopsies in men with a rising PSA and a background of a single set of previously negative TRUS guided

prostate biopsies. In addition, the vast majority of biopsy naïve men are also only considered for transperineal prostate biopsies. Now, the only group of men that we would currently consider for TRUS guided prostate biopsies are those who have a clearly palpably abnormal prostate on rectal examination requiring only a few biopsies.

Contribution: As the senior author, I devised the study, performed all of the biopsies and supervised the writing of the paper. I was also involved in recruiting the patients.

PUBLICATION 7: Nafie S, Berridge C, **Khan M**. Novel technique in performing standard Transperineal Template Prostate Biopsy under local anaesthetic.
Urology Journal, submitted 2021.

Summary/Abstract: Transperineal prostate biopsies have now become the standard method of obtaining histology for the diagnosis of prostate cancer. However, as these biopsies have traditionally been performed under general anaesthetic, it is not feasible for this to continue in the future due to the limited access we have to operating theatres in the NHS. We therefore devised a novel technique in performing such biopsies under local anaesthetic without compromising our cancer detection in 250 men. The study confirmed that the cancer detection rate was 62% in previously biopsy naïve men, which is very similar to the cancer detection rate when transperineal biopsies are performed under general anaesthetic.

Contribution: I devised this novel technique and performed all of the biopsies. I also assisted in the writing of the paper.

PUBLICATION 8: Cosma G, Acampora G, Brown D, Rees RC, **Khan M**, Pockley AG. Prediction of pathological stage in patients with prostate cancer: A Neuro-Fuzzy model.
PLoS ONE 2016, 11(6): e0155856. **doi:** 10.1371/journal.pone.0155856.
Citations: 42

Summary/Abstract: The prediction of cancer staging in prostate cancer is a process for estimating the likelihood that the cancer has spread before treatment is given to the patient. Although important for determining the most suitable treatment and optimal management strategy for patients, staging continues to present significant challenges to clinicians. Clinical test results such as the pre-treatment Prostate-Specific Antigen (PSA) level, the biopsy most common tumour pattern (Primary Gleason pattern) and the second most common tumour pattern (Secondary Gleason pattern) in tissue biopsies, and the clinical T stage can be used by clinicians to predict the pathological stage of cancer.

However, not every patient will return abnormal results in all tests. Herein we have developed a neuro-fuzzy computational intelligence model for classifying and predicting the likelihood of a patient having Organ-Confined Disease (OCD) or Extra-Prostatic Disease (ED) using a prostate cancer patient dataset obtained from The Cancer Genome Atlas (TCGA) Research Network. The system input consisted of the following variables: Primary and Secondary Gleason biopsy patterns, PSA levels, age at diagnosis, and clinical T stage. The performance of the neuro-fuzzy system was compared to other computational intelligence-based approaches, namely the Artificial Neural Network, Fuzzy C-Means, Support Vector Machine, the Naïve Bayes classifiers, and also the AJCC pTNM Staging Nomogram which is commonly used by clinicians. A comparison of the optimal Receiver Operating Characteristic (ROC) points that were identified using these approaches, revealed that the neuro-fuzzy system, at its optimal point, returns the largest Area Under the Curve (AUC), with a low number of false positives (FPR = 0.274, TPR = 0.789, AUC = 0.812). The proposed approach is also an improvement over the AJCC pTNM Staging Nomogram (FPR = 0.032, TPR = 0.197, AUC = 0.582).

Contribution: As a co-senior author and clinical lead, I supervised the study and the writing of the paper.

PUBLICATION 9: Cosma G, McArdle SE, Reeder S, Foulds G, Hood S, **Khan M**, Pockley AG. Identifying the presence of prostate cancer in individuals with PSA levels <20 ng/ml using computational data extraction analysis of high dimensional peripheral blood flow cytometric phenotyping data.
Frontiers in Immunology 2017, 8:1771. **doi:** 10.3389/fimmu.2017.01771.
Citations: 7

Summary/Abstract: Determining whether an asymptomatic individual with Prostate-Specific Antigen (PSA) levels below 20 ng/ml has prostate cancer in the absence of definitive, biopsy-based evidence continues to present a significant challenge to clinicians who must decide whether such individuals with low PSA values have prostate cancer.

Herein, we present an advanced computational data extraction approach which can identify the presence of prostate cancer in men with PSA levels <20 ng/ml based on peripheral blood immune cell profiles that have been generated using multi-parameter flow cytometry.

Statistical analysis of immune phenotyping datasets relating to the presence and prevalence of key leukocyte populations in the peripheral blood, as generated from individuals undergoing routine tests for prostate cancer (including tissue biopsy) using multi-parametric flow cytometric analysis, was unable to identify significant relationships between leukocyte population profiles and the presence of benign disease (no prostate cancer) or prostate cancer. By contrast, a Genetic Algorithm computational approach identified a subset of five flow cytometry features; $CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory Cells); $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA); $CD3^-CD19^+$ (B cells); $CD3^+CD56^+CD8^+CD4^+$ (NKT cells) from a set of twenty features, which could potentially discriminate between benign disease and prostate cancer. These features were used to construct a prostate cancer prediction model using the k-Nearest-Neighbour classification algorithm. The proposed model, which takes as input the set of flow cytometry features, outperformed the predictive model which takes PSA values as input. Specifically, the flow cytometry-based model achieved Accuracy = 83.33%, AUC = 83.40%, and optimal ROC points of FPR = 16.13%, TPR = 82.93%, whereas the PSA-based model achieved Accuracy = 77.78%, AUC = 76.95%, and optimal ROC points of FPR = 29.03%, TPR = 82.93%. Combining PSA and flow cytometry predictors achieved Accuracy = 79.17%, AUC = 78.17% and optimal ROC points of FPR = 29.03%, TPR = 85.37%.

The results demonstrate the value of computational intelligence-based approaches for interrogating immunophenotyping datasets and that combining peripheral blood phenotypic profiling with PSA levels improves diagnostic accuracy compared to using PSA test alone. These studies also demonstrate that the presence of cancer is reflected in changes in the peripheral blood immune phenotype profile which can be identified.

Contribution: As a senior co-author and the clinical lead, I recruited the patients to the study and collected all samples. In addition, I co-supervised the study and the writing of the paper.

PUBLICATION 10: Hood SP, Foulds GA, Imrie H, Reeder S, McArdle SE, **Khan MA**, Pockley AG. Phenotype and function of activated natural killer cells from patients with prostate cancer: Patient-dependent responses to priming and IL-2 activation.

Frontiers in Immunology 2019, 9:3169. **doi:** 10.3389/fimmu.2018.03169.

Citations: 3

Summary/Abstract: Although immunotherapy has emerged as the "next generation" of cancer treatments, it has not yet been shown to be successful in the treatment of patients with prostate cancer, for whom therapeutic options remain limited to radiotherapy and androgen (hormone) deprivation therapy. Previous studies have shown that priming natural killer (NK) cells isolated from healthy individuals via co-incubation with CTV-1 cells derived from an acute lymphoblastic leukaemia (ALL) enhances their cytotoxicity against human DU145 (metastatic) prostate cancer cells, but it remains unknown to what extent NK cells from patients with prostate cancer can be triggered to kill.

Herein, we explore the phenotype of peripheral blood NK cells in patients with prostate cancer and compare the capacity of CTV-1 cell-mediated priming and IL-2 stimulation to trigger NK cell-mediated killing of the human PC3 (metastatic) prostate cancer cell line.

The phenotype of resting, primed (co-incubation with CTV-1 cells for 17 h) and IL-2 activated (100 IU/ml IL-2 for 17 h) NK cells isolated from frozen-thawed peripheral blood mononuclear cell (PBMC) preparations from patients with benign disease (n = 6) and prostate cancer (n = 18) and their cytotoxicity against PC3 and K562 cells was determined by flow cytometry. Relationships between NK cell phenotypic features and cytotoxic potential were interrogated using Spearman Rank correlation matrices.

NK cell priming and IL-2 activation of patient-derived NK cells resulted in similar levels of cytotoxicity, but distinct NK cell phenotypes. Importantly, the capacity of priming and IL-2 stimulation to trigger cytotoxicity was patient-dependent and mutually exclusive, in that NK cells from ~50% of patients preferentially responded to priming whereas NK cells from the remaining patients preferentially responded to cytokine stimulation. In addition to providing more insight into the biology of primed and cytokine-stimulated NK cells, this study supports the use of autologous NK cell-based immunotherapies for the treatment of prostate cancer. However, our findings also indicate that patients will need to be stratified according to their potential responsiveness to individual therapeutic approaches.

Contribution: As a senior co-author and the clinical lead, I recruited the patients to the study and collected all the samples. In addition, I co-supervised the study and the writing of the paper.

Declaration: Material contained in this publication formed part of a submission for the award of PhD entitled 'Characterising the phenotype and function of natural killer cells

in patients with prostate cancer' which was awarded to Simon P. Hood by Nottingham Trent University(Hood 2016).

PUBLICATION 11: Hood SP, Cosma G, Foulds GA, Johnson C, Reeder S, McArdle SE, **Khan MA**, Pockley AG. Identifying prostate cancer and its clinical risk in asymptomatic men using Machine Learning of high dimensional peripheral blood flow cytometric natural killer cell subset phenotyping data

eLife 2020, e50936. **doi:** 10.7554/eLife.50936.

Citations: 9

Summary/Abstract: We demonstrate that prostate cancer can be identified by flow cytometric profiling of blood immune cell subsets. Herein, we profiled natural killer (NK) cell subsets in the blood of 72 asymptomatic men with Prostate-Specific Antigen (PSA) levels < 20 ng/ml, of whom 31 had benign disease (no cancer) and 41 had prostate cancer. Statistical and computational methods identified a panel of eight phenotypic features: $CD56^{dim}CD16^{high}$, $CD56^{+}DNAM-1^{-}$, $CD56^{+}LAIR-1^{+}$, $CD56^{+}LAIR-1^{-}$, $CD56^{bright}CD8^{+}$, $CD56^{+}NKp30^{+}$, $CD56^{+}NKp30^{-}$, $CD56^{+}NKp46^{+}$ that, when incorporated into an Ensemble machine learning prediction model, distinguished between the presence of benign prostate disease and prostate cancer. The machine learning model was then adapted to predict the D'Amico Risk Classification using data from 54 patients with prostate cancer and was shown to accurately differentiate between the presence of low-/ intermediate-risk disease and high-risk disease without the need for additional clinical data. This simple blood test has the potential to transform prostate cancer diagnostics.

Contribution: As a senior co-author and clinical lead, I recruited the patients to the study and collected all the samples. In addition, I co-supervised the study and the writing of the paper.

Background and Context

The published papers being considered for this PhD encompass work that has taken place over the past two decades. Since the early 1990's PSA has been regarded as the primary diagnostic tool for the detection of prostate cancer. However, due to its lack of specificity, PSA has been found to be of limited value. Although PSA is still used as a starting point in determining whether a person has prostate cancer we now rely on other factors as such pro-PSA in 'fine tuning' our diagnostic skills. In addition, prior to a decade ago biopsies obtained for the histological confirmation of prostate cancer relied wholly on the transrectal approach. This is not only associated with the risk of urosepsis but lacks specificity for detecting prostate cancer. With the publication of evidence that the transperineal template approach to prostate biopsies is not only safer for patients, but significantly more likely to detect prostate cancer, today, the vast majority of prostate biopsies are performed transperineally.

Despite our improved ability to detect prostate cancer by performing transperineal template prostate biopsies, we have not been able to differentiate between men harbouring low-risk disease from men with intermediate/high-risk prostate cancer. This is extremely important as low-risk prostate cancers are very unlikely to impact adversely on men and, therefore, do not need to be diagnosed. As such, with a background research experience of greater than 15 years in investigating our ability to improve the diagnose prostate cancer, I have collaborated with the John van Geest Cancer Centre (JvGCRC) at Nottingham Trent University since 2012. Our aim over the past decade has not only been to improve our ability to detect prostate cancer in men with an elevated PSA, but also improve our ability to differentiate, amongst men with an elevated PSA, those harbouring clinically significant prostate cancer prior to performing prostate biopsies. Our collaborative work has resulted in a number of publications revealing the role of peripheral immune markers and computational models in improving our ability to detect clinically significant prostate cancer. We expect that our findings will, in the future, lead to a dramatic change in our approach to the diagnosis of prostate cancer in men presenting with an elevated PSA test.

Introduction / Critical Study

Incidence and prevalence of prostate cancer

Prostate cancer is the most common non-skin cancer in older males (> 70 years of age) in Europe. It is a major health concern, especially in developed countries with a greater proportion of older men in the general population. Although the incidence is currently the highest in Northern and Western Europe (> 200 per 100,000 men), rates in Eastern and Southern Europe have shown a continuous increase (Arnold, Karim-Kos et al. 2015). There is a survival difference between men diagnosed in Eastern Europe and those in the rest of Europe (De Angelis, Sant et al. 2014). Overall, during the past decade, the 5-year relative survival percentages for prostate cancer has steadily increased from 73.4% in 1999-2001 to 83.4% in 2005-2007 (De Angelis, Sant et al. 2014). With the expected increase in the life expectancy of men and the subsequent rise in the incidence of prostate cancer, the economic burden of prostate cancer in Europe is also expected to increase. It is estimated that the total economic costs of prostate cancer in Europe exceed 8.43 billion Euros (Luengo-Fernandez, Leal et al. 2013), with a high proportion of the costs occurring in the first year after diagnosis. In European countries with available data (UK, Germany, France, Italy, Spain, the Netherlands), this amounted to 106.7-179.0 million Euros for all prostate cancer patients diagnosed in 2006. As such, prostate cancer is an important disease and accounts for over 45,000 new diagnosed cases yearly in the UK (WHO 2014). It is also responsible for greater than 10,000 deaths yearly in the UK (WHO 2014).

Genetic pre-disposition to prostate cancer

Epidemiological studies have shown strong evidence for a genetic predisposition to prostate cancer, based on two of the most important factors, racial/ethnic background and family history (Hemminki 2012, Jansson, Akre et al. 2012). Genome-wide association studies have identified 100 common susceptibility loci that contribute to the risk for prostate cancer (Al Olama, Kote-Jarai et al. 2009). A small subpopulation of men with prostate cancer (about 9%) have true hereditary prostate cancer. This is defined as three or more affected relatives, or at least two relatives who have developed early-onset disease, i.e. before the age of 55 (Hemminki 2012). More than 70 prostate cancer susceptibility loci, explaining ~30% of the familial risk for this disease, have been identified (Eeles, Al Olama et al. 2013). Patients with hereditary prostate cancer usually have a disease onset six to seven years earlier than spontaneous cases, but do not differ

in other ways (Hemminki 2012). The frequency of incidentally- and autopsy-detected cancers is roughly the same in different parts of the world (Haas, Delongchamps et al. 2008). This finding is in sharp contrast to the incidence of clinical prostate cancer, which varies widely between different geographical areas, being high in the USA and Northern Europe and low in South-East Asia. If Japanese men move from Japan to Hawaii, their risk of prostate cancer increases. If they move to California their risk increases even more, approaching that of American men (Breslow, Chan et al. 1977). These findings indicate that exogenous factors affect the risk of progression from so-called latent prostate cancer to clinical prostate cancer. Factors such as diet, sexual behaviour, alcohol consumption, exposure to ultraviolet radiation, chronic inflammation (Nelson, De Marzo et al. 2003, Leitzmann and Rohrmann 2012) and occupational exposure have all been discussed as being aetiologically important (Leitzmann and Rohrmann 2012).

Prostate cancer biomarkers

The National Cancer Institute (NCI) defines a biomarker as being a '*biological molecule found in blood, other body fluids, or tissues that can be objectively measured and evaluated as a sign of a normal/abnormal biological process and a pathogenic condition/disease*'. A biomarker may be used for screening purposes, for disease diagnosis and prognosis, for the evaluation of disease disposition and for the prediction/monitoring of treatment responses to various therapeutic interventions (Ilyin, Belkowski et al. 2004, Sawyers 2008).

Prostate specific acid phosphatase (PSAP) / prostatic acid phosphatase (PAP)

Until the introduction of prostate specific antigen (PSA) as biomarker for the detection of prostate cancer in 1986, the detection of prostate cancer relied on digital rectal examination (DRE) and the prostate specific acid phosphatase (PSAP, also known as prostatic acid phosphatase, PAP) blood test. However, the major drawback of DRE is that it is unreliable and dependent on the individual undertaking the test, in that one person may exam a prostate and regard it as being benign, whereas another might examine the same prostate and consider it to be suspicious for prostate cancer. In addition, a benign feeling prostate does not necessarily equate to a diagnosis of benign disease and men suspected of having prostate cancer based on an abnormal feeling prostate are more likely to harbour advanced prostate cancer. Hence, such prostate cancers are more likely to have spread beyond the prostate capsule and even develop metastatic disease. Unfortunately,

when such groups of men are diagnosed with prostate cancer, the prognosis is limited, and it is too late to be considered for radical treatment with curative intent in the majority of individuals.

The first prostate cancer biomarker, prostatic specific acid phosphatase (PSAP), was described in the 1930s and has since then been used as a clinical marker for the progression of prostate cancer given that serum PSAP levels were found to be elevated in cases of metastatic disease (Griffiths 1980, Epstein and Eggleston 1984). Elevated serum PSAP levels not only indicate the presence of prostate cancer, but also that the individual has developed metastatic disease having a limited prognosis and is therefore not suitable for radical treatment with curative intent. As such, PSAP has been of limited value when used for detecting prostate cancer and is not suitable for the screening of prostate cancer.

Prostate specific antigen (PSA)

Prostate specific antigen (PSA) is a kallikrein-related serine protease produced by the epithelial cells of the prostate gland (Romero Otero, Garcia Gomez et al. 2014). PSA is present in normal prostatic secretions and has the sole purpose of liquifying semen. PSA therefore plays an important role in fertility and, as such, the prostate gland is regarded as one of the four accessory glands of sexual function. PSA was first isolated by Wang et al in 1979 (Wang, Valenzuela et al. 1979) and initially used by the legal world in rape trials to confirm the presence of semen, as PSA is very highly concentrated in semen. Initial reports suggested a role for PSA as a biomarker for monitoring the progression of patients already diagnosed with prostate cancer, or for monitoring recurrence following curative therapy for organ-confined disease (Griffiths 1980, Epstein and Eggleston 1984). In a landmark study, Stamey et al performed the first large-scale analysis of serum PSA as a prostate cancer biomarker in 1987, convincingly demonstrating that PSA was more sensitive than PSAP for monitoring the disease (Stamey, Yang et al. 1987). They showed that PSA level increased with advancing clinical stage and was useful for detecting disease recurrence after curative therapy (Stamey, Yang et al. 1987). Subsequent studies shifted the focus of PSA towards early detection of prostate cancer. In 1986, the U.S. Food and Drug Administration (FDA) approved PSA as an adjunctive test to the DRE for the detection of prostate cancer in men over the age of 50. In 1991, Catalona et al demonstrated that the combination of a serum PSA measurement ≥4.0 ng/ml with other clinical findings, such as the results of a DRE, improved detection of prostate cancer in a

prospective study of 1653 healthy men with no history of cancer (Catalona, Smith et al. 1991). Numerous groups confirmed that PSA was useful as a diagnostic test for prostate cancer (Parkes, Wald et al. 1995).

Unfortunately, although PSA can effectively be regarded as organ-specific, in that it is almost exclusively secreted by the prostate, with minute amounts secreted by breast and pancreas, it is not disease-specific. As such, PSA can be elevated in men with urinary tract infection, catheterised men, men with benign prostatic hyperplasia (BPH - a natural and benign growth of the prostate in men above the age of 40 years), excessive cycling, recent sexual activity, and prostatitis which can be infective or inflammatory. In addition, prostatic massage and urethral instrumentation such as performing a cystoscopy can all lead to a rise in PSA levels. It is, therefore, not possible to distinguish a man with prostate cancer from another with a benign prostate solely based on an elevated PSA. In view of this, all men with an elevated PSA are considered to potentially harbour prostate cancer and are considered for prostate biopsy. As such, the specificity of prostate cancer detection in men with an elevated PSA is low at approximately 30% (Naughton, Miller et al. 2000). In addition, this rate continues to decline with repeated prostate biopsies due to persistently rising PSA (Djavan, Ravery et al. 2001). Recent findings from a decade-long study involving 415,000 British men (The Cluster Randomized Trial of PSA Testing for Prostate Cancer (CAP) Randomized Clinical Trial) have not supported single PSA testing for population-based screening and suggest that asymptomatic men should not be routinely tested to avoid unnecessary anxiety and treatment (Martin, Donovan et al. 2018). Due to its lack of specificity, PSA is therefore not regarded as a suitable biomarker for screening for prostate cancer and as such, unfortunately, at present there is no prostate cancer screening test or programme.

Although PSA has significantly improved our ability to detect prostate cancer at an earlier stage, in that most men diagnosed with prostate cancer today have a benign feeling prostate (Jang, Han et al. 2006), a further flaw with PSA is that it cannot distinguish low-risk prostate cancer from clinically significant prostate cancer. This is extremely important, as men diagnosed with low-risk prostate cancer will be very unlikely to ever experience any problems or adverse outcome due to prostate cancer. Nonetheless, such a group of men will require long-term follow-up and must live with the psychological consequences of knowing that they have a cancer. Such patients will not only experience

invasive investigations which can be life-threatening but will also be financially disadvantaged as their diagnosis of prostate cancer will have a negative impact when obtaining life/travel insurance. Furthermore, such a group of men adds unnecessary financial burden to the NHS as, irrespective of their low-risk disease, they still require regular Urology clinic follow-up with current PSA test values and DRE examinations. Hence, it is imperative that when we diagnose prostate cancer, it is clinically significant disease and not low-risk prostate cancer.

Traditionally, prostate cancer diagnosis has relied on performing transrectal prostate biopsies. Initially this was performed blindly with a very low detection rate. However, to visualise the prostate and improve our ability to obtain biopsies from the appropriate areas of the prostate gland, live ultrasound scan is now routinely used. Until recently, transrectal ultrasound guided prostate biopsies (TRUS) have been regarded as the gold standard approach to the histological diagnosis of men with prostate cancer. Furthermore, sextant biopsies (six areas) from the right and left posterior peripheral zones with two biopsies each from the base, mid and apex areas of the prostate are taken (Norberg, Egevad et al. 1997, Presti 2003). As such, 12 cores are traditionally taken to determine whether prostate cancer is present. Unfortunately, despite such a technique, which has been modified and improved over the past many decades, including the use of a local anaesthetic, thereby making the procedure more tolerable, the cancer detection rate is still very low. Less than 30% of men with an elevated PSA and a benign feeling prostate will have a histological diagnosis of prostate cancer. This rate reduces even further when repeated biopsies are performed at later dates due to a rising PSA (Djavan, Ravery et al. 2001).

In order to improve our ability to diagnose prostate cancer in men with an elevated PSA various approaches to PSA has been made. This included examining age-specific PSA, PSA density, PSA velocity and percent-free PSA. The prostate increases in size from the age of 40 years under the influence of testosterone. It, therefore, makes sense that as a man ages his prostate will continue to increase in size with a naturally resultant rise in detectable PSA levels. Accordingly, the normal PSA level is likely to rise with age. Despite this, previously a PSA level of 4.0 was regarded as a cut-off to distinguish normal from an elevated PSA level. To address the issue of a naturally rising PSA with age, age-

specific normal PSA levels have been established (Oesterling, Jacobsen et al. 1993). This has helped us to avoid unnecessary prostate biopsies in the older population.

PSA density is total PSA divided by the volume of the prostate. The rationale for this is that men with a high PSA density are thought to have prostate cancer due to the cancer area generating significantly more PSA than the benign prostate tissue. As such, a PSA density 0.15 or greater has been shown to modestly improve our ability to detect prostate cancer (Catalona, Southwick et al. 2000). Hence, men with a PSA density of greater than 0.15 are more likely to be considered for a prostate biopsy rather than continue with PSA follow-up/monitoring.

PSA velocity refers to the rate of rise in PSA per year. It has previously been reported that a PSA velocity of greater than 0.75 may indicate the presence of prostate cancer (Carter, Pearson et al. 1992). As such, men with a rapidly rising PSA are more likely to be considered for prostate biopsies over men with a slow-rising PSA.

In 1991 it was discovered that serum contains two distinct, major, immunodetectable forms of PSA. One form covalently binds primarily to a serum protease inhibitor (α1-antichymotrypsin) as well as other less dominant inhibitors and is known as complexed PSA, whereas the other form is a "free", nonactive, noncomplexed form known as free-PSA (fPSA) (Lilja, Christensson et al. 1991, Stenman, Leinonen et al. 1991). The measurement of the ratio of free and total PSA has led to a modest, but significant improvement in the discrimination of prostate cancer from benign disease in men with an elevated PSA (Veltri and Miller 1999, Catalona, Southwick et al. 2000). Despite this, it is estimated that only 30% to 50% of men with percent free PSA less than 15% will have a cancer diagnosis at biopsy (Catalona, Partin et al. 1998).

Therefore, differences within the molecular sub-populations of fPSA have been extensively investigated to further discriminate prostate cancer from benign disease. It is now known that fPSA is composed of at least 3 distinct forms of enzymatically inactive PSA. One of the forms is called benignPSA (bPSA), an internally cleaved or degraded form of PSA that is more highly associated with benign prostatic hyperplasia (benign disease, BPH) (Mikolajczyk, Millar et al. 2000, Linton, Marks et al. 2003). Another form is thought to contain a number of minor variants but appears to be composed largely of

intact PSA that is similar to native, active PSA except for structural or conformational changes that have rendered it enzymatically inactive (Mikolajczyk, Marks et al. 2002). The third has been identified as the proenzyme or precursor form of PSA, also known as proPSA (Mikolajczyk, Marks et al. 2002). Truncated proPSA refers to proPSA in which any of the normal 7 amino acids in the pro-leader peptide have been removed as a result of post-translational proteolytic cleavage of the pro-leader peptide (Khan and James 1998). All forms of the truncated proPSA containing any of the pro-leaded amino acids remain enzymatically inactive. It is these truncated forms of proPSA that have generated a great deal of interest since there is strong evidence to support the association of proPSA with prostate cancer (Kumar, Mikolajczyk et al. 1997, Mikolajczyk, Millar et al. 2000, Mikolajczyk, Marker et al. 2001, Peter, Unverzagt et al. 2001, Mikolajczyk, Marks et al. 2002). To this end, examining prostate tissue from patients undergoing radical prostatectomy, Mickolajczyk et al (Kumar, Mikolajczyk et al. 1997) reported that proPSA was increased in cancer areas compared to noncancerous tissue. The authors subsequently developed monoclonal antibodies to detect proPSA with Western blot analysis in the PSA purified from the serum of five men with biopsy proven prostate cancer (tPSA range 6 to 24 ng/ml, mean 13.4) and 3 men with negative prostate biopsies (tPSA 7 to 12, mean 9.7) (Mikolajczyk, Marker et al. 2001). Multiple proPSA forms were detected at higher levels in serum from patients with cancer whereas the truncated [-2]proPSA form was found to be the most consistently associated with prostate cancer in the five samples tested when compared to samples from individuals with benign disease. Peter et al (Peter, Unverzagt et al. 2001) conducted a similar study using mass spectrometry to detect PSA forms in which they detected [-7], [-5], [-4], [-2] and [-1]proPSA forms in serum from five patients with prostate cancer. The [-2]proPSA was present in four of five samples, and was the highest pro form in two of the samples. The mass spectrometry technique used by Peter et al necessitated the use of serum with much higher serum PSA values (for example, tPSA was 1890 ng/ml in one patient and PSA values were greater than 6000 ng/ml the other four samples). The development of [-2], [-4] and [-5,-7]proPSA immunoassays has enabled us to expand the sample population significantly and to determine whether serum proPSA measurement provides any additional information in the detection of prostate cancer in men with clinically relevant serum tPSA levels ranging from 4 to 10 ng/ml. We therefore assayed for percent-free PSA (%fPSA), total PSA (tPSA) and the three forms of proPSA in archival serum from 93 men who underwent a systematic 12-core prostate biopsy with a total PSA of between 4.0 and 10.0 ng/ml. Free PSA, the cumulative sum of

individual proPSA forms ([-2], [-4] and [-7], or sum-proPSA) and derivatives were determined. Our study revealed that sum-proPSA with the addition of tPSA and %fPSA improved the specificity of prostate cancer detection by greater than 10% at 90% sensitivity compared with any of these serum biomarkers measured individually (Mikolajczyk, Marker et al. 2001). Thus, we concluded that the combination of serum sum-proPSA, tPSA and %fPSA is likely to provide added clinical value resulting in fewer men with a tPSA ranging 4 to 10 ng/ml being unnecessarily biopsied, decreasing repeat biopsy frequency, and could possibly be of use for men undergoing active surveillance for low-risk prostate cancer (**PUBLICATION 1:** (Khan, Partin et al. 2003)).

We subsequently investigated the clinical utility of proPSA and bPSA in improving prostate cancer diagnosis in men with a %fPSA level of less than 15%. This study included archived sera from 161 consecutive men who were prospectively enrolled in the Early Detection Research Network prostate cancer early detection biomarker programme at Johns Hopkins Hospital, Baltimore, USA, with a %fPSA value of less than 15%. TotalPSA, fPSA, proPSA and bPSA were measured for each sample. Our study revealed that the ratio of proPSA/bPSA significantly differentiated between cancer and non-cancer groups with an Area Under Curve-Receiver Operator Characteristic (AUC-ROC) of 0.72. Furthermore, using a cut-off of 0.61 for the proPSA/bPSA ratio achieved a specificity of 46% and sensitivity of 90% for prostate cancer detection (Peter, Unverzagt et al. 2001). In addition, bPSA alone achieved the same overall AUC-ROC as proPSA/bPSA but at a significantly lower specificity of 20% at 90% sensitivity. We also determined that within the subgroup of men presenting with a tPSA level in the range of 4.0 to 10.0 ng/ml, bPSA and the proPSA/bPSA ratio were also able to distinguish between the two groups with near identical results to those of the overall study (**PUBLICATION 2:** (Khan, Sokoll et al. 2004)). Our study therefore revealed that the use of bPSA and the ratio of proPSA/bPSA in men presenting with a low %fPSA level below 15% can accurately distinguish between prostate cancer and non-cancerous / benign disease.

Since the publication of our two studies investigating the potential benefit of proPSA in improving our ability to diagnose prostate cancer in 2003 and 2004 many further studies have further proven the usefulness of proPSA in improving our ability to detect prostate cancer (Stephan, Kahrs et al. 2009, Sokoll, Sanda et al. 2010, Hori, Blanchet et al. 2013). This has led to the subsequent development of the Prostate Health Index (PHI). PHI is a

diagnostic blood test that combines free and total PSA and the (-2) pro-PSA isoform ([-2] proPSA/fPSA) $\times$ PSA$^{1/2}$). This test was developed by Beckman Coulter in partnership with the National Cancer Institute Early Detection Research Network. In prospective multicentre studies, the PHI test has outperformed free and total PSA for detection of prostate cancer and has improved prediction of clinically significant prostate cancer in men with an elevated PSA of up to 10 ng/ml. As such, using the PHI test has revealed that at the 90% sensitivity cut point (a score less than 28.6), 30 % of men could have been spared an unnecessary biopsy for benign disease or insignificant prostate cancer compared to 21% using the percent free PSA approach (Loeb, Sanda et al. 2015). Accordingly, the PHI test is now FDA approved.

Most prostate cancers arise from the peripheral zone with the minority from the transition zone. The major drawback in performing TRUS prostate biopsies is that it is only possible to accurately biopsy the posterior peripheral and transition zone due to limitations in mobility of the ultrasound probe. A negative TRUS biopsy of the prostate does not therefore necessarily equate to a cancer-free prostate, as prostate cancer may be present in the anterior parts of the peripheral or transition zone that are inaccessible via such a route. As such, unfortunately a negative TRUS biopsy could falsely be reassuring to the patient who then subsequently presents later with advanced/metastatic prostate cancer. In addition, the rectum is a high bacteria colonizing organ due to the presence of faeces. Accordingly, approximately 3-5% of men who undergo TRUS guided prostate biopsies will experience urosepsis (Raaijmakers, Kirkels et al. 2002) with many such patients requiring ITU care. Furthermore, a minority of men experiencing urosepsis from prostate biopsies will not survive the experience. Worryingly, the risk of developing urosepsis has increased over the past decade due to the development of multi-drug resistant faecal bacteria (Carlson, Bell et al. 2010).

Despite the risks of developing urosepsis, in an attempt to improve the diagnostic yield of TRUS guided prostate biopsies, saturation TRUS guided prostate biopsies have been performed in the past. This involves far greater numbers of prostate biopsies being taken. The rationale for this was that increasing the number of biopsies taken may, in turn, increase the chance of hitting the area of prostate involved with prostate cancer. However, transrectal saturation biopsies have not increased the prostate cancer detection rate (Jones, Patel et al. 2006). This is not surprising, as repeatedly biopsy of the same area is very

unlikely to give a different result. In view of this, TRUS guided saturation prostate biopsies are not recommended.

To avoid taking biopsies through the rectum and to be in a position to biopsy any part of the prostate, the transperineal approach has gained traction over the past decade. It is now considered as the preferred mode of biopsy for the majority of men presenting with an elevated PSA. Transperineal template prostate biopsies (TPTPB) involve the insertion of a number of needles under sterile condition through the perineal skin (area between the scrotum and anus) to obtain tissue samples from the prostate gland. This procedure is traditionally and currently in the vast majority of instances performed under a general anaesthetic. An ultrasound probe is inserted rectally to visualise the whole of the prostate and stabilised using an articulated arm clamped on to the theatre trolley. This enables us to accurately biopsy any part of the prostate.

I was the first person to perform TPTPB in the East Midlands in January 2008. This was borne out of the necessity to improve the diagnostic yield of prostate cancer in men with persistently rising PSA despite two previous sets of negative TRUS guided prostate biopsies. There was, at that time, a dilemma as to how to manage the growing population of men repeatedly attending the Urology Department due to a continuing rise in PSA with the background of multiple previous negative TRUS guided prostate biopsies. It is not surprising that biopsying the same area repeatedly did not bring about an improvement in the prostate cancer diagnosis rate. Indeed, it has been previously reported that repeated TRUS guided prostate biopsies diminish the cancer detection rate (Djavan, Ravery et al. 2001). As such, we determined the efficacy and safety of TPTPB in 40 men with rising PSA tests despite two previous sets of negative TRUS guided prostate biopsies. For this study, a total of 36 cores were taken under general anaesthetic, six each from the right and left anterior, mid and posterior areas. This method permitted sampling of the anterior peripheral and transition zones that were previously inaccessible via the TRUS route. Our study revealed that 68% of men were diagnosed with prostate cancer (**PUBLICATION 3:** (Pal, Elmussareh et al. 2012)). This was a very large and significant improvement in the cancer detection rate compared to TRUS guided prostate biopsies in such a group of men which has previously reported a less than 10% cancer detection rate (Djavan, Ravery et al. 2001). Furthermore, there were no cases of urosepsis confirming that this approach is significantly safer than TRUS guided prostate biopsies. This study brought about a

paradigm shift in the management of men who presented to our department with a rising PSA despite a previous set of negative TRUS guided prostate biopsies as such men were no longer considered for further TRUS biopsies, but instead TPTPB.

Having published the benefit of TPTPB in a small group of men we subsequently determined whether TPTPB could continue to demonstrate such benefit in a larger group of men. We therefore published our findings of 122 men who underwent TPTPB due to a rising PSA and two previous sets of negative TRUS guided prostate biopsies. This larger study revealed that 58% of men were diagnosed with prostate cancer (**PUBLICATION 4:** (Nafie, Pal et al. 2014)). In addition, in 80% of those in which cancer was detected, the cancer was either in the anterior peripheral zone (59%) or mid transition zone (21%) - the two areas that could not have been accessed via the TRUS route.

Having established the superiority of TPTPB over TRUS guided prostate biopsies in detecting prostate cancer in men with a rising PSA despite previous negative TRUS guided prostate biopsies we determined whether this was still the case for biopsy naïve men with an elevated PSA. As such, we performed a study in which men with an elevated PSA below 20 ng/ml and a benign feeling prostate underwent TRUS guided prostate biopsies and TPTPB on the same occasion. TRUS guided prostate biopsies were performed first followed by TPTPB. Biopsies were performed in this order to avoid criticism that if we had performed TPTPB in the first instance then a negative TRUS guided prostate biopsies may have been due to the TPTPB picking up the posterior peripheral zone cancer. This study investigated 50 men and found that TPTPB detected prostate cancer in 60% of men *versus* 32% of men for the TRUS guided prostate biopsies (**PUBLICATION 5:** (Nafie, Mellon et al. 2014)). The 32% cancer detection rate via the TRUS guided route mirrors well the published data in the world literature regarding this group of men (Djavan, Ravery et al. 2001). Furthermore, if only TRUS guided prostate biopsies were performed, then prostate cancer would have been missed in 28% of men. In addition, no prostate cancer was detected solely on TRUS guided prostate biopsies. This study consolidated our previous findings that TPTPB significantly out-performs TRUS guided prostate biopsies in the detection of prostate cancer.

As the final part of a series of studies we determined whether TPTPB is superior to TRUS guided prostate biopsies in detecting prostate cancer in men with a rising PSA despite a

previous single set of negative TRUS guided prostate biopsies. For this study, we simultaneously performed TRUS guided prostate biopsies and TPTPB in 42 men with a rising PSA, but below 20 ng/ml and a benign feeling prostate on rectal examination, with the background of a previous single set of negative TRUS guided prostate biopsies. TPTPB still demonstrated a significantly higher prostate cancer detection rate compared to TRUS guided prostate biopsies (12% *versus* 45%, P < 0.01) (**PUBLICATION 6**; (Nafie, Wanis et al. 2017))

As a result of the above studies, we now very infrequently perform TRUS guided prostate biopsies. The only instances on which we would now consider performing TRUS guided prostate biopsies is in men with an abnormal feeling prostate which only requires a few targeted biopsies of the abnormal area. With the increasing use of PSA, greater than 80% of prostate cancer today is detected in men with a benign feeling prostate. All such men are now undergoing TPTPB rather than TRUS guided prostate biopsies.

To further improve the detection of prostate cancer in men with a rising PSA, the use of MRI scans has become increasingly common. With advancements in the quality of MRI scans, radiologists can more accurately interpret abnormal areas within a prostate. The PROMIS study was conducted to determine the role of MRI scan in prostate cancer detection. PROMIS was a paired validating confirmatory study determining the role of multi-parametric MRI scan along with TPTPB and TRUS guided prostate biopsies in the detection of prostate cancer in men with an elevated PSA of 15.0 or less with the background of a benign feeling prostate (Ahmed, El-Shater Bosaily et al. 2017). The study revealed that in 25% of men with an MRI detected abnormal area in the prostate there was no significant prostate cancer found on biopsies. In addition, in 20% of men with an underlying unremarkable prostate gland on an MRI scan there was significant prostate cancer confirmed on biopsies. As such, although MRI scanning of the prostate may help us to target abnormal appearing appears, this does not necessarily equate to an area of cancer. Furthermore, a normal appearing area does not rule out prostate cancer. As such, the current guideline is that targeted along with saturation TPTPB should be considered in a man with and elevated PSA and a benign feeling prostate.

Due to the steep rise in the demand for TPTPB there has been a move towards performing such biopsies under local anaesthetic. This has posed significant challenges, as TPTPB

requires the patient to be in a lithotomy position with a transrectal probe to visualise the prostate for anything between 10 to 30 minutes. In addition, the prostate has a rich nerve supply making such a procedure very difficult under a local anaesthetic. Nonetheless, I am the first person in the East Midlands to have performed this procedure under local anaesthetic using a novel technique. I have now performed over 250 TPTPB using local anaesthetic over the past two years without any significant issues and with a high cancer detection rate, thereby demonstrating that local anaesthetic TPTPB is safe and feasible. This has led to over 90% of men now having their TPTPB under local anaesthetic at my institution. Only the minority of men who do not tolerate a DRE are considered for TPTPB under general anaesthetic. Our findings and technique have now been submitted for peer-reviewed publication (**PUBLICATION 7:** (Nafie, Berridge et al. 2021)).

Despite the improved utilisation of serum biomarkers and performing TPTPB leading to a significant increase in the detection of prostate cancer, the major problem that we are still facing today is our inability to predict low/intermediate risk prostate cancer versus high prostate cancer. The large, UK based prospective ProtecT (Prostate Testing for Cancer and Treatment) study randomised low, intermediate, and high-risk prostate cancer, all detected by TRUS guided prostate biopsies, to active surveillance or radical radiotherapy with neoadjuvant and adjuvant hormones or radical prostatectomy (Lane, Donovan et al. 2014). This study has revealed that all low risk and most intermediate risk prostate cancers can be safely managed on long-term active surveillance. In other words, it is very unlikely that anyone diagnosed with low-risk prostate cancer and most men diagnosed with intermediate risk prostate cancer will ever need any active treatment to address their prostate cancer. It is, therefore, important to only consider performing prostate biopsies in men highly likely to harbour high risk prostate cancer (clinically significant disease).

Our ability to do this will not only result in a significant reduction in the financial burden to the NHS as far fewer biopsies will need to be performed in the future, but also reduce unnecessary emotional and financial burden to our patients. As men diagnosed with prostate cancer, irrespective of the category, will experience financial implications when obtaining travel and life insurance. In addition, biopsies are not without their side-effects and risks. Furthermore, even those diagnosed with low-risk prostate cancer will have to

endure a life-long follow-up to rule out the very small risk of eventually developing clinically significant disease requiring active treatment.

When diagnosing prostate cancer, it is also important to have insight into the likely staging of the cancer as this impact on the management the individual and the disease. Cancer staging prediction is a process for estimating the likelihood that the disease has spread before treatment is given to the patient. The evaluation of cancer staging occurs before (i.e. at the prognosis stage) and after (i.e. at the diagnosis stage) the tumour is removed - the 'clinical' and 'pathological' stages respectively. The clinical stage evaluation is based on data gathered from clinical tests that are available prior to treatment or following the surgical removal of the tumour. There are three primary clinical stage tests for prostate cancer: the PSA test; a biopsy which is used to detect the presence of cancer in the prostate and to evaluate the degree of cancer aggressiveness (results are usually given in the form of the Primary and Secondary Gleason patterns); and a physical examination, namely the DRE which can determine the existence of disease and possibly provide sufficient information to predict the stage of the cancer. A limitation of the PSA test is that abnormally high PSA levels may not necessarily indicate the presence of prostate cancer, nor might normal PSA levels reflect the absence of prostate cancer. Pathological staging can be determined following surgery and the examination of the removed tumour tissue and is likely to be more accurate than clinical staging, as it allows a direct insight into the extent and nature of the disease.

Given the potential prognostic power of the clinical tests, a variety of prostate cancer staging prediction systems have been developed. The ability to predict the pathological stage of a patient with prostate cancer is important, as it enables clinicians to better determine the optimal treatment and management strategies. This is to the patient's considerable benefit, as many of the therapeutic options can be associated with significant short- and long- term side-effects. For example, radical prostatectomy, the surgical removal of the prostate gland, offers the best chance for curing the disease when prostate cancer is localised, and the accurate prediction of pathological stage is fundamental to determining which patients would benefit most from this approach (Blute, Nativ et al. 1989, Epstein, Pizov et al. 1993, Epstein, Walsh et al. 1994). Currently, as clinicians we use nomograms to predict a prognostic clinical outcome for prostate cancer, and these are based on statistical methods such as logistic regression (Dotan and Ramon 2009).

However, cancer staging continues to present significant challenges to the clinical community.

The prostate cancer staging nomograms which are used to predict the pathological stage of the cancer are based on results from the clinical tests. However, the accuracy of the nomograms is debatable (Chun, Karakiewicz et al. 2007, Briganti, Karakiewicz et al. 2009). Briganti et al. (Briganti, Karakiewicz et al. 2009) has stated that nomograms are accurate tools and that "Personalized medicine recognizes the need for adjustments, according to disease and host characteristics. It is time to embrace the same attitude in other disciplines of medicine. This includes urologic oncology where nomograms, regression-trees, lookup tables and neural networks represent the key tools capable of providing individualized predictions". However, Dr Joniau (Briganti, Karakiewicz et al. 2009) has expressed that the data used for devising the nomograms are subjective and, to a certain extent, biased by institutional protocols on which patients are selected for a given treatment. Dr Joniau also states that one of the drawbacks of nomograms is that various nomograms have been devised for risk estimation and it is difficult to determine which nomogram will provide the most reliable risk estimation for a particular patient. He emphasises that although nomograms allow for more accurate risk assessment, this risk estimation is a "snapshot in a risk continuum". Although this might allow personalized predictions, it also makes treatment decisions difficult (Briganti, Karakiewicz et al. 2009).

Cancer prediction systems which consider various variables for the prediction of an outcome require computational intelligent methods for efficient prediction outcomes (Tewari, Porter et al. 2001). Although computational intelligence approaches have been used to predict prostate cancer out-comes, very few models for predicting the pathological stage of prostate cancer exist. In essence, classification models based on computational intelligence are utilised for prediction tasks. Classification is a form of data analysis which extracts classifier models describing data classes, and uses these models to predict categorical labels (classes) or numeric values (Han 2005). When the classifier is used to predict a numeric value, as opposed to a class label, it is referred to as a predictor. Classification and numeric prediction are both types of prediction problems (Han 2005), and classification models are widely adopted to analyse patient data and extract a prediction model in the medical setting.

Computational intelligence approaches, and in particular fuzzy-based approaches, are based on mathematical models that are specially developed for dealing with the uncertainty and imprecision which is typically found in the clinical data that are used for prognosis and the diagnosis of diseases in patients. These characteristics make these algorithms a suitable platform on which to base new strategies for diagnosing and staging prostate cancer. For example, not everyone diagnosed with prostate cancer will exhibit abnormal results in all tests, as a consequence of which, different test result combinations can lead to the same outcome.

The capacity of fuzzy, and especially neuro-fuzzy approaches, to predict the pathological stage of prostate cancer has not been as widely evaluated as the more commonly used Artificial Neural Network (ANN) and other approaches. However, fuzzy approaches have been applied to other prostate cancer scenarios. Benechi et al. (Benecchi 2006) have applied the Co-Active Neuro-Fuzzy Inference System (CANFIS) to predict the presence of prostate cancer; Keles et al. (Keles, Hasiloglu et al. 2007) proposed a neuro-fuzzy system for predicting whether an individual has cancer or Benign Prostatic Hyperplasia (BPH). Çinar et al. (Çinar, Mehmet et al. 2009) designed a classifier-based expert system for the early diagnosis of prostate cancer, thereby aiding the decision-making process and informing the need for a biopsy. Castanho et al. (Castanho, Hernandes et al. 2013) developed a genetic-fuzzy expert system which combines pre-operative serum PSA, clinical stage, and Gleason grade of a biopsy to predict the pathological stage of prostate cancer (i.e. whether it was confined or not- confined).

Saritas et al. (Saritas, Ozkan et al. 2010) devised an ANN approach for the prognosis of cancer which can be used to assist clinical decisions relating to the necessity for a biopsy. Shariat et al. (Shariat, Kattan et al. 2009)have performed a critical review of prostate cancer prediction tools and concluded that predictive tools can help during the complex decision-making processes, and that they can provide individualised, evidence-based estimates of disease status in patients with prostate cancer.

Furthermore, Tsao et al. (Tsao, Liu et al. 2014) developed an ANN model to predict prostate cancer pathological staging in 299 patients prior to radical prostatectomy and found that the ANN model was superior at predicting Organ Confined Disease in prostate cancer than a Logistic Regression model. Tsao et al. (Tsao, Liu et al. 2014) also compared

their ANN model with Partin Tables and found that the ANN model more accurately predicted the pathological stage of prostate cancer.

We therefore proposed a neuro-fuzzy model for predicting the pathological stage of prostate cancer. The system inputs comprised the following variables: the most common tumour pattern (Primary Gleason pattern), the second most common pattern (Secondary Gleason pattern), PSA levels, age at diagnosis, and clinical T stage. The neuro-fuzzy model automatically constructed fuzzy rules via a training process which applied to existing and known patient records and status. These rules were then used to predict the prostate cancer stage of patients in a validation set. The model made use of the Adaptive Neuro-Fuzzy Inference System which was also used to optimise the predictive performance. The outcome for each patient record was a numerical prediction of the 'degree of belongingness' of each patient in the Organ-Confined Disease and Extra-Prostatic Disease classes. The performance of the neuro-fuzzy system was compared to other computational intelligence-based approaches, namely the Artificial Neural Network, Fuzzy C-Means, Support Vector Machine, the Naive Bayes classifiers, and also the AJCC pTNM Staging Nomogram which is commonly used by clinicians. A comparison of the optimal Receiver Operating Characteristic (ROC) points that were identified using these approaches, revealed that the neuro-fuzzy system, at its optimal point, returns the largest Area Under the ROC Curve (AUC), with a low number of false positives (FPR = 0.274, TPR = 0.789, AUC = 0.812). The proposed approach is also an improvement over the AJCC pTNM Staging Nomogram (FPR = 0.032, TPR = 0.197, AUC = 0.582) (**PUBLICATION 8:** (Cosma, Acampora et al. 2016)). Currently, the proposed framework has been implemented as a research tool, and once more evaluations are conducted, the tool will be developed as a simple to use application which can be made accessible to clinicians. The tool will take the clinical test results (i.e. age at diagnosis, PSA, biopsy Primary and Secondary Gleason patterns, and clinical T stage) of an individual patient and predict his likelihood of having extra-prostatic cancer, and thereby aid the clinical decision-making process. Ongoing work is applying the proposed neuro-fuzzy predictor to a larger dataset, examining other computational intelligence approaches, and continuing the development of novel algorithms for predicting disease status in patients with prostate cancer.

As stated previously, it is essential that men with low-risk prostate cancer are not diagnosed as having cancer, as they do not require any active treatment. However, this group of men, will, nonetheless, require life-long surveillance. This can have profound adverse psychological and financial consequences, not only to the individual but also to the healthcare system. We ,therefore, determined whether it is possible to predict with a high level of diagnostic accuracy ,for asymptomatic men with an elevated PSA <20 ng/l, between those with clinically significant prostate cancer from those with either benign disease or low-risk prostate cancer. The development of such approaches will spare men with benign disease or low-risk cancer from unnecessary invasive diagnostic procedures such as TRUS-guided prostate biopsies or TPTPB. Given the reciprocal interactions between tumours and the immune system, we hypothesized that the presence of disease, disease recurrence, and therapeutic resistance may be influenced, reflected in, or predicted by tumour-related immunoregulatory events that can be identified by changes in immune phenotypes in the periphery. Accordingly, we proposed that the analysis of immune phenotyping datasets using multi-parametric flow cytometric analysis can identify features that reflect the presence of disease and/or predict disease progression (Rajwa, Wallace et al. 2017). Although flow cytometry provides a vital tool for exploring, explaining, and understanding complex cellular dynamics and processes in a variety of experimental and clinical settings (Pockley, Foulds et al. 2015), key challenges with multi-parametric flow cytometry include the analysis and interpretation of the complex and increasingly multidimensional data and its conversion into biologically and clinically useful information. This study attempts to address and resolve some of these challenges using computational intelligence methods. Computational intelligence methods comprise evolutionary algorithms (also known as metaheuristic optimization, or nature-inspired optimization algorithms) coupled with machine learning methods, and hybrids of these.

A type of machine learning method, supervised learning, is used to derive prediction models which can be very effective in dealing with uncertainty, noise, and dimensionality in data. Supervised learning methods can learn from existing data to make informed predictions using new patient data and have been widely adopted for prostate cancer prediction tasks when using clinical and biomedical data (Cosma, Brown et al. 2017). It is now time to embrace computational intelligence methods for the analysis of flow cytometry data, since statistical methods alone may not be sufficient for the task of analysing and modelling such complex data (Cosma, Brown et al. 2017). Herein, we

assess whether advanced computational analysis of peripheral blood flow cytometry immunophenotyping data from a selected cohort of individuals can generate prediction models with potential clinical value and identify the presence of prostate cancer in asymptomatic individuals with a PSA level <20 ng/ml. The computational models and algorithms are trained to make predictions on new and previously unseen data using existing data. Significantly, this approach has identified a novel prostate cancer immunophenotyping "finger- print" which could potentially be used to identify the presence of prostate cancer in asymptomatic men having PSA levels <20 ng/ml; and which outperforms the predictive value of the PSA test alone. We have also shown that combining flow cytometry predictors with PSA levels improves diagnostic accuracy. Taken together, these studies demonstrate that the presence of cancer is reflected in changes in the peripheral blood immune phenotype profile which can be identified using computational analysis and interpretation of complex flow cytometry datasets, and the value of computational intelligence-based approaches for interrogating immunophenotyping datasets.

The prediction model was developed using a selected subset of flow cytometry features and the k-Nearest Neighbor (kNN) classification algorithm. The Genetic Algorithm proposed by Ludwig and Nunes (Ludwig and Nunes 2010) was utilized for the feature selection stage, and this algorithm returned the best combination of flow cytometry features (i.e. predictors) for discriminating between patients with benign disease and patients with cancer. These predictors were then input into the kNN classification algorithm. The kNN classifier was used to predict the disease status of an individual using new and previously unseen patient records. Feature selection was important because it enabled only the best subset of features (i.e., predictors) to be selected for the prediction task and, thus, removes the "noisy" features that are not useful in identifying cancer.

Our advanced computational data extraction approach identified the presence of prostate cancer in men with PSA levels <20 ng/ml based on peripheral blood immune cell profiles that have been generated using multi-parameter flow cytometry. Statistical analysis of immune phenotyping datasets relating to the presence and prevalence of key leukocyte populations in the peripheral blood, as generated from individuals undergoing routine tests for prostate cancer (including tissue biopsy) using multi-parametric flow cytometric analysis, was unable to identify significant relationships between leukocyte population

profiles and the presence of benign disease (no prostate cancer) or prostate cancer. By contrast, a Genetic Algorithm computational approach identified a subset of five flow cytometry features ($CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells); $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA); $CD3^-CD19^+$ (B cells); $CD3^+CD56^+CD8^+CD4^+$ (NKT cells)) from a set of twenty features, which could potentially discriminate between benign disease and prostate cancer. These features were used to construct a prostate cancer prediction model using the k-Nearest-Neighbor classification algorithm. The proposed model, which takes as input the set of flow cytometry features, outperformed the predictive model which takes PSA values as input. Specifically, the flow cytometry-based model achieved Accuracy = 83.33%, AUC = 83.40%, and optimal ROC points of FPR = 16.13%, TPR = 82.93%, whereas the PSA-based model achieved Accuracy = 77.78%, AUC = 76.95%, and optimal ROC points of FPR = 29.03%, TPR = 82.93%. Combining PSA and flow cytometry predictors achieved Accuracy = 79.17%, AUC = 78.17% and optimal ROC points of FPR = 29.03%, TPR = 85.37%. (**PUBLICATION 9:** (Cosma, McArdle et al. 2017)). Taken together, these studies demonstrate the value of computational intelligence-based approaches for interrogating immunophenotyping datasets and that combining peripheral blood phenotypic profiling with PSA levels improves diagnostic accuracy compared to using PSA test alone. These studies also demonstrate that the presence of cancer is reflected in changes in the peripheral blood immune phenotype profile which can be identified using computational analysis and interpretation of complex flow cytometry datasets.

Based on the reciprocal interaction between cancer and the immune system, we have proposed that immunological signatures within the peripheral blood (the peripheral blood 'immunome') can discriminate between men with benign prostate disease and those with prostate cancer and thereby reduce the dependency of diagnosis on invasive biopsies. To this end, we have shown above that the incorporation of a peripheral blood immune phenotyping-based feature set comprising five phenotypic features into a computation-based prediction tool enables the better detection of prostate cancer and strengthens the accuracy of the PSA test in asymptomatic men having PSA levels < 20 ng/ml (Cosma, McArdle et al. 2017). We then extended this new approach to determine if phenotypic profiling of peripheral blood NK cell subsets can also discriminate between the presence

of benign prostate disease and prostate cancer in the same cohort of asymptomatic men (Hood, Cosma et al. 2020). We also investigated the potential of the peripheral blood dataset to discriminate between low- or intermediate-risk prostate cancer and high-risk prostate cancer in those men having prostate cancer. Using statistical and computational methods, our study identified a panel of eight phenotypic features (CD56^{dim}CD16^{high}, CD56$^+$DNAM-1$^-$, CD56$^+$LAIR-1$^+$, CD56$^+$LAIR-1$^-$, CD56^{bright}CD8$^+$, CD56$^+$NKp30$^+$, CD56$^+$NKp30$^-$, CD56$^+$NKp46$^+$) that, when incorporated into an Ensemble machine learning prediction model, distinguished between the presence of benign prostate disease and prostate cancer. The machine learning model was then adapted to predict the D'Amico Risk Classification using data from 54 men with prostate cancer and was shown to accurately differentiate between the presence of low-/ intermediate-risk disease and high-risk disease without the need for additional clinical data (**PUBLICATION 10: (Hood, Cosma et al. 2020)**). This simple blood test therefore has the potential to transform prostate cancer diagnostics, as it may be feasible in the future to predict in men with an elevated PSA those who are likely to harbour clinically significant prostate cancer.

The primary treatment for advanced metastatic prostate cancer is androgen (hormone) deprivation therapy (ADT), with upfront chemotherapy if medically fit and with good renal function. Although the majority of patients initially respond to ADT, as evidenced by disease regression and disease stability (Kalina, Neilson et al. 2017), it is inevitable that the disease will progress and become hormone-resistant. At this point, second-line hormone therapy followed by further hormone manipulation therapy is considered, but will typically deliver only a very limited effect. As such, as we have demonstrated above that the immune system may be able to differentiate between clinically significant and non-significant prostate cancer, we determined whether immune modulation could play a role in the management of hormone-resistant prostate cancer.

Immunotherapy involving stimulating the patient's own immune system to retarget their cancer is emerging as the next generation of cancer treatment (Farkona, Diamandis et al. 2016). Currently, the only approved immunotherapy for treating castration-resistant prostate cancer is Sipuleucel-T immunotherapy which has been shown to improve the median overall survival by 4.1 months compared to a placebo group (Kantoff, Higano et al. 2010). Although preventing tumour-mediated immunoregulation using immune checkpoint inhibitors such as Ipilimumab has shown some success in treating

immunogenic cancers such as melanoma and non-small cell lung cancer, their use in patients with prostate cancer has not been shown to improve overall survival (Kwon, Drake et al. 2014). However, some evidence of beneficial effects have been observed and clinical trials testing Ipilimumab in combination with other standard prostate cancer treatments (e.g., ADT) are ongoing (Modena, Ciccarese et al. 2016). NK cells were first identified on the basis of their natural cytotoxicity toward cancerous cells and a number of NK cell-based immunotherapies are now in development (Trinchieri 1989, Miller 2013, Specht, Ahrens et al. 2015, Shevtsov and Multhoff 2016, Shevtsov and Multhoff 2016, Chiossone, Vienne et al. 2017, Multhoff, Seier et al. 2020). As reviewed by Sabry and Lowdell (Sabry and Lowdell 2013), the cytotoxic function of NK cells is controlled by the balance of signals transduced via activating and inhibitory receptors following ligation with stress ligands and MHC class I molecules, respectively (Dynamic Equilibrium Theory). Bryceson et al. demonstrated that natural cytotoxicity requires the co-engagement of multiple activating receptors (Bryceson, March et al. 2006, Bryceson, Ljunggren et al. 2009). Furthermore, work by Lowdell et al. led to the hypothesis that the natural cytotoxicity mechanism can be divided into two discrete stages; "priming" and "triggering" (North, Bakhsh et al. 2007, Sabry, Tsirogianni et al. 2011, Sabry and Lowdell 2013). For this, they hypothesized that the "priming" signal can be delivered either by the ligation of the appropriate number and combination of activating receptors with their target ligands or via an activating cytokine (e.g., IL-2). The "triggering" signal requires the ligation of at least one additional activating receptor to its target ligand that is specific to stressed cells (Sabry and Lowdell 2013).

Tumor primed NK cells (TpNK) can be generated *in vitro* by co-incubating resting NK cells with the acute lymphoblastic leukemia (ALL) cell line CTV-1 (North, Bakhsh et al. 2007). Phenotypically, tumour primed NK cells appear distinct from resting NK cells in that they exhibit reduced expression of activating receptors (e.g., CD16, NKG2D, NKp46), both in terms of intensity and proportion, whereas both the proportion and intensity of expression of co-receptors (e.g., CD69 and CD25) are up-regulated (North, Bakhsh et al. 2007, Sabry, Tsirogianni et al. 2011). Priming NK cells from healthy volunteers in this way has been reported to enhance their cytotoxicity against NK cell-resistant tumour cell lines such as the human metastatic prostate cancer cell line DU145 (Sabry, Tsirogianni et al. 2011).

The therapeutic potential of an autologous NK cell-based therapy requires that patient-derived NK cells can be appropriately triggered. We therefore determined whether activation of NK cells isolated from thawed peripheral blood mononuclear cell (PBMC) preparations derived from patients with prostate cancer by either co-incubation with mitomycin C treated CTV-1 cells or stimulation with IL-2 enhanced their capacity to kill the human metastatic disease-derived prostate cancer cell line PC3.

Tumour priming and IL-2 stimulation of patient-derived NK cells resulted in similar levels of cytotoxicity, but distinct NK cell phenotypes. Importantly, the capacity of priming and IL- 2 stimulation to trigger cytotoxicity was patient-dependent and mutually exclusive, in that NK cells from ~50% of patients preferentially responded to tumour priming, whereas NK cells from the remaining patients preferentially responded to IL- 2 stimulation. (**PUBLICATION 11:** (Hood, Foulds et al. 2018)). In addition to providing more insight into the biology of tumour primed and cytokine-stimulated NK cells, this study supports the use of autologous NK cell-based immunotherapies for the treatment of prostate cancer. However, our findings also indicate that patients will need to be stratified according to their potential responsiveness to individual therapeutic approaches (Hood 2016, Hood, Foulds et al. 2018).

In summary, over the past two decades I have improved the diagnostic yield of prostate cancer. This has not only been in improving PSA as a biomarker but also the modality of performing prostate biopsies by revealing that TPTPB is significantly superior to performing TRUS guided prostate biopsies. My collaboration with JvGCRC has led to the discovery of potential new immune markers that are not only significantly more accurate in detecting prostate cancer than PSA alone but is able to differentiate between clinically significant prostate cancer and benign disease/low-risk prostate cancer.

References

Ahmed, H. U., A. El-Shater Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, A. P. Kirkham, R. Oldroyd, C. Parker, M. Emberton and P. s. group (2017). "Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study." Lancet **389**(10071): 815-822.

Al Olama, A. A., Z. Kote-Jarai, G. G. Giles, M. Guy, J. Morrison, G. Severi, D. A. Leongamornlert, M. Tymrakiewicz, S. Jhavar, E. Saunders, J. L. Hopper, M. C. Southey, K. R. Muir, D. R. English, D. P. Dearnaley, A. T. Ardern-Jones, A. L. Hall, L. T. O'Brien, R. A. Wilkinson, E. Sawyer, A. Lophatananon, A. Horwich, R. A. Huddart, V. S. Khoo, C. C. Parker, C. J. Woodhouse, A. Thompson, T. Christmas, C. Ogden, C. Cooper, J. L. Donovan, F. C. Hamdy, D. E. Neal, R. A. Eeles and D. F. Easton (2009). "Multiple loci on 8q24 associated with prostate cancer susceptibility." Nat Genet **41**(10): 1058-1060.

Arnold, M., H. E. Karim-Kos, J. W. Coebergh, G. Byrnes, A. Antilla, J. Ferlay, A. G. Renehan, D. Forman and I. Soerjomataram (2015). "Recent trends in incidence of five common cancers in 26 European countries since 1988: Analysis of the European Cancer Observatory." Eur J Cancer **51**(9): 1164-1187.

Benecchi, L. (2006). "Neuro-fuzzy system for prostate cancer diagnosis." Urology **68**(2): 357-361.

Blute, M. L., O. Nativ, H. Zincke, G. M. Farrow, T. Therneau and M. M. Lieber (1989). "Pattern of failure after radical retropubic prostatectomy for clinically and pathologically localized adenocarcinoma of the prostate: influence of tumor deoxyribonucleic acid ploidy." J Urol **142**(5): 1262-1265.

Breslow, N., C. W. Chan, G. Dhom, R. A. Drury, L. M. Franks, B. Gellei, Y. S. Lee, S. Lundberg, B. Sparke, N. H. Sternby and H. Tulinius (1977). "Latent carcinoma of prostate at autopsy in seven areas. The International Agency for Research on Cancer, Lyons, France." Int J Cancer **20**(5): 680-688.

Briganti, A., P. I. Karakiewicz, S. Joniau and H. Van Poppel (2009). "Open to debate. The motion: Nomograms should become a routine tool in determining prostate cancer prognosis." Eur Urol **55**(3): 743-747.

Bryceson, Y. T., H. G. Ljunggren and E. O. Long (2009). "Minimal requirement for induction of natural cytotoxicity and intersection of activation signals by inhibitory receptors." Blood **114**(13): 2657-2666.

Bryceson, Y. T., M. E. March, H. G. Ljunggren and E. O. Long (2006). "Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion." Blood **107**(1): 159-166.

Carlson, W. H., D. G. Bell, J. G. Lawen and R. A. Rendon (2010). "Multi-drug resistant E.coli urosepsis in physicians following transrectal ultrasound guided prostate biopsies - three cases including one death." Can J Urol **17**(2): 5135-5137.

Carter, H. B., J. D. Pearson, E. J. Metter, L. J. Brant, D. W. Chan, R. Andres, J. L. Fozard and P. C. Walsh (1992). "Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease." JAMA **267**(16): 2215-2220.

Castanho, M. J. P., F. Hernandes, A. M. De Rè, S. Rautenberg and A. Billis (2013). "Fuzzy expert system for predicting pathological stage of prostate cancer." Expert Systems with Applications **40**(2): 466–470.

Catalona, W. J., A. W. Partin, K. M. Slawin, M. K. Brawer, R. C. Flanigan, A. Patel, J. P. Richie, J. B. deKernion, P. C. Walsh, P. T. Scardino, P. H. Lange, E. N. Subong, R. E. Parson, G. H. Gasior, K. G. Loveland and P. C. Southwick (1998). "Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial." JAMA **279**(19): 1542-1547.

Catalona, W. J., D. S. Smith, T. L. Ratliff, K. M. Dodds, D. E. Coplen, J. J. Yuan, J. A. Petros and G. L. Andriole (1991). "Measurement of prostate-specific antigen in serum as a screening test for prostate cancer." N Engl J Med **324**(17): 1156-1161.

Catalona, W. J., P. C. Southwick, K. M. Slawin, A. W. Partin, M. K. Brawer, R. C. Flanigan, A. Patel, J. P. Richie, P. C. Walsh, P. T. Scardino, P. H. Lange, G. H. Gasior, K. G. Loveland and K. R. Bray (2000). "Comparison of percent free PSA, PSA density, and age-specific PSA cutoffs for prostate cancer detection and staging." Urology **56**(2): 255-260.

Chiossone, L., M. Vienne, Y. M. Kerdiles and E. Vivier (2017). "Natural killer cell immunotherapies against cancer: checkpoint inhibitors and more." Semin Immunol **31**: 55-63.

Chun, F. K., P. I. Karakiewicz, A. Briganti, J. Walz, M. W. Kattan, H. Huland and M. Graefen (2007). "A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer." BJU Int **99**(4): 794-800.

Çinar, M., E. Mehmet, Z. E. Erkan and Z. Y. Ateşçi (2009). "Early prostate cancer diagnosis by using artificial neural net- works and support vector machines." Expert Systems with Applications **36**(3): 6357–6361.

Cosma, G., G. Acampora, D. Brown, R. C. Rees, M. Khan and A. G. Pockley (2016). "Prediction of pathological stage in patients with prostate cancer: A Neuro-Fuzzy Model." PLoS ONE **11**(6): e0155856.

Cosma, G., D. Brown, M. Archer, M. Khan and A. G. Pockley (2017). "A survey on computational intelligence approaches for predictive modeling in prostate cancer." Expert Systems with Applications **70**(Supplement C): 1-19.

Cosma, G., S. E. McArdle, S. Reeder, G. A. Foulds, S. Hood, M. Khan and A. G. Pockley (2017). "Identifying the presence of prostate cancer in individuals with PSA levels <20 ng ml^{-1} using computational data extraction analysis of high dimensional peripheral blood flow cytometric phenotyping data." Frontiers in Immunology **8:1771**.

De Angelis, R., M. Sant, M. P. Coleman, S. Francisci, P. Baili, D. Pierannunzio, A. Trama, O. Visser, H. Brenner, E. Ardanaz, M. Bielska-Lasota, G. Engholm, A. Nennecke, S. Siesling, F. Berrino, R. Capocaccia (2014). "Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE--5-a population-based study." Lancet Oncol **15**(1): 23-34.

Djavan, B., V. Ravery, A. Zlotta, P. Dobronski, M. Dobrovits, M. Fakhari, C. Seitz, M. Susani, A. Borkowski, L. Boccon-Gibod, C. C. Schulman and M. Marberger (2001). "Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop?" J Urol **166**(5): 1679-1683.

Dotan, Z. A. and J. Ramon (2009). "Nomograms as a tool in predicting prostate cancer prognosis." <u>European Urology Supplements</u> **8**(9): 721–724.

Eeles, R. A., A. Al Olama, S. Benlloch, E. J. Saunders, D. A. Leongamornlert, M. Tymrakiewicz, M. Ghoussaini, C. Luccarini, J. Dennis, S. Jugurnauth-Little, T. Dadaev, D. E. Neal, F. C. Hamdy, J. L. Donovan, K. Muir, G. G. Giles, G. Severi, F. Wiklund, H. Gronberg, C. A. Haiman, F. Schumacher, B. E. Henderson, L. Le Marchand, S. Lindstrom, P. Kraft, D. J. Hunter, S. Gapstur, S. J. Chanock, S. I. Berndt, D. Albanes, G. Andriole, J. Schleutker, M. Weischer, F. Canzian, E. Riboli, T. J. Key, R. C. Travis, D. Campa, S. A. Ingles, E. M. John, R. B. Hayes, P. D. Pharoah, N. Pashayan, K. T. Khaw, J. L. Stanford, E. A. Ostrander, L. B. Signorello, S. N. Thibodeau, D. Schaid, C. Maier, W. Vogel, A. S. Kibel, C. Cybulski, J. Lubinski, L. Cannon-Albright, H. Brenner, J. Y. Park, R. Kaneva, J. Batra, A. B. Spurdle, J. A. Clements, M. R. Teixeira, E. Dicks, A. Lee, A. M. Dunning, C. Baynes, D. Conroy, M. J. Maranian, S. Ahmed, K. Govindasami, M. Guy, R. A. Wilkinson, E. J. Sawyer, A. Morgan, D. P. Dearnaley, A. Horwich, R. A. Huddart, V. S. Khoo, C. C. Parker, N. J. Van As, C. J. Woodhouse, A. Thompson, T. Dudderidge, C. Ogden, C. S. Cooper, A. Lophatananon, A. Cox, M. C. Southey, J. L. Hopper, D. R. English, M. Aly, J. Adolfsson, J. Xu, S. L. Zheng, M. Yeager, R. Kaaks, W. R. Diver, M. M. Gaudet, M. C. Stern, R. Corral, A. D. Joshi, A. Shahabi, T. Wahlfors, T. L. Tammela, A. Auvinen, J. Virtamo, P. Klarskov, B. G. Nordestgaard, M. A. Roder, S. F. Nielsen, S. E. Bojesen, A. Siddiq, L. M. Fitzgerald, S. Kolb, E. M. Kwon, D. M. Karyadi, W. J. Blot, W. Zheng, Q. Cai, S. K. McDonnell, A. E. Rinckleb, B. Drake, G. Colditz, D. Wokolorczyk, R. A. Stephenson, C. Teerlink, H. Muller, D. Rothenbacher, T. A. Sellers, H. Y. Lin, C. Slavov, V. Mitev, F. Lose, S. Srinivasan, S. Maia, P. Paulo, E. Lange, K. A. Cooney, A. C. Antoniou, D. Vincent, F. Bacot, D. C. Tessier, Z. Kote-Jarai and D. F. Easton (2013). "Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array." <u>Nat Genet</u> **45**(4): 385-391, 391e381-382.

Epstein, J. I. and J. C. Eggleston (1984). "Immunohistochemical localization of prostate-specific acid phosphatase and prostate-specific antigen in stage A2 adenocarcinoma of the prostate: prognostic implications." <u>Hum Pathol</u> **15**(9): 853-859.

Epstein, J. I., G. Pizov and P. C. Walsh (1993). "Correlation of pathologic findings with progression after radical retropubic prostatectomy." <u>Cancer</u> **71**(11): 3582-3593.

Epstein, J. I., P. C. Walsh, M. Carmichael and C. B. Brendler (1994). "Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer." JAMA **271**(5): 368-374.

Farkona, S., E. P. Diamandis and I. M. Blasutig (2016). "Cancer immunotherapy: the beginning of the end of cancer?" BMC Med **14**: 73.

Griffiths, J. C. (1980). "Prostate-specific acid phosphatase: re-evaluation of radioimmunoassay in diagnosing prostatic disease." Clin Chem **26**(3): 433-436.

Haas, G. P., N. Delongchamps, O. W. Brawley, C. Y. Wang and G. de la Roza (2008). "The worldwide epidemiology of prostate cancer: perspectives from autopsy studies." Can J Urol **15**(1): 3866-3871.

Han, J. (2005). Data Mining: concepts and techniques. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc.

Hemminki, K. (2012). "Familial risk and familial survival in prostate cancer." World J Urol **30**(2): 143-148.

Hood, S. P. (2016). Characterising the phenotype and function of natural killer cells in patients with prostate cancer. PhD, Nottingham Trent University.

Hood, S. P., G. Cosma, G. A. Foulds, C. Johnson, S. Reeder, S. E. McArdle, M. A. Khan and A. G. Pockley (2020). "Identifying prostate cancer and its clinical risk in asymptomatic men using machine learning of high dimensional peripheral blood flow cytometric natural killer cell subset phenotyping data." Elife **9**.

Hood, S. P., G. A. Foulds, H. Imrie, S. Reeder, S. E. B. McArdle, M. Khan and A. G. Pockley (2018). "Phenotype and function of activated natural killer cells from patients with prostate cancer: Patient-dependent responses to priming and IL-2 activation." Front Immunol **9**: 3169.

Hori, S., J. S. Blanchet and J. McLoughlin (2013). "From prostate-specific antigen (PSA) to precursor PSA (proPSA) isoforms: a review of the emerging role of proPSAs in the detection and management of early prostate cancer." BJU Int **112**(6): 717-728.

Ilyin, S. E., S. M. Belkowski and C. R. Plata-Salaman (2004). "Biomarker discovery and validation: technologies and integrative approaches." Trends Biotechnol **22**(8): 411-416.

Jang, T. L., M. Han, K. A. Roehl, S. A. Hawkins and W. J. Catalona (2006). "More favorable tumor features and progression-free survival rates in a longitudinal prostate cancer screening study: PSA era and threshold-specific effects." Urology **67**(2): 343-348.

Jansson, K. F., O. Akre, H. Garmo, A. Bill-Axelson, J. Adolfsson, P. Stattin and O. Bratt (2012). "Concordance of tumor differentiation among brothers with prostate cancer." Eur Urol **62**(4): 656-661.

Jones, J. S., A. Patel, L. Schoenfield, J. C. Rabets, C. D. Zippe and C. Magi-Galluzzi (2006). "Saturation technique does not improve cancer detection as an initial prostate biopsy strategy." J Urol **175**(2): 485-488.

Kalina, J. L., D. S. Neilson, A. P. Comber, J. M. Rauw, A. S. Alexander, J. Vergidis and J. J. Lum (2017). "Immune modulation by androgen deprivation and radiation therapy: Implications for prostate cancer immunotherapy." Cancers (Basel) **9**(2).

Kantoff, P. W., C. S. Higano, N. D. Shore, E. R. Berger, E. J. Small, D. F. Penson, C. H. Redfern, A. C. Ferrari, R. Dreicer, R. B. Sims, Y. Xu, M. W. Frohlich, P. F. Schellhammer and I. S. Investigators (2010). "Sipuleucel-T immunotherapy for castration-resistant prostate cancer." N Engl J Med **363**(5): 411-422.

Keles, A., A. S. Hasiloglu, A. Keles and Y. Aksoy (2007). "Neuro-fuzzy classification of prostate cancer using NEFCLASS-J." Comput Biol Med **37**(11): 1617-1628.

Khan, A. R. and M. N. James (1998). "Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes." Protein Sci **7**(4): 815-836.

Khan, M. A., A. W. Partin, H. G. Rittenhouse, S. D. Mikolajczyk, L. J. Sokoll, D. W. Chan and R. W. Veltri (2003). "Evaluation of proprostate specific antigen for early detection of prostate cancer in men with a total prostate specific antigen range of 4.0 to 10.0 ng/ml." J Urol **170**(3): 723-726.

Khan, M. A., L. J. Sokoll, D. W. Chan, L. A. Mangold, P. Mohr, S. D. Mikolajczyk, H. J. Linton, C. L. Evans, H. G. Rittenhouse and A. W. Partin (2004). "Clinical utility of

proPSA and "benign" PSA when percent free PSA is less than 15%." Urology **64**(6): 1160-1164.

Kumar, A., S. D. Mikolajczyk, A. S. Goel, L. S. Millar and M. S. Saedi (1997). "Expression of pro form of prostate-specific antigen by mammalian cells and its conversion to mature, active form by human kallikrein 2." Cancer Res **57**(15): 3111-3114.

Kwon, E. D., C. G. Drake, H. I. Scher, K. Fizazi, A. Bossi, A. J. van den Eertwegh, M. Krainer, N. Houede, R. Santos, H. Mahammedi, S. Ng, M. Maio, F. A. Franke, S. Sundar, N. Agarwal, A. M. Bergman, T. E. Ciuleanu, E. Korbenfeld, L. Sengelov, S. Hansen, C. Logothetis, T. M. Beer, M. B. McHenry, P. Gagnier, D. Liu, W. R. Gerritsen and C. A. Investigators (2014). "Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial." Lancet Oncol **15**(7): 700-712.

Lane, J. A., J. L. Donovan, M. Davis, E. Walsh, D. Dedman, L. Down, E. L. Turner, M. D. Mason, C. Metcalfe, T. J. Peters, R. M. Martin, D. E. Neal, F. C. Hamdy (2014). "Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial." Lancet Oncol **15**(10): 1109-1118.

Leitzmann, M. F. and S. Rohrmann (2012). "Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates." Clin Epidemiol **4**: 1-11.

Lilja, H., A. Christensson, U. Dahlen, M. T. Matikainen, O. Nilsson, K. Pettersson and T. Lovgren (1991). "Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin." Clin Chem **37**(9): 1618-1625.

Linton, H. J., L. S. Marks, L. S. Millar, C. L. Knott, H. G. Rittenhouse and S. D. Mikolajczyk (2003). "Benign prostate-specific antigen (BPSA) in serum is increased in benign prostate disease." Clin Chem **49**(2): 253-259.

Loeb, S., M. G. Sanda, D. L. Broyles, S. S. Shin, C. H. Bangma, J. T. Wei, A. W. Partin, G. G. Klee, K. M. Slawin, L. S. Marks, R. H. van Schaik, D. W. Chan, L. J. Sokoll, A. B. Cruz, I. A. Mizrahi and W. J. Catalona (2015). "The prostate health index selectively identifies clinically significant prostate cancer." J Urol **193**(4): 1163-1169.

Ludwig, O. and U. Nunes (2010). "Novel maximum-margin training algorithms for supervised neural networks." IEEE Trans Neural Netw **21**(6): 972-984.

Luengo-Fernandez, R., J. Leal, A. Gray and R. Sullivan (2013). "Economic burden of cancer across the European Union: a population-based cost analysis." Lancet Oncol **14**(12): 1165-1174.

Martin, R. M., J. L. Donovan, E. L. Turner, C. Metcalfe, G. J. Young, E. I. Walsh, J. A. Lane, S. Noble, S. E. Oliver, S. Evans, J. A. C. Sterne, P. Holding, Y. Ben-Shlomo, P. Brindle, N. J. Williams, E. M. Hill, S. Y. Ng, J. Toole, M. K. Tazewell, L. J. Hughes, C. F. Davies, J. C. Thorn, E. Down, G. Davey Smith, D. E. Neal, F. C. Hamdy (2018). "Effect of a low-intensity PSA-based screening intervention on prostate cancer mortality: The CAP Randomized Clinical Trial." JAMA **319**(9): 883-895.

Mikolajczyk, S. D., K. M. Marker, L. S. Millar, A. Kumar, M. S. Saedi, J. K. Payne, C. L. Evans, C. L. Gasior, H. J. Linton, P. Carpenter and H. G. Rittenhouse (2001). "A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer." Cancer Res **61**(18): 6958-6963.

Mikolajczyk, S. D., L. S. Marks, A. W. Partin and H. G. Rittenhouse (2002). "Free prostate-specific antigen in serum is becoming more complex." Urology **59**(6): 797-802.

Mikolajczyk, S. D., L. S. Millar, T. J. Wang, H. G. Rittenhouse, L. S. Marks, W. Song, T. M. Wheeler and K. M. Slawin (2000). "A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue." Cancer Res **60**(3): 756-759.

Mikolajczyk, S. D., L. S. Millar, T. J. Wang, H. G. Rittenhouse, R. L. Wolfert, L. S. Marks, W. Song, T. M. Wheeler and K. M. Slawin (2000). ""BPSA," a specific molecular form of free prostate-specific antigen, is found predominantly in the transition zone of patients with nodular benign prostatic hyperplasia." Urology **55**(1): 41-45.

Miller, J. S. (2013). "Therapeutic applications: natural killer cells in the clinic." Hematology Am Soc Hematol Educ Program **2013**: 247-253.

Modena, A., C. Ciccarese, R. Iacovelli, M. Brunelli, R. Montironi, M. Fiorentino, G. Tortora and F. Massari (2016). "Immune Checkpoint Inhibitors and Prostate Cancer: A New Frontier?" Oncol Rev **10**(1): 293.

Multhoff, G., S. Seier, S. Stangl, W. Sievert, M. Shevtsov , C. Werner, A. G. Pockley, C. Blankenstein, M. Hildebrandt, R. Offner, N. Ahrens, K. Kokowski, M. Hautmann, C. Rödel, R. Fietkau, D. Lubgan, R. Huber, H. Hautmann, T. Duell, M. Molls, H. Specht, B. Haller, M. Devecka, A. Sauter and S. E. Combs (2020). "Targeted Natural Killer cell based adoptive immunotherapy for the treatment of patients with NSCLC after radiochemotherapy – a randomized phase II clinical study." Clinical Cancer Research **26**: 5368-5379.

Nafie, S., C. Berridge and M. Khan (2021). "Novel technique in performing standard Transperineal Template Prostate Biopsy under local anaesthetic." Urology Journal **submitted**.

Nafie, S., J. K. Mellon, J. P. Dormer and M. A. Khan (2014). "The role of transperineal template prostate biopsies in prostate cancer diagnosis in biopsy naive men with PSA less than 20 ng ml(-1.)." Prostate Cancer Prostatic Dis **17**(2): 170-173.

Nafie, S., R. P. Pal, J. P. Dormer and M. A. Khan (2014). "Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies." World J Urology **32**(4): 971-975.

Nafie, S., M. Wanis and M. Khan (2017). "The efficacy of Transrectal Ultrasound Guided Biopsy *versus* Transperineal Template Biopsy of the Prostate in diagnosing prostate cancer in men with previous negative Transrectal Ultrasound Guided Biopsy." Urology Journal **14**(2): 3008-3012.

Naughton, C. K., D. C. Miller, D. E. Mager, D. K. Ornstein and W. J. Catalona (2000). "A prospective randomized trial comparing 6 *versus* 12 prostate biopsy cores: impact on cancer detection." J Urol **164**(2): 388-392.

Nelson, W. G., A. M. De Marzo and W. B. Isaacs (2003). "Prostate cancer." N Engl J Med **349**(4): 366-381.

Norberg, M., L. Egevad, L. Holmberg, P. Sparen, B. J. Norlen and C. Busch (1997). "The sextant protocol for ultrasound-guided core biopsies of the prostate underestimates the presence of cancer." Urology **50**(4): 562-566.

North, J., I. Bakhsh, C. Marden, H. Pittman, E. Addison, C. Navarrete, R. Anderson and M. W. Lowdell (2007). "Tumor-primed human natural killer cells lyse NK-resistant tumor targets: evidence of a two-stage process in resting NK cell activation." J Immunol **178**(1): 85-94.

Oesterling, J. E., S. J. Jacobsen, C. G. Chute, H. A. Guess, C. J. Girman, L. A. Panser and M. M. Lieber (1993). "Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges." JAMA **270**(7): 860-864.

Pal, R. P., M. Elmussareh, M. Chanawani and M. A. Khan (2012). "The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies." BJU Int **109**(3): 367-371.

Parkes, C., N. J. Wald, P. Murphy, L. George, H. C. Watt, R. Kirby, P. Knekt, K. J. Helzlsouer and J. Tuomilehto (1995). "Prospective observational study to assess value of prostate specific antigen as screening test for prostate cancer." BMJ **311**(7016): 1340-1343.

Peter, J., C. Unverzagt, T. N. Krogh, O. Vorm and W. Hoesel (2001). "Identification of precursor forms of free prostate-specific antigen in serum of prostate cancer patients by immunosorption and mass spectrometry." Cancer Res **61**(3): 957-962.

Pockley, A. G., G. A. Foulds, J. A. Oughton, N. I. Kerkvliet and G. Multhoff (2015). "Immune cell phenotyping using flow cytometry." Curr Protoc Toxicol **66**: 18 18 11-34.

Presti, J. C., Jr. (2003). "Prostate biopsy: how many cores are enough?" Urol Oncol **21**(2): 135-140.

Raaijmakers, R., W. J. Kirkels, M. J. Roobol, M. F. Wildhagen and F. H. Schrder (2002). "Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program." Urology **60**(5): 826-830.

Rajwa, B., P. K. Wallace, E. A. Griffiths and M. Dundar (2017). "Automated assessment of disease progression in acute myeloid leukemia by Probabilistic Analysis of flow cytometry data." IEEE Trans Biomed Eng **64**(5): 1089-1098.

Romero Otero, J., B. Garcia Gomez, F. Campos Juanatey and K. A. Touijer (2014). "Prostate cancer biomarkers: an update." Urol Oncol **32**(3): 252-260.

Sabry, M. and M. W. Lowdell (2013). "Tumor-primed NK cells: waiting for the green light." Front Immunol **4**: 408.

Sabry, M., M. Tsirogianni, I. A. Bakhsh, J. North, J. Sivakumaran, K. Giannopoulos, R. Anderson, S. Mackinnon and M. W. Lowdell (2011). "Leukemic priming of resting NK cells is killer Ig-like receptor independent but requires CD15-mediated CD2 ligation and natural cytotoxicity receptors." J Immunol **187**(12): 6227-6234.

Saritas, I., I. A. Ozkan and I. U. Sert (2010). "Prognosis of prostate cancer by artificial neural networks." Expert Systems with Applications **37**(9): 6646–6650.

Sawyers, C. L. (2008). "The cancer biomarker problem." Nature **452**(7187): 548-552.

Shariat, S. F., M. W. Kattan, A. J. Vickers, P. I. Karakiewicz and P. T. Scardino (2009). "Critical review of prostate cancer predictive tools." Future Oncol **5**(10): 1555-1584.

Shevtsov, M. and G. Multhoff (2016). "Heat shock protein-peptide and HSP-based immunotherapies for the treatment of cancer." Front Immunol **7**: 171.

Shevtsov, M. and G. Multhoff (2016). "Immunological and translational aspects of NK cell-based antitumor immunotherapies." Front Immunol **7**: 492.

Sokoll, L. J., M. G. Sanda, Z. Feng, J. Kagan, I. A. Mizrahi, D. L. Broyles, A. W. Partin, S. Srivastava, I. M. Thompson, J. T. Wei, Z. Zhang and D. W. Chan (2010). "A prospective, multicenter, National Cancer Institute Early Detection Research Network study of [-2]proPSA: improving prostate cancer detection and correlating with cancer aggressiveness." Cancer Epidemiol Biomarkers Prev **19**(5): 1193-1200.

Specht, H. M., N. Ahrens, C. Blankenstein, T. Duell, R. Fietkau, U. S. Gaipl, C. Gunther, S. Gunther, G. Habl, H. Hautmann, M. Hautmann, R. M. Huber, M. Molls, R. Offner, C. Rodel, F. Rodel, M. Schutz, S. E. Combs and G. Multhoff (2015). "Heat shock protein 70

(Hsp70) peptide activated natural killer (NK) cells for the treatment of patients with non-small cell lung cancer (NSCLC) after radiochemotherapy (RCTx) - From preclinical studies to a clinical Phase II Trial." Front Immunol **6**: 162.

Stamey, T. A., N. Yang, A. R. Hay, J. E. McNeal, F. S. Freiha and E. Redwine (1987). "Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate." N Engl J Med **317**(15): 909-916.

Stenman, U. H., J. Leinonen, H. Alfthan, S. Rannikko, K. Tuhkanen and O. Alfthan (1991). "A complex between prostate-specific antigen and alpha 1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostatic cancer: assay of the complex improves clinical sensitivity for cancer." Cancer Res **51**(1): 222-226.

Stephan, C., A. M. Kahrs, H. Cammann, M. Lein, M. Schrader, S. Deger, K. Miller and K. Jung (2009). "A [-2]proPSA-based artificial neural network significantly improves differentiation between prostate cancer and benign prostatic diseases." Prostate **69**(2): 198-207.

Tewari, A., C. Porter, J. Peabody, E. D. Crawford, R. Demers, C. C. Johnson, J. T. Wei, G. W. Divine, C. O'Donnell, E. J. Gamito and M. Menon (2001). "Predictive modeling techniques in prostate cancer." Mol Urol **5**(4): 147-152.

Trinchieri, G. (1989). "Biology of natural killer cells." Adv Immunol **47**: 187-376.

Tsao, C. W., C. Y. Liu, T. L. Cha, S. T. Wu, G. H. Sun, D. S. Yu, H. I. Chen, S. Y. Chang, S. C. Chen and C. Y. Hsu (2014). "Artificial neural network for predicting pathological stage of clinically localized prostate cancer in a Taiwanese population." J Chin Med Assoc **77**(10): 513-518.

Veltri, R. W. and M. C. Miller (1999). "Free/total PSA ratio improves differentiation of benign and malignant disease of the prostate: critical analysis of two different test populations." Urology **53**(4): 736-745.

Wang, M. C., L. A. Valenzuela, G. P. Murphy and T. M. Chu (1979). "Purification of a human prostate specific antigen." Invest Urol **17**(2): 159-163.

WHO (2014). United Kingdom Cancer Country Profile: World Health Organisation.

Original Articles

EVALUATION OF PROPROSTATE SPECIFIC ANTIGEN FOR EARLY DETECTION OF PROSTATE CANCER IN MEN WITH A TOTAL PROSTATE SPECIFIC ANTIGEN RANGE OF 4.0 TO 10.0 NG/ML

MASOOD A. KHAN,* ALAN W. PARTIN,† HARRY G. RITTENHOUSE,‡
STEPHEN D. MIKOLAJCZYK,‡ LORI J. SOKOLL,‡ DANIEL W. CHAN‡ AND ROBERT W. VELTRI§

From the James Buchanan Brady Urological Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland, and Hybritech Incorporated, a subsidiary of Beckman Coulter, Inc., San Diego, California (HGR, SDM)

ABSTRACT

Purpose: In contemporary screening populations a major drawback of prostate specific antigen (PSA) is its relative lack of specificity, especially in the range of 4 to 10 ng/ml, where prostate cancer is found 25% of the time. ProPSA is a derivative of free PSA (fPSA) consisting of the truncated forms (eg [-2]proPSA, [-4]proPSA or the full-length [-7]proPSA). There is increasing evidence that proPSA is associated preferentially with prostate cancer. The objective of this study was to determine whether proPSA can influence the detection of early prostate cancer.

Materials and Methods: Archival serum samples obtained from 93 men who underwent a systematic 12-core prostate biopsy (total PSA range 4.0 to 10.0 ng/ml) were assayed for percent free PSA, total PSA and the 3 forms of proPSA (Hybritech Tandem Assays Beckman Coulter Access, Beckman Coulter, Inc., Brea, California). Free PSA, the cumulative sum of individual proPSA forms ([-2], [-4] and [-7], or sum-proPSA) and derivatives were determined. Of the 93 men assessed 41 (44%) had evidence of prostate cancer (76% Gleason 5/6, 19% Gleason 7 and 5% Gleason 8). Prostate volume was measured at systematic 12-core biopsy for the detection of prostate cancer. Results were analyzed using univariate and multivariate logistic regression (LR) nonparametric statistical methods.

Results: Using univariate LR, fPSA, percent fPSA (%fPSA), percent sum-proPSA and prostate volume significantly ($p < 0.05$) differentiated men with prostate cancer from those with benign disease. However, applying stepwise backward multivariate LR, total PSA, %fPSA and sum-proPSA were retained and generated a receiver operator characteristic curve with an area under the curve of 76.6%. At 90% sensitivity these 3 variables collectively achieved a specificity of 44% for the detection of prostate cancer. Individually, the 3 retained variables had a specificity of 23% (total PSA), 33% (%fPSA) and 13% (sum-proPSA).

Conclusions: Sum-proPSA, total PSA and %fPSA in combination improve the specificity of early prostate cancer detection in men with a total PSA of 4 to 10 ng/ml compared with the results of individual PSA molecular forms measured.

KEY WORDS: adenocarcinoma, prostate, prostate-specific antigen

The development and subsequent routine use of prostate specific antigen (PSA) during the last decade have revolutionized the management of prostate cancer. Total PSA

Editor's Note: This article is the first of 5 published in this issue for which category 1 CME credits can be earned. Instructions for obtaining credits are given with the questions on pages 960 and 961.

(tPSA) has increased our ability to detect and, in turn, treat early prostate cancer. However, the major drawback of tPSA is its relative lack of specificity. This drawback is especially important in the critical diagnostic range of 4 to 10 ng/ml where only 25% of patients biopsied will demonstrate prostate cancer.[1] To improve the specificity of tPSA further various methods have been introduced including age specific PSA,[2] PSA density,[3] percent free PSA (%fPSA)[4] and complexed PSA.[5] The continually changing natural history of prostate cancer in terms of age at detection, grade, stage and tumor size during the PSA era of the 1990s has also had an impact on the diagnostic performance of these assays in different age groups with regard to specificity in the range of 2 to 10 ng/ml tPSA.[6]

The free PSA (fPSA) form and its derivatives represent a heterogeneous group of at least 3 different subforms of inactive PSA. One form has been identified as the proenzyme or precursor form of PSA (proPSA), which contains a 7 amino

723

acid pro-leader peptide, and is preferentially associated with prostate cancer.[7,8] Total PSA is normally found in seminal fluid in a mature, active form with 237 amino acids lacking the pro-leader peptide. The pro-leader peptide is removed extracellularly to produce this active, mature form of PSA. Human kallikrein 2 has been proposed as a possible in vivo activator of proPSA.[9] The proPSA secreted from epithelial cells is truncated as a result of posttranslational proteolytic cleavage of the pro-leader peptide. Truncated pro-leader peptides containing 1 to 5 amino acids have been reported, all of which remain enzymatically inactive.[9–11] These truncated forms of proPSA have recently generated a great deal of clinical interest since they may have a diagnostic role in the early detection of prostate cancer, especially in the new, extended gray zone of 2.5 to 10 ng/ml. Another investigation has already revealed the positive impact of proPSA in the tPSA range of 2.5 to 4.0 ng/ml.[12] Therefore, we determined whether proPSA could positively influence the detection of early prostate cancer through decreasing false positives in patients with tPSA ranging from 4 to 10 ng/ml.

MATERIALS AND METHODS

Patient selection. Between July 1999 and October 2000 as part of a screening program which involved digital rectal examination (DRE) and serum tPSA measurement, 93 men were noted to have an increased tPSA between 4.0 and 10.0 ng/ml. Each of these men subsequently underwent transrectal, ultrasound guided, systematic, 12-core biopsies that included sampling of the lateral peripheral zone for prostate cancer detection. No prior PSA values were available on any of these patients and none had undergone a previous prostate biopsy. DRE was classified as suspicious or nonsuspicious for prostate cancer, and the biopsy result was classified as with or without evidence of prostate cancer. The urologist who performs the biopsy conducts the DRE examination before the procedure. The classification of suspicious means evidence of induration or a nodule and/or significant gland asymmetry is noted during the digital rectal examination.

Detection of total, free and proPSA. Archival serum samples from the 93 men obtained before biopsy were stored at −80C. These samples were analyzed for tPSA, fPSA (Hybritech Tandem Assays, Beckman Coulter Access, Beckman Coulter, Inc.), and [-2]proPSA, [-4p]proPSA, and [-5,-7]proPSA (the sum of [-5]proPSA and [-7]proPSA). Percent free PSA, sum-proPSA (the cumulative sum of individual proPSA forms) and percent sum-proPSA (sum-proPSA/fPSA) were calculated.[12] The Hybritech Access free and total PSA assays are 2-site immuno-enzymatic assays that use mouse monoclonal antibody in alkaline phosphatase conjugate and paramagnetic particles coated with a second mouse monoclonal antibody. After unbound particles are removed by washing, a chemiluminescent substrate, Lumi-Phos 530 (Lumigen, Inc., Southfield, Michigan), is added to produce light directly proportional to the amount of analyte in the sample as determined from a stored calibration curve. The individual [-2]proPSA, [-4p]proPSA and [-5,-7]proPSA forms (measured in ng/ml) were analyzed using individual time resolved fluorescence immunoassays (Beckman Coulter, Inc., San Diego, California, research use only, not for diagnostic use). The immunoassay format was a dual monoclonal sandwich assay in a microliter plate under standard conditions using a biotinylated capture anti-PSA mAb, and Europium labeled proPSA-specific mAbs for detection with a Victor 1420 multilabel counter (Wallac, EG&G PerkinElmer Life Sciences Inc., Boston, Massachusetts). These immunoassays have less than 0.2% cross-reactivity with mature PSA and other proPSA forms. The analytical detection limit was 0.01 ng/ml and the assays had a clinical detection limit in serum of 0.025 ng/ml.

Statistical analysis. All statistical analyses were performed using the STATA v7.0 software package (Stata Corporation, College Station, Texas). Differences in average test results between benign disease and prostate cancer samples were examined using the Student's t-test. Logistic regression (LR) analysis was also used to fit a model of the variables to predict the disease outcome of benign vs prostate cancer. The areas under the receiver operator characteristic curve (ROC-AUC), as well as sensitivity and specificity, were used to assess the diagnostic performance of the assays. Before starting the study informed consent was obtained from each patient.

RESULTS

Of the 93 men 85 (93%) assessed were white. Prostate cancer was detected using a 12-core biopsy procedure in 44% and benign disease was detected in 56% of the men (41 of 93 and 52 of 93, respectively). Even though this biopsy sampling strategy optimizes prostate cancer detection achieving rates of 44%, clearly as the gland increases in size the ability to detect small cancers may be compromised. Men with prostate cancer and those with benign disease were similarly distributed for age, DRE findings and tPSA (table 1). Of the men diagnosed with prostate cancer 76% had Gleason score (GS) 5 or 6 disease (1 and 30 men, respectively), 19% had GS 7 disease (8 men) and 5% had GS 8 disease (2). Of the 41 men diagnosed with prostate cancer, 14 have had anatomical radical retropubic prostatectomy performed to date. None of these men have evidence of disease recurrence after a mean followup of 2 years (range 1.5 to 2.5).

Using univariate LR, fPSA, %fPSA, percent sum-proPSA and prostate volume all significantly (p <0.05) differentiated men with prostate cancer from those with benign disease (table 2). Applying stepwise backward logistic regression and accepting variables that have a Pz ≤0.05, a multivariate model for predicting men with PCa was created. The model retained tPSA, %fPSA and sum-proPSA optimally to differentiate the men with prostate cancer from those with benign disease (p <0.001). The ROC-AUC for this model solution was 76.6% (see figure). This ROC curve has an AUC that is statistically significantly different than those generated for tPSA and sum-proPSA but not %fPSA. Table 3 compares the specificity for the individual variables as well as the complete LR model at 3 different sensitivities. Note that at 90% sensitivity these 3 variables collectively achieved a specificity of 44% for the detection of prostate cancer. Individually, the 3 retained independent variables had a specificity of 23% 33% and 13%, respectively. There was a consistent improvement in specificity in the complete LR model at the various sensitivities assessed.

DISCUSSION

In 1991 it was discovered that serum contains 2 distinct, major, immunodetectable forms of PSA. One form covalently binds primarily to a serum protease inhibitor (α_1-antichymotrypsin) as well as other less dominant inhibitors and is known as complexed PSA, while the other form is a "free," nonactive, noncomplexed form known as free PSA.[13,14] The measurement of the ratio of free and total PSA has led to a modest but significant improvement in the discrimination of prostate cancer from benign prostatic hyperplasia (BPH)

TABLE 1. *Comparison of age (years) total PSA and DRE in men with and without prostate cancer*

	Benign Disease		PCa		p Value
Mean age (range)	63	(46–76)	64	(47–77)	0.4
DRE stage (range)	T1c	(T1c–T2a)	T1c	(T1c–T2b)	<0.06
Mean tPSA (range)	6.8	(4.1–9.9)	6.3	(4.0–9.1)	<0.08

TABLE 2. *The ability of various serum markers to differentiate prostate cancer from benign disease*

Variable	Mean (range)			% ROC-AUC	p Value (LR)
	Benign Disease		PCa		
tPSA (ng/ml)	6.8	(4.1–9.9)	6.3 (4.0–9.1)	60.37	<0.072
fPSA (ng/ml)	1.13 (0.51–2.50)		0.79 (0.28–1.76)	70.80	<0.0002
%fPSA	16.7	(5.6–27.5)	12.7 (6.4–24.2)	70.59	<0.0005
Sum-proPSA (ng/ml)	0.28 (0.01–0.91)		0.26 (0.05–0.88)	54.03	0.5
%sum-proPSA	25.3	(2.7–80.7)	32.7 (4.6–75.4)	66.32	<0.03
Gland vol (ml)	57.6	(18–144)	39.63 (15–94)	71.06	<0.0002
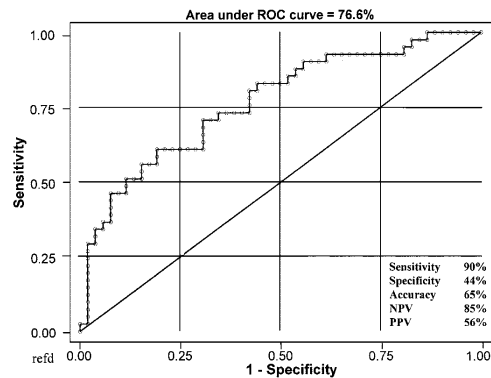
ROC curve for total PSA, free PSA and sum-proPSA to detect prostate cancer in 93 men. *NPV*, negative predictive value. *PPV*, positive predictive value.

in men with a PSA between 4.0 and 10.0 ng/ml.[15, 16] This improvement is due to the association of BPH with higher levels of fPSA compared with prostate cancer.[15, 16] As such, measurement of the free-to-total PSA ratio has been shown to improve specificity by avoiding 20% of unnecessary biopsies and yet is able to detect 95% of cancers when the %fPSA cutoff of 25% is used.[15, 16] Catalona et al[16] also recently reported on a large, multicenter study which determined the ability of %fPSA and age specific PSA to detect prostate cancer in 773 men with histologically confirmed diagnosis (379 with prostate cancer and 394 with benign prostatic disease). The study demonstrated that in the presence of tPSA ranging from 4 to 10 ng/ml, %fPSA maintained significantly higher sensitivities than age specific PSA cutoffs for the detection of prostate cancer in men older than 60 years.

In an attempt to discriminate prostate cancer from benign disease further, differences within molecular subpopulations of fPSA have been extensively investigated. It is now known that fPSA is composed of at least 3 distinct forms of enzymatically inactive PSA. One form is called BPSA, an internally cleaved or degraded form of PSA that is more highly associated with BPH.[9] Another form is thought to contain a number of minor variants but appears to be composed largely of intact PSA that is similar to native, active PSA except for structural or conformational changes that have rendered it enzymatically inactive.[7] The third has been identified as the proenzyme or precursor form of PSA, also known as proPSA.[7] Truncated proPSA refers to proPSA in which any of the normal 7 amino acids in the pro-leader peptide have been removed as a result of posttranslational proteolytic cleavage of the pro-leader peptide.[17] All forms of the truncated proPSA containing any of the pro-leader amino acids remain enzymatically inactive. It is these truncated forms of proPSA that have recently generated a great deal of interest since there is strong evidence to support the association of proPSA with prostate cancer.[7, 8, 10, 11, 18]

Examining prostate tissue from patients undergoing radical prostatectomy, Mikolajczyk et al[18] reported that proPSA was increased in cancer areas compared to noncancerous tissue. The authors subsequently developed monoclonal antibodies to detect proPSA with Western blot analysis in the PSA purified from the serum of 5 men with biopsy proven prostate cancer (tPSA range 6 to 24 ng/ml, mean 13.4) and 3 men with negative biopsies (tPSA range 7 to 12 ng/ml, mean 9.7).[10] Multiple proPSA forms were detected at higher levels in the cancer serum but the truncated [-2]proPSA was found to be the most consistently associated with cancer in the 5 samples tested when compared to benign serum. Peter et al[11] conducted a similar study using mass spectrometry to detect PSA forms where they obtained serum from 5 patients with prostate cancer and detected [-7], [-5], [-4], [-2] and [-1]proPSA forms. The [-2]proPSA was present in 4 of 5 samples, and was the highest pro form in 2 of the samples. The mass spectrometry technique used by Peter et al[11] necessitated the use of serum with much higher serum tPSA values (ie tPSA was 1,890 ng/ml in 1 patient and the other 4 samples had PSA values greater than 6,000 ng/ml). The recent development of [-2], [-4] and [-5,-7]proPSA immunoassays has enabled us to expand the sample population significantly and to determine whether serum proPSA measurement provides any additional information in the detection of prostate cancer in patients with clinically relevant serum tPSA levels ranging from 4 to 10 ng/ml.

Our study has shown that sum-proPSA (the total of the [-2]proPSA, [-4]proPSA, and [-5,-7]proPSA) with the addition of tPSA and %fPSA improves the specificity of prostate cancer detection by more than 10% at 90% sensitivity compared with any of these serum biomarkers measured individually. Thus, the combination of serum sum-proPSA, tPSA, and %fPSA may provide added clinical value resulting in fewer men with tPSA ranges from 4 to 10 ng/ml being unnecessarily biopsied, decreasing repeat biopsy frequency, and could possibly be of use for men undergoing a watchful waiting protocol for small tumors.[15, 16, 19] Clearly a larger multicenter study is needed to confirm our preliminary findings and to assess economic impact.

If it is confirmed that measurement of combined proPSA forms provides the most diagnostic value, a single assay for all of these forms has been developed and would have to be

TABLE 3. *Comparison of specificity and accuracy to detect PCa at 95%, 90% and 85% sensitivity for individual variables vs multivariate LR model*

	% Sensitivity					
	95		90		85	
Variable	Specificity	Accuracy	Specificity	Accuracy	Specificity	Accuracy
tPSA	15	51	23	53	25	52
%fPSA	27	57	33	58	37	57
Sum-proPSA	10	47	13	47	25	52
Gland vol	17	52	27	55	33	56
LR multivariate model	37	61	44	63	48	65

approved by the Food and Drug Administration. However, even with only 1 additional PSA test, we would have to determine if the economics of health care savings in PCa management outweigh additional PSA testing fees. In a study involving 2,138 men from 7 institutions, Ellison et al[20] recently performed such an economic evaluation of Food and Drug Administration approved PSA molecular assays and concluded that complexed PSA was most cost-effective. The use of combinations of markers to optimize disease management for clinical decisions will be a challenge in the future as more biomarkers surface through advances in proteomics technology. Since current evidence suggests that no single serum marker is likely to achieve the desired level of diagnostic specificity to detect cancer, true advances in prostate cancer detection may depend on the development of multiplexed assays that can simultaneously measure multiple analytes in a single serum sample. Multiplexed assays could include all of the markers described in this study in addition to existing markers such as human kallikrein 2, interleukin-6 and transforming growth factor-β insulin growth factor-1 currently under investigation for PCa.

CONCLUSIONS

The usefulness of %fPSA in the tPSA range of 4 to 10 ng/ml, especially in the background of normal DRE (clinical stage T1c), is widely accepted considering its significant increase in specificity for prostate cancer detection compared with the use of only tPSA. To improve cancer detection further and avoid unnecessary biopsies, a great deal of interest has recently been generated in the proPSA component of fPSA, which has been shown to be increased in the serum of patients with prostate cancer. Our study demonstrates that sum-proPSA, tPSA and %fPSA in combination significantly improve the specificity of early prostate cancer detection in men with a tPSA of 4 to 10 ng/ml when compared with any of these individual PSA molecular forms measured alone.

REFERENCES

1. Catalona, W. J., Richie, J. P., Ahmann, F. R., Hudson, M. A., Scardino, P. T., Flanigan, R. C. et al: Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. J Urol, **151:** 1283, 1994
2. Oesterling, J. E., Jacobsen, S. J., Chute, C. G., Guess, H. A., Girman, C. J., Panser, L. A. et al: Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges. JAMA, **270:** 860, 1993
3. Seaman, E., Whang, M., Olsson, C. A., Katz, A., Cooner, W. H. and Benson, M. C.: PSA density (PSAD). Role in patient evaluation and management. Urol Clin North Am, **20:** 653, 1993
4. Catalona, W. J., Smith, D. S., Wolfert, R. L., Wang, T. J., Rittenhouse, H. G., Ratliff, T. L. et al: Evaluation of percentage of free serum prostate-specific antigen to improve specificity of prostate cancer screening. JAMA, **274:** 1214, 1995
5. Brawer, M. K., Cheli, C. D., Neaman, I. E., Goldblatt, J., Smith, C., Schwartz, M. K. et al: Complexed prostate specific antigen provides significant enhancement of specificity compared with total prostate specific antigen for detecting prostate cancer. J Urol, **163:** 1476, 2000
6. Veltri, R. W., Miller, M. C., O'dowd, G. J. and Partin, A. W.: Impact of age on total and complexed prostate-specific antigen cut-offs in a contemporary referral series of men with prostate cancer. Urology, suppl., **60:** 47, 2002
7. Mikolajczyk, S. D., Marks, L. S., Partin, A. W. and Rittenhouse, H. G.: Free prostate-specific antigen in serum is becoming more complex. Urology, **59:** 797, 2002
8. Kumar, A., Mikolajczyk, S. D., Goel, A. S., Millar, L. S. and Saedi, M. S.: Expression of pro form of prostate-specific antigen by mammalian cells and its conversion to mature, active form by human kallikrein 2. Cancer Res, **57:** 3111, 1997
9. Mikolajczyk, S. D., Millar, L. S., Wang, T. J., Rittenhouse, H. G., Wolfert, R. L., Marks, L. S. et al: "BPSA," a specific molecular form of free prostate-specific antigen, is found predominantly in the transition zone of patients with nodular benign prostatic hyperplasia. Urology, **55:** 41, 2000
10. Mikolajczyk, S. D., Marker, K. M., Miller, L. S., Kumar, A., Saedi, M. S., Payne, J. K. et al: A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. Cancer Res, **61:** 6958, 2001
11. Peter, J., Unverzagt, C., Krogh, T. N., Vorm, O. and Hoesel, W.: Identification of precursor forms of free prostate-specific antigen in serum of prostate cancer patients by immunosorption and mass spectrometry. Cancer Res, **61:** 957, 2001
12. Sokoll, L. J., Chan, D. W., Mikolajczyk, S. D., Rittenhouse, H. G., Evans, C. L., Linton, H. J. et al: Proenzyme psa for the early detection of prostate cancer in the 2.5-4.0 ng/ml total psa range: preliminary analysis. Urology, **61:** 274, 2003
13. Lilja, H., Christensson, A., Dahlen, U., Matikainen, M. T., Nilsson, O., Pettersson, K. et al: Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin. Clin Chem, **37:** 1618, 1991
14. Stenman, U. H., Leinonen, J., Alfthan, H., Rannikko, S., Tuhkanen, K. and Alfthan, O.: A complex between prostate-specific antigen and alpha 1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostate cancer: assay of the complex improves clinical sensitivity for cancer. Cancer Res, **51:** 222, 1991
15. Veltri, R. W. and Miller, M. C.: Free/total PSA ratio improves differentiation of benign and malignant disease of the prostate: critical analysis of two different test populations. Urology, **53:** 736, 1999
16. Catalona, W. J., Southwick, P. C., Slawin, K. M., Partin, A. W., Brawer, M. K., Flanigan, R. C. et al: Comparison of percent free PSA, PSA density, and age-specific PSA cutoffs for prostate cancer detection and staging. Urology, **56:** 255, 2000
17. Khan, A. R. and James, M. N.: Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. Protein Sci, **7:** 815, 1998
18. Mikolajczyk, S. D., Millar, L. S., Wang, T. J., Rittenhouse, H. G., Marks, L. S., Song, W. et al: A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue. Cancer Res, **60:** 756, 2000
19. Carter, H. B., Walsh, P. C., Landis, P. and Epstein, J. I.: Expectant management of nonpalpable prostate cancer with curative intent: preliminary results. J Urol, **167:** 1231, 2002
20. Ellison, L., Cheli, C. D., Bright, S., Veltri, R. W. and Partin, A. W.: Cost-benefit analysis of total, free/total, and complexed prostate-specific antigen for prostate cancer screening. Urology, suppl., **60:** 42, 2002

ADULT UROLOGY
CME ARTICLE

CLINICAL UTILITY OF proPSA AND "BENIGN" PSA WHEN PERCENT FREE PSA IS LESS THAN 15%

MASOOD A. KHAN, LORI J. SOKOLL, DANIEL W. CHAN, LESLIE A. MANGOLD, PHAEDRE MOHR,
STEPHEN D. MIKOLAJCZYK, HARRY J. LINTON, CINDY L. EVANS, HARRY G. RITTENHOUSE, AND
ALAN W. PARTIN

ABSTRACT

Objectives. To investigate the clinical utility of the subforms of free prostate-specific antigen (PSA), namely proPSA and "benign" PSA (BPSA), to improve cancer detection when the percent free PSA level is less than 15%. Percent free PSA, while maintaining sensitivity, has greatly improved the specificity of PSA for the early detection of prostate cancer. A low percent free PSA value indicates a greater risk of cancer, but only 30% to 50% of men with percent free PSA levels of less than 15% actually have cancer at biopsy.

Methods. Archived sera from 161 consecutive men who were prospectively enrolled in our Early Detection Research Network prostate cancer early detection biomarker program with a percent free PSA value of less than 15% were included in the study. Total PSA, free PSA, proPSA, and BPSA were measured for each sample.

Results. The mean total PSA was 6.1 ng/mL (range 1.8 to 24.0). The mean age of the study group was 62 $\pm$ 7 years. Prostate cancer was detected in 66 (41%) of 161 men. The area under the curve–receiver operating characteristic for total and percent free PSA was 0.51 and 0.54, respectively. BPSA and proPSA/BPSA both improved cancer detection compared with percent free PSA alone; the improvement was statistically significant (P <0.001) . The area under the curve–receiver operating characteristic for proPSA/BPSA was 0.72, giving a sensitivity and specificity of 90% and 46%, respectively.

Conclusions. The results of our preliminary studies have suggested that the ratio of proPSA and BPSA can distinguish cancer with greater accuracy when the percent free PSA value is very low (less than 15%), and may, therefore, provide better clinical utility in this lower range of percent free PSA. UROLOGY **64:** 1160–1164, 2004. © 2004 Elsevier Inc.

The development and subsequent routine use of prostate-specific antigen (PSA) during the past decade has revolutionized the management of prostate cancer. PSA measurement has increased our ability to detect and, in turn, treat early prostate cancer. However, the major drawback of PSA is its relative lack of specificity. This is especially important in the critical diagnostic range of 4.0 to 10.0 ng/mL, in which an elevated PSA level may reflect either prostate cancer or benign disease such as benign prostatic hyperplasia (BPH). At this range of total PSA, it is associated with a specificity of approximately 25% for the detection of prostate cancer.[1,2] This limited specificity has led to unnecessary prostate biopsies, with associated anxiety, cost, and potential morbidity.

It was discovered in 1991 that serum contains two distinct major forms of PSA. One form covalently bound to endogenous serum protease inhibitors such as alpha$_1$-antichymotrypsin is known as complexed PSA, the other form is present as free, noncomplexed PSA and is known as free PSA.[3,4] The measurement of the ratio of free and total PSA (ie, percent free PSA) has led to a modest but statistically significant improvement (approximately 25%) in the discrimina-

TABLE I. *Ability of total serum PSA, percent free PSA, and subforms of free PSA to predict prostate cancer in 161 men with percent free PSA level less than 15%*

Variable	Noncancer	Cancer	AUC-ROC	*P* Value
Age (yr)	62 ± 7	62 ± 8	0.53	NS
Total PSA (ng/mL)	6.0 ± 3.0	6.4 ± 4.0	0.51	NS
Percent free PSA (%)	10.0 ± 3.0	10.0 ± 3.0	0.54	NS
proPSA (ng/mL)	0.3 ± 0.2	0.25 ± 0.2	0.55	NS
BPSA (ng/mL)	0.4 ± 0.4	0.2 ± 0.2	0.72	<0.001
proPSA/BPSA	0.9 ± 0.7	1.33 ± 0.7	0.72	<0.001
Abnormal DRE	21/66	19/95	0.56	NS
Prostate volume (cm^3)	45 ± 20	37 ± 19	0.63	<0.05

KEY: *AUC-ROC = area under the curve receiver operating characteristic; NS = not significant; PSA = prostate-specific antigen; BPSA = benign PSA; DRE = digital rectal examination.*
Data presented as mean ± standard deviation.

tion of prostate cancer from benign prostatic hyperplasia (BPH) in men with PSA levels between 4.0 and 10.0 ng/mL.[5,6] This is because of the association of BPH with high levels of free PSA compared with the levels associated with prostate cancer. Despite this, it is estimated that only 30% to 50% of men with percent free PSA levels less than 15% will have cancer at biopsy.[5]

The free PSA subforms represent a heterogeneous group consisting of at least three different subforms of inactive PSA.[7] One form has been identified as the proenzyme or precursor form of PSA (proPSA), which is normally expressed with a 7 amino acid pro-leader peptide. However, proPSA with truncated pro-leader peptides have been identified that appear to be more associated with prostate cancer.[7–9] Immunoassays have been developed that are specific for proPSA forms containing pro-leader peptides with 2, 4, and 5 plus 7 amino acids, [−2], [−4], and [−5,−7]proPSA, respectively (Beckman Coulter, for research use only; not for diagnostic procedures). Studies with these assays have shown that proPSA, within the total PSA range of 2.5 to 10 ng/mL, can significantly improve the specificity of prostate cancer detection.[10–12] In contrast to proPSA, "benign" PSA (BPSA) is a component of the free PSA that has been shown to be associated with BPH.[13,14] Serum levels of BPSA have been shown to increase in men with BPH.[12] However, the presence of prostate cancer does not alter the relative proportions of BPSA in sera.[14] We, therefore, investigated whether proPSA and BPSA could be used in men with free PSA levels less than 15% to improve the detection rate of prostate cancer.

MATERIAL AND METHODS

STUDY GROUP

A total of 161 men (mean age 62 ± 7 years; range 46 to 80) with a percent free PSA level less than 15% were prospectively enrolled in our Early Detection Research Network prostate cancer early detection biomarker program. These men were

selected on the basis of voluntary presentation to our institution for prostate cancer screening and were subsequently noted to have a percent free PSA level of less than 15%. Before undertaking transrectal ultrasound-guided biopsies (12 or more cores), serum was obtained and stored at −80°C until ready for analysis. Biopsies were taken from all men, irrespective of the digital rectal examination (DRE) findings. On histologic examination, 66 (41%) of 161 had a diagnosis of prostate cancer and 95 (59%) had no evidence of cancer. Of the 66 cancer cases, 63 (95%) were clinically significant with a Gleason score of 6 or greater (49 [74%] with a Gleason score of 6, 12 [17%] with a Gleason score of 7, and 2 [3%] with a Gleason score of 8 to 9).

All data, as well as the serum samples, were collected under an institutional review board-approved protocol that passed HIPAA compliance. All men provided written informed consent.

MEASUREMENT OF TOTAL PSA, FREE PSA, proPSA, AND BPSA

Total and free PSA were measured using Hybritech Tandem assays (Beckman Coulter Access, Beckman Coulter, Brea, Calif). ProPSA, defined here as the cumulative sum of the truncated [−2]pPSA, [−4]pPSA, [−5]pPSA forms, as well as the native proPSA containing a 7-amino acid pro-peptide, [−7]pPSA,[9] was measured using individual time-resolved fluorescence research immunoassays at Johns Hopkins and Beckman Coulter. These assays have less than 1% cross-reactivity with other PSA forms. The BPSA assay used alkaline phosphate-labeled detect antibody and also had less than 1% cross-reactivity with other forms of PSA.[14] Because of the newness of the proPSA and BPSA assays, limited stability studies have been performed.[15] However, the current results have not indicated the presence of stability problems.[15] In addition, specimens for this study had been frozen at −80°C before analysis.

STATISTICAL ANALYSIS

All statistical analyses were performed using the STATA, version 8.0, software package (Stata, College Station, Tex). The areas under the curve–receiver operating characteristic (AUC-ROCs), as well as the sensitivity and specificity, were used to assess the diagnostic performance of the subforms of free PSA.

RESULTS

The differences in the serum total PSA, percent free PSA, and proPSA levels were not statistically
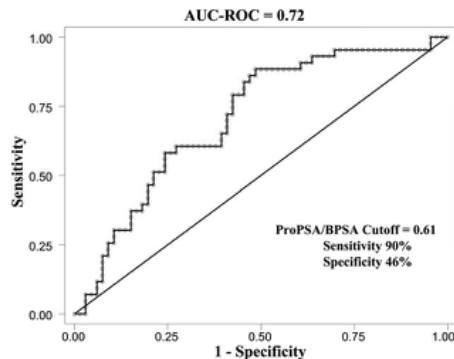
FIGURE 1. *ROC curve for ability of proPSA/BPSA to detect prostate cancer in 161 men with percent free PSA less than 15%.*

significant between the cancer and noncancer groups (Table I). The prostate volume was significantly greater statistically (P <0.05) in the noncancer than in the cancer groups (45 ± 20 g versus 37 ± 19 g). Although a greater number of abnormal DRE findings were detected in the cancer than in the noncancer groups (21 [32%] of 66 versus 18 [19%] of 95, respectively), this did not reach statistical significance. BPSA was statistically significantly greater (P <0.001) and ratio of proPSA/BPSA was statistically significantly lower (P <0.001) in the noncancer group (Table I). The clinical utility of proPSA and BPSA were compared with that of percent free PSA using ROC analysis. The AUC-ROCs for percent free PSA and proPSA were very similar (0.54 and 0.55, respectively; Table I). To this end, unsurprisingly, the percent proPSA level (proPSA/free PSA) also failed to distinguish between the two groups (data not shown). However, the AUC-ROC for BPSA and proPSA/BPSA was significantly greater than the percent free PSA (P <0.001 for both), with both producing a value of 0.72 (Table I). Furthermore, using a proPSA/BPSA cutoff of 0.61, a sensitivity and specificity of 90% and 46%, respectively, was achieved for the detection of prostate cancer (Fig. 1). Although the AUC-ROCs of BPSA and proPSA/BPSA were identical at 0.72, BPSA alone had a statistically significantly lower specificity of 20% (P <0.05) when the sensitivity was maintained at 90% for the detection of prostate cancer.

The racial demographic of the study population with prostate cancer was as follows: 58 men (88%) were white, 5 (8%) were African American, and 3 (4%) other. Of those with a negative biopsy, 79 (83%) were white, 8 (8.5%) African American, and 8 (8.5%) other.

COMMENT

The availability of serum PSA as a biomarker for the early detection of prostate cancer has dramatically changed our approach to the management of this common disease. Before its availability, we relied primarily on DRE findings to determine the presence of prostate cancer. However, DRE is not only unreliable, even by experienced examiners, but is also frequently associated with understaging of prostate cancer.[16] To this end, the widespread use of serum PSA measurement has resulted in a 20% increase in the detection of clinically localized prostate cancer that are, in turn, amenable to definitive therapy.[17] However, the major drawback associated with serum PSA is its relative lack of specificity, particularly with total serum PSA levels in the "gray" zone (ie, 4.0 to 10.0 ng/mL) in which an elevated serum PSA value may reflect either prostate cancer or benign disease, such as BPH. Accordingly, within this range, serum PSA is associated with a specificity of approximately 25% for the detection of prostate cancer.[1,2] In an attempt to improve specificity and, thereby reduce the number of unnecessary biopsies, attention during the past decade on the role of the free (ie, unbound) component of serum PSA. Even though percent free PSA determination significantly improves the specificity of prostate cancer detection, it is noteworthy that most men with percent free PSA levels less than 15% will have negative prostate biopsies.[5] Therefore, in an attempt to improve the specificity further, components of free PSA have recently received great interest.

Free PSA is thought to be composed of at least three components, namely proPSA, BPSA, and "inactive" PSA.[7] PSA is normally found in seminal fluid in a mature, active form with 237 amino acids lacking the pro-leader peptide. The pro-leader peptide is removed extracellularly, possibly by human glandular kallikrein 2, to produce the active mature form of PSA.[7] Intracellularly, the proPSA form is itself truncated as a result of post-translational proteolytic cleavage of the pro-leader peptide in which 2 to 6 of the 7 pro-leader amino acids are removed. All forms of truncated pPSA remain enzymatically inactive. It is these truncated forms of pPSA that have recently generated a great deal of clinical interest as they may have a diagnostic role in the early detection of prostate cancer.[7,18] In support of this, we have recently reported that proPSA can result in the detection of 75% of cancer while sparing 59% of unnecessary biopsies in men with a total serum PSA level in the range of 2.5 to 4.0 ng/mL.[10] However, within this total serum PSA range, the percent free PSA would have resulted in sparing only 33% of unnecessary biopsies. We have also recently demonstrated that proPSA signifi-

cantly improves the specificity of prostate cancer detection in men presenting with total serum PSA levels in the range of 4.0 to 10.0 ng/mL.[11] In addition, a recent two-site study of 1091 men by Catalona *et al.*[12] has shown that percent proPSA determination significantly improved detection throughout the 2 to 10-ng/mL PSA range, and that the [−2]pPSA form of proPSA was particularly useful in the 2 to 4-ng/mL PSA range. That large study confirmed earlier preliminary studies that proPSA enhances prostate cancer detection. Our current study, unsurprisingly, failed to demonstrate the ability of percent proPSA to distinguish between the two groups because this study was biased toward men with low percent free PSA, which, along with proPSA, did not differ significantly between the two groups.

In contrast to proPSA, BPSA is a degraded form of free PSA that contains peptide bond cleavages at Lys145 and Lys182. BPSA was initially described in nodular tissue samples from the transition zone of BPH tissue,[14] subsequently in seminal plasma,[19] and finally in the serum[14] of men with benign prostatic disease. Tissue samples obtained from enlarged prostates secondary to BPH and from prostate cancer demonstrated significantly greater BPSA expression in the transition zone.[13] Using purified BPSA as an immunogen, a specific assay for measuring BPSA was developed that revealed that an estimated 15% to 50% of free PSA in men with benign disease is BPSA and was undetectable in the sera from normal controls.[14]

Our current study has revealed that men presenting with similar age and total serum PSA levels, with a mean total serum PSA range within the "gray zone," along with low percent free PSA values (less than 15%), BPSA and the ratio of proPSA/BPSA can significantly differentiate between the cancer and non-cancer groups with AUC-ROC values of 0.72. Furthermore, using a cutoff of 0.61 for the proPSA/BPSA ratio, a specificity of 46% while maintaining sensitivity at 90% for prostate cancer detection was achieved. BPSA alone had the same overall AUC-ROC as proPSA/BPSA (Table I), but had a significantly lower specificity of 20% at 90% sensitivity. We also determined that within the subgroup of men presenting with a total serum PSA level in the range of 4.0 to 10.0 ng/mL (43 men with cancer and 66 men with no evidence of cancer), BPSA and the pro/BPSA ratio were also able to distinguish between the two groups with near identical results to those of the overall study (data not shown). We, therefore, in the future, hope to develop an algorithm applying proPSA and BPSA data in men presenting with low percent free PSA and total serum PSA levels in the "gray" zone to predict accurately the likelihood of harboring prostate cancer, and thereby reduce the number of unnecessary prostate biopsies.

The findings from this study required four PSA tests: total PSA, free PSA, proPSA, and BPSA. A single, pan-proPSA assay has been developed that can measure all the proPSA forms (data not shown). Although the application of four or more tests may seem financially unrealistic in common medical practice, it is becoming evident that no single marker can be expected to provide sufficient value owing to the heterogeneous nature of prostatic disease. The challenges for prostatic disease diagnosis include cancer detection, distinguishing significant cancer from indolent cancer, identifying metastasis, treatment options for PSA failure after prostatectomy, and monitoring and treatment for benign conditions such as BPH that can interfere with prostate cancer detection. Thus, panels of serum markers, combined with algorithms, may be the future of prostate diagnostics to achieve statistically significant improvements over current tests. Additional studies with proPSA and BPSA are in progress to investigate other aspects of prostate disease.

CONCLUSIONS

Serum PSA has dramatically improved our ability to detect early prostate cancer. However, it is limited by its low specificity, particularly in the critical diagnostic range of 4.0 to 10.0 ng/mL in which an elevated serum PSA level may reflect either prostate cancer or benign disease. Despite the improvements in specificity associated with the use of percent free PSA, most men presenting with a percent free PSA level of less than 15% will have negative biopsies. Our study has demonstrated that the use of BPSA and the ratio of proPSA/BPSA in men presenting with a low percent free PSA level (less than 15%) and mean total serum PSA level within the "gray zone" can accurately distinguish prostate cancer from noncancer cases. On the basis of our preliminary study, we hope to develop an algorithm that can be implemented to predict accurately the likelihood of harboring prostate cancer and, in turn, result in a reduction in the number of unnecessary prostate biopsies performed.

REFERENCES

1. Woodrum DL, Brawer MK, Partin AW, *et al:* Interpretation of free prostate specific antigen clinical research studies for the detection of prostate cancer. J Urol **159**: 5–12, 1998.
2. Levine MA, Ittman M, Melamed J, *et al:* Two consecutive sets of transrectal ultrasound guided sextant biopsies of the prostate for the detection of prostate cancer. J Urol **159**: 471–476, 1998.
3. Lilja H, Christensson A, Dahlan U, *et al:* Prostate-specific antigen in serum occurs predominantly in complex with α_1-antichymotrypsin. Clin Chem **37**: 1618–1625, 1991.

4. Stenman UH, Leinonen J, Alfthan H, *et al:* A complex between prostate specific antigen and α_1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostate cancer: assay of the complex improves clinical sensitivity for cancer. Cancer Res **51:** 222–226, 1991.

5. Catalona WJ, Partin AW, Slawin KM, *et al:* Use of the percentage of free PSA to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. JAMA **279:** 1542–1547, 1998.

6. Martinez-Pineiro L, Tabernero A, Contreras T, *et al:* Determination of the percentage of free prostate-specific antigen helps to avoid unnecessary biopsies in men with normal rectal examination and total prostate-specific antigen of 4–10 ng/ml. Eur Urol **37:** 289–296, 2000.

7. Mikolajczyk SD, Marks LS, Partin AW, *et al:* Free prostate-specific antigen in serum is becoming more complex. Urology **59:** 797–802, 2002.

8. Mikolajczyk SD, Grauer LS, Miller LS, *et al:* A precursor form of PSA (pPSA) is a component of the free PSA in prostate cancer serum. Urology **50:** 710–714, 1997.

9. Mikolajczyk SD, Marker KM, Miller LS, *et al:* A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. Cancer Res **61:** 6958–6963, 2001.

10. Sokoll LJ, Chan DW, Mikolajczyk SD, *et al:* Proenzyme PSA for the early detection of prostate cancer in the 2.5–4.0 ng/ml total PSA range: preliminary analysis. Urology **61:** 274–276, 2003.

11. Khan MA, Partin AW, Rittenhouse HG, *et al:* Evaluation of pro-PSA for early detection of prostate cancer in men with a total PSA range of 4.0–10.0 ng/ml. J Urol **170:** 723–726, 2003.

12. Catalona WJ, Bartsch G, Rittenhouse HG, *et al:* Serum proPSA improves cancer detection compared to free and complexed PSA in men with PSA values from 2 to 4 ng/ml. J Urol **170:** 2181–2185, 2003.

13. Mikolajczyk SD, Millar LS, Wang TJ, *et al:* "BPSA," a specific molecular form of free prostate-specific antigen, is found predominantly in the transition zone of patients with nodular benign prostatic hyperplasia. Urology **55:** 41–45, 2000.

14. Linton HJ, Marks LS, Millar LS, *et al:* Benign prostate-specific antigen (BPSA) in serum is increased in benign prostatic disease. Clin Chem **49:** 253–259, 2003.

15. Mikolajczyk SD, Catalona WJ, Evans CL, *et al:* Proenzyme forms of prostate-specific antigen in serum improve the detection of prostate cancer. Clin Chem **50:** 1017–1025, 2004.

16. Partin AW, Yoo JK, Carter HB, *et al:* The use of prostate-specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. J Urol **150:** 110–114, 1993.

17. Catalona WJ, Smith DS, Ratliff TL, *et al:* Detection of organ-confined prostate cancer is increased through prostate-specific antigen based screening. JAMA **270:** 948–954, 1993.

18. Mikolajczyk SD, Miller LS, Wang TJ, *et al:* A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue. Cancer Res **60:** 756–759, 2000.

19. Mikolajczyjk SD, Millar LS, Marker KM, *et al:* Seminal plasma contains "BPSA" a molecular form of prostate specific antigen that is associated with benign prostatic hyperplasia. Prostate **45:** 271–276, 2000.

The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography–guided prostate biopsies

Raj P. Pal, Muhammad Elmussareh, Malek Chanawani and Masood A. Khan
Department of Urology, University Hospitals of Leicester, Leicester General Hospital, UK

Study Type – Diagnostic (exploratory cohort)
Level of Evidence   2b

What's known on the subject? and What does the study add?
Template assisted transperineal biopsy of the prostate has become increasingly popular over the past decade. Several studies have demonstrated that transperineal prostate biopsy (TPB) is associated with an increased rate of cancer detection, increased histological concordance with final prostatectomy samples and an increase in anterior and apical prostate cancers than standard TRUS biopsy. However, interpretation of the literature is difficult due to considerable variation between studies in terms of technique and equipment.

We examined a small cohort ($n = 40$) of patients using a standardized 36 core template assisted TPB technique. We show that utilising this technique is associated with high cancer (68%) detection rate in patients with two previous negative TRUS biopsies. Of patients were found to have anterior gland tumours which would not have been detected by standard TRUS guided biopsy.

OBJECTIVE

• To determine the efficacy and safety of a standardized 36 core template-assisted transperineal biopsy technique for detecting prostate cancer in patients with previously negative transrectal ultrasonography-guided prostate biopsies and elevated prostate-specific antigen (PSA) levels.

PATIENTS AND METHODS

• Between April 2008 to September 2010, a total of 40 patients with a mean (range) age of 63 (49–73) years, a mean (range) elevated PSA level of 21.9 (4.7–87) ng/mL and two previous sets of negative TRUS-guided prostate biopsies underwent standardized 36 core template-assisted transperineal prostate biopsies under general anaesthetic as a day case procedure.
• The cancer detection rate and complications for all cases were evaluated.

RESULTS

• In total, 27 of 40 (68%) patients were found to have adenocarcinoma of the prostate, two patients (5.0%) had atypical small acinar proliferation, one had high-grade prostatic intraepithelial neoplasia (2.5%), four (10%) had chronic active inflammation and six (15%) had benign histology.
• Gleason scores were in the range 6–9, with a median Gleason score of 7.
• There were no cases of urosepsis, urinary tract infections or haematuria. A single patient experienced acute urinary retention, with a subsequent succesful trial without a catheter, and haematospermia was common, although minor.

CONCLUSIONS

• Our standardized 36 core template-assisted transperineal prostate biopsy technique is safe and associated with a high detection rate of prostate cancer.
• This technique should be considered in patients with elevated PSA levels and previously negative TRUS-guided prostate biopsies.

KEYWORDS

INTRODUCTION

Prostate cancer is the most commonly diagnosed cancer and the second most common cause of death from cancer in men [1]. The widespread use of PSA testing has contributed to a dramatic increase in the number of men undergoing TRUS-guided prostate biopsies [2,3]. Data from the USA suggest that more than 1.2 million needle prostate biopsies are performed each year [4].

TRUS was introduced in 1968 as an imaging tool to assist in the diagnosis of prostate cancer [5]. TRUS allows biopsies to be accurately guided towards the peripheral zone where cancers predominate, therefore achieving reasonable sampling of the

367

PAL *ET AL.*

prostate. It has a sensitivity of 39–52% and a specificity of 82% [6]. Additionally, it carries low morbidity and can be performed in the office setting. Hence, systematic TRUS-guided prostate biopsies remain the gold standard for detecting prostate cancer. However, a limitation of this technique is that men with an initial negative biopsy are often found to have subsequent prostate cancer. Almost one-quarter of prostate cancers are identified after an initial negative biopsy [7]. Furthermore, the cancer detection rate decreases with an increasing number of biopsy sessions, with yields of 10–20% for the second biopsy and below 10% for subsequent biopsies [8,9]. Repeated TRUS-guided biopsy results in sampling of the same prostatic areas and other potential tumour sites can be missed. The question of whether to pursue further repeat TRUS-guided biopsy for patients with a rising PSA level subsequent to an initial negative biopsy is a common clinical dilemma and remains a diagnostic challenge.

Many studies have proposed a number of different biopsy approaches in these patients, including repeated standard biopsies and transrectal saturation biopsies [10,11]. More recently, transperineal prostate biopsy (TPB) utilizing a brachytherapy template grid has become increasingly popular but, currently, a standardized TPB technique does not exist [12–15]. In the present study, we determined the role of a standardized 36 core transperineal template prostate biopsy technique in the detection of prostate cancer in men with an elevated PSA levels and at least two previous negative TRUS-guided prostate biopsies.

PATIENTS AND METHODS

Between Januray 2008 and September 2010, a total of 40 men were selected to undergo template-assisted TPB at the Leicester General Hospital. Men were selected if they had had at least two previous sets of TRUS-guided prostate biopsies yielding a non-cancerous diagnosis and an elevated PSA level. All patients had a minimum life expectancy of 10 years. Additionally, a further five patients with known prostate adenocarcinoma who were managed with active surveillance were selected for template-assisted TPB as part of their surveillance regime.

All patients underwent template-assisted TPB, which was performed by the same surgeon (M.A.K.) using the same biopsy technique in every case. The procedure was performed under general anaesthetic. At the time of induction, a combination of intravenous co-amoxiclav and gentamicin were used as antibiotic prophylaxis. A 14-F urethral catheter was inserted and the patient was placed in an extended lithotomy position. A DRE was performed. This was followed by insertion of the transrectal probe (BK Medical Pro-Focus 2202; BK Medical, Mileparken, Denmark) and prostate volume measurement. A STEPPER (Galil Medical; Crawley, Sussex, UK) with an articulated arm and a stabilizer was used to fix the ultrasonographic probe. Ater prepping and draping the perineum, a standard 0.5-cm brachytherapy template grid was attached to the STEPPER and positioned over the perimeum.

With the prostate at its widest in the transverse plane, the gland was divided on the TRUS screen into six zones (right anterior, left anterior, right mid, left mid, right posterior and left posterior) (Fig. 1a). Next, six needles were inserted at a time into a single zone, and biopsies through an 18-gauge needle were taken from the apex of the prostate to the base. The ultrasonographic probe was then switched to the sagittal plane and the needles were withdrawn (Fig. 1b). A total of 36 biopsies were taken from each case for histological analysis. In patients with larger prostates (>50 mL) where the anterior part of the gland could not be reached, biopsy needles were manually manipulated to ensure the whole gland was sampled. The urinary catheters were removed at the end of the procedure and the patients were discharged home the same day after voiding. Patients who experienced post-procedure retention were recatheterized and re-attended for a trial without catheter the next day.

RESULTS

Mean (range) patient age was 62.9 (49–73) years. Mean (range) PSA level was 21.9 (4.7–119) ng/mL at the time of template-assisted TPB compared to 18.7 (3.7–122) ng/mL at the time of the previous TRUS biopsy session. The PSA and previous histological results for each patient before template-assisted TPB are shown in Table 1. The mean

FIG. 1. *a*, Transverse plane view of prostate subdivided into six zones (RA, right anterior; LA, left anterior; RM, right mid; LM, left mid; RP, right posterior; LP, left posterior). In total, six biopsies were taken from each zone. *b*, Sagittal view of the prostate showing biopsy needles placed in the anterior part of the gland (arrows). B, catheter balloon; c, catheter; P, prostate; R, rectum.
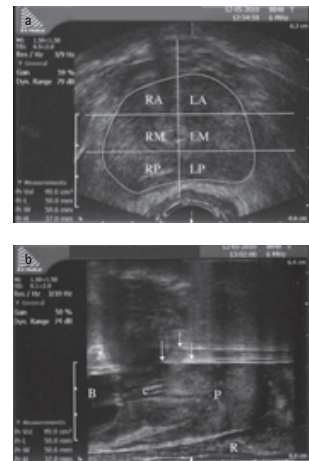
(range) number of biopsies taken at the most recent TRUS biopsy session was 11.6 (10–18). The median (range) time from the most recent TRUS biopsy session to TPB was 16 (2–58) months.

Table 1 summarizes the clinical and histological findings for all patients after template-assisted TPB. In the present study cohort of patients, 27 of 40 (68%) patients had prostate adenocarcinoma diagnosed after template-assisted TPB. The remaining histological findings were: two patients (5%) had atypical small acinar proliferation, one had high-grade prostatic intraepithelial neoplasia (2.5%), four (10%) had chronic active inflammation and six (15%) had benign histology.

In the patients with prostate cancer, the median (range) number of positive cores was 8 (1–22). In total, 21 of 27 (71.4%) patients who had prostate adenocarcimona on the template-assisted TPB had cancer involving five or more cores. Out of 27 patients, 16 (59.3%) patients had a Gleason

TABLE 1 *Clinical parameters for patients with benign histology before transperineal prostate biopsy (TPB)*

Case number	Number of past biopsy sessions	Histology from previous biopsy	PSA at most recent TRUS biopsy (ng/mL)	PSA at TPB (ng/mL)	TRUS volume (mL)	TPB histology	Number of cores containing cancer
1	2	Benign	8.8	5.6	35	chronic active inflammation	
2	2	Benign	7.0	7.4	37	ASAP	
3	2	HGPIN	7.7	8.7	30	Adenocarcinoma (3 + 3 = 6)	17
4	2	Benign	4.2	8.9	78	HPIN	
5	2	HGPIN	6.2	9.5	45	Adenocarcinoma (3 + 3 = 6)	8
6	2	HGPIN	9.4	10	98	Benign	
7	3	Benign	11	11	31	Benign	
8	3	Benign	6.7	11	50	Adenocarcinoma (3 + 3 = 6)	1
9	2	Benign	4.9	12	21	Adenocarcinoma (4 + 5 = 9)	22
10	3	Benign	11	14	42	Adenocarcinoma (3 + 3 = 6)	9
11	5	ASAP	20	19	32	Adenocarcinoma (3 + 3 = 6)	6
12	2	Benign	8.7	20	29	Adenocarcinoma (4 + 5 = 9)	9
13	3	Benign	13.1	21	110	Benign	
14	2	HGPIN	8.7	21	55	Benign	
15	2	Benign	9.4	22	71	chronic active inflammation	
16	2	Benign	23	23	28	Adenocarcinoma (3 + 3 = 6)	5
17	2	Benign	27	25	22	Adenocarcinoma (3 + 4 = 7)	12
18	2	HGPIN	44	28	41	Adenocarcinoma (4 + 3 = 7)	6
19	2	Benign	15	29	40	Adenocarcinoma (4 + 4 = 8)	10
20	2	Benign	20	32	40	Adenocarcinoma (3 + 3 = 6)	6
21	2	Benign	18	36	50	Adenocarcinoma (3 + 4 = 7)	10
22	2	Benign	40	40	66	Adenocarcinoma (3 + 3 = 6)	1
23	2	Benign	40	41	29	Adenocarcinoma (3 + 3 = 6)	6
24	2	HGPIN	44	41	47	Adenocarcinoma (4 + 5 = 9)	16
25	2	Benign	72	87	30	Adenocarcinoma (4 + 3 = 7)	20
26	2	Benign	24	24	37	Adenocarcinoma (3 + 3 = 6)	10
27	2	ASAP	6.9	6.9	30	Adenocarcinoma (3 + 3 = 6)	2
28	2	ASAP	5.6	5.6	32	Adenocarcinoma (3 + 3 = 6)	1
29	2	ASAP	4.7	4.7	26	Adenocarcinoma (3 + 3 = 6)	9
30	2	Benign	122	119	60	Adenocarcinoma (3 + 5 = 8)	15
31	2	Benign	5.8	8.8	49	Adenocarcinoma (3 + 3 = 6)	11
32	2	HGPIN	13	13	72	chronic active inflammation	
33	4	Benign	5.6	9.8	60	Benign	
34	3	HGPIN	17	33	31	Adenocarcinoma (4 + 4 = 8)	7
35	2	Benign	7.8	11	85	Adenocarcinoma (3 + 3 = 6)	2
36	2	HGPIN + ASAP	11	10	89	Adenocarcinoma (3 + 4 = 7)	2
37	2	HGPIN	8.6	11	61	ASAP	
38	2	Benign	3.7	6.2	20	Adenocarcinoma (3 + 3 = 6)	5
39	2	Benign	Not known	24	32	Benign	
40	2	HGPIN	9.7	4.7	28	chronic active inflammation	

ASAP, atypical small acinar proliferation; HGPIN, high-grade prostatic intraepithelial neoplasia.

score of 6, five (18.5%) patients had a Gleason score of 7, three (11.1%) patients had a Gleason score 8 and three (11.1%) patients had a Gleason score of 9. Cancer involving the anterior zone was present in 12 of 27 (44.4%) patients. In four of these patients, the tumour was localized to the anterior zone, and did not affect the remainder of the gland. The mean PSA level for patients diagnosed with prostate adenocarcinoma was 26.1 vs 13.0 ng/mL for those with non-cancerous histology.

The TRUS biopsy and template-assisted TPB results of those patients already diagnosed with prostate cancer and treated with active surveilance are outlined in Table 2. Of five patients, four experienced an increase in Gleason grade, and three out of five patients had a higher percentage of positive cores on repeat template-assisted TPB compared to the initial TRUS biopsy. The mean (range) time interval between TRUS and template-assisted TPB was 11.6 (4–15) months.

There were no cases of urosepsis, urinary tract infection or haematuria. Although

haematospermia was common, it was minor because our patients did not raise any concerns. In the one patient who developed urinary retention, a catheter was replaced and the patient was re-admitted the next week and had a successful trial without catheter.

DISCUSSION

TRUS-guided biopsy is the standard approach for men with suspected prostate cancer. However, the use of TPB has become increasingly popular over the past decade. Several studies have now investigated this technique as a diagnostic alternative to TRUS-guided biopsy in both men with and without previous biopsy. These studies have shown a reduction in upstaging at the time of radical prostatectomy and more concordance with final prostatectomy specimen biopsy results [14]. Additionally, studies sampling a higher number of biopsy yield an increased cancer detection rate, particularly in the apical and anterior regions of the prostate [12,16,17].

However, a difficulty in the interpretation of the current literature results from a lack of standardization of a TPB technique. Both within and between individual studies, there is a significant variation in patient selection and the number of biopsies sampled, as well as with respect to whether a template grid is used [12–20]. In the present study, we report data obtained using a standardized 36 core biopsy template-assisted TPB technique for all patients. All patients selected for biopsy at our institute had at least two previous negative TRUS-guided biopsies and all had elevated PSA levels.

Although several series of TPB have now been published, only one randomized control trial has compared TPB against TRUS biopsy, and this was carried out in a cohort without previous prostatic biopsy [15]. Although no significant difference in cancer detection rates were noted between both techniques, that study utilized only a 12 core biopsy technique for both TPB and TRUS biopsy. However, in comparison, a recent study investigating a template-assisted TPB technique sampling a mean core biopsy number of 54 in patients without previous prostate biopsy reported far higher cancer detection rates in patients without previous biopsy (79% vs 48%), and this was in a

cohort of patients with lower mean PSA levels and a higher prostate volume, where fewer cancers would be expected [16]. Cancer detection rates after one or more benign sextant or octant biopsies have been reported to be in the range 13–34% when using TRUS-guided saturation biopsy [10,11]. Pinkstaff *et al.* [20] and Demura *et al.* [13] reported cancer detection rates of 36% and 37%, respectively, when sampling between 20 to 21 cores using a template-assisted TPB technique. In a study conducted by Bott *et al.* [18] in men with high-grade prostatic intraepithelial neoplasia or atypical small acinar proliferation on previous TRUS biopsy, the cancer detection rate was 38% when a median of 24 cores were sampled using a template-assisted TPB technique. When template-assisted TPB was performed using a mean core biopsy technique of 53.8 in men with negative TRUS biopsy, Bittner *et al.* [17] showed that cancer detection rates were 44.7%. Studies sampling a higher number of biopsy cores at transperineal prostate have yielded high prostate cancer detection rates. Accordingly, it is evident that prostate cancer detection rates are dependent on the number of biopsy cores sampled. However, exactly how many cores are required for the optimum biopsy strategy, as well as when this should be performed, remains to be determined.

Our current method of TTMB is based on a 36 core biopsy technique that samples three times more biopsies than our standard TRUS biopsy technique. In the present study, we report a cancer detection rate of 68%, which is higher than previous studies, including those sampling a larger number of cores, although the mean PSA level for our study group is greater than that reported in previous studies [13,17,18,20]. A potential concern of the increasing detection rates by more substantial sampling is the increase in clinically insignificant cancers. However,

previous studies have shown that this is unlikely to be the case, with only a minority of clinically insignificant cancers being detected when using a transperineal technique [18,20]. Of the 27 patients diagnosed with cancer in the present study, 24 patients had more than one core involved, and 21 patients had more than five cores involved, indicating that the most cancers could potentially be clinically significant tumours. However, it should be noted that five positive cores on template-assisted TPB equates to 13.8% of the cores being positive for cancer. If translated as a percentage to a 12-core TRUS guided biopsy technique (although this is not an entirely accurate comparison), between one and two cores would be positive, equating to a relative low volume of disease.

Additionally, other studies have investigated the utility of TPB as the initial biopsy technique in patients with an elevated PSA level or abnormal DRE [15,16]. Although associated with potentially higher cancer detection rates [16] because of the need for a general anaesthetic, offering TPB as an initial biopsy strategy to all patients may not be feasible as a result of time and financial constraints in many centres.

A limitation of the present study is the sample size employed. In particular, it is difficult to assess the safety of a standardized 36 core technique in this small cohort. However, only one patient reported transient mild haematuria, one patient experienced urinary retention and haematospermia, although mild, was common. Other complications were not reported. The safety profile of this TPB is very similar to TRUS-guided biopsy. Only one randomized study has investigated the complication rates between these two approaches [15]. Although lower rates

TABLE 2 *Clinical parameters of active surveillance patients undergoing repeat template-assisted transperineal prostate biopsy (TPB)*

Initial TRUS biopsy			Template-assisted TPB		
PSA (ng/mL)	Gleason score	Number of cores	PSA (ng/mL)	Gleason score	Number of cores
4	3 + 3 = 6	1/10	8.6	3 + 4 = 7	4/36
9.1	3 + 3 = 6	1/10	12	3 + 4 = 7	2/36
4.2	3 + 3 = 6	1/10	5.1	3 + 3 = 6	7/36
4.0	3 + 3 = 6	2/10	6.1	3 + 4 = 7	11/36
5.3	3 + 3 = 6	1/10	8.0	3 + 4 = 7	1/36

of retention and sepsis were reported in patients undergoing TPB compared to a TRUS biopsy, these were statistically insignificant. A further limitation of the present study is that the time interval between template-assisted TPB and TRUS biopsy was greater than 1 year in some instances. It may be that the disease progressed in some of these men during this period, explaining why cancer was detected on template-assisted TPB rather than previous TRUS biopsy sessions. However, identifying the time when to perform repeat biopsy remains a dilemma for many clinicians. Not only is this heavily influenced by patient choice after a negative biopsy, but also repeat biopsy within a short time period may result in unnecessary discomfort and patient morbidity.

In conclusion, we present our early data on a standardized 36 template-assisted TPB technique. In the present study cohort, we report a high cancer detection rate. A current difficulty in determining the utility of template-assisted TPB is related to the lack of a standardized technique for all patients. The present study supports the role of a 36 core biopsy technique for patients with elevated PSA levels despite previous negative TRUS-guided biopsies. The finding of the present study support the need for further investigations in the form of a prospective multicentre randomized control trial.

CONFLICT OF INTEREST

None declared.

REFERENCES

1 **Cancer Research UK.** News and Resources Centre Cancer Stats. Prostate Cancer 2008. Available at: http://info.cancerresearchuk.org/cancerstats/. Accessed July 2010

2 **Barry MJ.** Clinical practice. Prostate-specific-antigen testing for early diagnosis of prostate cancer. *N Engl J Med* 2001; **344**: 1373–7

3 **Frankel S, Smith GD, Donovan J, Neal D.** Screening for prostate cancer. *Lancet* 2003; **361**: 1122–8

4 **Haker SJ, Mulkern RV, Roebuck JR** *et al.* Magnetic resonance guided prostate interventions. *Top Magn Reson Imaging* 2005; **16**: 355–68

5 **Heijmink SW, van Moerkerk H, Kiemeney LA, Witjes JA, Frauscher F, Barentsz JO.** A comparison of diagnostic performance of systematic versus ultrasound guided biopsies of prostate cancer. *Eur Radiol* 2006; **16**: 927–38

6 **Norberg M, Egevad L, Bolmberg L, Sparen P, Norlen BJ, Busch C.** The sextant protocol for ultrasound guided core biopsies of the prostate underestimates the presence of cancer. *Urology* 1997; **50**: 562–6

7 **Roehl KA, Antenor JA, Catalona WJ.** Serial biopsy results in prostate cancer screening study. *J Urol* 2002; **167**: 2435–9

8 **Djavan B, Ravery V, Zlotta A** *et al.* Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? *J Urol* 2001; **166**: 1679–83

9 **Roehrborn CG, Pickens GJ, Sanders JS.** Diagnostic yield of repeated transrectal ultrasound-guided biopsies stratified by specific histopathologic diagnoses and prostate specific antigen levels. *Urology* 1996; **47**: 347–52

10 **Stewart CS, Leibovich BC, Weaver AL, Lieber MM.** Prostate cancer diagnosis using a saturation needle biopsy technique after previous negative sextant biopsies. *J Urol* 2001; **166**: 86–91

11 **Fleshner N, Klotz L.** Role of 'saturation biopsy' in the detection of prostate cancer among difficult diagnostic cases. *Urology* 2002; **60**: 93–7

12 **Moran BJ, Braccioforte MH.** Stereotactic transperineal prostate biopsy. *Urology* 2009; **73**: 386–8

13 **Demura T, Hioka T, Furuno T** *et al.* Differences in tumor core distribution between palpable and nonpalpable prostate tumors in patients diagnosed using extensive transperineal ultrasound-guided template prostate biopsy. *Cancer* 2005; **103**: 1826–32

14 **Emiliozzi P, Maymone S, Paterno A** *et al.* Increased accuracy of biopsy Gleason score obtained by extended needle biopsy. *J Urol* 2004; **172**: 2224–6

15 **Hara R, Jo Y, Fujii T** *et al.* Optimal approach for prostate cancer detection as initial biopsy: prospective randomized study comparing transperineal versus transrectal systematic 12-core biopsy. *Urology* 2008; **71**: 191–5

16 **Taira AV, Merrick GS, Galbreath RW** *et al.* Performance of transperineal template-guided mapping biopsy in detecting prostate cancer in the initial and repeat biopsy setting. *Prostate Cancer Prostatic Dis* 2010; **13**: 71–7

17 **Bittner N, Merrick GS, Andreini H** *et al.* Prebiopsy PSA velocity not reliable predictor of prostate cancer diagnosis, Gleason score, tumor location, or cancer volume after TTMB. *Urology* 2009; **74**: 171–6

18 **Bott SR, Henderson A, Halls JE, Montgomery BS, Laing R, Langley SE.** Extensive transperineal template biopsies of prostate: modified technique and results. *Urology* 2006; **68**: 1037–41

19 **Li H, Yan W, Zhou Y** *et al.* Transperineal ultrasound-guided saturation biopsies using 11-region template of prostate: report of 303 cases. *Urology* 2007; **70**: 1157–61

20 **Pinkstaff DM, Igel TC, Petrou SP, Broderick GA, Wehle MJ, Young PR.** Systematic transperineal ultrasound-guided template biopsy of the prostate: three-year experience. *Urology* 2005; **65**: 735–9

Correspondence: Raj P Pal, Department of Urology, Leicester General Hospital, Gwendolen Road, Leicester LE5 4 PW, UK. e-mail: rppal@doctors.org

Abbreviation: **TPB**, transperineal prostate biopsy.

ORIGINAL ARTICLE

Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies

Shady Nafie · Raj P. Pal · John P. Dormer · Masood A. Khan

Abstract

Background The possibility of prostate cancer as a cause for steadily rising PSA despite previously negative transrectal ultrasound (TRUS)-guided prostate biopsies is a major concern. An initial negative TRUS-guided prostate biopsy does not necessarily exclude the presence of clinically significant prostate cancer. We determined the role of transperineal template prostate biopsy (TPTPB) in prostate cancer detection in men with raised PSA despite two previous sets of negative TRUS biopsies.

Methods Between January 2008 and August 2012, a total of 122 men's records were reviewed after having 36-core TPTPB following two previous sets of negative TRUS biopsies despite raised PSA. A retrospective record of PSA levels, clinicopathological parameters and histological outcomes was made.

Results Mean age was 63 years (range 49–77), and mean PSA was 18.0 (range 2.0–119.0). A total of 71/122 (58 %) men were diagnosed with prostate cancer on TPTPB. Of these, 28 (39 %), 34 (48 %), 5 (7 %), and 4 (6 %) had Gleason score 6, 7 (3 + 4), 7 (4 + 3), and 9 (4 + 5), respectively. The mean number of positive cores was 7 (range 1–22). Of these, only 15 (21 %) had ≤2 cores positive and Gleason score of 6. Of the 51 (42 %) men with a negative histology on TPTPB, 11 (22 %), 10 (19 %), and 30 (59 %) had atypical small acinar proliferation, high-grade prostatic intraepithelial neoplasia, or benign pathology.

Conclusion TPTPB is associated with a high rate of clinically significant prostate cancer diagnosis (58 %) in men with raised PSA despite two previous sets of negative TRUS biopsies.

Keywords Prostate · TRUS · Transperineal · Biopsy · Cancer · Detection

S. Nafie (✉) · R. P. Pal · M. A. Khan
Department of Urology, University Hospitals of Leicester NHS Trust, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, UK
e-mail: shady.nafie@me.com

R. P. Pal
e-mail: rpp3@le.ac.uk

M. A. Khan
e-mail: Masood.Khan@uhl-tr.nhs.uk

J. P. Dormer
Department of Cellular Pathology, University Hospitals of Leicester NHS Trust, Leicester Royal Infirmary, Infirmary Square, Leicester LE5 4PW, UK
e-mail: john.dormer@uhl-tr.nhs.uk

Introduction

The possibility of prostate cancer as a cause for steadily rising PSA despite previously negative transrectal ultrasound (TRUS)-guided prostate biopsies is a major concern. In addition, previous studies have reported that TRUS-guided prostate biopsies has a sensitivity of only 39–52 % [1], cancer detection rate of around 25 % on initial biopsies [2], and 18–32 % in the repeated biopsies [3, 4]. Therefore, an initial negative TRUS-guided prostate biopsies does not necessarily exclude the presence of clinically significant prostate cancer. Hence, modification of the standard biopsy technique along with TRUS-guided saturation biopsies has been performed in an attempt to increase the specificity of prostate cancer detection [5–8]. Unfortunately, variations of the standard TRUS-guided prostate biopsy techniques have failed to demonstrate a significant increase in the diagnosis of prostate cancer [9].

Springer

Transperineal template prostate biopsies (TPTPB) is a technique which is gaining popularity and is associated with a high prostate cancer detection rates in men with a background of a rising PSA and previously negative TRUS-guided prostate biopsies. We have previously reported our preliminary findings of the prostate cancer detection rate in 40 men who underwent a standardized 36-core TPTPB due to rising PSA despite two previous sets of negative TRUS-guided prostate biopsies [10]. Our preliminary data revealed prostate cancer detection rates approaching 70 %, but this study was limited by a relatively low sample size. In this present study, we now present retrospective data in a cohort of 122 men who have undergone the same TPTPB technique due to a persistently rising PSA after two sets of negative TRUS-guided prostate biopsies.

Patients and methods

Between January 2008 and September 2012, a total of 122 men were selected to undergo TPTPB. Men were selected if they had two previous sets of standard 12-core TRUS-guided prostate biopsies yielding a non-cancerous diagnosis and an increased PSA. All patients had a minimum life expectancy of 10 years.

All patients underwent TPTPB that was performed by the same surgeon (MAK) using the same biopsy technique in every case as previously described [10]. The procedure was performed under general anesthetic. A 14-Fr urethral catheter was inserted and the patient was placed in the extended lithotomy position. This was followed by insertion of the transrectal probe (BK Medical Pro-Focus 2202, BK Medical, Mileparken, Denmark) and prostate volume measurement. A STEPPER (Galil Medical; Crawley, Sussex, UK) with an articulated arm and a stabilizer was used to fix the ultrasound probe. After prepping and draping the perineum, a standard 0.5 cm brachytherapy template grid was attached to the STEPPER and positioned over the perineum. With the prostate at its widest in the transverse plane, the gland was divided on the TRUS screen into 6 sectors (right anterior, left anterior, right mid, left mid, right posterior and left posterior). Six 18-gauge needles were inserted at a time into a single sector in the transverse plane and biopsies were taken from the apex of the prostate to the base in the sagittal plane. A total of 36 biopsies were taken from each case for histological analysis. The urinary catheters were removed at the end of the procedure unless the prostate volume was >60 cc, in which case a trial without catheter (TWOC) was performed 5 days later. All patients were discharged home the same day. A retrospective record of clinicopathological parameters and histological outcomes was made. PSA velocity was calculated as per Carter et al. using the first and last PSA values (FL method) [11, 12]. Student's t test was used for statistical analysis and $p < 0.05$ was considered significant. As TPTPB was offered to this cohort as part of a routine practice in our unit, we did not seek ethical approval.

Results

The mean age of the 122 patients was 63 years (range 49–73) with mean prostate volume of 54.0 cc (range 16.0–140.0). Mean PSA level was 18.0 ng/mL (range 2.0–119.0) at the time of TPTPB, which was higher compared to 14.0 ng/mL (range 1.4–122.0) at the time of the previous TRUS biopsy ($p < 0.01$). The mean time interval between the previous TRUS biopsy and the TPTPB was 24 months (range 2–128) with mean PSA velocity of 2.3 ng/mL per year (range 0–33.6).

Out of the 122 men, a total of 71 (58 %) were found to have a malignant pathology, while 51 (42 %) men were found to have non-cancer histology. Mean PSA density (PSAD) was 0.5 ng/mL/cc (range 0–2.9) in the cancer cohort compared to 0.3 ng/mL/cc (range 0.06–1.36) in the non-cancer cohort ($p = 0.9$). Table 1 demonstrates the histological findings along with Gleason scores. Figure 1 shows the distribution of the cancer positive cores.

Out of the 71 patients with prostate cancer, 28 (39 %) had Gleason score of 6 (3 + 3), 34 (48 %) had Gleason score of 7 (3 + 4), 5 (7 %) had Gleason score of 7 (4 + 3), and 4 (6 %) had Gleason score of 9 (4 + 5). The mean number of positive cores was 7 (range 1–22), with mean tumor volume of 9 % of the positive cores (range 0.3–54 %). A total of 15 (21 %) patients had ≤3 positive cores, Gleason score of 6, and tumor volume <50 %. In total 221 cancer positive cores were found, of these, 108 (49 %) were found in the anterior zone, 69 (31 %) in the middle zone, and 44 (20 %) in the posterior zone. Out of the 51 (42 %) patients with a non-malignant diagnosis, 11 (22 %) men had atypical small acinar proliferation, 10 (19 %) men had high-grade prostatic intraepithelial neoplasia, and 30 (59 %) men had a benign pathology.

There was no significant difference in the mean PSA levels at the time TPTPB between the 71 patients with prostate cancer (20.2 ng/mL, range 4.6–119.0) and the 51 men with non-cancer histology (14.9 ng/mL, range 1.6–70.0) ($p = 0.09$). Furthermore, there was no significant difference in the mean PSA velocity of the both groups (2.9 vs. 1.1 ng/mL/year, $p = 0.17$).

There were no cases of urosepsis. Urinary retention occurred in 5 % of the cases and hematuria in 30 % of the cases. Minor hematospermia was common but did not raise any patient concerns. A minority of patients experienced mild perineal ecchymosis and/or swelling.

Table 1 Pathology findings in 122 patients who underwent the TPTPB

TPTP pathology	No. of patients	Age			Prostate volume			TPTP PSA			No. of +ve cores		
		Mean	Range		Mean	Range		Mean	Range		Mean	Range	
3 + 3	28	62	76	52	51	140	20	15	48	4.6	5	15	1
3 + 4	34	64	77	51	41	89	16.1	22	119	5.2	9	20	2
4 + 3	5	69	77	61	42	52	31	36	93	16	13	19	8
4 + 5	4	64	68	56	39	60	21.2	25	41	12	13	22	8
Benign	51	63	77	49	67	132	25	15	70	4.6	0	0	0
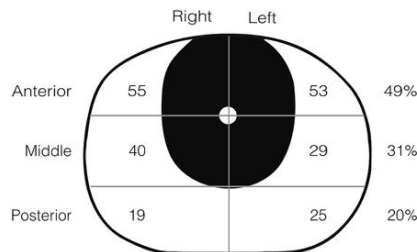
Fig. 1 Distribution of cancer positive cores found by the TPTPB in the prostate sectors

Discussion

Since the introduction of the original sextant biopsy by Hodge et al. [13], TRUS-guided prostate biopsy has been regarded as the gold standard approach to the diagnosis of prostate cancer in men with an increased serum PSA. However, the probability of missing a cancer with prostate biopsy under TRUS guidance is estimated at 25 %, even with saturation biopsies [4, 5, 7, 9, 14]. Furthermore, with each successive biopsy, the risk of prostate cancer detection falls [15]. This has in turn resulted in many men being initially falsely reassured by a negative biopsy.

TPTPB is a technique that has evolved and gained more popularity over the last decade. A study [16] assessed TPTPB in the initial and repeated biopsy setting and concluded that it provides high rates of cancer detection of 77.9 % in biopsy naïve men and a rate of 46.9 % as a repeat biopsy. Although different authors use varying transperineal biopsy techniques, cancer detection rates following TPTPB in men with previously negative TRUS biopsy is higher than reported with saturation biopsies via the transrectal route [5, 17, 18]. This has partly been attributed to the fact that TPTPB detects, in particular, tumors located in the apical and anterior zones of the prostate, often poorly sampled by TRUS biopsy [19, 20]. We previously reported high cancer detection rates using a standardized 36-core transperineal biopsy technique. In our initial study, 68 % of patients who had had two previously negative TRUS biopsies underwent TPTPB, all having 36 cores sampled in a standardized manner. However, our initial study was limited by a small cohort size, and this present study was performed to evaluate this technique in a larger number of patients. By sampling 36 cores at the time of TPTPB in 122 men with two negative prior TRUS biopsies, we again report high cancer detection rates of 58 %, again which compares favorably to other studies [16–19, 21–23].

Earlier data regarding the cancer detection yield of TPTPB, as with TRUS biopsy, demonstrates that the cancer detection rate appears to vary depending on the number of cores sampled. In studies sampling 20 cores, the cancer detection rate in men with prior negative TRUS biopsy is around 35 % [24], whereas in those studies sampling almost 50 cores report higher cancer detection rates approaching 50 % [18]. However, more recently, Ekweume et al. [25] noted similarly high cancer detection rates of 54 %—as in our study—in a cohort of patients undergoing a median of 28 biopsies following a median of two set of negative TRUS biopsies. Furthermore, Bittner et al. [23] reported a cancer detection rate of 46.6 % for transperineal biopsy with 86.7 % clinically significant cancers in a cohort of patients undergoing a median of 59 biopsies following at least one negative set of TRUS biopsies. Together with our data, this indicates that studies sampling higher numbers of cores may be unnecessary and does not increase the diagnostic yield. It must be borne in mind that clinical parameters and the selection criteria for transperineal biopsies (i.e., PSA, PSA velocity and number of previous TRUS biopsy sessions) vary from study to study and this will have a clear bearing on the cancer detection rates reported [10, 16–20, 24].

As reported by previous authors [19, 25, 26], in our study a high proportion of biopsies containing cancer were located in the anterior zone of the prostate. Almost 50 % of cancer-containing cores were found in the anterior zone of the prostate, a region poorly sampled during TRUS biopsy, hence, the high proportion of anterior tumors is to be expected. As previous studies have shown that TPTPB is

able to adequately sample anterior tumors, and TRUS biopsy is able to adequately sample posterior tumors, and in our cohort all patients had two previous sets of TRUS biopsies, one would expect that TRUS biopsies would have detected the majority of the posterior located tumors. Nonetheless, a substantial number of cancers (20 %) were located in the posterior zone, highlighting further limitations of TRUS biopsy in comparison to TPTPB.

Following TPTPB, concerns have been raised regarding the clinical significance of the cancers diagnosed through this route [18, 24]. However, in our present study, 43 (61 %) had a Gleason grade score ≥7, automatically placing them into the intermediate or high-risk categories when using established risk stratification criteria [21]. Furthermore, in our cohort only one-fifth of patients were noted to have clinically insignificant cancer when using the Epstein criteria, which has been described in studies evaluating saturation TRUS biopsy [27]. Hence, in our cohort, TPTPB appeared to detect tumors that require treatment. Despite detecting a number of patients with moderately or poorly differentiated disease in our and other studies [10, 25], studies have demonstrated that in comparison to TRUS biopsy, TPTPB identifies tumors of a smaller size and earlier stage [28]. This data could suggest that TPTPB identifies more clinically significant tumors at an earlier stage, making it a far more ideal diagnostic test for localized prostate cancer than TRUS biopsy.

One limitation in this study, however, is that the time difference between the previous TRUS biopsy, and TPTPB was an average of 2 years; hence, natural progression to cancer could have happened during this period accounting for the high detection rate of cancer following TPTPB. Although the mean PSA value was higher in cancer cases, our data showed that there was no significant difference in rate of cancer diagnosis in patients who had the TPTPB biopsy within 1 year from last TRUS biopsy and those who had it after this period. A further question mark arises as to the timing of TPTPB in patients with suspected prostate cancer. Several authors have shown that TPTPB in biopsy naïve patients is associated with high cancer detection rates [16], but due to the need for a general anesthetic, offering TPTPB as an initial biopsy strategy to all patients may not be feasible as a result of time and financial constraints in many centers.

We predict that in the future such a cohort will initially undergo a multiparametric MRI scan of the prostate and then proceed to TPTPB, which may or may not be focused. However, further studies are needed before this can be confirmed.

Conclusion

In conclusion, we report a high cancer detection rate of 58 % by TPTPB in a cohort of 122 men with rising PSA despite previous two sets of negative TRUS-guided biopsies. Most cancers involved the anterior zone with vast majority of Gleason score ≥7. Hence, TPTPB should be regarded as the gold standard investigation in such cases, and consideration must be given to performing this technique in men with one negative TRUS biopsy in whom cancer is suspected.

Conflict of interest   Authors declare no conflict of interest.

References

1. Norberg M, Egevad L, Holmberg L et al (1997) The sextant protocol for ultrasound-guided core biopsies of the prostate underestimates the presence of cancer. Urology 50:562–566. doi:10.1016/S0090-4295(97)00306-3
2. Naughton CK, Miller DC, Mager DE et al (2000) A prospective randomized trial comparing 6 versus 12 prostate biopsy cores: impact on cancer detection. J Urol 164:388–392
3. Yuasa T, Tsuchiya N, Kumazawa T et al (2008) Characterization of prostate cancer detected at repeat biopsy. BMC Urol 8:14. doi:10.1186/1471-2490-8-14
4. Aganovic D, Prcic A, Kulovac B, Hadziosmanovic O (2011) Prostate cancer detection rate and the importance of premalignant lesion in rebiopsy. Med Arh 65:109–112
5. Stewart CS, Leibovich BC, Weaver AL, Lieber MM (2001) Prostate cancer diagnosis using a saturation needle biopsy technique after previous negative sextant biopsies. J Urol 166:86–91 discussion 91–92
6. Stamey TA (1995) Making the most out of six systematic sextant biopsies. Urology 45:2–12
7. Fleshner N, Klotz L (2002) Role of "saturation biopsy" in the detection of prostate cancer among difficult diagnostic cases. Urology 60:93–97
8. Walz J, Graefen M, Chun FKH et al (2006) High incidence of prostate cancer detected by saturation biopsy after previous negative biopsy series. Eur Urol 50:498–505. doi:10.1016/j.eururo.2006.03.026
9. Jones JS, Patel A, Schoenfield L et al (2006) Saturation technique does not improve cancer detection as an initial prostate biopsy strategy. J Urol 175:485–488. doi:10.1016/S0022-5347(05)00211-9
10. Pal RP, Elmussareh M, Chanawani M, Khan MA (2012) The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies. BJU Int 109:367–371. doi:10.1111/j.1464-410X.2011.10355.x
11. Carter HB, Pearson JD, Metter EJ et al (1992) Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. JAMA 267:2215–2220
12. Connolly D, Black A, Murray LJ et al (2007) Methods of calculating prostate-specific antigen velocity. Eur Urol 52:1044–1051. doi:10.1016/j.eururo.2006.12.017
13. Hodge KK, McNeal JE, Terris MK, Stamey TA (1989) Random systematic versus directed ultrasound guided transrectal core biopsies of the prostate. J Urol 142:71–74 Discussion 74–75
14. Shariat SF, Semjonow A, Lilja H et al (2011) Tumor markers in prostate cancer I: blood-based markers. Acta Oncol 50(Suppl 1):61–75. doi:10.3109/0284186X.2010.542174
15. Djavan B, Ravery V, Zlotta A et al (2001) Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? J Urol 166:1679–1683

16. Taira AV, Merrick GS, Galbreath RW et al (2009) Performance of transperineal template-guided mapping biopsy in detecting prostate cancer in the initial and repeat biopsy setting. Prostate Cancer Prostatic Dis 13:71–77. doi:10.1038/pcan.2009.42

17. Bittner N, Merrick GS, Andreini H et al (2009) Prebiopsy PSA velocity not reliable predictor of prostate cancer diagnosis, Gleason score, tumor location, or cancer volume after TTMB. Urology 74:171–176. doi:10.1016/j.urology.2008.12.013

18. Bott SRJ, Henderson A, Halls JE et al (2006) Extensive transperineal template biopsies of prostate: modified technique and results. Urology 68:1037–1041. doi:10.1016/j.urology.2006.05.033

19. Mabjeesh NJ, Lidawi G, Chen J et al (2012) High detection rate of significant prostate tumours in anterior zones using transperineal ultrasound-guided template saturation biopsy. BJU Int 110:993–997. doi:10.1111/j.1464-410X.2012.10972.x

20. Dimmen M, Vlatkovic L, Hole KH et al (2012) Transperineal prostate biopsy detects significant cancer in patients with elevated prostate-specific antigen (PSA) levels and previous negative transrectal biopsies. BJU Int 110:E69–E75. doi:10.1111/j.1464-410X.2011.10759.x

21. D'Amico AV, Whittington R, Malkowicz SB et al (1998) Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 280:969–974. doi:10.1001/jama.280.11.969

22. Emiliozzi P, Longhi S, Scarpone P et al (2001) The value of a single biopsy with 12 transperineal cores for detecting prostate cancer in patients with elevated prostate specific antigen. J Urol 166:845–850

23. Bittner N, Merrick GS, Butler WM et al (2013) Incidence and pathological features of prostate cancer detected on transperineal template guided mapping biopsy after negative transrectal ultrasound guided biopsy. J Urol 190:509–514. doi:10.1016/j.juro.2013.02.021

24. Pinkstaff DM, Igel TC, Petrou SP et al (2005) Systematic transperineal ultrasound-guided template biopsy of the prostate: three-year experience. Urology 65:735–739. doi:10.1016/j.urology.2004.10.067

25. Ekwueme K, Simpson H, Zakhour H, Parr NJ (2013) Transperineal template-guided saturation biopsy using a modified technique: outcome of 270 cases requiring repeat prostate biopsy. BJU Int 111:E365–E373. doi:10.1111/bju.12134

26. Pepe P, Aragona F (2013) Prostate biopsy: results and advantages of the transperineal approach—twenty-year experience of a single center. World J Urol. doi:10.1007/s00345-013-1108-1

27. Epstein JI, Sanderson H, Carter HB, Scharfstein DO (2005) Utility of saturation biopsy to predict insignificant cancer at radical prostatectomy. Urology 66:356–360. doi:10.1016/j.urology.2005.03.002

28. Hossack T, Patel MI, Huo A et al (2012) Location and pathological characteristics of cancers in radical prostatectomy specimens identified by transperineal biopsy compared to transrectal biopsy. J Urol 188:781–785. doi:10.1016/j.juro.2012.05.006

ORIGINAL ARTICLE

The role of transperineal template prostate biopsies in prostate cancer diagnosis in biopsy naïve men with PSA less than 20 ng ml^{-1}

S Nafie[1], JK Mellon[1], JP Dormer[2] and MA Khan[1]

BACKGROUND: To compare prostate cancer detection rates between transrectal ultrasound (TRUS) prostate biopsy and transperineal template prostate biopsy (TPTPB) in biopsy naïve men. TRUS biopsy is still regarded as gold standard for prostate cancer diagnosis. TPTPB has been shown to improve prostate cancer detection in men with rising PSA and previous negative TRUS biopsies. We carried out a prospective study performing both biopsies in the same group of men with a benign feeling digital rectal examination (DRE), PSA <20 ng ml^{-1} and no previous prostate biopsies.
METHODS: A total of 50 patients with mean age of 67 years (range: 54–84), mean prostate volume 58 cc (range: 19–165) and mean PSA 8 ng l^{-1} (range: 4–18) underwent standard 12-core TRUS biopsy followed immediately by 36-core TPTPB under general anaesthetic. We determined the prostate cancer detection rate between the two diagnostic modalities.
RESULTS: In total, 20/50 (40%) had benign pathology. Of 30/50 (60%) diagnosed with prostate cancer, 16 (32%) had positive results in both TRUS and TPTPB, whereas 14 (28%) had negative TRUS but positive TPTPB. No cancers were detected solely by TRUS biopsy. TRUS biopsy detected cancer in 32% versus 60% with TPTPB. In total, 19/30(63%) cancers detected by TPTPB had Gleason score $\geqslant 7$. 2 (4%) experienced urosepsis, 7 (14%) temporary urinary retention, 16 (32%) mild haematuria and 19 (38%) haematospermia.
CONCLUSIONS: TPTPB is associated with significantly higher prostate cancer detection rate than TRUS biopsies in biopsy naïve men with a benign feeling DRE and PSA <20 ng ml^{-1}. PSA appears to be better biomarker than previously thought.

Prostate Cancer and Prostatic Disease (2014) **17,** 170–173; doi:10.1038/pcan.2014.4; published online 4 March 2014

Keywords: prostate; cancer; biopsy; TRUS; transperineal; detection

INTRODUCTION

Since the introduction of serum PSA in 1986 as a biomarker for prostate cancer diagnosis, not only there has been a steady world-wide increase in the diagnosis of prostate cancer, there has also been a shift toward diagnosing this disease earlier in its pathway. Hence, the majority of men diagnosed with prostate cancer today are done so purely on the basis of an elevated PSA,[1] whereas prior to the advent of PSA, the majority were diagnosed on the basis of an abnormal digital rectal examination (DRE) and therefore more advanced disease at diagnosis. However, PSA is criticised for its poor specificity of $\sim 30\%$ in diagnosing prostate cancer.[2]

The technique of transrectal ultrasound (TRUS)-guided prostate biopsy has been in existence since the 1980s with subsequent refinements made to the technique in taking biopsies[3] as well as the number of biopsies that should be taken.[3,4] It is currently regarded as the gold standard procedure for prostate cancer diagnosis. However, the yield from TRUS biopsies is low, especially when men have already undergone previous negative TRUS biopsies.[2]

In view of the ability of transperineal template prostate biopsy (TPTPB) to detect prostate cancer at significantly higher rates than TRUS biopsies,[5–8] we questioned whether we should move away from TRUS biopsies to TPTPB and whether PSA is actually a more

specific biomarker for prostate cancer detection than previously thought. To this end, we performed a prospective study to have a head-to-head comparison between TRUS biopsy and TPTPB whereby the same cohort of biopsy naïve men with a benign feeling prostate on a DRE and an elevated PSA <20 ng ml^{-1} underwent simultaneous TRUS biopsies and TPTPB under general anaesthetic. Hence, these patients acted as their own controls.

MATERIALS AND METHODS

Between August 2012 and August 2013, a total of 50 men with at least 10 years life expectancy along with a benign feeling prostate on DRE and elevated serum PSA <20 ng ml^{-1} were enroled in our study. Research protocol was registered and approved by the National Research Ethics Service committee of East Midlands and by the research and development department in the University Hospitals of Leicester NHS trust. All of our participants were given information sheets explaining the nature of the study and they all signed their informed consent forms.

After performing the first 40 cases, a significant difference in cancer detection rates between the two techniques was demonstrated. Power analysis was conducted using a power model based on a one-proportion Z, χ^2 test within STATISTICA (StatSoft, Tulsa, OH, USA). This analysis indicated that to obtain a power of 0.9 (using alpha value of 0.05, a TRUS frequency of 0.32 and a TPTPB frequency of 0.6) would require 30 cases. Furthermore, a power analysis was undertaken for a two-way two-

[1]Department of Urology, Leicester General Hospital, University Hospitals of Leicester NHS Trust, Leicester, UK and [2]Department of Cellular Pathology, Leicester Royal Infirmary, University Hospitals of Leicester NHS Trust, Leicester, UK. Correspondence: S Nafie, Department of Urology, Leicester General Hospital, University Hospitals of Leicester NHS Trust, Gwendolen Road, Leicester LE5 4PW, UK.
E-mail: shady.nafie@me.com

proportion Z-test, this analysis indicated that to obtain a power of 0.8 (using the same alpha value and the same frequencies) would require 50 cases. This was on the basis of a null hypothesis that the proportions of positive cases detected were equal. In order to ensure that we had an appropriate power across these analysis methods, 50 cases were chosen as the optimum.

Each patient was given a dose of Ciprofloxacin 500 mg orally at least 30 min prior to being anaesthetised. At induction, 120 mg of Gentamicin and 1.2 g of Augmentin were administered intravenously unless the patient was penicillin allergic, in which case 400 mg of Teicoplanin was administered intravenously instead. After being placed in the left lateral position, the prostate volume was calculated, 12 TRUS-guided core biopsies were taken with 6 each from the right and left peripheral zones as previously described by Presti *et al.*[4] The patient was subsequently placed in the extended lithotomy position. After shaving the perineal area and securing the scrotum away from the biopsy area using mepore tape, an ultrasound probe (BK Medical Pro-Focus 2202; BK Medical, Mileparken, Denmark) was placed in the rectum to visualise the prostate. Thereafter, the perineum and the genital areas were prepped and draped. A 14-Fr urethral catheter was subsequently inserted in order to visualise the urethra and determine the degree of haematuria at the end of the procedure.

TPTPB was then performed as previously described.[9,10] In short, on the ultrasound screen using a marker pen the prostate, at its largest diameter on the transverse plane, was divided into six sectors, namely, right and left anterior, mid and posterior. In each sector, six 18-gauge biopsy needles (Pro-Mag Biopsy Needle, 18G × 20 cm, MCXS1820AX) were placed into the prostate in the transverse plane using a Brachytherapy template grid. Once all six needles were inserted, the probe was switched to the sagittal plane and the needle gently withdrawn, one at a time, to the apex of the prostate and biopsies taken from the apex towards the base of the prostate. In every case, the biopsies were performed in exactly the same systematic manner starting with the right anterior followed by left anterior and then right mid and so on ending with the left posterior. At the end of the procedure, the catheter was removed unless there was significant haematuria, in which case the catheter was removed after 1–2 h Furthermore, if the prostate volume was measured to be ⩾60cc then the catheter was removed after 5 days. All cases were discharged home the same day. All biopsies were performed by the same surgeon (MAK) and histological analysis was undertaken by the same pathologist (JPD), using standard haematoxylin and eosin-stained, formalin-fixed and paraffin-embedded sections. Standard 4-μm sections were examined over three levels from each core. Where necessary immunoperoxidase to p63, 34betaE12 and AMACR (p504s) antigens were also employed to render a diagnosis. Statistical analysis was carried out using Fisher's exact test to evaluate the association of nominal variables. All calculated P-values were two-sided, considering $P < 0.05$ statistically significant.

RESULTS

Our 50 participants had mean age of 67 years (range: 54–84) with mean prostate volume of 58 cc (range: 19–165) and mean serum PSA level of 8 ng ml^{-1} (range: 4–18). In total, 20/50 (40%) men had benign pathology with no cancer detected in either TRUS biopsies or TPTPB. Out of the 30 patients with detected cancers, 16 (32%) had positive results in both TRUS biopsies and TPTPB, whereas 14 (28%) had positive results in the TPTPB only. There was no cancer detected by the TRUS biopsies solely. Patients with detected cancers had a mean PSA density of 0.19 ng ml^{-1} per cc (range: 0.05–0.57), whereas patients with no detected cancers had a mean PSA density of 0.16 ng ml^{-1} per cc (range: 0.03–0.53) with no significant difference between both ($P = 0.3$). In total, TRUS biopsy detected cancers in 16/50 patients, whereas TPTPB detected cancer in 30/50 patients with detection rates of 31 and 60%, respectively ($P < 0.0001$). Considereing TPTPB as the reference standard, TRUS biopsy had a negative predictive value of 58.8% and a sensitivity of 53.3%. Table 1 demonstrates the details of cancer detection rates along with the pathology findings of both biopsy types.

Of the 14/50 patients who had malignant pathology exclusively detected by the TPTPB, six patients (43%) had Gleason score of 3 + 3, seven (50%) had 3 + 4 and one (7%) had 4 + 3 denoting that 57% of these patients had clinically significant cancers; meanwhile

Table 1. Cancer detection rates and pathology types of TRUS biopsy versus TPTPB specimens in the total cohort of patients ($n = 50$)

	TRUS Biopsy		TPTP Biopsy	
	Number	Percentage (%)	Number	Percentage (%)
Cancer	16	32	30	60
G3 + 3 = 6	6	12	11	22
G3 + 4 = 7	6	12	14	28
G4 + 3 = 7	3	6	3	6
G4 + 4 = 8	0	0	0	0
G4 + 5 = 9	1	2	2	4
Benign	34	68	20	40
HGPIN	19	38	8	16
Atypia	10	20	8	16
Benign	5	10	4	8

Abbreviations: TPTPB, transperineal template prostate biopsy; TRUS, transrectal ultrasound.

Table 2. Post procedure (TRUS biopsy and TPTPB) complications ($n = 50$)

Complication	Number	Percentage (%)
Haematuria	16	32
PR bleeding	8	16
Hematospermia	16	32
Ecchymosis	15	30
AUR	7	14
Urosepsis	2	4

Abbreviations: TPTPB, transperineal template prostate biopsy; TRUS, transrectal ultrasound.

TRUS biopsy showed eight (56%) patients with high-grade PIN, three (21%) with atypia and three (21%) with benign prostate tissue in the same cohort. In total, there was no significant difference in the Gleason scores of cancers detected by both biopsy modalities, as 19/30 (63%) cancers detected by the TPTPB and 10/16 (63%) of cancers detected by TRUS biopsy had Gleason score ⩾7 ($P = 0.9$). In total, 64/600 (11%) cancer-positive cores were detected by TRUS biopsy, whereas 226/1800 (12.5%) positive cores were detected by the TPTPB ($P = 0.2$). In total, 92/226 (40%) cores were detected in the anterior zone, 68 (31%) in the middle zone and 66 (29%) in the posterior zone of the prostate gland.

Post procedure complications are summarised in Table 2. The seven (14%) patients who had temporary urine retention were catheterised and all of them had a successful trial without catheter after 5 days. Those who had urosepsis were admitted and treated by intravenous antibiotics. Other complications were mild, not of worrisome to our patients and required no intervention.

DISCUSSION

It is well accepted that the introduction of PSA as a biomarker has resulted in the detection of prostate cancer at an earlier stage of its natural history.[1] Hence, the majority of men are diagnosed today with organ-confined disease, which allows for radical intervention with curative intent. Furthermore, TRUS biopsies not only allow us to visualise the prostate but to also focus our biopsies to areas within the prostate likely to harbour cancer. Despite advancements in our approach to TRUS prostate biopsies, the majority of men undergoing such a procedure have negative biopsies. This creates great dilemma both for the physician and patient alike, which is further compounded when there is a persistent rise in PSA despite multiple sets of previous negative

172

TRUS biopsies. This has led to a loss in confidence in PSA as an effective biomarker.

Another tact has been to alter our approach to taking prostate biopsies as it is increasingly appreciated that TRUS biopsy does not permit the whole prostate to be scrutinised. Hence, the potential risk of missing cancers arising from the anterior areas of the prostate, and a false negative TRUS biopsy result. In view of this, the role of TPTPB is slowly gaining momentum. To this end, TPTPB has been shown to improve our ability to diagnose prostate cancer.[5–7,11] We have previously demonstrated that TPTPB can detect prostate cancer in the majority of men with a raised PSA despite two prior sets of negative TRUS biopsies.[9,12] This led us to consider the role of TPTPB in biopsy naïve men with a benign feeling prostate on a DRE and PSA <20 ng ml^{-1}. Our study has demonstrated that TPTPB is associated with a 60% prostate cancer detection rate in this particular group of patients. Furthermore, only 32% would have been given the diagnosis of prostate cancer if only TRUS biopsies were performed. This figure is consistent with the previously reported prostate cancer detection rate for such groups of men.[2,8] Hence, it appears that our ability to diagnose prostate cancer by taking TRUS biopsies is not inferior to the published literature.

Very few studies have previously performed a direct comparison between TRUS and transperineal biopsies in the same patients. Emiliozzie *et al*[13] reported a higher cancer detection rate of 38% in 6-core transperineal biopsy (using the fan scheme) compared with 32% in 6-core transrectal biopsy. Kawakami *et al.*[14] compared 14-core transperineal versus 12-core transrectal biopsies, using a transrectal ultrasound-guided three-dimensional 26-core biopsy designed to sample the peripheral zone (regardless of a history of negative biopsy, DRE-finding, PSA level or prostate volume) showed that both biopsies had nearly the same cancer detection rate of 36%. However, the heterogeneity of the cohort criteria, the techniques used and the relatively low number of transperineal biopsy cores could have contributed to their result. Earlier data suggest that the number of biopsy cores taken might affect the cancer detection rate, as studies sampling 20 cores, the cancer detection rate is around 35%[15], whereas others sampling almost 50 cores, the detection rate approaches 50%.[16] Shen *et al*[17] reported no significant difference in cancer detection rate between both biopsies in a meta-analysis of seven studies. However, this included a heterogeneous cohort and moreover the comparison included sextant, extensive and saturation biopsy procedures.

Equally important, our study bucks the general notion that PSA is a poor biomarker for prostate cancer diagnosis. This study instead shows clearly that although PSA is still far from perfect, it is a significantly better biomarker than has previously been given credit for. As such, the low-cancer yield in men with an elevated PSA is more likely to be related to the way we currently perform prostate biopsies rather than PSA itself. Hence, our study has revealed that men with a background of no previous biopsies, elevated PSA and benign feeling DRE, should undergo TPTPB as TRUS biopsy is likely to miss prostate cancer in 28% of cases. In view of this, TRUS biopsies should be abandoned in this group of men, especially as we did not experience any case where cancer was only detected in TRUS biopsies. However, we appreciate that before this paradigm shift in our approach to diagnosing prostate cancer is widely accepted, a larger multi-centre, multi-national study may be required. Furthermore, we accept that men with an abnormal DRE should still continue to be considered for TRUS biopsies as such a procedure is unlikely to miss the diagnosis of prostate cancer in this group of men, which reassuringly is becoming increasingly uncommon.

Our study has also demonstrated that the majority of cancers detected on TPTPB (63%) were Gleason score $\geqslant 7$ indicating that these were clinically significant disease. Hence, TPTPB did not appear to increase the risk of prostate cancer over-diagnosis. Furthermore, although TPTPB does require the need for general anaesthesia, there is some evidence that such a procedure might be feasible under local anaesthetic, pudendal and periprostatic nerve block.[18–20] Nonetheless, the additional cost of a general anaesthetic can offset the need for further TRUS biopsies due to further rise in PSA at a future date. In addition, to date we have performed over 400 cases of TPTPB without any associated urosepsis. In comparison, TRUS biopsy is associated with an $\sim 5\%$ risk of urosepsis requiring hospital admission.[21]

TPTPB is not without its limitations. Men with large prostates can pose a problem in adequately accessing the anterior areas of the prostate due to potential pubic arch interference. However, this can be overcome in the majority by manipulating the angle of the TRUS probe placed in the rectum. In addition, TPTPB was associated with a 14% risk of temporary urinary retention in our study, however this is consistent with previously published data on TPTPB.[22] Nonetheless, this can be addressed by pharmacotherapy if needed. Furthermore, men with larger prostates may need to be considered for a greater number of biopsies in order to adequately sample the whole prostate. Future work will be needed to address the optimum number of biopsies. In our study, in order to keep a uniform approach, all cases underwent 36-core TPTPB irrespective of the prostate volume.

There is growing evidence that multiparametric magnetic resonance imaging (MRI)-guided focal TPTPB is likely to have a greater role in the future. However, evidence for that is still emerging and not confirmed. In addition, there are concerns relating to false negative results associated with MRIs. It is of note that MRI is currently a privilege of the developed world. As such, it is impractical financially and logistically for it to be considered as the initial modality of investigation for all patients with an elevated PSA and a benign feeling prostate. However, it is likely that in the future we will have a better panel of biomarkers, which may or may not include PSA, in predicting the presence of clinically significant prostate cancer with a greater accuracy. At that time, we might be able to consider such patients for an MRI scan prior to undergoing focal TPTPB.

CONCLUSION

TPTPB is associated with a significantly greater prostate cancer detection rate than TRUS biopsies in biopsy naïve men with an elevated PSA <20 ng ml^{-1} and a benign feeling prostate. We propose that TPTPB should be regarded as the biopsy technique of choice in such cases. We do accept that a larger multi-centre, multi-national study might be required before our recommendation is widely accepted. In addition, PSA appears to be a better biomarker for the detection of prostate cancer than previously thought. Hence, in the future, PSA is still likely to have a prominent role as part of a multi-panel biomarker in aiding the diagnosis of prostate cancer.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1 Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* 1987; **317**: 909–916.

2 Welch HG, Fisher ES, Gottlieb DJ, Barry MJ. Detection of prostate cancer via biopsy in the Medicare-SEER population during the PSA era. *J Natl Cancer Inst* 2007; **99**: 1395–1400.

3 Stamey TA. Making the most out of six systematic sextant biopsies. *Urology* 1995; **45**: 2–12.

4 Presti JC, Chang JJ, Bhargava V, Shinohara K. The optimal systematic prostate biopsy scheme should include 8 rather than 6 biopsies: results of a prospective clinical trial. *J Urol* 2000; **163**: 163–166.

5 Dimmen M, Vlatkovic L, Hole KH, Nesland JM, Brennhovd B, Axcrona K. Transperineal prostate biopsy detects significant cancer in patients with elevated prostate-specific antigen (PSA) levels and previous negative transrectal biopsies. *BJU int* 2012; **110**: E69–E75.

6 Takeshita H, Numao N, Kijima T, Yokoyama M, Ishioka J, Matsuoka Y *et al.* Diagnostic performance of initial transperineal 14-core prostate biopsy to detect significant cancer. *Int Urol Nephrol* 2013; **45**: 645–652.

7 Emiliozzi P, Longhi S, Scarpone P, Pansadoro A, DePaula F, Pansadoro V. The value of a single biopsy with 12 transperineal cores for detecting prostate cancer in patients with elevated prostate specific antigen. *J Urol* 2001; **166**: 845–850.

8 Djavan B, Ravery V, Zlotta A, Dobronski P, Dobrovits M, Fakhari M *et al.* Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? *J Urol* 2001; **166**: 1679–1683.

9 Pal RP, Elmussareh M, Chanawani M, Khan MA. The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies. *BJU int* 2012; **109**: 367–371.

10 Chanawani M, Elmussareh M, Khan MA. Transperineal Template Biopsy: Technique and Role in the Diagnosis of Prostate Cancer 2011, 1–10. at < http://www.bjui.org/BJUI_IMAGES/SurgicalAtlas/766transperineal%20template%20biopsyfor%20PDF.pdf > .

11 Mabjeesh NJ, Lidawi G, Chen J, German L, Matzkin H. High detection rate of significant prostate tumours in anterior zones using transperineal ultrasound-guided template saturation biopsy. *BJU int* 2012; **110**: 993–997.

12 Nafie S, Pal RP, Dormer JP, Khan MA. Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies. *World Journal of Urology* 2013; 1–5 (doi:10.1007/s00345-013-1225-x) e-pub ahead of print.

13 Emiliozzi P, Corsetti A, Tassi B, Federico G, Martini M, Pansadoro V. Best approach for prostate cancer detection: a prospective study on transperineal versus transrectal six-core prostate biopsy. *Urology* 2003; **61**: 961–966.

14 Kawakami S, Yamamoto S, Numao N, Ishikawa Y, Kihara K, Fukui I. Direct comparison between transrectal and transperineal extended prostate biopsy for the detection of cancer. *Int J Urol* 2007; **14**: 719–724.

15 Pinkstaff DM, Igel TC, Petrou SP, Broderick GA, Wehle MJ, Young PR. Systematic transperineal ultrasound-guided template biopsy of the prostate: three-year experience. *Urology* 2005; **65**: 735–739.

16 Bott SRJ, Henderson A, Halls JE, Montgomery BSI, Laing R. Langley SEM. Extensive transperineal template biopsies of prostate: modified technique and results. *Urology* 2006; **68**: 1037–1041.

17 Shen P-F, Zhu Y-C, Wei W-R, Li Y-Z, Yang J, Li Y-T *et al.* The results of transperineal versus transrectal prostate biopsy: a systematic review and meta-analysis. *Asian J Androl* 2011; **14**: 310–315.

18 Galfano A, Novara G, Iafrate M, Cosentino M, Cavalleri S, Artibani W *et al.* Prostate biopsy: the transperineal approach. *EAU-EBU Update Series* 2007; **5**: 241–249.

19 Kubo Y, Kawakami S, Numao N, Takazawa R, Fujii Y, Masuda H *et al.* Simple and effective local anesthesia for transperineal extended prostate biopsy: application to three-dimensional 26-core biopsy. *Int Journal Urol* 2009; **16**: 420–423.

20 Iremashvili VV, Chepurov AK, Kobaladze KM, Gamidov SI. Periprostatic local anesthesia with pudendal block for transperineal ultrasound-guided prostate biopsy: a randomized trial. *Urology* 2010; **75**: 1023–1027.

21 Loeb S, Carter HB, Berndt SI, Ricker W, Schaeffer EM. Complications after prostate biopsy: data from SEER-Medicare. *J Urol* 2011; **186**: 1830–1834.

22 Willis S, Bott S, Montgomery B. Urinary retention following transperineal template prostate biopsy - study of risk factors. *J Clin Urol* 2013; **6**: 55–58.

RESULTS

DISCUSSION

	Right	Left	
Anterior	43	39	(82) 58%
Middle	23	16	(39) 28%
Posterior	10	10	(20) 14%

Figure 1. Site of cancer positive cores detected by TPTPB (n = 141)

TTTPB after one negative TRUS biopsy-Nafie et al.

radical prostatectomy specimens identified by
transperineal biopsy compared to transrectal
biopsy. J Urol. 2012; 188: 781-5.

6. Optimal sampling sites for repeat prostate
biopsy: a recursive partitioning analysis of
three-dimensional 26-core systematic biopsy.
Eur Urol. 2007; 51: 675-83.

7. Ong WL, Weerakoon M, Huang S, et al.
Transperineal biopsy prostate cancer detection
in first biopsy and repeat biopsy after negative
transrectal ultrasound-guided biopsy: the
Victorian Transperineal Biopsy Collaboration
experience. BJU Int. 2015; 116: 568-76.

8. Bittner N, Merrick GS, Butler WM, Bennett
A, Galbreath RW. Incidence and pathological
features of prostate cancer detected on
transperineal template guided mapping biopsy
after negative transrectal ultrasound guided
biopsy. J Urol. 2013; 190: 509-14.

9. Kawakami K, Hyuchi N, Yonese J, et al. Three-
dimensional combination of transrectal and
transperineal biopsies for efficient detection
of stage T1c prostate cancer. International
journal of clinical oncology. 2006; 11: 127-
32.

10. Dimmen M, Vlatkovic L, Hole KH, Nesland
JM, Brennhovd B, Axcrona K. Transperineal
prostate biopsy detects significant cancer in
patients with elevated prostate-specific antigen
(PSA) levels and previous negative transrectal
biopsies. BJU Int. 2012; 110: 69-75.

11. Mabjeesh NJ, Lidawi G, Chen J, German L,
Matzkin H. High detection rate of significant
prostate tumours in anterior zones using
transperineal ultrasound-guided template
saturation biopsy. BJU Int 2012; 110: 993-7.

12. Kawakami K, Yonese J, Igari T, et al. Optimal
sampling sites for repeat prostate biopsy:
a recursive partitioning analysis of three-
dimensional 26-core systematic biopsy. Eur
Urol. 2007; 51: 675-82.

13. Grummet JP, Weerakoon M, Huang S, et al.
Sepsis and 'superbugs': should we favour the
transperineal over the transrectal approach for
prostate biopsy? BJU Int. 2014; 114: 384-8.

14. Chang DTS, Challacombe B, Lawrentschuk
N. Transperineal biopsy of the prostate — is
this the future? Nat Rev Urol. 2013; 10: 690-
702.

15. Shen P-F, Zhu Y-C, Wei W-R, et al. The
results of transperineal versus transrectal
prostate biopsy: a systematic review and meta-
analysis. Asian Journal of Andrology. 2012;
14: 310-5.

16. Takenaka A, Hara R, Ishimura T, et al.
A prospective randomized comparison of
diagnostic efficacy between transperineal and
transrectal 12-core prostate biopsy. Prostate
Cancer Prostatic Dis. 2008; 11: 134-8.

17. Gao L-H, Wu R, Xu H-X, et al. Comparison
between Ultrasound Guided Transperineal and

18. Hara R, Jo Y, Fujii T, et al. Optimal Approach
for Prostate Cancer Detection as Initial Biopsy:
Prospective Randomized Study Comparing
Transperineal Versus Transrectal Systematic
12-Core Biopsy. Urology. 2008; 71: 191-5.

19. Abdollah F, Novara G, Briganti A, et al.
Trans-rectal Versus Trans-Perineal Saturation
Rebiopsy of the Prostate: Is There a Difference
in Cancer Detection Rate? Urology. 2011; 77:
921-5.

20. Zaytoun OM, Jones JS. Prostate cancer
detection after a negative prostate biopsy:
lessons learnt in the Cleveland Clinic
experience. Int J Urol. 2011; 18: 557-68.

21. Siddiqui MM, Rais-Bahrami S, Turkbey B,
et al. Comparison of MR/ultrasound fusion-
guided biopsy with ultrasound guided biopsy
for the diagnosis of prostate cancer. JAMA.
2015; 313: 390-7.

22. Ekwueme K, Simpson H, Zakhour H, Parr
NJ. Transperineal template-guided saturation
biopsy using a modified technique: outcome
of 270 cases requiring repeat prostate biopsy.
BJU Int. 2013; 111: 365-73.

Urological Oncology [2011]

Novel Technique in Performing Standard Transperineal Template Prostate Biopsies Under Local Anaesthetic

Shady Nafie [1], Christopher Berridge[1] & Masood A Khan [1]

1. Department of Urology, Leicester General Hospital, University Hospitals of Leicester NHS Trust, Gwendolen Road, Leicester, LE5 4PW, United Kingdom.

ABSTRACT

Purpose: Transperineal template prostate biopsies (TPTPB) are now increasingly commonly performed for the diagnosis of prostate cancer. TPTPB are traditionally performed under general anaesthetic. However, this poses a significant strain on hospital theatre capacity. As such, local anaesthetic (LA) TPTPB are becoming more popular. We describe a novel technique in performing the standard TPTPB under LA in the outpatient setting.

Materials and Methods: Between February 2019- February 2021, 254 consecutive men (median age 69; range: 44-80 years) with a median PSA of 8.7 ng/ml (range: 2.2-76) underwent L/A TPTPB using our novel technique. This is whereby 50mls of 1% prilocaine was injected partially around the perineal skin and partially deep bilateral periprostatic areas. Multiple simultaneous prostate biopsies were then taken with the standard template grid and stepper.

Results: A total of 250/254 (98.4%) men underwent successful L/A TPTPB with a median visual analogue pain score of 4 (range: 2-8). The median prostate volume was 49cc (range: 14-240cc). The median number of cores taken were 18 (range: 14-24). A total of 163/250 men (65%) had a positive histology for prostate cancer with a median of 5 cores being involved with prostate cancer (range: 1-18). In addition, 101/163 men (62%) diagnosed with prostate cancer had either Gleason score 3+4=7 or greater. None experienced urosepsis and only 2/250 men (0.8%) had temporary urinary retention.

Conclusion: Our novel LA technique in performing the standard TPTPB is safe, feasible and well tolerated and associated with a high rate of prostate cancer detection.

Keywords: anaesthetic; biopsy; local; perineum; prilocaine; prostate

INTRODUCTION

Over the past two decades the role of transrectal ultrasound guided prostate biopsies, in the detection of prostate cancer, has been steadily falling out of favour due to its limited ability to accurately biopsy the prostate, in particular the the anterior gland, [1,2] along with its associated risk of urosepsis in up to 5% of cases [3,4]. As such, transperineal template prostate biopsies (TPTPB) have increasingly become the standard diagnostic tool in obtaining adequate and accurate prostate tissue [5-7]. In addition, multi-parametric magnetic resonance imaging (mpMRI) of the prostate now plays an important role in directing biopsies to the most suspicious areas within the prostate; thereby increasing our ability to detect prostate cancer [8]. However, as up to 20% of men with a negative MRI scan of the prostate can still harbour clinically significant prostate cancer[8,9], it is important not only to be able to accurately take biopsies from the MRI suspicious areas but also perform saturation biopsies from all other areas of the prostate.

TPTPB have traditionally been performed under general anaesthetic (GA). However, the rapid rise, over the past two decades, in the number of men undergoing such a procedure has resulted in a great strain in gaining access to the operating theatres. To address this dilemma, a local anaesthetic (LA) approach in performing TPTPB has been increasingly desired. To date, the commonly used technique in performing LA TPTPB is by using PrecisionPoint[TM]. Unfortunately, not only does this technique add to the cost of performing such a procedure but there is some evidence that this technique may not be as accurate in detecting prostate cancer as the standard TPTPB performed under GA[10,11]. We, therefore, determined whether it is feasible to perform the standard TPTPB using the template grid and a stepper by modifying the LA technique.

PATIENTS AND METHODS

Between February 2019 and February 2021, a total of 254 consecutive men with no selection bias underwent TPTPB under LA in our unit by a single surgeon (MAK). As previously described, all men had pre-procedure antibiotic cover with oral Ciprofloxacin 500 mg, Co-Amoxicalv 625 mg and intravenous Gentamicin 120 mg. The men were placed in the lithotomy position and the perineal area shaved and cleaned with Betadine®. The scrotum was lifted away and fixed in place using a Mepore® sticky tape. A total of 50mls of Prilocaine 1% was used as LA. Of the 50mls of 1% prilocaine approximately 25mls is applied to the skin and subcutaneous tissue covering a wide area around the right of the anal margin. Thereafter, approximately 5 mls of LA is injected immediately above the anal margin at a midline and a slim strip of area to the left of the anal margin in a distorted/skewed horse-shoe distribution using a 23G needle as shown in Figure 1. The reason for this distribution of the LA is that all biopsies (right and left) are exclusively taken from the widely infiltrated area on the right. The rationale for adding LA to the other sites is that subsequently the ultrasound probe (BK Medical Pro-Focus 2202, BK Medical, Mileparken, Denmark) with Endocavity Balloon™ is inserted in the rectum and held in place using a mechanical stepper arm (Galil Medical; Crawley, Sussex, UK). Under ultrasound guidance (both sagittal and transverse views), a spinal needle (19G) is inserted to inject 10 mls each of Prilocaine 1% in the peri-prostatic area between the rectum and Denonvellier's fascia (at the posterior lateral area of the apex of the prostate) on the right and left. Hence, the spinal needle is inserted in the areas above and right of the anal margin previously injected by LA (Figure1). In total, 50 mls of 1% prilocaine LA is used. Thereafter, the prostate volume is measured by ultrasound, the gland divided on the ultrasound screen into six areas (Right Anterior, Right Mid, Right Posterior, Left Anterior, Left Mid, Left Posterior). A standard 0.5 cm brachytherapy template grid is fixed to the mechanical arm (stepper) and placed over the right perineal area (Figure 2). The stepper not only stabilizes the ultrasound probe but also permits the probe to be tilted in various angles thereby gaining access to the whole of the prostate for biopsy solely from the single area to the right of the anus (Figure 2). TPTPB is then carried out using the template grid and 18G needles. This enable multiple needles to be inserted at the same time ensuring that a wide spread of biopsies are taken (Figure 2). Cognitive MRI fusion is used to take 4 biopsies from the MRI suspicious areas and two biopsies each from all other areas. In cases where the pre-biopsy MRI scan excluded any suspicious lesions within the prostate (i.e. negative MRI), four biopsies are taken from each of the six areas to ensure that thorough saturation biopsies are

undertaken. As such, between 14 and 24 biopsies were taken from each men. At the end of the procedure, each men were asked to complete a visual analogue score, where 1 is no pain and 10 is unbearable/severe pain. Subsequently, the men returned to the waiting area, and were discharged home after voiding urine.

RESULTS

Out of the 254 consecutive men, 4 (1.6%) did not tolerate the ultrasound probe under LA, therefore the procedure was abandoned and referred for TPTPB under GA. The 250 men who underwent the procedure had a median age of 69 years (range: 44-80) with a median PSA of 8.7 ng/mL (range: 2.2-76) and mean Prostate volume of 49 cc (14-240). A median of 18 cores (range: 24-14) were taken from each men with a median of 5 positive cores (range: 1-18). A total of 163/250 men (65%) had positive histology for prostate cancer. Of these, 62(38%) had Gleason score (GS) 3+3, 68(42%) had GS 3+4, 15(9%) had GS 4+3, 10(6%) had G4+4 and 8(5%) had GS 4+5 as shown in Figure 3. The median pain score of 4 (range: 2-8). None of the men experienced urosepsis and 2 had temporary post-procedure urinary retention (0.8%).

DISCUSSION

Over the past two decades TPTPB has gained greater momentum in being the procedure of choice in adequately sampling the prostate. It carries the advantages over conventional TRUS guided prostate biopsies in being associated with a lower risk of urinary sepsis as well as having a higher yield in detecting prostate cancer due to being able to access all areas of the prostate. However, the rapid rise in the volume of TPTPB being performed has added significantly to the burden on operative theatre utilisation as it is commonly performed under GA. In order to move TPTPB out of the theatre environment and towards an out-patient scenario, great effort has been invested in developing LA techniques for this increasingly common procedure.

To our knowledge, our LA technique this the first of its type to be described in the world literature, whereby we have been able to perform the standard TPTPB using the template grid and stepper along with multiple simultaneous needle insertions; which is currently the technique used when performing this procedure under GA. Our experience has confirmed that it is feasible and safe to perform TPTPB under LA without compromising the areas sampled or the numbers of biopsies taken when compared with the standard TPTPB

under GA. In addition, our yield in detecting prostate cancer of 65% is consistent with the standard TPTPB performed under GA[6,7].

At present, there are few readily available systems for performing transperineal biopsies under LA such as PrecisionPoint™ access system [10,11] and the CamProbe [12]. These systems have been shown to be effective in performing transperineal biopsies under LA in the outpatient setting. However, they come with additional financial costs and are free-hand held devices, thus affecting the accuracy in taking biopsies from the appropriate/suspicious areas [13]. In addition, as only a single needle can be inserted on each occasion a biopsy is taken, these two techniques are less predictable in ensuring that appropriate biopsies including saturation biopsies are taken. Therefore, it can be debated whether such biopsy techniques are truly template biopsies.

However, our technique using the template grid permitting the simultaneous insertion of multiple needles offers better needle distribution and ensures that the same point of the prostate is not sampled twice. Furthermore, using the mechanical arm offers motion stability and precision in taking the biopsies. In addition, if a needle is not deemed to be in the correct position, it can easily be withdrawn and reinserted to a appropriate position as per cognitive MRI fusion.

CONCLUSION

Our LA technique, which is feasible, safe and well tolerated, enables us to perform the standard TPTPB without compromising cancer detection rates. By avoiding the need for additional specialist equipment and moving this procedure to the out-patient setting will also have a significant cost and time saving benefit.

CONFLICT OF INTEREST

Authors report no conflict of interest.

REFRENCES

1.   King CR, McNeal JE, Gill H, Presti JC, Jr. Extended prostate biopsy scheme improves reliability of Gleason grading: implications for radiotherapy patients. Int J Radiat Oncol Biol Phys. 2004;59:386-91.

2.   Han M, Chang D, Kim C, et al. Geometric evaluation of systematic transrectal ultrasound guided prostate biopsy. J Urol. 2012; 188: 2404-9.

3.   Grummet JP, Weerakoon M, Huang S, et al. Sepsis and 'superbugs': should we favour the transperineal over the transrectal approach for prostate biopsy? BJU Int. 2014; 114: 384-8.

4.   Anastasiadis E, van der Meulen J, Emberton M. Hospital admissions after transrectal ultrasound-guided biopsy of the prostate in men diagnosed with prostate cancer: a database analysis in England. Int J Urol. 2015; 22: 181-6.

5.   Chang DT, Challacombe B, Lawrentschuk N. Transperineal biopsy of the prostate--is this the future? Nat Rev Urol. 2013; 10: 690-702.

6.   Nafie S, Mellon JK, Dormer JP, Khan MA. The role of transperineal template prostate biopsies in prostate cancer diagnosis in biopsy naive men with PSA less than 20 ng/ml. Prostate Cancer Prostatic Dis. 2014; 17: 170-3.

7.   Nafie S, Pal RP, Dormer JP, Khan MA. Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies. World J Urol. 2014; 32: 971-5.

8.   Brown LC, Ahmed HU, Faria R, et al. Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study. Health Technol Assess. 2018; 22: 1-176.

9. Hanna N, Wszolek MF, Mojtahed A, et al. Multiparametric MRI/Ultrasound Fusion Biopsy Improves but Does Not Replace Standard Template Biopsy for the Detection of Prostate Cancer. J Urol. 2019; 202: 944-51.

10. Meyer AR, Joice GA, Schwen ZR et al. Initial Experience Performing In-office Ultrasound-guided Transperineal Prostate Biopsy Under Local Anesthesia Using the PrecisionPoint Transperineal Access System. Urology. 2018; 115: 8-13.

11.   Kum F, Elhage O, Maliyil J et al. Initial outcomes of local anaesthetic freehand transperineal biopsies in the outpatient setting. BJU Int. 2020; 125: 244-52.

12.   Thurtle D, Starling L, Leonard K, Stone T, Gnanapragasam VJ. Improving the safety and tolerability of local anaesthetic outpatient transperineal prostate biopsies: A pilot study of the CAMbridge PROstate Biopsy (CAMPROBE) method. J Clin Urol. 2018; 11: 192-9.

13.   Dundee PE, Grummet JP, Murphy DG. Transperineal prostate biopsy: template-guided or freehand? BJU Int. 2015; 115: 681-3.

Figure 1: The marked region demonstrates the skewed horse-shoe area anaesthetised. X: represents the approximate places where the spinal needle is inserted to inject the deep peri-prostatic local anaesthetic around the right and left apical areas.

A

B

C

D
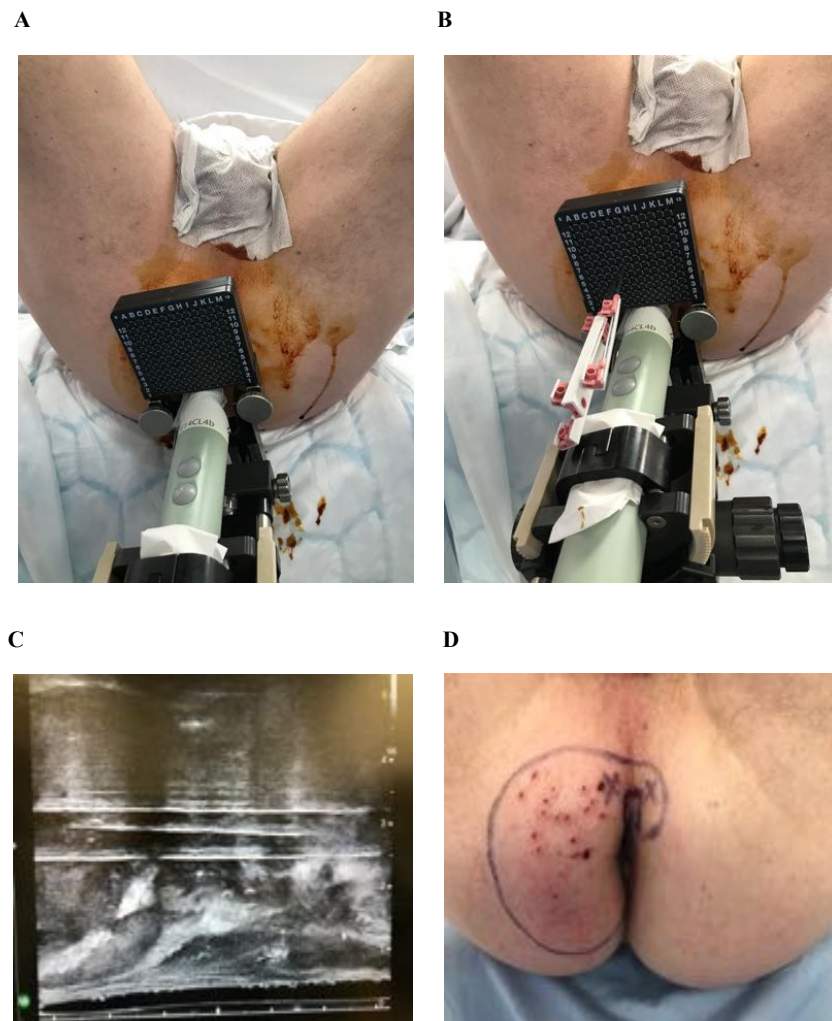
Figure 2A-D:

A: Demonstrating the angle of the ultrasound probe with the template grid, stabilised by the stepper, required to take biopsies from the right areas of the prostate. The probe is then adjusted slightly to take biopsies from the left areas of the prostate.

B: Showing simultaneous insertions of multiple biopsy needles.

C: Sagittal ultrasound image of the multiple biopsy needles insertions.

D: Biopsy needle insertion sites around the anaesthetised perineal area (post-procedure)
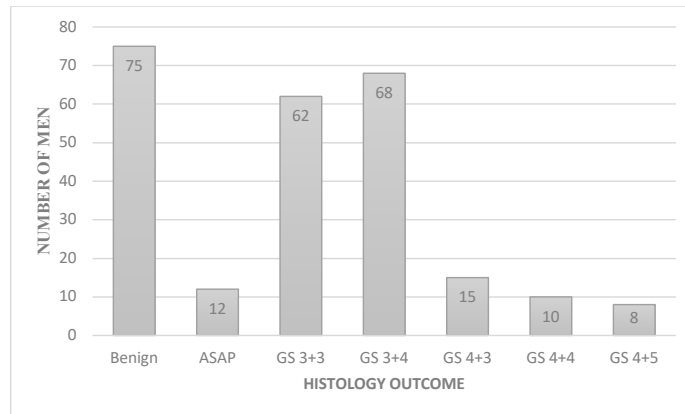
Figure 3. Distribution of the various histology types in men undergoing LA TPTPB (GS=Gleason Score)

PLOS | ONE

Prediction of Pathological Stage in Patients with Prostate Cancer: A Neuro-Fuzzy Model

Georgina Cosma[1]☯*, **Giovanni Acampora**[1]‡, **David Brown**[1]‡, **Robert C. Rees**[2], **Masood Khan**[3]☯, **A. Graham Pockley**[2]☯

1 Computing and Technology, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom, **2** John van Geest Cancer Research Centre, Nottingham Trent University, Nottingham, United Kingdom, **3** Department of Urology, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* georgina.cosma@ntu.ac.uk

Abstract

The prediction of cancer staging in prostate cancer is a process for estimating the likelihood that the cancer has spread before treatment is given to the patient. Although important for determining the most suitable treatment and optimal management strategy for patients, staging continues to present significant challenges to clinicians. Clinical test results such as the pre-treatment Prostate-Specific Antigen (PSA) level, the biopsy most common tumor pattern (Primary Gleason pattern) and the second most common tumor pattern (Secondary Gleason pattern) in tissue biopsies, and the clinical T stage can be used by clinicians to predict the pathological stage of cancer. However, not every patient will return abnormal results in all tests. This significantly influences the capacity to effectively predict the stage of prostate cancer. Herein we have developed a neuro-fuzzy computational intelligence model for classifying and predicting the likelihood of a patient having Organ-Confined Disease (OCD) or Extra-Prostatic Disease (ED) using a prostate cancer patient dataset obtained from The Cancer Genome Atlas (TCGA) Research Network. The system input consisted of the following variables: Primary and Secondary Gleason biopsy patterns, PSA levels, age at diagnosis, and clinical T stage. The performance of the neuro-fuzzy system was compared to other computational intelligence based approaches, namely the Artificial Neural Network, Fuzzy C-Means, Support Vector Machine, the Naive Bayes classifiers, and also the AJCC pTNM Staging Nomogram which is commonly used by clinicians. A comparison of the optimal Receiver Operating Characteristic (ROC) points that were identified using these approaches, revealed that the neuro-fuzzy system, at its optimal point, returns the largest Area Under the ROC Curve (AUC), with a low number of false positives (FPR = 0.274, TPR = 0.789, AUC = 0.812). The proposed approach is also an improvement over the AJCC pTNM Staging Nomogram ($FPR = 0.032$, $TPR = 0.197$, $AUC = 0.582$).

Introduction

Cancer staging prediction is a process for estimating the likelihood that the disease has spread before treatment is given to the patient. Cancer staging evaluation occurs before (i.e. at the prognosis stage) and after (i.e. at the diagnosis stage) the tumor is removed—the clinical and pathological stages respectively. The clinical stage evaluation is based on data gathered from clinical tests that are available prior to treatment or the surgical removal of the tumor. There are three primary clinical stage tests for prostate cancer: the Prostate Specific Antigen (PSA) test which measures the level of PSA in the bloodstream; a biopsy which is used to detect the presence of cancer in the prostate and to evaluate the degree of cancer aggressiveness (results are usually given in the form of the Primary and Secondary Gleason patterns); and a physical examination, namely the Digital Rectal Examination (DRE) which can determine the existence of disease and possibly provide sufficient information to predict the stage of the cancer. A limitation of the PSA test is that abnormally high PSA levels may not necessarily indicate the presence of prostate cancer, nor might normal PSA levels reflect the absence of prostate cancer. Pathological staging can be determined following surgery and the examination of the removed tumor tissue, and is likely to be more accurate than clinical staging, as it allows a direct insight into the extent and nature of the disease. More information on the clinical tests is provided in the next subsection *Medical Background*.

Given the potential prognostic power of the clinical tests, a variety of prostate cancer staging prediction systems have been developed. The ability to predict the pathological stage of a patient with prostate cancer is important, as it enables clinicians to better determine the optimal treatment and management strategies. This is to the patient's considerable benefit, as many of the therapeutic options can be associated with significant short- and long- term side-effects. For example, radical prostatectomy (RP)—the surgical removal of the prostate gland—offers the best chance for curing the disease when prostate cancer is localised, and the accurate prediction of pathological stage is fundamental to determining which patients would benefit most from this approach [1–3]. Currently, clinicians use nomograms to predict a prognostic clinical outcome for prostate cancer, and these are based on statistical methods such as logistic regression [4]. However, cancer staging continues to present significant challenges to the clinical community.

The prostate cancer staging nomograms which are used to predict the pathological stage of the cancer are based on results from the clinical tests. However, the accuracy of the nomograms is debatable [5, 6]. Briganti et al. [5] argues that nomograms are accurate tools and that "Personalized medicine recognizes the need for adjustments, according to disease and host characteristics. It is time to embrace the same attitude in other disciplines of medicine. This includes urologic oncology where nomograms, regression-trees, lookup tables and neural networks represent the key tools capable of providing individualized predictions". Dr Joniau in [5] argues that the data used for devising the nomograms are subjective and, to a certain extent, biased by institutional protocols on which patients are selected for a given treatment. Dr Joniau states that one of the drawbacks of nomograms is that various nomograms have been devised for risk estimation and it is difficult to determine which nomogram will provide the most reliable risk estimation for a particular patient. He emphasises that although nomograms allow for more accurate risk assessment, this risk estimation is a "snapshot in a risk continuum". Although this might allow personalized predictions, it also makes treatment decisions difficult [5].

Cancer prediction systems which consider various variables for the prediction of an outcome require computational intelligent methods for efficient prediction outcomes [7]. Although computational intelligence approaches have been used to predict prostate cancer outcomes, very few models for predicting the pathological stage of prostate cancer exist. In

essence, classification models based on computational intelligence are utilised for prediction tasks. Classification is a form of data analysis which extracts classifier models describing data classes, and uses these models to predict categorical labels (classes) or numeric values [8]. When the classifier is used to predict a numeric value, as opposed to a class label, it is referred to as a predictor. Classification and numeric prediction are both types of prediction problems [8], and classification models are widely adopted to analyse patient data and extract a prediction model in the medical setting.

Computational intelligence approaches, and in particular fuzzy-based approaches, are based on mathematical models that are specially developed for dealing with the uncertainty and imprecision which is typically found in the clinical data that are used for prognosis and the diagnosis of diseases in patients. These characteristics make these algorithms a suitable platform on which to base new strategies for diagnosing and staging prostate cancer. For example, not everyone diagnosed with prostate cancer will exhibit abnormal results in all tests, as a consequence of which, different test result combinations can lead to the same outcome.

The capacity of fuzzy, and especially neuro-fuzzy approaches, to predict the pathological stage of prostate cancer has not been as widely evaluated as the more commonly used Artificial Neural Network (ANN) and other approaches. However, fuzzy approaches have been applied to other prostate cancer scenarios. Benechi et al. [9] have applied the Co-Active Neuro-Fuzzy Inference System (CANFIS) to predict the presence of prostate cancer; Keles et al.[10] proposed a neuro-fuzzy system for predicting whether an individual has cancer or Benign Prostatic Hyperplasia (BPH, a benign enlargement of the prostate). Çinar [11] designed a classifier-based expert system for the early diagnosis of prostate cancer, thereby aiding the decision-making process and informing the need for a biopsy. Castanho et al. [12] developed a genetic-fuzzy expert system which combines pre-operative serum PSA, clinical stage, and Gleason grade of a biopsy to predict the pathological stage of prostate cancer (i.e. whether it was confined or not-confined).

Saritas et al. [13] devised an ANN approach for the prognosis of cancer which can be used to assist clinical decisions relating to the necessity for a biopsy. Shariat et al. [14] have performed a critical review of prostate cancer prediction tools and concluded that predictive tools can help during the complex decision-making processes, and that they can provide individualised, evidence-based estimates of disease status in patients with prostate cancer.

Finally, Tsao et al. [15] developed an ANN model to predict prostate cancer pathological staging in 299 patients prior to radical prostatectomy, and found that the ANN model was superior at predicting Organ Confined Disease in prostate cancer than a Logistic Regression model. Tsao et al. [15] also compared their ANN model with Partin Tables, and found that the ANN model more accurately predicted the pathological stage of prostate cancer.

Herein we propose a neuro-fuzzy model for predicting the pathological stage of prostate cancer. The system inputs comprise the following variables: the most common tumor pattern (Primary Gleason pattern), the second most common pattern (Secondary Gleason pattern), PSA levels, age at diagnosis, and clinical T stage. The neuro-fuzzy model automatically constructs fuzzy rules via a training process which is applied to existing and known patient records and status. These rules are then used to predict the prostate cancer stage of patients in a validation set. The model makes use of the Adaptive Neuro-Fuzzy Inference System which is also used to optimise the predictive performance. The outcome for each patient record is a numerical prediction of the 'degree of belongingness' of each patient in the Organ-Confined Disease and Extra-Prostatic Disease classes.

Medical Background

This section describes the variables used for diagnosis.

Prostate Specific Antigen (PSA). The Prostate Specific Antigen (PSA) test is a blood test that measures the level of prostate-specific antigen in the bloodstream. Although having limitations, the PSA test is currently the best method for identifying an increased risk of localised prostate cancer. PSA values tend to rise with age, and the total PSA levels (ng/ml) recommended by the Prostate Cancer Risk Management Programme are as follows [16]: 50–59 years, $PSA \geq 3.0$; 60–69 years, $PSA \geq 4.0$; and 70 and over, $PSA > 5.0$. Abnormally high and raised PSA levels may, but does not necessarily, indicate the presence of prostate cancer. The European Study of Screening for Prostate Cancer revealed that screening significantly reduces death from prostate cancer, and that a man who undergoes PSA testing will have his risk of dying from prostate cancer reduced by 29% [17, 18], and [19]. However, it should also be noted that a normal PSA test does not necessarily exclude the presence of prostate cancer.

Primary and Secondary Gleason Patterns. A tissue sample (biopsy) is used to detect the presence of cancer in the prostate and to evaluate its aggressiveness. The results from a prostate biopsy are usually provided in the form of the Gleason grade score. For each biopsy sample, pathologists examine the most common tumor pattern (Primary Gleason pattern) and the second most common pattern (Secondary Gleason pattern), with each pattern being given a grade of 3 to 5. These grades are then combined to create the Gleason score (a number ranging from 6 to 10) which is used to describe how abnormal the glandular architecture appears under a microscope. For example, if the most common tumor pattern is grade 3, and the next most common tumor pattern is grade 4, the Gleason score is 3 + 4, or 7. A score of 6 is regarded as low risk disease, as it poses little danger of becoming aggressive; and a score of 3 + 4 = 7 indicates intermediate risk. Because the first number represents the majority of abnormal tissue in the biopsy sample, a 3 + 4 is considered less aggressive than a 4 + 3. Scores of 4 + 3 = 7, or 8 to 10 indicate that the glandular architecture is increasingly more abnormal and associated with high risk disease which is likely to be aggressive.

Clinical and Pathological Stages. The clinical stage is an estimate of the prostate cancer stage, and this is based on the results of the digital rectal examination (DRE). The pathological stage can be determined if a patient has had surgery and hence is based on the examination of the removed tissue. Pathological staging is likely to be more accurate than clinical staging, as it can provide a direct insight into the extent of the disease. At the clinical stage, there are four categories for describing the local extent of a prostate tumor (T1 to T4). Clinical and pathological staging use the same categories, except that the T1 category is not used for pathological staging. In summary, stages T1 and T2 describe a cancer that is probably organ-confined, T3 describes cancer which is beginning to spread outside the prostate, and T4 describes a cancer that has likely begun to spread to nearby organs. Category T1 is when the tumor cannot be felt during the DRE or be seen with imaging such as transrectal ultrasound (TRUS). Category T1 has three subcategories: T1a cancer is found incidentally during a transurethral resection of the prostate (TURP) which will have been performed for the treatment of Benign Prostatic Hyperplasia, and the cancer is present in no more than 5% of the tissue removed; T1b cancer is found during a TURP, but is present in more than 5% of the tissue removed, and T1c cancer is found in a needle biopsy which has been performed due to an elevated PSA level. Category T2 is when the tumor can be felt during a DRE or seen with imaging, but still appears to be confined to the prostate gland. Category T2 has three subcategories: T2a cancer is in one half or less of only one side (left or right) of the prostate; T2b cancer is in more than half of only one side (left or right) of the prostate; and T2c cancer is in both sides of the prostate. Category T3 has two subcategories: T3a cancer extends outside the prostate, but not to the seminal vesicles; and T3b

cancer has spread to the seminal vesicles. Finally, category T4 cancer has grown into tissues next to the prostate (other than the seminal vesicles), such as the urethral sphincter, the rectum, the bladder, and/or the wall of the pelvis.

The TNM staging is the most widely used system for prostate cancer staging and aims to determine the extent of:

- primary tumor (T stage),

- the absence or presence of regional lymph node involvement (N stage), and

- the absence or presence of distant metastases (M stage)

The TNM system has been accepted by the Union for International Cancer Control (UICC) and the American Joint Committee on Cancer (AJCC). Most medical facilities use the TNM system as their main method for cancer reporting. The clinical TNM and pathological TNM are provided in Tables 1 and 2 respectively. Once the T, N, and M are determined, a stage of I, II, III, or IV is assigned, with stage I being early and stage IV being advanced disease. Upon determining the T, N, and M stages, a prognosis can be made about the anatomic stage of cancer using the groupings shown in Table 3 where a stage of I, II, III, or IV is assigned to a patient, with stage I being early and stage IV being advanced disease [20]. Stages I, II, are organ confined cancer stages, whereas Stages III and IV are extra-prostatic stages. TNM systems have gone through several refinements in order to "improve the uniformity of patient evaluation and to maintain a clinically relevant evaluation" [20]. In the most recent American Joint

Table 1. Definitions of clinical TNM according AJCC 2010 [21].

Primary tumor (pT)	
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Clinically inapparent tumor neither palpable nor visible by imaging (T1)	
T1a	Tumor incidental histologic finding in $\leq$ 5% of tissue resected
T1b	Tumor incidental histologic finding in > 5% of tissue resected
T1c	Tumor identified by needle biopsy (e.g. because of elevated PSA)
Tumor confined within prostate (T2)	
T2a	Tumor involves one-half of one lobe or less
T2b	Tumor involves more than one-half of one lobe but not both lobes
T2c	Tumor involves both lobes
Tumor extends through the prostate capsule (T3)	
T3a	Extracapsular extension (unilateral or bilateral)
T3b	Tumor invades seminal vesicle(s)
T4	Tumor is fixed or invades adjacent structures other than seminal vesicles such as external sphincter, rectum, bladder, levator muscles, and/or pelvic wall
Regional lymph nodes (pN)	
NX	Regional lymph nodes were not assessed
N0	No regional lymph node metastasis
N1	Metastasis in regional lymph node(s)
Distant metastasis (pM)	
M0	No distant metastasis
M1	Distant metastasis
M1a	Non-regional lymph node(s)
M1b	Bone(s)
M1c	Other site(s) with or without bone disease

doi:10.1371/journal.pone.0155856.t001

Table 2. Pathological TNM according AJCC 2010 [21]. There is no pT1 classification.

Organ confined (pT2)	
pT2a	Unilateral, one-half of one side or less
pT2b	Unilateral, involving more than one-half of one side, but not both sides
pT2c	Bilateral disease
Extraprostatic extension (pT3)	
pT3a	Extraprostatic extension or microscopic bladder neck invasion
pT3b	Seminal vesicle invasion
pT4	Invasion of rectum levator muscles, and/or pelvic wall
Regional lymph nodes (pN)	
pNX	Regional lymph nodes not sampled
pN0	No positive regional lymph nodes
pN1	Metastasis in regional lymph node(s)
Distant metastasis (pM)	
pM1	Distant metastasis
pM1a	Non-regional lymph node(s)
pM1b	Bone(s)
pM1c	Other site(s) with or without bone disease

doi:10.1371/journal.pone.0155856.t002

Committee on Cancer (AJCC) [21], the Gleason score and PSA have been incorporated in the cancer stage/prognostic groups 3.

Methods I—Neuro-Fuzzy Model

Fuzzy logic is an extension of multivalued logic that deals with approximate, rather than fixed and exact reasoning. Fixed reasoning is the traditional binary logic where variables may take on true or false values. Fuzzy logic starts with the concept of a fuzzy set [22], which is a set without a crisp, clearly defined boundary. A fuzzy set can contain elements with only a partial degree of membership, and hence allows for degrees of truth, making fuzzy logic applicable to medical scenarios which are considered to involve complexity, uncertainty and vagueness.

Table 3. Anatomic stage/prognostic groups (from AJCC 2010) [21].

Group	T	N	M	PSA	Gleason score (GS)
I	T1a–c	N0	M0	$PSA < 10$	$GS \leq 6$
	T2a	N0	M0	$PSA < 10$	$GS \leq 6$
	T1–2a	N0	M0	PSA X	GS X
IIA	T1a–c	N0	M0	$PSA < 20$	GS 7
	T1a–c	N0	M0	$PSA \geq 10 < 20$	$GS \leq 6$
	T2a	N0	M0	$PSA < 20$	$GS \leq 7$
	T2b	N0	M0	$PSA < 20$	$GS \leq 7$
	T2b	N0	M0	PSA X	GS X
IIB	T2c	N0	M0	Any PSA	Any GS
	T1–2	N0	M0	$PSA \geq 20$	Any GS
	T1–2	N0	M0	Any PSA	$GS \geq 8$
III	T3a–b	N0	M0	Any PSA	Any GS
IV	T4	N0	M0	Any PSA	Any GS
	Any T	N1	M0	Any PSA	Any GS
	Any T	Any N	M1	Any PSA	Any GS

doi:10.1371/journal.pone.0155856.t003

Fuzzy logic has been combined with various soft computing methodologies, including neuro-computing, thereby leading to powerful neuro-fuzzy systems.

The neuro-fuzzy system proposed herein (a combination of fuzzy logic-based algorithms that are illustrated in Fig 1) predicts the pathological stage of cancer (i.e. diagnosis outcome), using data that are obtained from pre-operative clinical tests that are conducted at the prognosis stage. As with the TNM system, our proposed neuro-fuzzy system predicts whether a patient has organ-confined disease (OCD, pathological stage pT2) or extra-prostatic disease (ED, pathological stage $>$ pT2).
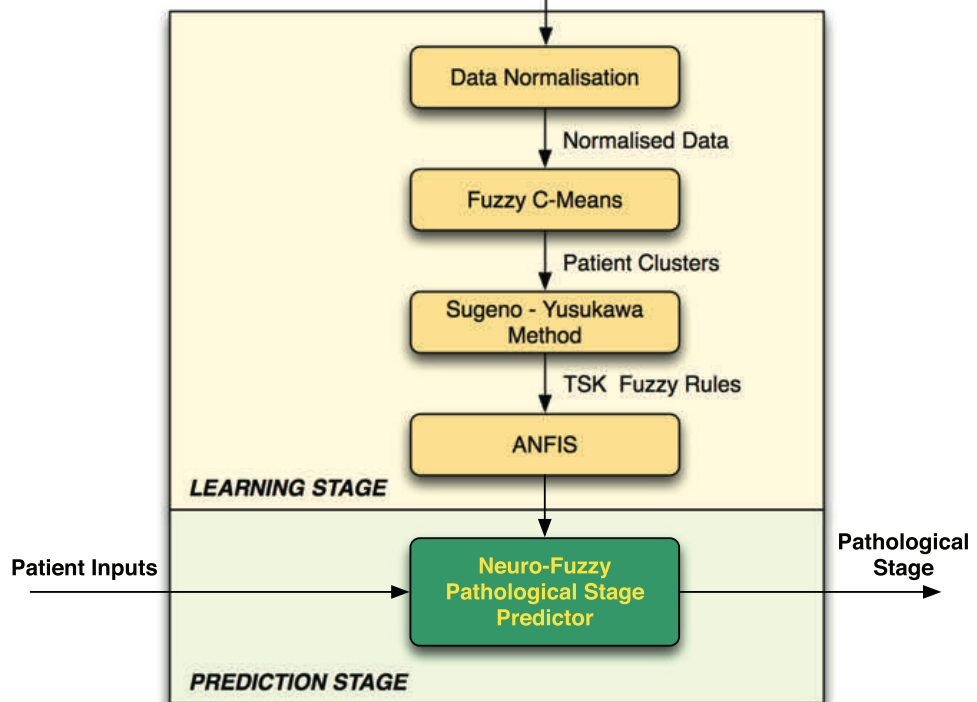
Fig 1. Neuro-Fuzzy Prostate Cancer Pathological Stage Predictor.

doi:10.1371/journal.pone.0155856.g001

The clinical data used for pathological cancer stage prediction are typically affected by imprecision, primarily due to the fact that not all patients exhibit abnormal results in all clinical tests. This poses a problem when trying to predict the progression of the cancer and therefore deciding on the best treatment strategy for patients. Hence, fuzzy logic is a suitable approach for this type of clinical prediction because it can be used to model human reasoning—in real scenarios the clinician would consider the data and give an estimation rather than a definite answer. The neuro-fuzzy system will make a prediction about a particular patient and return a value representing the 'degree of membership' of the patient's cancer in the extra-prostatic set. The proposed framework is illustrated in Fig 1 and described in the subsections that follow.

The neuro-fuzzy model comprises two main stages: learning and prediction. At the *learning stage*, the model trains itself using patient records for which the pathological stage is known, and at the *prediction stage* the model predicts the pathological stage using the knowledge which has been obtained during the learning stage. The following subsections describe the processes that are involved during the learning and prediction stages.

System Inputs

At the *learning stage*, the neuro-fuzzy predictor learns (i.e. trains itself) using existing patient record data in order to create the knowledge which will be used (during the *prediction stage*) to make predictions on new, and previously unseen, data. During the learning stage, the system takes as input data relating to each patient's clinical features (i.e. age at diagnosis, PSA, biopsy Primary Gleason pattern, biopsy Secondary Gleason pattern, and clinical T stage) and known pathological stage results (i.e. known outputs) that have been obtained during diagnosis. The system represents the inputs as a matrix A of size $n \times m$, where n is the total number of patient records, and m is the total number of clinical features (i.e. system inputs, $m = 5$). The system represents the targets in the form of a $n \times 1$ vector T, where each cell t_i holds the pathological T stage (pT) value for each patient record.

At the *prediction stage*, the system only requires as input an $1 \times m$ vector holding the results of a patient's clinical features (i.e. age at diagnosis, PSA, biopsy Primary Gleason pattern, biopsy Secondary Gleason pattern, and clinical T stage), and the system will return a value representing the likelihood of the patient having Extra-Prostatic Disease (i.e. pathological stage results).

Data Normalisation

The age, PSA level, clinical T stage and pathological stage (pT) variables must be grouped before they are input into the fuzzy predictor. The normalisation of the values is described in the *Results* Section. The normalisation process is performed in order to ensure a balanced distribution among the data and to remove any outliers from the data which could affect the performance of the predictor algorithm.

Fuzzy C-Means

Formally, let $A = [v_1, v_2, v_3, \ldots, v_n]$ be the [patient record cases]-by-[clinical features] matrix and let $2 \leq c < n$ be an integer, where c is the number of clusters (i.e. classes) and n is the total number of patient record cases. In this particular prostate cancer application, $c = 2$ since we have two clusters: Organ-Confined Disease (OCD) and Extra-Prostatic Disease (ED). The Fuzzy C-Means (FCM) algorithm returns a list of cluster centers $X = x_1, \ldots, x_c$ and a membership matrix $U = \mu_{i,k} \in [0, 1]; i = 1, \ldots, n; k = 1, \ldots, c$, where each element μ_{ik} holds the total membership of a data point v_k (i.e. patient record) belonging to cluster c_i. FCM updates the centers of clusters Organ-Confined Disease and Extra-Prostatic Disease, and the membership

grades for each data point, representing a patient record, by iteratively moving the cluster centers to the correct location within a data set. Essentially, this iteration process is based on minimizing an objective function which represents the distance from any given data point to a cluster center weighted by that data point's membership grade. The objective function for FCM is a generalisation of Eq (1)

$$J(U, c_1, \ldots, c_c) = \sum_{i=1}^{c} \sum_{k=1}^{N} \mu_{ik}^{m} ||v_k - x_i||^2, 1 \leq m \leq \infty \qquad (1)$$

where μ_{ik} represents the degree of membership of patient record v_i in the ith cluster; x_i is the cluster centre of fuzzy group i; $|| * ||$ is the Euclidean distance between ith cluster and jth data point; and $m \in [1, \infty]$ is a weighting exponent. The necessary conditions for function (1) to reach its minimum are shown in functions (2) and (3).

$$c_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m v_k}{\sum_{k=1}^{N} \mu_{i,k}^m}, \qquad (2)$$

$$\mu_{ik} = \frac{1}{\sum_{k=1}^{c} \left(\frac{||v_k - x_i||}{||v_k - x_i||} \right)^{2/(m-1)}}, \qquad (3)$$

Sugeno-Yusukawa Method

A collection of Takagi-Sugeno-Kang (TSK) rules [23], one for each cluster, for determining the membership of a patient record to a particular cluster are generated. This Sugeno-type Fuzzy Inference System (FIS) is generated using the FCM clustering algorithm. The number of clusters derived from the clustering process determines the number of rules and membership functions in the generated FIS. The FIS structure maps inputs through input membership functions and associated parameters, and then through output membership functions and associated parameters to outputs. The output FIS is passed into the Adaptive-Neuro Fuzzy Inference System (ANFIS) model [24] which then tunes the FIS parameters using the input/output training data in order to optimise the prediction model.

Adaptive-Neuro Fuzzy Inference System

The Adaptive Neuro-Fuzzy Inference System (ANFIS) [24] combines Artificial Neural Networks and Fuzzy Logic algorithms. ANFIS creates a fuzzy inference system with membership functions that are generated by adaptive backpropagation learning. The architecture of a Type-3 ANFIS, which is the ANFIS used in the proposed model, is explained in [24]. The following is a brief description of ANFIS and is based on [25]. ANFIS consists of five layers. In layer 1, each node generates a membership grade of a linguistic variable (in the prostate cancer scenario, linguistic variables are the staging classes, i.e. Organ-Confined Disease and Extra-Prostatic Disease) using a membership function. The Gaussian membership function is used within the neuro-fuzzy model. Layer 2 calculates the firing strength of each rule, and layer 3 calculates the ratio of each rule's firing strength to the total of all firing strengths. At layer 4, the contribution of each rule toward the overall output is computed, and, finally, layer 5 calculates the overall output as the summation of the contribution from each rule. During the learning process, ANFIS adapts the parameters associated with the membership functions and tunes them using a gradient vector which, given a set of parameters, measures the performance of the system on the basis of how well it models input and output data. ANFIS has been used in conjunction

with FCM, and thus the FIS returned from FCM clustering is input into the ANFIS, and the FIS parameters are tuned using the input/output training data in order to optimise the prediction model.

The training process stops whenever the designated epoch number is reached, or the error goal is achieved. The performance of ANFIS is evaluated using the array of root mean square errors (difference between the FIS output and the training data output) at each epoch. Thus, the membership degree of a patient record into a particular cluster (e.g. Extra-Prostatic Disease), determines how close a prediction is to the next cluster (e.g. Organ-Confined Disease). In simple terms, let c_a and c_b be cluster Extra-Prostatic Disease and cluster Organ-Confined Disease respectively, a patient record v_k can belong to cluster c_a such that $v_k \in c_a$, or it can belong in the intersection area between two clusters such that $v_k \in c_a \wedge v_k \in c_b$.

Neuro-Fuzzy Predictor

The neuro-fuzzy predictor takes as input a vector X_i of size $1 \times m$, where m is the total number of clinical features, hence 1×5 and the patient's record is clustered as Organ-Confined Disease or Extra-Prostatic Disease, via the FCM clustering algorithm [26]. The predetermined Takagi-Sugeno-Kang (TSK) rules [23] are then applied in order to evaluate the degree of membership of the patient's record to a particular cluster. The output is a numerical value representing the likelihood of a patient belonging to the Extra-Prostatic Disease cluster. This value is particularly useful when deciding on the suitable treatment to be offered to the patient. For example, treatment might be different if a patient is predicted as having Organ-Confined Disease with a value which leans more toward Extra-Prostatic Disease.

Methods II—Other Computational Intelligence Approaches

Artificial Neural Network Classifier

An Artificial Neural Network (ANN) can be trained to recognise patterns in data and this is a suitable approach for solving classification problems involving two or more classes. For the prostate cancer staging prediction problem, the ANN is trained to recognise the patients which have Organ-Confined Disease or Extra-Prostatic Disease. The pattern recognition neural network used was a two-layer feedforward network, in which the first layer has a connection from the network input and is connected to the output layer which produces the network's output. A *log-sigmoid transfer function* was embedded in the hidden layer, and a *softmax transfer function* was embedded in the output layer.

A neuron has R number of inputs where R is the number of elements in an input vector. Let an input vector X be a patient record X_i belonging to a class Organ-Confined Disease or Extra-Prostatic Disease. Each input X_i is weighted with an appropriate weight w. The sum of the weighted inputs and the bias forms the input to the transfer function f. Neurons can use a differentiable transfer function f to generate their output. The Log-Sigmoid function which generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity was used. The Softmax neural transfer function was used to calculate a layer's output from its net input. Softmax functions convert a raw value into a posterior probability and this provides a measure of certainty. The number of hidden neurons is set to 5 in order to match the number of inputs. The number of output neurons is set to 2, which is equal to the number of elements in the target vector (the number of classes, Organ-Confined Disease and Extra-Prostatic Disease). The maximum number of epochs (repetitions) was set to $\epsilon = 200$ and in order to avoid over-fitting, training stops when the maximum number of epochs is reached. The ANN was trained using the Scaled Conjugate Gradient (SCG) for Fast Supervised Learning which is suitable for large-scale problems [27]. The process of training the ANN involves tuning the values

of the weights and biases of the network in order to optimise network performance which is measured by the mean squared error network function.

Naive Bayes Classifier

Although the Naive Bayes classifier is designed for use when predictors within each class are independent of one another within each class, it is known to work well even when that independence assumption is not valid. The Naive Bayes classifies data in two steps. The first step is the training (i.e. learning) step which uses the training data, which are patient cases and their corresponding pathological cancer stage (i.e. Organ-Confined Disease or Extra-Prostatic Disease), to estimate the parameters of a probability distribution, assuming predictors are conditionally independent given the class. The second step is the prediction step, during which the classifier predicts any unseen test data and computes the posterior probability of that sample belonging to each class. It subsequently classifies the test data according to the largest posterior probability. The following Naive Bayes description is based on that presented by Han et. al [8].

Let $P(c_i|X)$ be the posterior probability that a patient record X_i will belong to a class c_i (class can be Organ-Confined Disease or Extra-Prostatic Disease), given the attributes of vector X_i. Let $P(c_i)$ be the prior probability that a patient's record will fall in a given class regardless of the record's characteristics; and $P(X)$ is the prior probability of record X, and hence the probability of the attribute values of each record. The Naive Bayes classifier predicts that a record X_i belongs to the class c_i having the *highest posterior probability*, conditioned on X_i if and only if $P(c_i|X) > P(c_j|X)$ for $1 \leq j \leq m, j \neq i$, maximising $P(c_i|X)$. The class c_i for which $P(c_i|X)$ is maximised is called the *maximum posteriori hypothesis* and estimated using formula (4)

$$P(c_i|X) = \frac{P(X|c_i)P(c_i)}{P(X)}. \qquad (4)$$

To predict the class label of a given record X_i, $P(X|c_i)P(c_i)$ is evaluated for each class c_i. The classifier predicts that the class label of record X_i is the class c_i if and only if

$$P(X|c_i)P(c_i) > P(X|c_j)P(c_j) \qquad (5)$$

for $1 \leq j \leq m, j \neq i$.

The Naive Bayes outcome is that each patient's record, which is represented as a vector X_i, is mapped to exactly one class c_i, where $c_i = 1, \ldots, n$ where n is the total number of classes, i.e. $n = 2$. The Naive Bayes classification function can be tuned on the basis of an assumption regarding the distribution of the data. Experiments were conducted using two methods of density estimation: the first one assumes normality and models each conditional distribution with a single Gaussian; and the second uses nonparametric kernel density estimation. Hence, the Naive Bayes classifier was tuned using two functions: a Gaussian distribution (GD) and the Kernel Density Estimation (KDE). The Gaussian distribution assumes that the variables are conditionally independent given the class label and thereby exhibit a multivariate normal distribution, whereas *kernel density estimation* does not assume a normal distribution and hence it is a non-parametric technique.

Support Vector Machine Classifier

The Support Vector Machine (SVM) classification method uses nonlinear mapping to transform the original training data (i.e. the patient dataset) into a higher dimensional feature space. It then determines the best separating hyperplane, which serves as a boundary separating the data from two classes. The best separating hyperplane for a Support Vector Machine means the one with the largest margin between the two classes. The bigger the margin, the better the

generalisation error of the linear classifier is defined by the separating hyperplane. Support vectors are the points that reside on the canonical hyperplanes and are the elements of the training set that would change the position of the dividing hyper plane if removed. As with all supervised learning models, a support vector machine is initially trained on existing data records, after which the trained machine is used to classify (predict) new data. Various Support Vector Machine kernel functions can be utilised to obtain satisfactory predictive accuracy.

The Support Vector Machine finds the Maximum Marginal Hyperplane (MMH) and the support vectors using a Lagrangian formulation and solving the equation using the Karush-Kuhn-Tucker (TTK) conditions, details of which can be found in [28]. Once the Support Vector Machine has been trained, the classification of new unseen patient records is based on the Lagrangian formulation. For many 'real-world' practical problems, using the linear boundary to separate the classes may not reach an optimal separation of hyperplanes. However, Support Vector Machine kernel functions which are capable of performing linear and nonlinear hyperplane separation exist. The outcome of applying the Support Vector Machine for prediction is that each patient record, represented as a vector X_i, is mapped to exactly one class label y_i, where $y_i = \pm 1$, such that $(X_1, y_1), (X_2, y_2), \ldots (X_m, y_m)$, and hence y_i can take one of two values, either -1 or $+1$ corresponding to the classes Organ-Confined Disease and Extra-Prostatic Disease. Further details on the Support Vector Machine can be found in [29], and [30].

Results I: Dataset Analysis

Dataset Description

The Cancer Genome Atlas (TCGA) Research Network provides datasets for cancer patients which are made open to the public through the Data Coordinating Center and the TCGA Data Portal. The prostate cancer dataset obtained from the TCGA contains records collected from 399 patients diagnosed with a type of *Prostate Adenocarcinoma Acinar*, during the years 2000–2013. All patients had prostate needle core biopsies for diagnosis before they underwent prostatectomy, and all patients had undergone a prostatectomy. The variables selected from the dataset were those that are used for performing prostate cancer stage predictions by clinicians, and which are also required for undertaking staging prediction using the AJCC pTNM Nomogram [21]—namely, biopsy Primary and Secondary Gleason patterns, pre-treatment PSA level, patient's age at diagnosis, clinical T stage, and pathological T stage.

The age and PSA variables were categorically divided into groups that were chosen in order to ensure a balanced distribution between the data, as described later in this section. Table 4 provides statistics about the variables before they were categorised.

The Primary and Secondary Gleason pattern variables (see Table 5) did not require any modification, as they are already categorically divided into three groups.

Table 4. Dataset Statistics.

	Statistics of variables before categorisation			
	Minimum	Maximum	Mean	Standard deviation
Primary Gleason pattern	3	5	3.54	0.60
Secondary Gleason pattern	3	5	3.74	0.69
PSA level (ng/mL)	0.70	107.00	9.84	11.25
Age at Diagnosis	41.10	78.00	59.88	6.92
Clinical T	1.00	5.00	2.19	1.45
Pathological T stage	1.00	2.00	1.55	0.50

doi:10.1371/journal.pone.0155856.t004

Table 5. Primary and Secondary Gleason pattern groups.

Primary Gleason pattern groups	Frequency count	Proportion of patients(%)
3	205	51.4
4	173	43.4
5	21	5.3
Total	399	100.0
Secondary Gleason pattern groups	Frequency count	Proportion of patients(%)
3	159	39.8
4	185	46.4
5	55	13.8
Total	399	100.0

doi:10.1371/journal.pone.0155856.t005

The variables of pre-treatment total serum PSA levels and age were categorically divided into groups, as described in Tables 6 and 7. Table 6 shows how the PSA values have been grouped, the number of cases in each group (i.e. count), and their percentage. The histogram in Fig 2 illustrates the frequency distributions of the grouped PSA values.

Table 7 displays how the age values have been grouped, the number of patient cases in each group (i.e. count), and the percentage of cases. Although it is very unlikely for a patient to have prostate cancer under the age of 35, there is still a possibility, and for this reason, groupings 1 to 4 have been formed. However, none of the patient cases fall in this category in the particular dataset which was used in the current study. This does not affect the performance of the system

Table 6. PSA groups.

PSA group	PSA range	Frequency count	Proportion of patients (%)
1	0–2.5 ng/mL	16	4.01
2	2.6–4.0 ng/mL	33	8.27
3	4.1–6.0 ng/mL	124	31.08
4	6.1–9.9 ng/mL	124	31.08
5	10–19 ng/mL	67	16.79
6	$\geq$ 20 ng/mL	35	8.77

doi:10.1371/journal.pone.0155856.t006

Table 7. Age groups.

Age group	Age range	Frequency count	Proportion of patients (%)
1	< 25	0	0
2	25–29	0	0
3	30–34	0	0
4	35–39	0	0
5	40–44	5	1.25
6	45–49	22	5.51
7	50–54	68	17.04
8	55–59	97	24.31
9	60–64	100	25.06
10	65–69	76	19.05
11	> 70	31	7.77

doi:10.1371/journal.pone.0155856.t007
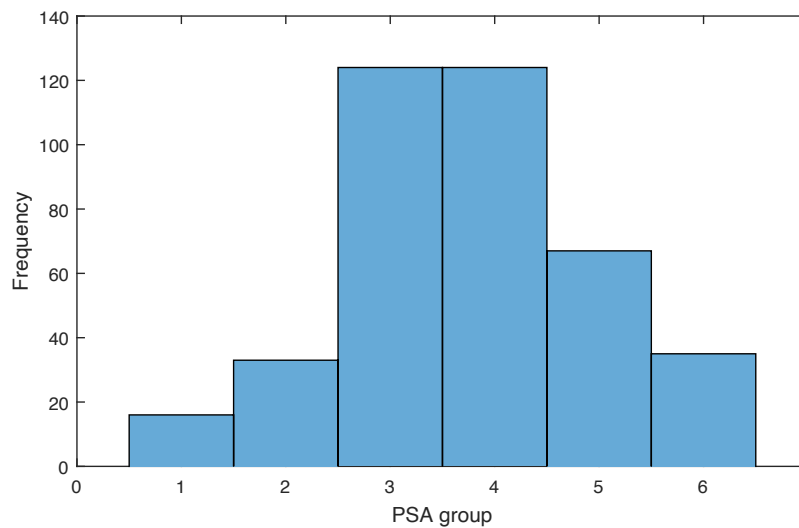
Fig 2. Histogram of grouped PSA values.

in any way, and these groupings were included in order to create a comprehensive prediction system.

The clinical T and pathological T stage variable values were grouped as shown in Tables 8 and 9 respectively. The clinical T stage variable values were grouped in such a way so as to match the groups that are presented on the TNM nomogram. Group 1 includes T stages T1a-c, which reflects the fact that tumor is present in one or both lobes by needle biopsy, but is not identifiable on the basis of palpation or is reliably visible by imaging. Group 2 includes clinical T stages in which the cancer is unilateral, meaning that it is located on one-half of one side or less; group 3 is unilateral and involves more than one-half of one side, but not both sides; group 4 is when the cancer is in the form of bilateral disease which is located on both sides of

Table 8. Clinical T stage groups.

Clinical T group	Clinical T stage	Frequency count	Proportion of patients (%)
1	T1(a-c)	204	51.13
2	T2a	53	13.28
3	T2b	53	13.28
4	T2c	42	10.53
5	T3a	29	7.27
5	T3b	16	4.01
5	T4	2	0.50
	Total	399	100.00

Table 9. Pathological T (pT) stage groups.

pT group	Pathological T (pT) stage	Frequency count	Proportion of patients (%)	OCD or ED
1	T2(unknown if a or b)	1	0.25	OCD
1	T2a	14	3.51	OCD
1	T2b	47	11.78	OCD
1	T2c	117	29.32	OCD
2	T3a	142	35.59	ED
2	T3b	72	18.05	ED
2	T4	6	1.50	ED
	Total	399	100.00	

doi:10.1371/journal.pone.0155856.t009

the prostate; and group 5 is an Extra-Prostatic Disease, meaning that the cancer has started to spread (or has spread) to the bladder neck, rectum, and/or nearby organs.

The independent variable (i.e. predictor variable) of the dataset is the pathological T (pT) stage, whose values have been categorised as 1 for Organ-Confined Disease (OCD) and 2 for Extra-Prostatic Disease (ED). There is no pathologic pT1 classification, and pathological T stage values in the pT2 range indicate an Organ-Confined Disease and pathological T stage values in the range of 3–4 indicate an Extra-Prostatic Disease.

Given that the aim of the model was to predict whether a patient has organ-confined disease (OCD, TNM pathological stage pT2) or extra-prostatic disease (ED, TNM pathological stage $> pT$2), and not the likelihood of cancer-related death (DOD, dead of disease), the pT3 and pT4 groupings were consolidated. T3a, T3b and T4 disease are all considered as being 'high-risk disease' and strongly considered for active treatment, whereas those with lower categories of disease might instead be considered for active surveillance. Table 9 includes the groupings of the pathological T stage values.

Tables 10 and 11 present a sample of the data before and after data normalisation (i.e. categorically divided into groups), respectively. As previously mentioned, the values of Primary and Secondary Gleason patterns did not require any transformation, as they were already grouped.

Age at Diagnosis and its Association with PSA Values

Evidence indicates that age could be a contributing factor to increased PSA levels [16], and for this reason it is informative to investigate whether there are any associations between PSA levels and age in the dataset. The histograms illustrating the frequency distributions of the grouped PSA levels and age are illustrated in Figs 2 and 3 respectively. The mean *age at diagnosis* of patients was 59.88 ± 6.92, and the mean *pre-treatment PSA level* was 9.84 ± 11.25. Table 12 shows the mean and standard deviation PSA values of each age group.

Table 10. Before data normalisation.

Case No.	Primary Gleason Pattern	Secondary Gleason Pattern	PSA	Age	Clinical T stage	Pathological (pT) stage
1	3	3	1.00	51.6	T2b	T2a
2	3	3	1.70	77.0	T2b	T2c
3	3	3	2.05	55.2	T2a	pT2b
4	3	3	2.09	61.1	T1c	pT2b
5	3	3	2.20	57.0	T1c	T3a
n	…	…	…	…	…	…

doi:10.1371/journal.pone.0155856.t010

Table 11. After data normalisation.

Case No.	Primary Gleason Pattern	Secondary Gleason Pattern	PSA group	Age group	Clinical T group	Pathological (pT) group
1	3	3	1	7	3	1
2	3	3	1	11	3	1
3	3	3	1	8	2	1
4	3	3	1	9	1	1
5	3	3	1	8	1	2
n	…	…	…	…	…	…

doi:10.1371/journal.pone.0155856.t011
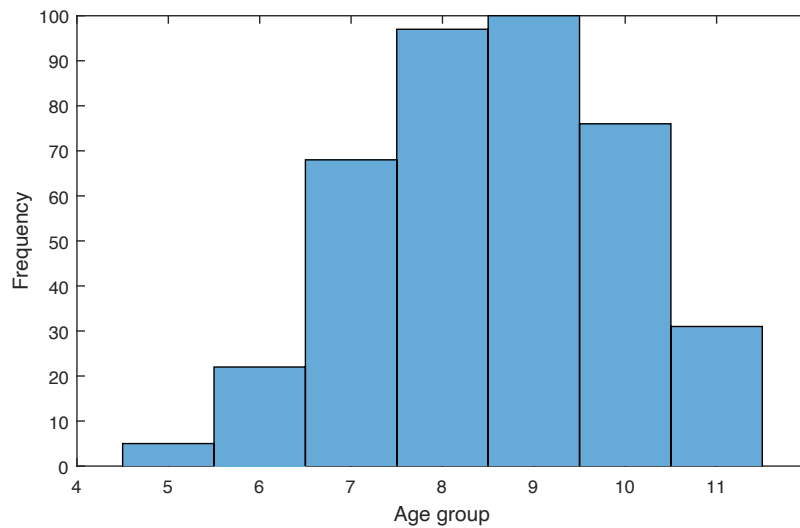
Fig 3. Histogram of grouped age values.

doi:10.1371/journal.pone.0155856.g003

Table 12. PSA levels categorised by age group.

Age group	Patient count	PSA mean	Standard deviation of PSA values
5	5	4.40	1.14
6	22	4.14	0.89
7	68	3.54	1.23
8	97	3.75	1.20
9	100	3.74	1.14
10	76	3.75	1.21
11	31	3.81	1.56
Total	399	3.75	1.21

doi:10.1371/journal.pone.0155856.t012

To investigate if there is a statistically significant association between age and PSA levels, a Spearman's rho correlation test was performed. The results revealed no significant associations among age and PSA levels ($r = 0.003$, $p = 0.948$), at least over the age ranges (40 to 78) that were included in the dataset which was used. A one-way ANOVA was also conducted to test for statistically significant differences between the PSA values of the various age groups. The test indicated that there were no statistically significant differences among any of the groups, $F(6,392) = 0.956$, $p = 0.455$, $p > 0.05$). Since age and PSA values are not associated, then both variables were used as inputs by the prediction system.

Analysis of the Clinical T Stage Values

The clinical T stage values are the results of a Digital Rectal Examination (DRE) test. Table 8 shows the total number of patients for each clinical T stage. As shown in Table 8, the clinical stage values of T1 ($n = 204$, 51.13%), T2a ($n = 53$, 13.28%), T2b ($n = 53$, 13.28%), T2c ($n = 42$, 10.53%), denote an Organ-Confined Disease cancer stage, and the T3a ($n = 29$, 7.27%), T3b ($n = 16$, 4.01%), T4 ($n = 2$, 0.50%) denote an Extra-Prostatic Disease cancer stage. At a clinical stage, a total of 352 patients (88.22%) exhibited Organ-Confined Disease, and a total of 47 patients exhibited Extra-Prostatic Disease (11.78%). Clearly, the results of the clinical T test alone is not as reliable as the pathological T (pT) test for determining the stage of prostate cancer. This is evident since, at the pathological stage (i.e. diagnosis stage), out of the 399 patients, a total of 44.86% ($n = 179$) had Organ-Confined Disease, and 55.14% ($n = 220$) had Extra-Prostatic Disease, meaning that 43.36% (n = 173) patients with Extra-Prostatic Disease were misdiagnosed as having Organ-Confined Disease. DRE is not always a reliable test, since the location of the tumor within the prostate might influence the capacity to feel it.

Analysis of the Pathological T (pT) Stage Values

Table 9 shows the total number of patients for each pathological T stage. As shown in Table 9, the pathological T stage values of T2a-c denote an Organ-Confined Disease cancer stage, and the pathological T stage values of T3a,b,T4 denote an Extra-Prostatic Disease cancer stage. Of the 399 patients, a total of 44.86% ($n = 179$) exhibited Organ-Confined Disease, and 55.14% (n = 220) exhibited Extra-Prostatic Disease.

Table 13 shows the relationship between the prediction variables and prostate cancer with Organ-Confined Disease and Extra-Prostatic Disease. A one-way ANOVA test revealed significant differences between the means of the two groups for the biopsy Primary Gleason pattern ($F(1,397) = 7.87$, $p = 0.005$, $p < 0.05$), biopsy Secondary Gleason pattern ($F(1,397) = 5.83$, $p = 0.016$, $p < 0.05$), and clinical T stage ($F(1,397) = 5.062$, $p = 0.025$, $p < 0.05$) variables. These

Table 13. Mean and Standard deviation values for Organ-Confined Disease (OCD) and Extra-Prostatic Disease (ED) groups diagnosed at the Pathological stage.

Variables	Groups		p
	OCD	**ED**	
	n = 179	**n = 220**	
Primary Gleason pattern	3.45 ± 0.52	3.61 ± 0.64	0.005
Secondary Gleason pattern	3.65 ± 0.64	3.81 ± 0.71	0.016
Pre-treatment PSA level (ng/mL)	3.76 ± 1.26	3.74 ± 1.17	0.848
Age at diagnosis (groups)	8.49 ± 1.37	8.60 ± 1.40	0.434
Clinical T stage	2.01 ± 1.36	2.33 ± 1.51	0.025

doi:10.1371/journal.pone.0155856.t013

results were expected, since Gleason 1 grading of the biopsy determines the aggressiveness of the cancer, and hence patients with Extra-Prostatic Disease are more likely to have a higher value in this category than patients with Organ-Confined Disease. The same applies for PSA levels, as these tend to increase with progressive and more extensive disease. In addition, the clinical T stage values determine the outcome of a physical examination and the higher the number of the clinical T stage value, the more the disease has progressed. Interestingly, in this particular dataset, no statistically significant differences among the means values of the pre-treatment PSA levels ($F(1,397) = 0.037$, $p = 0.848$, $p > 0.05$) and age ($F(1,397) = 0.614$, $p = 0.434$, $p > 0.05$) variables were apparent. Although the mean age of patients with Extra-Prostatic Disease was higher than that of patients with Organ-Confined Disease, this difference is not statistically significant. Also, there were no statistically significant differences among the mean PSA values of patients with Extra-Prostatic Disease and Organ-Confined Disease. In summary, the mean values of all but the PSA and age variables were significantly higher in the Extra-Prostatic Disease than in the Organ-Confined Disease groups.

Results II: Pathological Stage Prediction Using the Neuro-Fuzzy Model

Experiment Methodology

Having analysed the dataset and, as appropriate, grouped the data, the next step is to explain how the transformed data will be input into the neuro-fuzzy model and the other models which will be used for the comparison process. In particular, the performance of the neuro-fuzzy model is compared to other computational intelligence based approaches, namely the Artificial Neural Network, Fuzzy C-Means, Support Vector Machine, and the Naive Bayes classifier. All of these classifiers are suitable for solving prediction problems, as is the American Joint Committee on Cancer (AJCC) pTNM Nomogram [21] which is a statistical approach that is commonly adopted by clinicians for predicting prostate cancer staging. For predicting the pathological stage of cancer, the AJCC pTNM Nomogram uses all variables except the age at diagnosis variable. These variables are found in Table 13.

System Inputs

All classification models take as input a matrix A of size $n \times m$, where n is the total number of patient records and m is the total number of clinical features, hence 399×5; and a $n \times 1$ vector T, where $n = 399$ and each cell t_i holds the pathological T (pT) stage value for each patient record. Table 11 shows the first five records of matrix A after normalising the input values, in which the first 5 columns are the inputs and the last column pathological T (pT) stage holds the target output values. The dataset ($n = 399$) was separated into two subsets, a training subset and a validation subset, and the same subsets were used across the models undergoing evaluation in order to ensure a fair comparative evaluation. The training subset comprised 266 (66.6%) records, which were used for training each model. The validation subset comprised 133 (33.3%) records and these were used for determining the predictive accuracy of each model (i.e. validating its performance). Of the 133 records used for validation, 66 (49.62%) records corresponded to patients with Organ-Confined Disease, and 67 (50.38%) records corresponded to patients with Extra-Prostatic Disease. Fig 4 shows a set of Gaussian membership functions generated by the proposed system, for each input data. The input comprised of 5 inputs given by two external markers (Organ-Confined Disease, and Extra-Prostatic Disease). Observing the membership curves found in Fig 4, reveals a consistency among them—the results of tests corresponding to
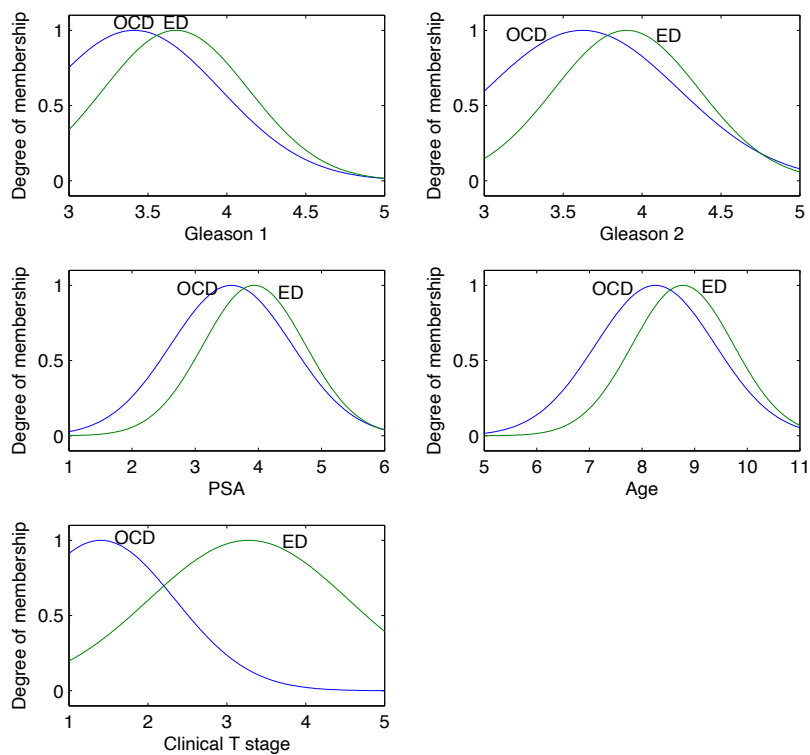
Fig 4. Neuro-Fuzzy System Membership Functions: Gleason 1 is Primary Gleason Pattern; Gleason 2 is Secondary Gleason pattern; PSA is Prostate Specific Antigen; Age represents the Age group; and clinical T stage is the result of the Digital Rectal Examination. OCD is Organ-Confined Disease and ED is Extra-Prostatic Disease.

patients with Extra-Prostatic Disease fall in a higher range than the results of patients with Organ-Confined Disease.

The performance of the proposed neuro-fuzzy model was compared to that of an Artificial Neural Network, Fuzzy C-Means, Support Vector Machine, and the Naive Bayes classifiers; and the AJCC pTNM Nomogram statistical approach [21] which can also be considered as a classifier. The subsections below give a brief introduction to each classification model and details on how the parameters of each model were appropriately tuned in order to report their best performance.

Performance Evaluation Measures

The evaluation measures that were adopted for assessing the performance of each approach for predicting the pathological stage of patients are *Sensitivity* and *Specificity*. These statistical measures are used for evaluating the performance of binary classification tests and are suitable since the aim is to measure the performance of each system in distinguishing Extra-Prostatic Disease from the Organ-Confined Disease.

Sensitivity (i.e. True Positive Rate) measures the proportion of actual positives which are correctly identified as such (e.g. the percentage of Extra-Prostatic Disease patients who are correctly identified as Extra-Prostatic Disease). Specificity (i.e. True Negative Rate) measures the proportion of negatives which are correctly identified as such (e.g. the percentage of patients with Organ-Confined Disease who are correctly identified as not having Extra-Prostatic Disease). A perfect system would return 100% sensitivity (e.g., all patients with Extra-Prostatic Disease are classed as Extra-Prostatic Disease) and 100% specificity (e.g. all patients with Organ-Confined Disease are not classed as Extra-Prostatic Disease). The following notation relates to the evaluation measures.

- Let $|TP|$ be the total number of patients with Extra-Prostatic Disease correctly classified as Extra-Prostatic Disease.

- Let $|TN|$ be total the number of patients with Organ-Confined Disease correctly classified as Organ-Confined Disease.

- Let $|FP|$ be the total number of patients with Organ-Confined Disease incorrectly classified as Extra-Prostatic Disease.

- Let $|FN|$ be the total number of patients with Extra-Prostatic Disease incorrectly classified as Organ-Confined Disease.

- Let $|P|$ be the total number of Extra-Prostatic Disease cases that exist in the dataset, where $|P| = |TP| + |FN|$.

- Let $|N|$ be the total number of patients with Organ-Confined Disease that exist in the dataset, where $|N| = |FP| + |TN|$.

The functions for the Sensitivity and Specificity evaluation measures are presented in Functions (6) and (7) respectively.

$$Sensitivity(k) = \frac{|TP|}{|TP| + |FN|}, \in [0, 1]. \tag{6}$$

$$Specificity(k) = \frac{|TN|}{|TN| + |FP|}, \in [0, 1]. \tag{7}$$

The closer the values of Sensitivity and Specificity are to 1.0, the better the detection performance of the system.

Evaluation measures based on the Receiver Operating Characteristic (ROC) curve analysis are fundamental in clinical research. ROC curves are used to determine the performance of the systems, and they can be used to establish a cutoff value for optimal performance of each system. The ROC curve is a graph of sensitivity (y-axis) against 1-specificity (x-axis) across different cut-off points. The area under the ROC curve (AUC) is a reflection of how good the system's performance is at distinguishing (or discriminating) between patients with and without Extra-Prostatic Disease—the larger the area, the better the performance. The aim is to determine the cutoff point for which the classifier returns the high number of true positives

Table 14. Performance evaluation.

	Performances based on ROC evaluation measurements					
	Neuro-Fuzzy (Our approach)	FCM	Quadratic-SVM	ANN	GB-NB	AJCC pTNM Nomogram
Area Under the Curve (AUC)	0.812	0.809	0.738	0.699	0.750	0.582
Optimal ROC point FPR	0.274	0.403	0.242	0.303	0.274	0.032
Optimal ROC point TPR	0.789	0.901	0.718	0.701	0.775	0.197
Asymp. Sig. (McNemars)	1.000	0.868	0.499	1.000	1.000	0.000

and the low number of false positives. Maximizing sensitivity corresponds to some large True Positive Rate (y-axis value) on the ROC curve, and maximizing specificity corresponds to a small False Positive Rate value (x-axis value) on the ROC curve. Thus, the optimal cutoff value is to the upper left corner of the chart, the higher the overall accuracy of the classifier. Hence, a system which perfectly discriminates Organ-Confined Disease and Extra-Prostatic Disease has 1.0 (or 100%) sensitivity and 1.0 (or 100%) specificity.

Comparison of the Neuro-Fuzzy Model with Other Methods

The aim of the evaluation is to measure the ability of each system to predict the pathological stage of patients. The results presented in this section are those for validating the system, as these determine the true ability of a system to discriminate Organ-Confined Disease from Extra-Prostatic Disease using the knowledge which has been acquired by the system during the training (i.e. learning) process. To perform these evaluations, the actual outputs returned by each system during the validation stage were compared against the targets (i.e. known) outputs.

The results of the comparisons are shown in Table 14 and illustrated in Fig 5. The ROC curves for all systems are shown in Fig 6. The cutoff points of each classifier are presented in
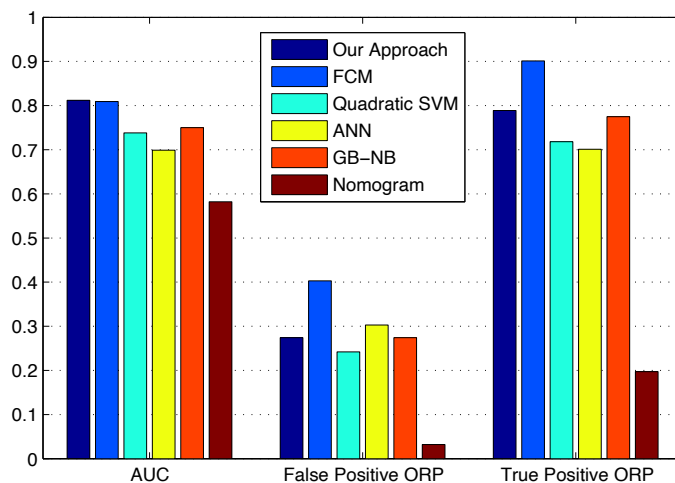
Fig 5. Performance Comparison.
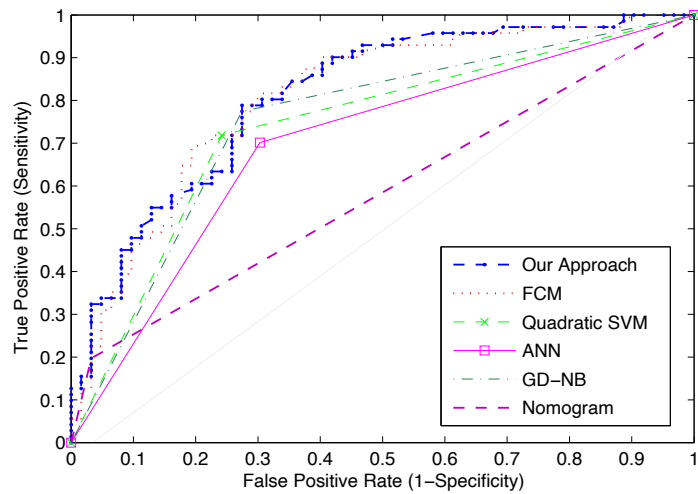
Fig 6. ROC Curves: Performance Comparison.

rows, Optimal ROC point False Positive Rate and Optimal ROC point True Positive Rate, of Table 14 and were computed with the alpha values set to $\alpha = 0.05$ (95% Confidence Interval). The best system would return the largest AUC, a high number of true positives, and a low number of false positives.

The Support Vector Machine was trained using the Linear kernel function, Quadratic, Gaussian Radial Basis (GRB), Multilayer Perceptron kernel (MP) functions. The results of testing the performance (i.e. validation) of the Support Vector Machine using the various kernel functions are presented in Table 15. The results show that the Quadratic-Support Vector Machine has the largest AUC ($AUC = 0.738$), thereby outperforming all other functions.

The Naive Bayes classifier results, when tuned using the Gaussian Distribution (i.e. normal distribution) and Kernel Density Estimation functions, are presented in Table 16. The results revealed that the Gaussian Naive Bayes (GD-NB) classifier returned a larger AUC ($AUC = 0.750$), thereby outperforming the Kernel Density Estimation Naive Bayes (KDE-NB) classifier ($AUC = 0.696$).

Table 15. Support Vector Machine(SVM) performance evaluation when applying various kernel functions.

Evaluation Measure	Kernel Function			
	Linear	Quadratic	GRB	MP
Specificity (TNR)	0.758	0.758	0.661	0.597
Sensitivity (TPR)	0.704	0.718	0.747	0.690
Area Under the Curve	0.731	0.738	0.704	0.644
Optimal ROC point FPR	0.242	0.242	0.339	0.403
Optimal ROC point TPR	0.704	0.718	0.747	0.690

Table 16. Naive Bayes(NB) performance evaluation using the Gaussian distribution and Kernel Density Estimation functions.

Evaluation Measure	Type of function	
	GD-NB	KDE-NB
Specificity (TNR)	0.726	0.645
Sensitivity (TPR)	0.745	0.747
Area Under the Curve	0.750	0.696
Optimal ROC point FPR	0.274	0.355
Optimal ROC point TPR	0.775	0.747

doi:10.1371/journal.pone.0155856.t016

The best Support Vector Machine and Naive Bayes classifiers are included in Table 14 which contains all of the results. The results of the comparison revealed that the proposed neuro-fuzzy system, at its optimal point, returns the largest AUC, with a low number of false positives ($FPR = 0.274$, $TPR = 0.789$, $AUC = 0.812$). Although the FCM classifier returned the highest number of true positives, it returned a very high number of false positives ($FPR = 0.403$, $TPR = 0.901$, $AUC = 0.809$). For these reasons, FCM cannot be considered as the optimum classifier. The Quadratic-Support Vector Machine, ANN, and GB-NB did not perform as well as the neuro-fuzzy approach—they returned a smaller AUC, and a lower optimal TPR. The AJCC pTNM Nomogram performed the worst, with the smallest AUC (0.582), and the lowest number of TPR (0.197) at the optimal ROC point. The proposed neuro-fuzzy system therefore outperformed all other systems.

Finally, Table 14 shows the results of the McNemar test which was applied to investigate whether any statistically significant differences exist among each system's target outputs and the actual outputs. The McNemar test is a statistical test which is applied on paired nominal data. It uses an approximate chi-square test of goodness to test the null hypothesis, i.e. there are no significant differences among targets and outputs. Each pair comprised of the actual and predicted values of each system. A good system will not return a statistically significant difference (i.e. $p > 0.05$) amongst its predicted outputs and the actual target outputs (known outputs). The results revealed that there were no statistically significant differences among the actual and predicted outputs of the proposed neuro-fuzzy approach ($p = 1.000$, $p > 0.05$); the FCM classifier, ($p = 0.868$, $p > 0.05$); the Quadratic-SVM, ($p = 0.499$, $p > 0.05$); the ANN ($p = 1.000$, $p > 0.05$); and the GB-NB, ($p = 1.000$, $p > 0.05$). However, there was a statistically significant difference among the outputs of the AJCC pTNM Nomogram against targets ($p = 0.00$, $p < 0.05$).

Discussion and Conclusion

At the clinical prostate cancer staging process, the patient undergoes various clinical tests for the prognosis of prostate cancer, and based on these tests, the clinician estimates (or predicts) how much the cancer has spread. It is only after surgery, and hence at the pathological stage, that it is possible to more accurately diagnose cancer and determine the extent of its spread beyond the prostate gland. The ability to predict that pathological stage of prostate cancer is important, as it allows clinicians to determine the best approach for treating and managing the disease.

Herein, we propose the application of a neuro-fuzzy based approach for the prediction of the pathological stage of prostate cancer. The algorithm is suitable for the particular problem due to the imprecision, and the uncertainty which is typically found in the results of the clinical tests which can be used for predicting the pathological stage of prostate cancer. The system

input comprised of variables Primary and Secondary Gleason patterns, PSA levels, age at diagnosis, and clinical T stage. The output is the pathological stage of the cancer which can be either Organ-Confined Disease or Extra-Prostatic Disease. Experiments were performed using an existing and validated prostate cancer patient dataset comprising n = 399 patient records obtained from The Cancer Genome Atlas (TCGA) Research Network. The performance of the proposed neuro-fuzzy system was compared to other classifiers: the Artificial Neural Network, Fuzzy C-Means, Support Vector Machines, the Naive Bayes, and the AJCC pTNM Nomogram [21]. The results revealed that the proposed neuro-fuzzy system outperformed all other classifiers. Our results appear to be consistent to those of Castanho et al. [12] who have also developed genetic-fuzzy expert system for predicting whether prostate cancer if confined or not-confined. Their results have also revealed that computational intelligence approaches based on fuzzy algorithms are suitable for prostate cancer staging prediction, and exceed the performance of nomograms.

The algorithm proposed by Castanho et al. [12] tunes the membership functions using a genetic algorithm, whereas we have used the Adaptive Neuro Fuzzy Inference System (ANFIS) to optimise the membership functions. Furthermore, Castanho et al.'s [12] and our proposed system both aim to predict whether a patient has organ-confined disease (OCD, pathological stage pT2) or extra-prostatic disease (ED, pathological stage $> pT2$). Although both systems use pre-operative serum PSA, clinical stage, and primary and secondary Gleason grades of a biopsy to predict the pathological stage of prostate cancer, our system considers age as an additional input variable. Castanho et al.'s [12] genetic-fuzzy system achieved an Area Under the Curve of 0.824 which they compared against Partin probability tables which have been proposed by Makarov et al. [31], and which only achieved an Area Under the Curve of 0.693. Our proposed neuro-fuzzy approach achieved an Area under the curve of 0.812, and the AJCC nomogram achieved an Area Under the Curve of 0.582. These results approximate to those reported by Castanho et al. [12], and reveal a high degree of consistency among the two outcomes of the two studies, despite the fact that different datasets were used for each study. The nomograms used by Castanho et al., and the AJCC nomogram both use the TNM Classification of Malignant Tumors grading system [32]. A major limitation of the AJCC nomogram is that the biopsy Gleason 7 values are not split into 3 + 4 = 7 vs. 4 + 3 = 7 which have drastically different clinical outcomes. The proposed neuro-fuzzy model considers the Gleason Grades 3 + 4 and 4 + 3, and this was one of the reasons that it performed better than the AJCC nomogram.

A recent study by Tsao et al. [15] has also reported similar AUC values, to those returned by our model and that of Castanho et al. [12], when using the Partin probability tables proposed which have been proposed by Makarov et al. [31] to predict the pathological stage of prostate cancer in patients prior to receiving radical prostatectomy. Tsao et al. [15] developed an artificial neural network (ANN) model to predict the pathological stage of prostate cancer and evaluated the model on 299 patients, of whom 109 (36.45%) displayed prostate cancer with extracapsular extension (ECE), and 190 (63.55%) displayed organ-confined disease (OCD). Overall, their results revealed that the ANN model (AUC = 0.795) significantly outperformed a Linear Regression statistical model (AUC = 0.746), and the Partin Tables (AUC = 0.695).

It should be noted that other predictors/nomograms consider features other than clinical stage, PSA, age, and biopsy Gleason grade. Some use the amount of tumor present, while others are starting to incorporate results of molecular analysis data, such as data from Prolaris [33] or oncotype measurements. Such models were not considered in the current study, as the relevant information is not available via the TCGA datasets and urologists predominantly use the Kattan preoperative nomogram [34] and the Partin Tables [35] for determining the likelihood of prostate cancer recurrence following radical prostatectomy, at least in the UK and Europe.

A study by Tamblyn et al. [36] which compared the Cancer of the Prostate Risk Assessment (CAPRA) score [37] against the Kattan (version 1998) [34] and Stephenson nomograms (version 2006) [38] revealed that the Kattan (version 1998) [34] tool was the best predictor of absolute risk of recurrence. Furthermore, a recent study by Boehm et al. [39] which compared three preoperative models, D'Amico [40], CAPRA [37] and Stephenson [38], revealed that these tools are reliable in North American patients, but have shortcomings for identifying patients at high risk of prostate cancer death in Europe. D'Amico [40] and CAPRA [37] include the amount of tumor detected on biopsy as part of their risk prediction algorithm and consider volume to have an influence on the risk of disease recurrence. However, from the evidence presented in Boehm et al. [39], this is not necessarily the case for non-US patients. Therefore, the precise influence of tumor volume on the risk of disease is inconclusive. Finally, although the volume of tumour per each core has been used to determine the significance of the tumor, Gleason 6 disease is still regarded as being a non-significant pathology, whereas Gleason 7 or greater is thought to be significant disease, irrespective of volume. As such, volume of disease adds very little to the decision-making process.

Currently, the proposed framework has been implemented as a research tool, and once more evaluations are conducted, the tool will be developed as a simple to use application which can be made accessible to clinicians. The tool will take the clinical test results (i.e. age at diagnosis, PSA, biopsy Primary and Secondary Gleason patterns, and clinical T stage) of an individual patient and predict his likelihood of having extra-prostatic cancer, and thereby aid the clinical decision-making process. Ongoing work is applying the proposed neuro-fuzzy predictor to a larger dataset, examining other computational intelligence approaches, and continuing the development of novel algorithms for predicting disease status in patients with prostate cancer.

Author Contributions

Conceived and designed the experiments: GC AGP MK DB. Performed the experiments: GC. Analyzed the data: GC. Contributed reagents/materials/analysis tools: GC. Wrote the paper: GC DB GA RR MK AGP. Performed the data analysis: GC. Designed the prediction model: GC. Performed the experiments: GC. Drafted the manuscript: GC. Contributed equally with GC in writing and revising the entire manuscript and also in writing the clinical elements of the manuscript: AGP. Contributed significantly in writing the clinical elements of the manuscript and in revising the manuscript: MK RR. Contributed significantly to the design and experimental part of the study and in revising the manuscript: DB. Contributed to the design of the prediction model: GA. Participated in the preparation of the final manuscript: GC GA DB RR MK AGP. Read and approved the final manuscript: GC GA DB RR MK AGP.

References

1. Blute M, Nativ O, Zincke H, Farrow G, Therneau T, Lieber M. Pattern of failure after radical retropubic prostatectomy for clinically and pathologically localized adenocarcinoma of the prostate: influence of tumor deoxyribonucleic acid ploidy. J Urol. 1989; 142:1262–1265. PMID: 2810503

2. Epstein J, Pizov G, Walsh P. Correlation of pathologic findings with progression after radical retropubic prostatectomy. Cancer. 1993; 71:3582–3593. doi: 10.1002/1097-0142(19930601)71:11%3C3582::AID-CNCR2820711120%3E3.0.CO;2-Y PMID: 7683970

3. Epstein J, Walsh P, Carmichael M, Brendler C. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA. 1994; 271:368–374. doi: 10.1001/jama.1994.03510290050036 PMID: 7506797

4. Dotan ZA, Ramon J. Nomograms as a tool in predicting prostate cancer prognosis. European Urology Supplements. 2009; 8(9):721–724. doi: 10.1016/j.eursup.2009.06.013

5. Briganti A, Karakiewicz PI, Joniau S, Van Poppel H. The Motion: nomograms should become a routine tool in determining prostate cancer prognosis. European Urology. 2009; 55(3):743–747. doi: 10.1016/j.eururo.2008.11.038 PMID: 19081177

6. Chun FKH, Karakiewicz PI, Briganti A, Walz J, Kattan MW, Huland H, et al. A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. BJU International. 2007; 99 (4):794–800. doi: 10.1111/j.1464-410X.2006.06694.x PMID: 17378842

7. Tewari A, Porter C, Peabody J, Crawford ED, Demers R, Johnson CC, et al. Predictive modeling techniques in prostate cancer. Molecular Urology. 2001; 5(4):147–152. doi: 10.1089/10915360152745812 PMID: 11790275

8. Han J. Data Mining: concepts and techniques. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2005.

9. Benecchi L. Neuro-fuzzy system for prostate cancer diagnosis. Urology. 2006; 68(2):357–361. doi: 10. 1016/j.urology.2006.03.003 PMID: 16904452

10. Keles A, Hasiloglu AS, Keles A, Aksoy Y. Neuro-fuzzy classification of prostate cancer using NEF-CLASS-J. Computers in Biology and Medicine. 2007; 37(11):1617–1628. doi: 10.1016/j.compbiomed. 2007.03.006 PMID: 17531217

11. Çinar M, Mehmet E, Erkan ZE, Ateşçi ZY. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. Expert Syst Appl. 2009; 36(3):6357–6361. doi: 10.1016/j.eswa. 2008.08.010

12. Castanho MJP, Hernandes F, De Rè AM, Rautenberg S, Billis A. Fuzzy expert system for predicting pathological stage of prostate cancer. Expert Systems with Applications. 2013; 40(2):466–470. doi: 10. 1016/j.eswa.2012.07.046

13. Saritas I, Ozkan IA, Sert IU. Prognosis of prostate cancer by artificial neural networks. Expert Systems with Applications. 2010; 37(9):6646–6650. doi: 10.1016/j.eswa.2010.03.056

14. Shariat SF, Kattan MW, Vickers AJ, Karakiewicz PI, Scardino PT. Critical review of prostate cancer predictive tools. Future Oncol. 2009; 5(10):1555–84. doi: 10.2217/fon.09.121 PMID: 20001796

15. Tsao CW, Liu CY, Cha TL, Wu ST, Sun GH, Yu DS, et al. Artificial neural network for predicting pathological stage of clinically localized prostate cancer in a Taiwanese population. Journal of the Chinese Medical Association. 2014; 77(10):513–518. doi: 10.1016/j.jcma.2014.06.014 PMID: 25175002

16. Burford D, Kirby M, Austoker J. Prostate cancer risk management programme: information for primary care PSA testing in asymptomatic men. NHS Cancer Screening Programmes; 2009.

17. Schröder FH, Hugosson J, Carlsson S, Tammela T, Määttänen L, Auvinen A, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). European Urology. 2012; 62(5):745–752. doi: 10.1016/j.eururo.2012.05.068 PMID: 22704366

18. Luján M, Páez À, Berenguer A, Rodríguez JA. Mortality due to prostate cancer in the Spanish arm of the European Randomized Study of Screening for Prostate Cancer (ERSPC). Results after a 15-year follow-up. Actas Urológicas Españolas (English Edition). 2012; 36(7):403–409.

19. Heijnsdijk E, Wever E, de Koning H. Cost-effectiveness of prostate cancer screening based on the European Randomised Study of Screening Pprostate Cancer. The Journal of Urology. 2012; 187(4, Supplement):e491. doi: 10.1016/j.juro.2012.02.1502

20. Falzarano SM, Magi-Galluzzi C. Staging prostate cancer and its relationship to prognosis. Diagnostic Histopathology. 2010; 16(9):432–438. doi: 10.1016/j.mpdhp.2010.06.010

21. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A. AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer; 2010.

22. Zadeh LA. Fuzzy sets. Information and Control. 1965; 8(3):338–353. doi: 10.1016/S0019-9958(65) 90241-X

23. Sugeno M, Yasukawa T. A fuzzy-logic-based approach to qualitative modeling. IEEE Transactions on Fuzzy Systems. 1993; 1(1):7–31. doi: 10.1109/TFUZZ.1993.390281

24. Jang JSR. ANFIS: adaptive-network-based fuzzy inference system. Systems, Man and Cybernetics, IEEE Transactions on. 2002; 23(3):665–685. doi: 10.1109/21.256541

25. Jang JSR. Input selection for ANFIS learning. In: Proceedings of the IEEE International Conference on Fuzzy Systems; 1996. p. 1493–1499.

26. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. Norwell, MA, USA: Kluwer Academic Publishers; 1981.

27. Möller MF. A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks. 1993; 6(4):525–533. doi: 10.1016/S0893-6080(05)80056-5

28. Hsu CW, Lin CJ. A comparison of methods for multiclass support vector machines. Neural Networks, IEEE Transactions on. 2002; 13(2):415–425. doi: 10.1109/72.991427

29. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. New York, NY, USA: Cambridge University Press; 2000.

30. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.; 2001.

31. Makarov DV, Trock BJ, Humphreys EB, Mangold LA, Walsh PC, Epstein JI, et al. Update nomogram to predict pathologic stage of prostate cancer given prostate-specific antigen level, clinical stage, and biopsy Gleason score (Partin Tables) based on cases from 2000 to 2005. Urology. 2007; 6(69):1095–1101. doi: 10.1016/j.urology.2007.03.042

32. Sobin L, Wittekind C. TNM classification of malignant tumors. 5th ed. John Wiley and Sons, Inc., New York; 1997.

33. Brawer K, Cuzick J, Cooperberg M, Swanson G, Freedland S, Reid J, et al. Prolaris: a novel genetic test for prostate cancer prognosis. Journal of Clinical Oncology. 2013; 31.

34. Kattan M, Eastham J, Stapleton A, Wheeler T, Scardino P. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. Journal of the National Cancer Institute. 1998; 10(90):766–71. doi: 10.1093/jnci/90.10.766

35. Partin AW, Mangold LA, Lamm DM, Walsh PC, Epstein JI, Pearson JD. Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. Urology. 2001; 58(6):843–848. doi: 10.1016/S0090-4295(01)01441-8 PMID: 11744442

36. Tamblyn D, Chopra S, Yu C, Kattan M, Pinnock C, Kopsaftis T. Comparative analysis of three risk assessment tools in Australian patients with prostate cancer. Journal of the British Association of Urological Surgeons. 2011; 108(2):51–66.

37. Cooperberg M, Pasta D, Elkin E, Litwin M, Latini D, Du CJ, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. The Journal of Urololgy. 2005; 173(6):1938–42. doi: 10.1097/01.ju.0000158155.33890.e7

38. Stephenson AJ, Scardino PT, Eastham JA, Bianco FJ, Dotan ZA, Fearn PA, et al. Preoperative Nomogram Predicting the 10-Year Probability of Prostate Cancer Recurrence After Radical Prostatectomy. Journal of the National Cancer Institute. 2006; 98:715–717. doi: 10.1093/jnci/djj190 PMID: 16705126

39. Boehm K, Larcher A, Beyer B, Tian Z, Tilki D, Steuber T, et al. Identifying the Most Informative Prediction Tool for Cancer-specific Mortality After Radical Prostatectomy: Comparative Analysis of Three Commonly Used Preoperative Prediction Models. European Urology. 2015;.

40. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998; 280(11):969–974. doi: 10.1001/jama.280.11.969 PMID: 9749478

Identifying the Presence of Prostate Cancer in Individuals with PSA Levels <20 ng ml⁻¹ Using Computational Data Extraction Analysis of High Dimensional Peripheral Blood Flow Cytometric Phenotyping Data

Georgina Cosma[1]*, Stéphanie E. McArdle[2], Stephen Reeder[2], Gemma A. Foulds[2], Simon Hood[2], Masood Khan[3] and A. Graham Pockley[2]*

[1] School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom, [2] John van Geest Cancer Research Centre, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom, [3] University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

Determining whether an asymptomatic individual with Prostate-Specific Antigen (PSA) levels below 20 ng ml⁻¹ has prostate cancer in the absence of definitive, biopsy-based evidence continues to present a significant challenge to clinicians who must decide whether such individuals with low PSA values have prostate cancer. Herein, we present an advanced computational data extraction approach which can identify the presence of prostate cancer in men with PSA levels <20 ng ml⁻¹ on the basis of peripheral blood immune cell profiles that have been generated using multi-parameter flow cytometry. Statistical analysis of immune phenotyping datasets relating to the presence and prevalence of key leukocyte populations in the peripheral blood, as generated from individuals undergoing routine tests for prostate cancer (including tissue biopsy) using multi-parametric flow cytometric analysis, was unable to identify significant relationships between leukocyte population profiles and the presence of benign disease (no prostate cancer) or prostate cancer. By contrast, a Genetic Algorithm computational approach identified a subset of five flow cytometry features ($CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells); $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA); $CD3^-CD19^+$ (B cells); $CD3^+CD56^+CD8^+CD4^+$ (NKT cells)) from a set of twenty features, which could potentially discriminate between benign disease and prostate cancer. These features were used to construct a prostate cancer prediction model using the k-Nearest-Neighbor classification algorithm. The proposed model, which takes as input the set of flow cytometry features, outperformed the predictive model which takes PSA values as input. Specifically, the flow cytometry-based model achieved Accuracy = 83.33%, AUC = 83.40%, and optimal ROC points of FPR = 16.13%, TPR = 82.93%, whereas the PSA-based model achieved Accuracy = 77.78%, AUC = 76.95%, and optimal ROC points of FPR = 29.03%, TPR = 82.93%. Combining

PSA and flow cytometry predictors achieved Accuracy = 79.17%, AUC = 78.17% and optimal ROC points of FPR = 29.03%, TPR = 85.37%. The results demonstrate the value of computational intelligence-based approaches for interrogating immunophenotyping datasets and that combining peripheral blood phenotypic profiling with PSA levels improves diagnostic accuracy compared to using PSA test alone. These studies also demonstrate that the presence of cancer is reflected in changes in the peripheral blood immune phenotype profile which can be identified using computational analysis and interpretation of complex flow cytometry datasets.

1. INTRODUCTION

The introduction of the serum Prostate-Specific Antigen (PSA) level as a biomarker for the presence of prostate cancer in 1986 prompted a progressive global increase in the diagnosis, and earlier diagnosis of the disease. The fact that most men are now diagnosed with organ-confined disease enables intervention with curative intent. However, although the initial diagnosis of prostate cancer in most men is based on a PSA test and digital rectal examination (DRE) (1), the PSA test has been criticized for its poor diagnostic specificity (30%) (2). Further investigations are, therefore, indicated in the event of an elevated PSA or abnormal DRE. These include a transrectal ultrasound (TRUS)-guided prostate biopsy and subsequent examination and reporting by a pathologist. However, TRUS-guided prostate biopsies have a documented sensitivity of only 39–52% (3), and cancer detection rates of around 25% on initial biopsies (4), and 18–32% on repeated biopsies (5, 6). This approach is also costly and rarely detects prostate cancers that an elevated PSA and/or DRE cannot predict. Although TRUS is commonly used to guide a biopsy, it is not, therefore, recommended for routine screening. An alternative approach to the TRUS is the Transperineal Template Prostate Biopsy (TPTPB), and we have previously shown that TPTPB can identify clinically significant prostate cancer in 71/122 (58%) of men with raised PSA, despite two previous sets of negative TRUS biopsies (7). An important element of these findings was that 61% of the patients in whom prostate cancer was diagnosed had a Gleason grade score ≥7 (most which were in the anterior zone), thereby automatically placing them into the "intermediate" or "high-risk" categories when applying established risk stratification criteria (7). The capacity of the TPTPB to identify more clinically significant tumors at an earlier stage, therefore, suggests that it is a better diagnostic test for localized prostate cancer than the TRUS biopsy. Given the ability of the TPTPB to detect prostate cancer at significantly higher rates than TRUS biopsies (8–12), we questioned whether we should move away from TRUS biopsies to TPTPB and whether PSA is actually a more specific biomarker for prostate cancer detection than had been previously thought. To this end, we performed a prospective study which directly compared the diagnostic potential of the TRUS and TPTPB approaches in the same cohort of biopsy naïve men with an elevated PSA <20 ng ml^{-1} and a benign feeling prostate on a DRE. These patients, therefore, served as their own controls (13). The study demonstrated that the TRUS biopsy detected cancer in 32 versus 60% with TPTP, and that TPTPB is associated with a significantly higher prostate cancer detection rate than TRUS biopsies in biopsy naïve men with PSA <20 ng ml^{-1} and a benign feeling DRE (13). However, given that TRUS guided prostate biopsies are associated with a 5% risk of urosepsis (which can be life-threatening), and that TPTPB is performed under general anesthetic and associated with a 5% risk of urinary retention, both procedures are associated with a significant cost and potential for complications. It is also essential that men with low-risk prostate cancer are not diagnosed as having cancer, as they do not require any active treatment and such individuals are "labeled" as having cancer. This can have profound adverse psychological and financial consequences, and assign them to life-long surveillance. The fundamental aim of this study is, therefore, to develop an approach which delivers a high level of diagnostic accuracy for asymptomatic men with an elevated PSA <20 ng l^{-1}. The development of such approaches will spare men with benign disease or low-risk cancer from unnecessary invasive diagnostic procedures such as TRUS-guided prostate biopsies or TPTPB. Given the reciprocal interactions between tumors and the immune system, we hypothesized that the presence of disease, disease recurrence, and therapeutic resistance may be influenced, reflected in, or predicted by tumor-related immunoregulatory events that can be identified by changes in immune phenotypes in the periphery. We, therefore, proposed that the analysis of immune phenotyping datasets using multi-parametric flow cytometric analysis can identify features that reflect the presence of disease and/or predict disease progression (14). Although flow cytometry provides a vital tool for exploring, explaining, and understanding complex cellular dynamics and processes in a variety of experimental and clinical settings (15), key challenges with multi-parametric flow cytometry include the analysis and interpretation of the complex and increasingly multidimensional data and its conversion into biologically and clinically useful information. This study attempts to address and resolve some of these challenges using computational intelligence methods. Computational intelligence methods comprise evolutionary algorithms (also known as metaheuristic optimization, or nature-inspired optimization algorithms) coupled with machine learning methods, and hybrids of these.

A type of machine learning method, supervised learning, is used to derive prediction models which can be very effective in dealing with uncertainty, noise, and dimensionality in data. Supervised learning methods can learn from existing data to make informed predictions using new patient data, and have been widely adopted for prostate cancer prediction tasks when using clinical and biomedical data (16). It is now time to embrace computational intelligence methods for the analysis of flow cytometry data, since statistical methods alone may not be sufficient for the task of analyzing and modeling such complex data (16). Herein, we assess whether advanced computational analysis of peripheral blood flow cytometry immunophenotyping data from a selected cohort of individuals can generate prediction models with potential clinical value and identify the presence of prostate cancer in asymptomatic individuals with a PSA level <20 ng ml⁻¹. The computational models and algorithms are trained to make predictions on new and previously unseen data using existing data. Significantly, this approach has identified a novel prostate cancer immunophenotyping "fingerprint" which could potentially be used to identify the presence of prostate cancer in asymptomatic men having PSA levels <20 ng ml⁻¹; and which outperforms the predictive value of the PSA test alone. We have also shown that combining flow cytometry predictors with PSA levels improves diagnostic accuracy. Taken together, these studies demonstrate that the presence of cancer is reflected in changes in the peripheral blood immune phenotype profile which can be identified using computational analysis and interpretation of complex flow cytometry datasets, and the value of computational intelligence-based approaches for interrogating immunophenotyping datasets.

2. MATERIALS

2.1. Data Collection

Patients with suspected prostate cancer attending the Urology Clinic at Leicester General Hospital (University Hospitals of Leicester NHS Trust, Leicester, UK) were examined by Professor Masood Khan (Consultant Urologist) and Mr. Shady Nafie (Registrar in Urology). Samples were obtained from a selected cohort of patients which met the following criteria—being biopsy naïve, with a PSA level of <20 ng ml⁻¹ and agreeing to undergo simultaneous TRUS biopsy (12 cores) and a transperineal template prostate biopsy (TPTPB) (36 cores) procedures under general anesthetic. Samples from the TPTPB cohort were collected from 24 October 2012 to 15 August 2014. Further details on how patients were recruited and treated are described in Nafie et al. (7). The cohort comprised samples from 72 males who had a TRUS-guided biopsy and then a TPTPB. The mean age for this cohort was 66 years old (age range of 50–84 years old). Given the more definitive diagnostic power of the TPTPB (7, 13), samples that were considered as being from individuals with benign disease were obtained from this cohort. A total of 41 (56.94%) patients were diagnosed with prostate cancer. The remaining 31 (43.06%) patients were classed as having benign disease following pathological examination and the application of established criteria. Of those patients diagnosed with benign disease, 10 patients were diagnosed with High Prostatic Intraepithelial Neoplasia

(High-Grade); 10 patients were diagnosed with Atypical Small Acinar Proliferation and 2 patients with Atypia. The remaining 9 patients were diagnosed as having benign disease. Patients with multi-focal high-grade PIN or ASAP commonly have a prostatic core biopsy showing a focus which is suspicious for, but not diagnostic of, cancer (17).

2.2. Ethics Statement

Research Protocols were registered and approved by the National Research Ethics Service (NRES) Committee East Midlands and by the Research and Development Department in the University Hospitals of Leicester NHS Trust. All participants were given information sheets explaining the nature of the study and all provided informed consent. All samples were collected by suitably qualified individuals using standard procedures. Ethical approval for the collection and use of samples from the TPTPB cohort (Project Title: Defining the role of Transperineal Template-guided prostate biopsy) was given by NRES Committee East Midlands-Derby 1 (NREC Reference number: 11/*EM*/3012; UHL11068). Ethical approval for the collection of peripheral blood from healthy volunteers was obtained from the Nottingham Trent University College of Science and Technology Human Ethics Committee (Application numbers 165 and 412).

2.3. Flow Cytometric Analysis

Peripheral blood (60 ml) was collected from all patients using standard clinical procedures. Aliquots (30 ml) were transferred into two sterile 50 ml polypropylene (Falcon) tubes containing 300 µl of sterilized Heparin (1000 U ml⁻¹, Sigma). Anti-coagulated samples were immediately transferred to the John van Geest Cancer Research Centre at Nottingham Trent University (Nottingham, UK) and were processed immediately upon receipt (as described in this section), and within 3 h of collection. 200 µl of whole blood was used to profile the key immune cell subsets in the periphery (Overview of the Immune System: "OVIS"—see **Table 1**).

Absolute cell counts in whole blood samples were determined by the inclusion of BD Trucount™ beads (BD Biosciences; Mountain View, CA, USA), as per the manufacturer's protocol. For the flow cytometric analysis, 100 µl of blood was mixed directly in the BD Trucount™ bead tube and T cell, B cell, and NK cell populations identified using the conjugated monoclonal antibodies (mAbs) detailed in **Table 1**. For the staining, cells

TABLE 1 | Monoclonal antibody panel.

Antibody	Fluorochrome	Clone No.	Supplier
CD8	FITC	SK1	Biolegend
CD19	PE	HIB19	Biolegend
CD28	PE-Texas Red (ECD)	CD28.2	Beckman Coulter
CD56	PE-Cy5	NCAM	Biolegend
CD3	PE-Cy7	HIT3a	Biolegend
CD45RA	Allophycocyanine (APC)	HI100	eBioscience
CD14	Alexa Fluor 700	HCD14	Biolegend
CD27	APC eFluor 780	O323	eBioscience
CD45	Pacific Blue	J33	Beckman Coulter
CD4	Krome Orange	13B8.2	Beckman Coulter

were incubated for 15 min at room temperature, protected from the light, after which erythrocytes were lysed by incubating samples for 15 min at room temperature in BD Pharm Lyse™ (BD Biosciences). Once staining was complete, cells were washed in phosphate buffered saline (PBS), resuspended in Coulter Isoton™ diluent. Data were acquired within 1 h using a 10-color/3-laser Beckman Coulter Gallios™ flow cytometer and analyzed using Kaluza™ v1.3 data acquisition and analysis software (Beckman Coulter). Controls used a "Fluorescence minus One," "FMO" approach. A typical gating strategy for the analyses is presented in **Figure 1**.

2.4. Data Normalization and Statistical Analysis

For this study, we considered a feature to be the grouped set of flow cytometry phenotypic variables shown in **Table 2**. The mean and Standard Deviation (SD) values of each flow cytometry feature shown in **Table 2** indicate clear variation, as a consequence of which data were normalized to put them on the same scale and enable the comparison of two or more variables (i.e., flow cytometry features). Let $X_{mxn} = [x_{ij}]$ be a m x n matrix with m rows and n columns. Z-score normalization was applied to each column n of matrix X. Applying normalization returned the z-score value for each matrix element x_{ij}, and each column j of matrix X was centered to have a mean value of 0 and scaled to have a SD value of 1. The standardized data set retains the shape properties of

the original data set (same skewness and kurtosis). The z-score normalization function is shown in Function (1):

$$z = \frac{(x_{ij} - \bar{x})}{\sigma} \tag{1}$$

where x_{ij} is a data point; $\bar{x}$ is the mean value of column j; σ is the SD; and z is the transformed value of data point x_{ij}.

Figure 2 illustrates the distribution of the flow cytometry features in the form of box plots, and allows for quick visualization of variability. Outliers were included in the analyses as it is important to consider those "out of range values" when creating a prediction model. **Figure 3** illustrates the flow cytometry values derived from individuals with benign disease and patients with prostate cancer before and after data normalization.

Table 3 provides descriptive statistics of the normalized dataset, and these are also illustrated in **Figure 2**. The Interquartile range (IQR) is an informative measure of variability and determined by computing the distance between the Upper Quartile (i.e., top) and Lower Quartile (i.e., bottom) of the box. The features with the smallest degree of variability are those with the smallest IQR values, and hence: $CD4^+CD45RA^+CD27^-CD28^-$ (ID 10, first smallest); $CD3^+CD56^+CD8^+CD4^+$ (ID 17, second smallest); $CD3^+CD56^+CD8^-CD4^+$ (ID 19, third smallest); $CD4^+CD45RA^-CD27^-CD28^-$ (ID 9, fourth smallest), $CD8^+CD45RA^-CD27^-CD28^-$ (ID 4, fifth smallest). These variables, therefore, appear to the best candidate predictors when considered independently.

FIGURE 1 | Representative gating strategies for the flow cytometric analysis of single cells. The Overview of the Immune System (OVIS) staining panel confirmed CD45 expression then determined cell populations as $CD14^+$ monocytes, $CD3^-CD56^+$ NK cells (with $CD56^{bright}$ and $CD56^{dim}$ subsets), $CD3^+CD56^+$ NKT cell subpopulations, $CD19^+$ B cells, $CD3^+CD4^+$ and $CD3^+CD8^+$ Naïve, Central Memory, Effector Memory, Terminally Differentiated Effector Memory Cells Expressing CD45RA T cells populations. The definition of monocytes based on $CD45^+CD4^+$ generated the same data as defining them based on $CD3^-CD14^+$ (data not shown).

TABLE 2 | Flow Cytometry features.

Feature ID	Flow cytometry feature	Mean	SD
1	$CD3^+CD8^+$	450.39	402.03
2	$CD8^+CD45RA^+CD27^+CD28^+$	96.92	75.99
3	$CD8^+CD45RA^-CD27^+CD28^-$	68.45	58.73
4	$CD8^+CD45RA^-CD27^-CD28^-$	45.37	104.69
5	$CD8^+CD45RA^+CD27^-CD28^-$	120.05	197.85
6	$CD3^+CD4^+$	877.88	468.35
7	$CD4^+CD45RA^+CD27^+CD28^+$	393.72	214.27
8	$CD4^+CD45RA^-CD27^+CD28^+$	311.24	211.16
9	$CD4^+CD45RA^-CD27^-CD28^-$	17.78	39.78
10	$CD4^+CD45RA^+CD27^-CD28^-$	14.19	35.48
11	$CD45^+CD14^+$	116.16	87.19
12	$CD3^-CD19^+$	257.70	251.40
13	$CD3^+CD56^+$ NKT	76.54	85.74
14	$CD3^-CD56^+$ NK	260.34	202.84
15	$CD3^-CD56^{low}$	253.20	192.27
16	$CD3^-CD56^{high}$	16.06	14.66
17	$CD3^+CD56^+CD8^+CD4^+$	5.67	16.16
18	$CD3^+CD56^+CD8^+CD4^-$	53.32	59.76
19	$CD3^+CD56^+CD8^-CD4^+$	10.96	20.41
20	$CD3^+CD56^+CD8^-CD4^-$	6.59	6.58

Mean and Standard Deviation (SD) values of raw data.

The Kruskal–Wallis test ("one-way ANOVA on ranks") tested for statistically significant differences between the mean ranks of the normalized flow cytometry variables observed in individuals with benign disease and patients with prostate cancer due to the presence of unequal variances, and demonstrated there to be no statistically significant differences at the alpha level of $\alpha = 0.05$ in the mean ranks of the flow cytometry features between these two groups (**Table 4**). A more sophisticated approach that has the potential to determine which features would better indicate the presence of disease was, therefore, adopted. For this, a Genetic Algorithm was used to explore the different combinations of features and return the optimal combination of features which indicate the presence of prostate cancer. As a final stage of the analysis, and prior to applying a Genetic Algorithm for feature selection, it is useful to determine whether any correlations among the flow cytometry features exist. For this, the non-parametric Spearman rank correlation assessed the degree of association between flow cytometry features. The rho values arising from this analysis were plotted in a heatmap graph (shown in **Figure 4**) in order to visualize those feature pairs having strong positive and strong negative correlations. **Figure 4** shows that many pairs have positive correlation values (color red). The p values were computed to determine which of these correlations were significant at $\alpha = 0.05$. The rho correlation values range from -1.0 to $+1.0$. A value of 0 suggests no correlation, a value of $+1.0$ suggests a strong positive correlation and a value of -1.0 suggests a strong negative correlation. A total of 141 unique pairs of features returned significant correlations with $p < 0.05$.

The large number of pairs having significant correlations presents significant challenges for identifying features which better identify the presence of disease. This is because if two features have a strong correlation, then only one of those features should be selected as a candidate predictor. A Genetic Algorithm evaluates these combinations and identifies those features that, as a combination, deliver the best subset of predictors.

3. RESULTS

3.1. Experiment Methodology

The aim of the experiments is to identify a suitable set of features which would, as a combination, deliver an immunophenotypic "fingerprint" for determining whether an individual with Prostate-Specific Antigen (PSA) levels below 20 ng ml⁻¹ has prostate cancer in the absence of definitive biopsy-based evidence. This fingerprint, or set of features, would then be utilized to construct a prediction model. Given that the optimum number of features was unknown, a Genetic Algorithm [18] was applied λ times, with $\lambda = 2, 3, \ldots, n$ where n is the total number of flow cytometry features. Therefore, each time the Genetic Algorithm was run a combination containing λ number of features was returned. A total of 19 subsets of features were returned by the Genetic Algorithm, with the first subset s_1 containing the best 2 selected features; subset s_2 the best 3 selected features, subset s_3 the best 4 selected features, and so forth. Each subset, s_i of selected features, was input into a kNN classifier. Experiments were conducted with kNN using various distance measures, as this would allow for it to be tuned for the specific problem at hand. The number of kNN neighbors was set to $k = 2$ and was chosen experimentally to be the best setting. The state-of-the-art Leave-One-Out Cross Validation (LOOCV) approach was adopted for evaluating the performance of the kNN classifier using various parameter settings. During LOOCV, the training and testing process is repeated m times and in every iteration, a different patient record is left out for testing until all records are left out [19]. To perform the evaluations, the actual outputs returned by the classification model during the validation stage were compared against the targets (i.e., known outputs). The Receiver Operating Characteristic (ROC) curves were created and the optimal cut-off points (optimal ROC point (ORP): False Positive Rate (FPR), True Positive Rate (TPR)) were computed with the alpha value set to $\alpha = 0.05$ (95% Confidence Interval). An efficient classification system (i.e., prediction model) would return the largest Area Under the Curve (AUC); a high number of True Positives; and a low number of False Positives. The methods of Hanley and McNeil [20, 21] were used for the calculation of the Standard Error of an Area Under the Curve (AUC (SE)), and the Binomial Exact Confidence Interval for an Area Under the Curve (AUC (BEC)) was also calculated.

3.2. Prostate Cancer Prediction Using Immunophenotyping Data

This section discusses the results of the experiments when tuning the kNN with various distance measures and when using each subset, s_i, of flow cytometry features which were returned by the Genetic Algorithm. **Table 5** shows the best results that were achieved after applying the kNN classifier using each subset of features and different distance measures. As shown in **Table 5**, the best performance was achieved using the FC-PM(Correlation(5)) which reached an AUC = 83.40% and Optimal ROC point of FPR = 16.13%, TPR = 82.93%. The FC-PM(Correlation(5)) utilized 5 flow cytometry features with IDs: 4, 9, 10, 12, 17 which correspond to flow cytometry features:
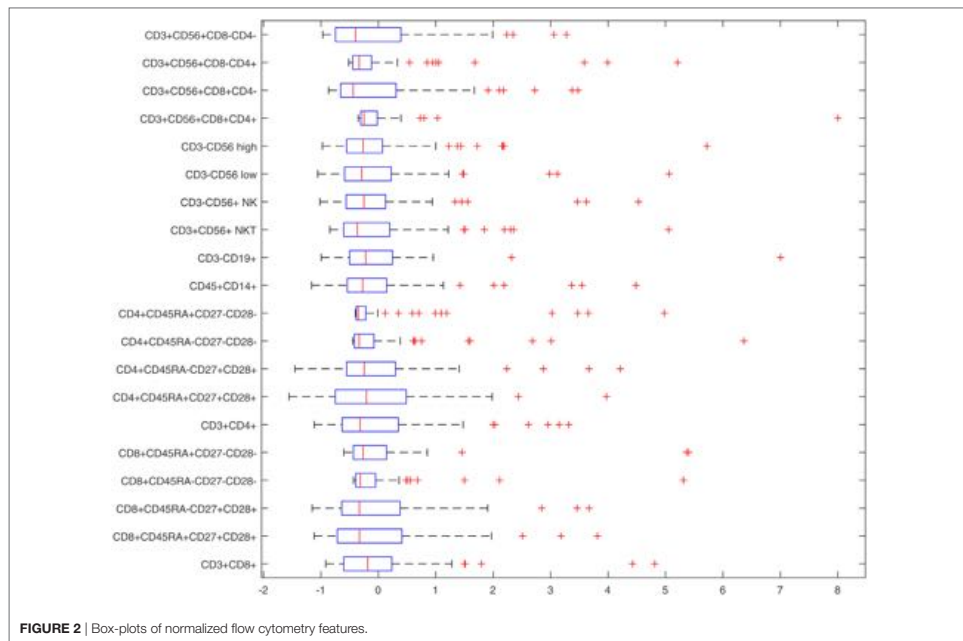
FIGURE 2 | Box-plots of normalized flow cytometry features.
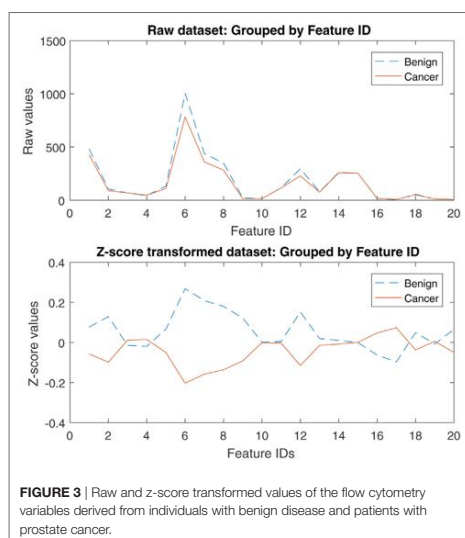
FIGURE 3 | Raw and z-score transformed values of the flow cytometry variables derived from individuals with benign disease and patients with prostate cancer.

$CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells); $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Effector Memory Cells); $CD4^+CD45RA^+CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA); $CD3^-CD19^+$ (B cells); $CD3^+CD56^+CD8^+CD4^+$ (NKT cells).

Given that this set contains the best combination of flow cytometry predictors, it can be used as a signature for distinguishing between the presence of benign disease and cancer. FC-PM(Cosine(6)) achieved the same value for Accuracy as FC-PM(Correlation(5)) using 6 features. Feature $CD45^+$ $CD14^+$ (ID 11) was included in the feature set used by FC-PM(Cosine(6)). **Figure 5** shows the AUCs and optimal ROC Points of the two flow cytometry-based prediction models, FC-PM(Correlation(5)) and FC-PM(Cosine(6)). FC-PM (Cosine(6)) achieved a 12.9% higher False Positive Rate than FC-PM (Correlation(5)) (**Table 5**), and lower Confidence Interval(CI) values shown in **Table 6**, which suggests that it has weaker ability than FC-PM(Correlation(5)) to discriminate between benign and cancer patients. In addition, **Table 6** shows the percentage of patients correctly classified in each group. FC-PM(Cosine(6)) achieved a lower predictive accuracy for benign patients compared to FC-PM(Correlation(5)), but correctly classified more cancer patients. The comparison suggests that FC-PM(Cosine(6)) is relatively more likely to misclassify benign patients as cancer patients, which is not a desirable outcome, and thus the model's confidence in identifying benign disease is lower.

TABLE 3 | Descriptive statistics of the normalized dataset.

	Flow cytometry feature	Range	Minimum	Maximum	IQR	Skewness
1	$CD3^+CD8^+$	5.73	−0.92	4.81	0.84	2.92
2	$CD8^+CD45RA^+CD27^+CD28^+$	4.93	−1.12	3.82	1.13	1.65
3	$CD8^+CD45RA^-CD27^+CD28^+$	4.82	−1.15	3.67	1.02	1.74
4	$CD8^+CD45RA^-CD27^-CD28^-$	5.75	−0.43	5.31	0.34	4.41
5	$CD8^+CD45RA^+CD27^-CD28^-$	6.00	−0.60	5.40	0.58	4.44
6	$CD3^+CD4^+$	4.43	−1.12	3.32	0.98	1.71
7	$CD4^+CD45RA^+CD27^+CD28^+$	5.53	−1.56	3.97	1.23	1.27
8	$CD4^+CD45RA^-CD27^+CD28^+$	5.66	−1.45	4.21	0.85	2.17
9	$CD4^+CD45RA^-CD27^-CD28^-$	6.81	−0.45	6.36	0.34	4.41
10	$CD4^+CD45RA^+CD27^-CD28^-$	5.38	−0.40	4.98	0.17	3.46
11	$CD45^+CD14^+$	5.65	−1.16	4.49	0.68	2.65
12	$CD3^-CD19^+$	8.00	−1.00	7.00	0.75	5.02
13	$CD3^+CD56^+$ NKT	5.90	−0.85	5.05	0.80	2.46
14	$CD3^-CD56^+$ NK	5.55	−1.02	4.53	0.69	2.62
15	$CD3^-CD56^{low}$	6.12	−1.06	5.06	0.81	2.67
16	$CD3^-CD56^{high}$	6.70	−0.97	5.72	0.63	3.18
17	$CD3^+CD56^+CD8^+CD4^+$	8.35	−0.35	8.00	0.29	7.28
18	$CD3^+CD56^+CD8^+CD4^-$	4.34	−0.87	3.48	0.97	1.80
19	$CD3^+CD56^+CD8^-CD4^+$	5.73	−0.52	5.21	0.32	3.64
20	$CD3^+CD56^+CD8^-CD4^-$	4.25	−0.97	3.28	1.14	1.48

TABLE 4 | Results of the Kruskal–Wallis test for testing for significant differences, at $\alpha < 0.05$, between the mean ranks of the normalized flow cytometry variables observed between patients with benign disease and patients with prostate cancer.

	Flow cytometry feature	Chi-Sq. χ^2	Asy. Sig. p value
1	$CD3^+CD8^+$	1.73	0.19
2	$CD8^+CD45RA^+CD27^+CD28^+$	0.82	0.37
3	$CD8^+CD45RA^-CD27^+CD28^+$	0.04	0.83
4	$CD8^+CD45RA^-CD27^-CD28^-$	0.06	0.81
5	$CD8^+CD45RA^+CD27^-CD28^-$	0.44	0.51
6	$CD3^+CD4^+$	3.72	0.05
7	$CD4^+CD45RA^+CD27^+CD28^+$	1.33	0.25
8	$CD4^+CD45RA^-CD27^+CD28^+$	1.79	0.18
9	$CD4^+CD45RA^-CD27^-CD28^-$	3.44	0.06
10	$CD4^+CD45RA^+CD27^-CD28^-$	0.88	0.35
11	$CD45^+CD14^+$	0.80	0.37
12	$CD3^-CD19^+$	0.74	0.39
13	$CD3^+CD56^+$ NKT	0.59	0.44
14	$CD3^-CD56^+$ NK	0.74	0.39
15	$CD3^-CD56^{low}$	0.96	0.33
16	$CD3^-CD56^{high}$	0.52	0.47
17	$CD3^+CD56^+CD8^+CD4^+$	0.61	0.44
18	$CD3^+CD56^+CD8^+CD4^-$	0.68	0.41
19	$CD3^+CD56^+CD8^-CD4^+$	2.85	0.09
20	$CD3^+CD56^+CD8^-CD4^-$	0.03	0.86

Revisiting the results which are presented in **Table 3**, features $CD8^+CD45RA^-CD27^-CD28^-$ (ID 4); $CD4^+CD45RA^-CD27^-CD28^-$ (ID 9); $CD4^+CD45RA^+CD27^-CD28^-$ (ID 10); $CD3^+CD56^+CD8^+CD4^+$ (ID 17) were among those flow cytometry features with the smallest IQR values (and, therefore, least variability in data) and which would potentially be good candidates for indicating the presence cancer. Furthermore, the Genetic Algorithm identified an additional flow cytometry feature as part of the selected features ($CD3^-CD19^+$ (ID 12)) which was not an obvious candidate during the initial statistical analysis. When feature ID12 is placed into a group with other features, it contributes to improving prediction performance. This reinforces the point

as to why it is important to examine combinations of features rather than individual features when choosing those which would make a cancer predictors (i.e., fingerprint). Importantly, not all flow cytometry features with a low IQR are needed to reach high predictive accuracy, and a subset containing the optimal combination of features was created using the Genetic Algorithm.

The heatmap in **Figure 4** shows that the correlation values between the five selected features range from +0.10 to +0.66, with six out of the ten pairs having a weak correlation value *rho* < 0.50 (ID 4, ID 10) = 0.43, (ID 4, ID 12) = 0.28, (ID 9, ID 12) = 0.18, (ID 10,ID 12) = 0.10, (ID 10, ID 17) = 0.47, (ID 12, ID 17) = 0.23 and the remaining four pairs having moderate correlation values (ID 4, ID 9) = 0.66, (ID 4, ID 17) = 0.57, (ID 9, ID 10) = 0.58, and (ID 9, ID 17) = 0.63, thereby suggesting that these five features are most suitable, since none of these pairs are highly correlated. Hence, we can conclude that the flow cytometry features: $CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells); $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Effector Memory Cells); $CD4^+CD45RA^+CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA); $CD3^-CD19^+$ (B cells); $CD3^+CD56^+CD8^+CD4^+$ (NKT cells) can be considered as an immunophenotyping profile which predicts the presence of prostate cancer in men with Prostate-Specific Antigen (PSA) levels below 20 ng ml^{-1}.

3.3. Prostate Cancer Prediction: Immunophenotyping versus Prostate-Specific Antigen (PSA) Data

The Prostate-Specific Antigen (PSA) test measures circulating levels of PSA and is currently considered to be the best method for identifying an increased risk of localized prostate cancer. However, elevated PSA levels do not necessarily indicate the presence of prostate cancer, and a normal PSA test does not necessarily exclude the presence of prostate cancer. PSA values tend to rise with age, and the total PSA levels (ng ml^{-1}) recommended by the Prostate Cancer Risk Management Programme are as follows
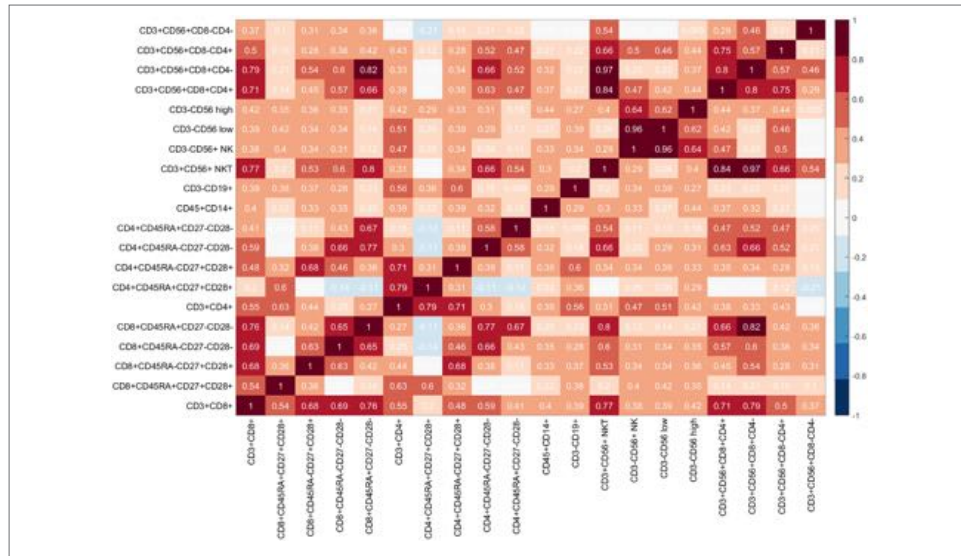
FIGURE 4 | Heatmap of flow cytometry features: Each cell of the heatmap provides a Spearman rho correlation value between two flow cytometry features.

TABLE 5 | FC-based prediction models using kNN classification and the selected flow cytometry features.

Prediction model name	Feature IDS	Accuracy (%)	AUC (%)	Optimal ROC Point (%)	
				FPR	TPR
FC-PM(Correlation(5))	4, 9, 10, 12, 17	83.33	83.40	16.13	82.93
FC-PM(Cosine(6))	4, 9, 10, 11, 12, 17	83.33	81.83	29.03	92.68
FC-PM(Chebychev(6))	4, 9, 10, 11, 12, 17	81.94	81.39	22.58	85.37
FC-PM(Minkowski(6))	4, 9, 10, 11, 12, 17	80.56	79.78	25.81	85.37
FC-PM(Euclidean(6))	4, 9, 10, 11, 12, 17	80.56	79.78	25.81	85.37
FC-PM(Seuclidean(6))	4, 9, 10, 11, 12, 17	80.56	79.78	25.81	85.37
FC-PM(Mahalanobis(6))	4, 9, 10, 11, 12, 17	77.78	76.55	32.26	85.37
FC-PM(Cityblock(7))	4, 9, 10, 11, 12, 16, 17	77.78	76.55	32.26	85.37
FC-PM(Spearman(8))	2, 4, 9, 10, 11, 12, 17, 19	83.33	70.89	38.71	80.49

The feature IDs map those presented in **Table 2**. The naming of the models includes the distance measure and number of features which were selected by the GA.

(22): 50–59 years, $PSA \geq 3.0$; 60–69 years, $PSA \geq 4.0$; and 70 and over, $PSA > 5.0$. According to a study by the European Study of Screening for Prostate Cancer, screening can significantly reduce death from prostate cancer by 29% (23–25). Herein, we compare the capacity of the proposed flow cytometry-based prostate cancer predictive model (FC-PM) and a predictive model based on PSA blood test results (PSA-PM) to discriminate between benign disease and prostate cancer. Since PSA values were already between 1 and 20, it was not necessary to apply z-score transformation. **Figure 6** shows the PSA values for individuals with benign disease and patients with cancer. A Kruskal–Wallis test sought significant differences between the mean rank PSA values of the benign disease and cancer groups. The test indicated that there were no significant differences in the mean rank PSA

values between the individuals with benign disease and patients with cancer, χ^2 (1, $N = 72$) = 0.03, $p = 0.955$.

PSA values were input into the kNN model and performance was evaluated using the LOOCV approach. Although experiments were performed with various distance measures, the Cityblock, Mahalanobis, Minkowski, Seuclidean, Euclidean and Chebychev returned exactly the same results, as shown in **Table 7**.

Figure 7 illustrates the AUCs and optimal ROC Points of PSA-PM and FC-PM (Correlation(5)). **Table 8** shows a comparison of AUC statistics using PSA-PM and FC-PM. The CI values shown in **Table 8** are higher for FC-PM(Correlation(5)) thereby meaning that the model is more capable of achieving higher prediction accuracies. Comparing the classification performances of FC-PM(Correlation(5)) (Accuracy = 83.33%) and the PSA-PM
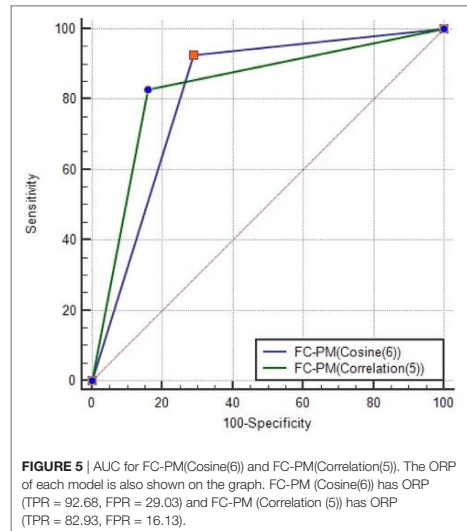
FIGURE 5 | AUC for FC-PM(Cosine(6)) and FC-PM(Correlation(5)). The ORP of each model is also shown on the graph. FC-PM (Cosine(6)) has ORP (TPR = 92.68, FPR = 29.03) and FC-PM (Correlation (5)) has ORP (TPR = 82.93, FPR = 16.13).

TABLE 6 | A comparison using FC-PM(Correlation(5)) and FC-PM(Cosine(6)).

	FC-PM(Correlation(5))	FC-PM(Cosine(6))
AUC%	83.40	81.83
AUC (SE)[a]	0.0514	0.0550
AUC 95% CI[b]	0.728–0.911	0.710–0.899
Benign (% of correctly classified)	83.87	70.97
Cancer (% of correctly classified)	82.93	92.68
Misclassified (%)	16.67	16.67

[a]Hanley and McNeil (20).
[b]Binomial exact.

(Accuracy = 77.78%), there is a 5.55% increase in accuracy when using the FC-PM. Furthermore, there is a 12.9% increase in False Positive Rate (FPR) when using PSA-PM, as opposed to when using the FC-PM(Correlation(5)). In conclusion, the FC-PM (Correlation (5)) which is based on immunophenotyping features provides a more accurate identification of prostate cancer than PSA-PM and is better able to discriminate between the presence of benign disease and cancer.

3.4. Does Adding the PSA Test Values to the Flow Cytometry Phenotyping Strengthen the Diagnostic Accuracy and Potential?

Given that current clinical practice uses the PSA test as an initial indicator of prostate cancer, we determined whether combining PSA test values with the selected flow cytometry predictors can strengthen diagnostic accuracy of the PSA test (the PSAFC

prediction model, PSAFC-PM). The PSA-PM was tuned using the Euclidean distance measure, whereas the PSAFC-PM was tuned with the Correlation distance measure. Although the PSA-PM performed exactly the same when tuned with distance measures other than Euclidean as shown in **Table 7**, the Euclidean distance measure was selected because it is the simplest to compute. Experiments with various distance measures revealed that PSAFC-PM achieved its highest predictive accuracy using the correlation distance measure. Results of the performance evaluation using the best models are presented in **Table 8** and illustrated in **Figure 7**. Comparing the predictive performance of the PSA-PM to the PSAFC-PM(Correlation(5)), an important observation is that the latter achieved 2.44% higher TPR than the PSA-PM, without increasing the FPR. Furthermore, the PSAFC-PM(Correlation(5)) returned an overall predictive accuracy of 79.17%, whereas the PSA-PM(Correlation(5)) returned 77.78% overall predictive accuracy, and thus an improvement of 1.39% when flow cytometry features were combined with PSA. It is useful to observe the impact of the predictors on the classification accuracy for each group of individuals, i.e., benign disease and cancer. **Table 8** holds these values and it also contains the values of FC-PM for comparison purposes. **Table 8** shows that the PSAFC-PM(Correlation(5)) performed better than the PSA-PM with regard to identifying benign disease (0.18% improvement), and it was also 2.44% more accurate at identifying cancer than the PSA-PM. In particular, the PSAFC-PM(Correlation(5)) achieved a 85.37% accuracy in detecting cancer, whereas the PSA-PM delivered 82.93% accuracy (a 2.44% difference).

The PSA-based prediction models, PSA-PM and PSAFC-PM, clearly suffer from higher FPRs than the FC-PM model, primarily because combining the PSA with the FC predictors inherits the disadvantage of PSA returning a high number of false positive cases. **Table 8** shows that combining PSA with flow cytometry predictors increases the confidence interval and reduces the Standard Error of the AUC (SE) of the prediction compared to using PSA predictors alone, meaning that fewer patients will be misdiagnosed when using the PSAFC-PM, as opposed to the PSA-PM model.

Herein, we propose a predictive model, PSAFC-PM, which improves the diagnostic capacity of the PSA test by combining PSA with flow cytometry features. A very important finding from the experiments is that if current clinical practice favors the continuation of the PSA test as an initial indicator of prostate cancer, then combining PSA predictor with a subset of flow cytometry predictors can increase the accuracy of the initial PSA test.

4. DISCUSSION

The results of this study demonstrate that the presence of prostate cancer in asymptomatic men with PSA levels <20 ng ml^{-1} can be better identified using immune cell profiles that have been generated using multiparametric flow cytometricanalysis of the peripheral blood. Prediction models were implemented using an advanced computational data extraction approach and a comprehensive statistical analysis. The computational approach comprised a metaheuristic optimization method, namely the Genetic Algorithm, which identified significant relationships between
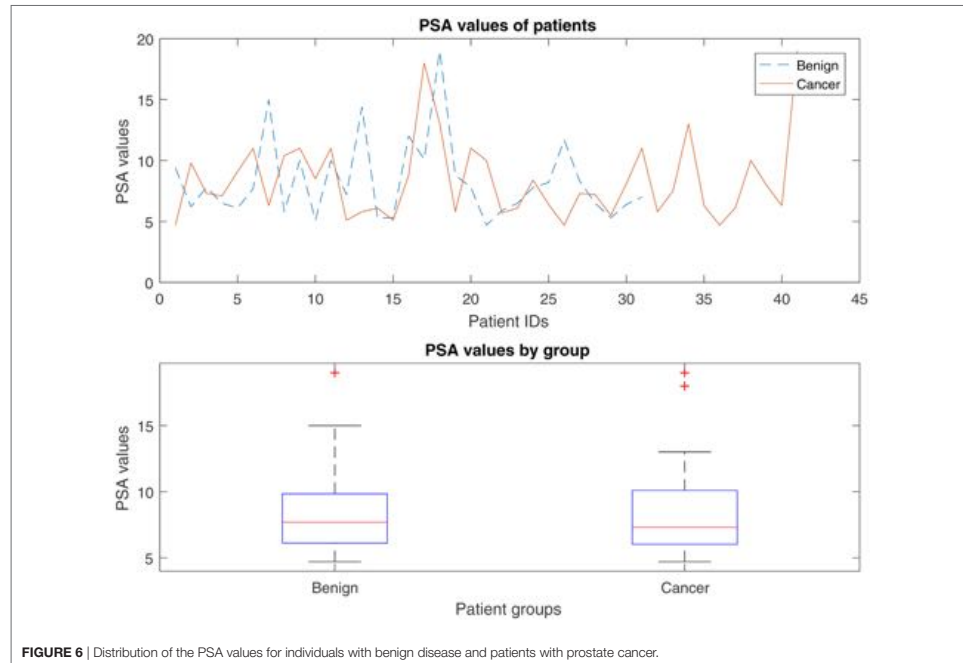
FIGURE 6 | Distribution of the PSA values for individuals with benign disease and patients with prostate cancer.

TABLE 7 | Prediction using PSA data as input into the kNN classification model.

Prediction model	Accuracy (%)	AUC (%)	Optimal ROC Point (%)	
			FPR	TPR
PSA-PM(Cityblock)	77.78	76.95	29.03	82.93
PSA-PM(Mahalanobis)	77.78	76.95	29.03	82.93
PSA-PM(Minkowski)	77.78	76.95	29.03	82.93
PSA-PM(Seuclidean)	77.78	76.95	29.03	82.93
PSA-PM(Euclidean)	77.78	76.95	29.03	82.93
PSA-PM(Chebychev)	77.78	76.95	29.03	82.93
PSA-PM(Correlation)	56.94	50.00	100.00	100.00
PSA-PM(Cosine)	43.06	50.00	100.00	100.00
PSA-PM(Spearman)	56.94	50.00	100.00	100.00

leukocyte population profiles and the presence of benign disease (no prostate cancer) or prostate cancer. A subset of five flow cytometry features was selected ($CD8^+CD45RA^-CD27^-CD28^-$; $CD4^+CD45RA^-CD27^-CD28^-$; $CD4^+CD45RA^+CD27^-CD28^-$; $CD3^-CD19^+$; $CD3^+CD56^+CD8^+CD4^+$) from a set of 20 features, which could potentially discriminate between the presence of benign disease and prostate cancer. A prostate cancer prediction model was constructed using the selected features and the k-Nearest Neighbor classification algorithm. The proposed model, which takes as input the abovementioned five flow cytometry features, outperformed the predictive model which takes

PSA values as input. In particular, the flow cytometry-based model achieved Accuracy = 83.33%, AUC = 83.40%, and optimal ROC points of FPR = 16.13%, TPR = 82.93%, whereas the PSA-based model achieved Accuracy = 77.78%, AUC = 76.95%, and optimal ROC points of FPR = 29.03%, TPR = 82.93%. Combining PSA and flow cytometry-based parameters as predictors achieved Accuracy = 79.17%, AUC = 78.17%, and optimal ROC points of FPR = 29.03% TPR = 85.37%.

Since current clinical practice favors the use of the PSA test as an initial indicator of prostate cancer, complementing the PSA prediction model with a subset of flow cytometry predictions can increase the accuracy of the initial prostate cancer test and reduce the misclassified patient cases. The proposed prediction model has the potential to improve outcomes of prostate cancer patients. Future studies will undertake further evaluations using the identified set of cancer predictors, and explore the use of deep learning algorithms for the analysis and interpretation of high dimensional flow cytometry data.

5. METHODS

The prediction model was developed using a selected subset of flow cytometry features and the k-Nearest Neighbor (kNN) classification algorithm. The Genetic Algorithm proposed by Ludwig and Nunes (18) was utilized for the feature selection stage, and
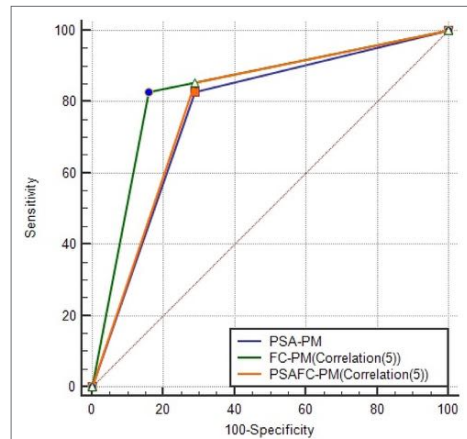
FIGURE 7 | AUCs and optimal ROC points of PSA-PM, FC-PM(Correlation(5)) and PSAFC-PM (Correlation(5)). This figure illustrates the differences among the models in predictive performance. FC-PM (Correlation(5)) was the best model in reducing the false positives. PSA-PM has ORP (TPR = 82.93, FPR = 29.03), FC-PM (Correlation (5)) has ORP (TPR = 82.93, FPR = 16.13), and PSAFC-PM (Correlation (5)) has ORP (TPR = 85.37, FPR = 29.03).

TABLE 8 | A comparison using PSA-PM, FC-PM, and PSAFC-PM.

	PSA-PM	FC-PM(Correlation(5))	PSAFC-PM
AUC%	76.95	83.40	78.17
AUC (SE[a])	0.0590	0.0514	0.0581
AUC 95% CI[b]	0.655–0.861	0.728–0.911	0.669 to 0.870
ORP TPR (%)	82.93	82.93	85.37
ORP FPR (%)	29.03	16.13	29.03
Accuracy (%)	77.78	83.33	79.17
Benign Accuracy (%)	70.79	83.87	70.97
Cancer Accuracy (%)	82.93	82.93	85.37
Misclassified (%)	22.22	16.67	20.83

[a]Hanley and McNeil (20).

[b]Binomial exact.

this algorithm returned the best combination of flow cytometry features (i.e., predictors) for discriminating between patients with benign disease and patients with cancer. These predictors were then input into the kNN classification algorithm. The kNN classifier is used to predict the disease status of an individual using new and previously unseen patient records. Feature selection is important because it enables only the best subset of features (i.e., predictors) to be selected for the prediction task and, thus, removes the "noisy" features that are not useful in identifying cancer.

The Genetic Algorithm is a powerful metaheuristic optimization method which aims to find optimal solutions to NP-hard optimization problems (26)—these are problems which require searching a space for the best solution (27). Let X be a m x n matrix with m rows and n columns, where m is the total number of patient records and n is the total number of flow cytometry

features. Each patient record, x_i, is represented by an n-dimensional feature vector, and it is given a corresponding known class label y_i, which has a value of either benign disease or cancer. The known labels were derived because of the highly accurate TPTP biopsy. The Genetic Algorithm is designed to take as input the m x n matrix X, and a m x 1 vector Y, where each element y_i contains the target output of each patient record. The Genetic Algorithm returns a set of indices of size λ containing the selected features. Importantly, the λ number of features returned are the best combination of features for discriminating the two groups of individuals (i.e., benign disease or cancer). It was important to use a Genetic Algorithm for the flow cytometry feature selection task for three main reasons:

- There were no significant differences between the mean flow cytometry values of the benign disease and cancer groups (**Table 4**), as a consequence of which a more sophisticated approach for identifying the best predictor features was needed.
- Searching for the best number of features is a combinatorial optimization problem, such that

$$\frac{n!}{2(n-\lambda)!},\qquad(2)$$

where n is the total number of flow cytometry features and λ is the desired number of features. Given that the value of λ is not known beforehand, experiments are needed with the number of features starting from $\lambda = 2, \ldots, 20$. The total possible number of combinations is 104,855,5 making this a computational expensive task, which is also impossible to be completed by basic statistical approaches. The Genetic Algorithm proposed by Ludwig and Nunes (18) was adapted and applied to extract the best set of flow cytometry features.

- When choosing the best subset of features for a predictive modeling task, it is important to take into consideration the interaction between features and the efficiency of these, as a group, for predicting an outcome (i.e., whether a patient belongs to the benign disease or cancer class), as opposed to choosing the best subset of features based on an analysis of each feature alone.

5.1. The k-Nearest Neighbor (kNN) Classification Algorithm

The subset of features returned by the Genetic Algorithm was input into the kNN classifier, and this was then used to construct a prediction model based on the particular subset of features. Nearest-neighbor classifiers are based on learning analogy, meaning that by comparing a given test case with training cases that are similar to the test cases. All training cases are represented as points in an n-dimensional space. The kNN classifier is a popular classification method, primarily due to its simplicity. It is a non-parametric approach and, hence, does not make any assumptions about the distribution of the data. When given an unknown case to classify, a kNN classifier searches the pattern space for the k training cases, i.e., "nearest neighbors" that are closest to the unknown case (i.e., the case that needs to be classified). Many distance measures exist,

including the Euclidean distance, the Minkowski distance, the Hamming distance, Pearson's correlation coefficient, and cosine similarity. The performance of the kNN classifier depends on the choice of k-nearest neighbors, and the distance measure d selected. The values selected for k and d depend on the dataset and the specification of the problem, and for this reason they are selected experimentally. Given a patient record (represented as a data point) x holding the flow cytometry values; a k number of neighbors; and a distance metric d, the kNN classifier first locates the k data points (i.e., k patient records) that are the closest to the data point x (i.e., patient record x) as the k-nearest neighbors to determine the target class of the data point. The proposed kNN approach uses the exhaustive search method, also known as the brute force method. The exhaustive search method finds the distance from each query point (i.e., a record to be classified), x, to every point in X, ranks them in ascending order, and returns the k points with the smallest distances. For the experiments reported in this paper, the kNN classifier can be tuned by selecting a distance measure d, and a k number of neighbors.

5.2. Performance Evaluation Measures

With regard to measuring performance, the aim was to adopt a variety of relevant evaluation metrics in order to get a more representative view of each classifier's performance. Let $|TP|$ be the total number of patients with cancer correctly classified as having cancer; $|TN|$ be total the number of benign patients correctly classified as benign; $|FP|$ be the total number of benign patients incorrectly classified as cancer patients; $|FN|$ be the total number of cancer patients incorrectly classified as benign; $|P|$ be the total number of cancer patients that exist in the dataset, where $|P| = |TP| + |FN|$; and $|N|$ be the total number of benign that exist in the dataset, where $|N| = |FP| + |TN|$. The following commonly used evaluation measures can be defined:

$$Accuracy = \frac{|TP| + |TN|}{|TP| + |FP| + |FN| + |TN|}, \in [0,1], \quad (3)$$

$$TPR = \frac{|TP|}{|TP| + |FN|}, \in [0,1], \quad (4)$$

$$TNR = \frac{|TN|}{|TN| + |FP|}, \in [0,1], \quad (5)$$

$$FNR = \frac{|FN|}{|TP| + |FN|} = 1 - Sensitivity, \in [0,1], \quad (6)$$

$$FPR = \frac{|FP|}{|FP| + |TN|} = 1 - Specificity, \in [0,1]. \quad (7)$$

The closer the values of Accuracy, True Positive Rate (i.e., TPR, Sensitivity) and True Negative Rate (i.e., TNR, Specificity) are to 1.0, then the better the classification performance of a system. The Receiver Operating Characteristic (ROC) is another important metric which can be used to evaluate the quality of a classifier's performance. The optimal operating point of the ROC curve is made up of the False Positive Rate (FPR) and True Positive Rate

(TPR) values. The optimal operating point for the ROC curve is computed by finding the slope, S, using function (8) and then identifying the optimal operating point by moving the straight line with slope S from the upper left corner of the ROC plot ($FPR = 0$, $TPR = 1$) down and to the right, until it intersects the ROC curve.

$$S = \frac{Cost(P \mid N) - Cost(N \mid N)}{Cost(N \mid P) - Cost(P \mid P)} \times \frac{N}{P}, \quad (8)$$

where $Cost(N|P)$ is the cost of misclassifying a positive class as a negative class; $Cost(P|N)$ is the cost of misclassifying a negative class, as a positive class; P and N are the total instance counts in the positive and negative class, respectively. The Area Under the ROC Curve (AUC) can be computed and reflects a system's performance at discriminating between the data obtained from individuals with benign disease and patients with cancer. The larger the AUC, the better the overall capacity of the classification system to correctly identify benign disease and cancer.

6. POTENTIAL IMPACT

It is essential that men with low-risk prostate abnormalities are not diagnosed as having prostate cancer, as even those with low-grade disease do not require active treatment, yet they become "labeled" as having cancer. This can have adverse psychological and financial consequences and assign these men to life-long surveillance. The strategies described herein have the potential to deliver new approaches for diagnosing asymptomatic men with an elevated PSA <20 ng l^{-1}. Inserting the data derived from the analysis of the peripheral blood from an individual into the algorithm will return a prediction about that individual. The algorithm could be retrained when more patient data are collected in order to learn patterns from a larger population, and it is possible that this will increase the accuracy of the approach. For example, re-training can occur every 50 new records. Such approaches will spare men with benign disease or low-risk cancer from unnecessary invasive diagnostic procedures such as TRUS guided prostate biopsies or TPTPB.

ETHICS STATEMENT

Research Protocols were registered and approved by the National Research Ethics Service (NRES) Committee East Midlands and by the Research and Development Department in the University Hospitals of Leicester NHS Trust. All participants were given information sheets explaining the nature of the study and all provided informed consent. All samples were collected by suitably qualified individuals using standard procedures. Ethical approval for the collection and use of samples from the TPTPB cohort (Project Title: Defining the role of Transperineal Template-guided prostate biopsy) was given by NRES Committee East Midlands-Derby 1 (NREC Reference number: 11/EM/3012; UHL11068). Ethical approval for the collection of peripheral blood from healthy volunteers

was obtained from the Nottingham Trent University College of Science and Technology Human Ethics Committee (Application numbers 165 and 412).

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

1. Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* (1987) 317(15):909–16. doi:10.1056/NEJM198710083171501

2. Welch HG, Fisher ES, Gottlieb DJ, Barry MJ. Detection of prostate cancer via biopsy in the Medicare seer population during the PSA era. *J Natl Cancer Inst* (2007) 99(18):1395. doi:10.1093/jnci/djm119

3. Norberg M, Egevad L, Holmberg L, Sparén P, Norlén BJ, Busch C. The sextant protocol for ultrasound-guided core biopsies of the prostate underestimates the presence of cancer. *Urology* (1997) 50(4):562–6. doi:10.1016/S0090-4295(97)00306-3

4. Naughton CK, Miller DC, Mager DE, Ornstein DK, Catalona WJ. A prospective randomized trial comparing 6 versus 12 prostate biopsy cores: impact on cancer detection. *J Urol* (2000) 164(2):388–92. doi:10.1097/00005392-200008000-00028

5. Aganovic D, Prcic A, Kulovac B, Hadziosmanovic O. Prostate cancer detection rate and the importance of premalignant lesion in rebiopsy. *Med Arh* (2011) 65(2):109–12.

6. Yuasa T, Tsuchiya N, Kumazawa T, Inoue T, Narita S, Saito M, et al. Characterization of prostate cancer detected at repeat biopsy. *BMC Urol* (2008) 8:14. doi:10.1186/1471-2490-8-14

7. Nafie S, Pal RP, Dormer JP, Khan MA. Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies. *World J Urol* (2014) 32(4):971–5. doi:10.1007/s00345-013-1225-x

8. Dimmen M, Vlatkovic L, Hole K-H, Nesland JM, Brennhovd B, Axcrona K. Transperineal prostate biopsy detects significant cancer in patients with elevated prostate-specific antigen (PSA) levels and previous negative transrectal biopsies. *BJU Int* (2012) 110(2b):E69–75. doi:10.1111/j.1464-410X.2011.10759.x

9. Takeshita H, Numao N, Kijima T, Yokoyama M, Ishioka J, Matsuoka Y, et al. Diagnostic performance of initial transperineal 14-core prostate biopsy to detect significant cancer. *Int Urol Nephrol* (2013) 45(3):645–52. doi:10.1007/s11255-013-0416-0

10. Emiliozzi P, Longhi S, Scarpone P, Pansadoro A, DePaula F, Pansadoro V. The value of a single biopsy with 12 transperineal cores for detecting prostate cancer in patients with elevated prostate specific antigen. *J Urol* (2001) 166(3):845–50. doi:10.1016/S0022-5347(05)65849-1

11. Djavan B, Ravery V, Zlotta A, Dobronski P, Dobrovits M, Fakhari M, et al. Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? *J Urol* (2001) 166(5):1679–83. doi:10.1016/S0022-5347(05)65652-2

12. Pal RP, Elmussareh M, Chanawani M, Khan MA. The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies. *BJU Int* (2012) 109(3):367–71. doi:10.1111/j.1464-410X.2011.10355.x

13. Nafie S, Mellon J, Dormer J, Khan M. The role of transperineal template prostate biopsies in prostate cancer diagnosis in biopsy naive men with PSA less than 20 ng ml⁻¹. *Prostate Cancer Prostatic Dis* (2014) 17(2):170–3. doi:10.1038/pcan.2014.4

14. Rajwa B, Wallace PK, Griffiths EA, Dundar M. Automated assessment of disease progression in acute myeloid leukemia by probabilistic analysis of flow cytometry data. *IEEE Trans Biomed Eng* (2017) 64(5):1089–98. doi:10.1109/TBME.2016.2590950

15. Pockley AG, Foulds GA, Oughton JA, Kerkvliet NI, Multhoff G. Immune cell phenotyping using flow cytometry. *Curr Protoc Toxicol* (2015) 66:18.8.1–34. doi:10.1002/0471140856.tx1808s66

16. Cosma G, Brown D, Archer M, Khan M, Pockley AG. A survey on computational intelligence approaches for predictive modeling in prostate cancer. *Exp Syst Appl* (2017) 70:1–19. doi:10.1016/j.eswa.2016.11.006

17. Lee KY, Choi Y, Lee K, Yun S, Choe G. Atypical small acinar proliferation of prostate: follow-up study of 114 patients. *Basic Appl Pathol* (2011) 4(4):116–9. doi:10.1111/j.1755-9294.2011.01115.x

18. Ludwig O, Nunes U. Novel maximum-margin training algorithms for supervised neural networks. *IEEE Trans Neural Netw* (2010) 21(6):972–84. doi:10.1109/TNN.2010.2046423

19. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence – Volume 2, IJCAI'95*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. (1995). p. 1137–43.

20. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* (1982) 143(1):29–36. doi:10.1148/radiology.143.1.7063747

21. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* (1983) 148(3):839–43. doi:10.1148/radiology.148.3.6878708

22. Burford DC, Kirby M, Austoker J. *Prostate Cancer Risk Management Programme Information for Primary Care; PSA Testing in Asymptomatic Men. Evidence Document*. NHS Cancer Screening Programmes (2010). Available from: http://www.cancerscreening.nhs.uk/prostate/pcrmp-guide-2.html

23. Schröder FH, Hugosson J, Carlsson S, Tammela T, Määttänen L, Auvinen A, et al. Screening for prostate cancer decreases the risk of developing metastatic disease: findings from the European randomized study of screening for prostate cancer (ERSPC). *Eur Urol* (2012) 62(5):745–52. doi:10.1016/j.eururo.2012.05.068

24. Luján M, Páez À, Berenguer A, Rodríguez J. Mortality due to prostate cancer in the Spanish arm of the European randomized study of screening for prostate cancer (ERSPC). Results after a 15-year follow-up. *Actas Urológicas Españolas* (2012) 36(7):403–9. doi:10.1016/j.acuroe.2011.10.004

25. Heijnsdijk E, Wever E, de Koning H. Cost-effectiveness of prostate cancer screening based on the European randomised study of screening prostate cancer. *J Urol* (2012) 187(Suppl 4):e491. doi:10.1016/j.juro.2012.02.1502

26. Mitchell M. *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press (1998).

27. Boussaïd I, Lepagnot J, Siarry P. A survey on optimization metaheuristics. *Info Sci* (2013) 237:82–117. doi:10.1016/j.ins.2013.02.041

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Identifying prostate cancer and its clinical risk in asymptomatic men using machine learning of high dimensional peripheral blood flow cytometric natural killer cell subset phenotyping data

Simon P Hood[1†‡], Georgina Cosma[2†*], Gemma A Foulds[1,3], Catherine Johnson[1,3], Stephen Reeder[1,3], Stéphanie E McArdle[1,3], Masood A Khan[4], A Graham Pockley[1,3†*]

[1]John van Geest Cancer Research Centre, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom; [2]Department of Computer Science, Loughborough University, Loughborough, United Kingdom; [3]Centre for Health, Ageing and Understanding Disease (CHAUD), School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom; [4]Department of Urology, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

*For correspondence: g.cosma@lboro.ac.uk (GC); graham.pockley@ntu.ac.uk (AGP)

[†]These authors contributed equally to this work

Present address: [‡]Cancer Research UK Manchester Institute, University of Manchester, Manchester, United Kingdom

Abstract We demonstrate that prostate cancer can be identified by flow cytometric profiling of blood immune cell subsets. Herein, we profiled natural killer (NK) cell subsets in the blood of 72 asymptomatic men with Prostate-Specific Antigen (PSA) levels < 20 ng ml^{-1}, of whom 31 had benign disease (no cancer) and 41 had prostate cancer. Statistical and computational methods identified a panel of eight phenotypic features ($CD56^{dim}CD16^{high}$, $CD56^{+}DNAM-1^{-}$, $CD56^{+}LAIR-1^{+}$, $CD56^{+}LAIR-1^{-}$, $CD56^{bright}CD8^{+}$, $CD56^{+}NKp30^{+}$, $CD56^{+}NKp30^{-}$, $CD56^{+}NKp46^{+}$) that, when incorporated into an Ensemble machine learning prediction model, distinguished between the presence of benign prostate disease and prostate cancer. The machine learning model was then adapted to predict the D'Amico Risk Classification using data from 54 patients with prostate cancer and was shown to accurately differentiate between the presence of low-/intermediate-risk disease and high-risk disease without the need for additional clinical data. This simple blood test has the potential to transform prostate cancer diagnostics.

Introduction

Early diagnosis and treatment increase curative rates for many cancers. The WHO considers that the burden of cancer on health services can be reduced by early detection and that this is achievable via three integrated steps: 1) awareness and accessing care, 2) clinical evaluation, diagnosis, and staging, 3) access to treatment (http://www.who.int/mediacentre/factsheets/fs297/en/). Although the clinical introduction of the Prostate-Specific Antigen (PSA) test in 1986 increased the early diagnosis of localized prostate cancer (*Catalona et al., 1991*; *Hankey et al., 1999*), elevated PSA levels are not necessarily indicative of prostate cancer because PSA levels can be raised by prostatitis, other localised infections, benign hyperplasia and/or factors such as physical stress. Contrastingly, 15% of men with 'normal' PSA levels typically have prostate cancer, with a further 15% of these cancers being high-grade (https://prostatecanceruk.org/prostate-information/prostate-tests/psa-test). The

eLife digest With an estimated 1.8 million new cases in 2018 alone, prostate cancer is the fourth most common cancer in the world. Catching the disease early increases the chances of survival, but this cancer remains difficult to detect.

The best diagnostic test currently available measures the blood level of a protein called the prostate-specific antigen (PSA for short). Heightened amounts of PSA may mean that the patient has cancer, but 15% of individuals with prostate cancer have normal levels of the protein, and many healthy people can have high amounts of PSA. This blood test is therefore not widely accepted as a reliable diagnostic tool.

Other methods exist to detect prostate cancer, yet their results are limited. A small piece of the prostate can be taken for analysis, but results from this invasive procedure are often incorrect. Scans can help to spot a tumor, but they are not accurate enough to be conclusive on their own. New tests are therefore urgently needed.

Prostate cancer is often associated with changes in the immune system that can be detected through a blood test. In particular, the appearance of a type of white blood (immune) cells called natural killer cells may be altered. Yet, it was unclear whether measurements based on these cells could help to detect prostate cancer and assess the severity of the disease.

Here, Hood, Cosma et al. collected and examined the natural killer cells of 72 participants with slightly elevated PSA levels and no other symptoms. Amongst these, 31 individuals had prostate cancer and 41 were healthy. These biological data were then used to produce computer models that could detect the presence of the disease, as well as assess its severity. The algorithms were developed using machine learning, where previous patient information is used to make prediction on new data. This work resulted in a new detection tool which was 12.5% more accurate than the PSA test in detecting prostate cancer; and in a detection tool that was 99% accurate in predicting the risk of the disease (in terms of clinical significance) in individuals with prostate cancer.

Although these new approaches first need to be validated in the clinic before being deployed, they could ultimately improve the detection and diagnosis of prostate cancer, saving lives and reducing the need for further tests.

reliable diagnosis of prostate cancer based on PSA levels alone is therefore not possible and confirmation using invasive biopsies is currently required. In 2011/12 approximately 32,000 diagnostic biopsies (28,000 TRUS and 4,000 TPTPB) were performed by the NHS in England (*NICE, 2014*). Although the transrectal ultrasound guided prostate (TRUS) biopsy is the most commonly used technique, it is limited to taking 10 to 12 biopsies primarily from the peripheral zone of the prostate and has a positive detection rate between 26% and 33% (*Aganovic et al., 2011*; *Nafie et al., 2014a*; *Naughton et al., 2000*; *Yuasa et al., 2008*). The Transperineal Template Prostate biopsy (TPTPB) is a 36 core technique that samples all regions of the prostate and delivers a better positive detection rate between 55% and 68% (*Dimmen et al., 2012*; *Nafie et al., 2014b*; *Pal et al., 2012*). However, invasive biopsies are painful and associated with a significant risk of potentially serious side-effects such as urosepsis and erectile dysfunction (*Chang et al., 2013*). Given the potential challenges of invasive tests and the risk of significant side-effects, considerable interest in the potential of non-invasive blood or urine-based tests/approaches ('liquid biopsies') for diagnosing disease has developed (*Quandt et al., 2017*). Liquid biopsies can provide information about both the tumour (e.g. circulating cells, cell-free and exosomal DNA and RNA) and the immune response (e.g. immune cell composition and their gene, protein, and exosome expression profiles). Liquid biopsies are minimally invasive and enable serial assessments and 'live' monitoring speedily and cost-effectively (*Quandt et al., 2017*).

Based on the reciprocal interaction between cancer and the immune system, we have proposed that immunological signatures within the peripheral blood (the peripheral blood 'immunome') can discriminate between men with benign prostate disease and those with prostate cancer and thereby reduce the dependency of diagnosis on invasive biopsies. To this end, we have previously shown that the incorporation of a peripheral blood immune phenotyping-based feature set comprising five phenotypic features $CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells),

$CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Effector Memory cells), $CD4^+CD45RA^+CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA), $CD3^-CD19^+$ (B cells), $CD3^+CD56^+CD8^+CD4^+$ (NKT cells) into a computation-based prediction tool enables the better detection of prostate cancer and strengthens the accuracy of the PSA test in asymptomatic men having PSA levels < 20 ng/ml (*Cosma et al., 2017*). Herein, we have extended this new approach to determine if phenotypic profiling of peripheral blood natural killer (NK) cell subsets can also discriminate between the presence of benign prostate disease and prostate cancer in the same cohort of asymptomatic men. We also investigate the potential of the peripheral blood dataset to discriminate between low- or intermediate-risk prostate cancer and high-risk prostate cancer in those men having prostate cancer.

Results

Distinguishing between benign prostate disease and prostate cancer: statistical analysis of NK cell phenotypic features and PSA levels

Herein, we consider a 'feature' to be a single phenotypic variable (as determined using flow cytometry) or a pre-grouped set of phenotypic variables, as shown in *Table 1*. It was not possible to discriminate between men with benign prostate disease and men with prostate cancer based on differences between phenotypic features/profiles due to their similarity (*Table 1*, *Figure 1*, *Figure 2*).

These findings highlight the difficulty in identifying combinations of features that can best identify the presence of cancer. These difficulties are compounded by the challenge of identifying the best combination of predictors which comprise n number of features, and that features within a combination, ideally, should not correlate. It is important to evaluate correlations between features, because if two features are highly correlated, then only one of these could serve as a candidate predictor. However, there may be occasions where both features are needed and besides the impact of this on the dimensionality of the dataset, there is no other negative impact. Furthermore, when two features are highly correlated and are important, it may be difficult to decide which feature to remove. *Figure 3* shows the correlations between features, where +1.0 indicates a strong positive correlation between two features, and −1.0 indicates a strong negative correlation between two features.

The Kolmogorov-Smirnov and Shapiro-Wilk tests of normality were carried out to determine whether the dataset is normally distributed, as this would determine the choice of statistical tests, that is whether to use parametric (for normally distributed datasets), or non-parametric (for not normally distributed datasets) tests. The results of the normality tests are shown in *Table 2*. The results revealed that only 7–8 features (depending on the normality test) were normally distributed (with $p > 0.05$), and for the remaining features the p value was less than 0.05 ($p < 0.05$) which indicates that there is a statistically significant difference between the distribution of the data of those features and the normal distribution. Based on the results of the test, we can conclude that the dataset is not normally distributed.

Given that most features in the dataset are not normally distributed, the Kruskal-Wallis (also called the 'one-way ANOVA on ranks', a rank-based non-parametric test) tests were used to check for statistically significant differences between the mean ranks of the NK cell phenotypic features in men with benign prostate disease and patients with prostate cancer rather than its parametric equivalent (one-way analysis of variance, ANOVA). Although the Kruskal-Wallis test did not return any significant differences in the mean PSA values between individuals with benign disease and those with prostate cancer ($\chi^2 = 0$; p=0.949, *Figure 4*), statistically significant differences at the alpha level of $\alpha = 0.05$ in the mean ranks of the $CD56^{bright}CD8^+$ (ID14, p=0.007), $CD56^+NKp30^+$ (ID15, p=0.008), $CD56^+NKp30^-$ (ID16, p=0.031), $CD56^+NKp46^+$ (ID17, p=0.023) populations in men with benign prostate disease and those with prostate cancer (*Table 3*) were observed.

This initial analysis provided insight into which phenotypic features might be good candidates for distinguishing between the presence of benign disease and prostate cancer. The next step was to examine whether using these as inputs into a machine learning algorithm can achieve this. An Ensemble Subspace kNN classifier was developed for the task at hand. The section which follows explains the approaches that were used to compare the diagnostic accuracy of the classifier when using the subset of features derived from the statistical analysis, and those features which were selected as a combination using the Genetic Algorithm (GA) for feature selection.

Table 1. Descriptive statistics of the dataset.

		Min.		Max.		Mean		Std.		IQR		Range		Diff.
		Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	
	PSA	4.70	4.70	19.00	19.00	8.26	8.34	3.31	3.28	3.30	4.08	14.30	14.30	−0.08
$CD56^{dim}$ %														
1	$CD16^+$	83.85	73.04	96.61	96.98	90.98	90.64	3.35	5.46	4.13	5.02	12.76	23.94	0.34
2	$CD16^{high}$	24.38	49.66	87.46	89.33	72.88	73.32	11.74	10.22	15.00	10.45	63.08	39.67	−0.44
3	$CD16^{low}$	5.17	6.57	64.22	44.00	17.74	16.84	10.40	7.45	8.76	7.66	59.05	37.43	0.90
4	$CD16^-$	1.41	1.25	11.11	18.06	4.83	4.89	2.45	3.48	2.58	2.68	9.70	16.81	−0.06
5	$CD56^{dim} total$	91.29	87.24	98.70	98.70	95.81	95.53	2.02	2.58	2.96	3.02	7.41	11.46	0.28
$CD56^{bright}$ %														
6	$CD16^+$	0.46	0.65	5.10	5.88	1.91	1.83	1.06	1.04	1.64	0.92	4.64	5.23	0.08
7	$CD16^{high}$	0.09	0.12	1.97	1.15	0.60	0.47	0.44	0.25	0.50	0.40	1.88	1.03	0.13
8	$CD16^{low}$	0.34	0.40	3.11	4.95	1.27	1.35	0.72	0.86	0.97	0.63	2.77	4.55	−0.07
9	$CD16^-$	0.61	0.58	5.78	9.09	2.28	2.64	1.14	1.82	1.42	1.75	5.17	8.51	−0.36
10	$CD56^{bright} total$	1.30	1.30	8.71	12.76	4.19	4.47	2.02	2.58	2.95	3.01	7.41	11.46	−0.28
CD8%														
11	$CD56^+CD8^+$	21.88	9.20	86.70	80.47	46.43	40.71	15.64	14.66	24.03	20.05	64.82	71.27	5.72
12	$CD56^+CD8^-$	13.30	19.53	78.12	90.80	53.57	59.29	15.64	14.66	24.03	20.05	64.82	71.27	−5.72
13	$CD56^{dim}CD8^+$	19.63	8.60	82.38	77.47	45.18	39.11	15.31	14.10	24.72	19.36	62.75	68.87	6.07
14	$CD56^{bright}CD8^+$	0.37	0.25	4.75	6.64	1.41	1.70	1.07	1.41	0.70	1.60	4.38	6.39	−0.29
NKp30 %														
15	$CD56^+NKp30^+$	40.69	56.80	96.74	98.43	79.78	88.56	16.42	10.41	21.80	10.44	56.05	41.63	−8.78
16	$CD56^+NKp30^-$	3.26	1.57	58.34	44.59	20.05	11.43	16.22	10.46	20.54	10.49	55.08	43.02	8.61
NKp46 %														
17	$CD56^+NKp46^+$	38.11	45.37	86.52	95.82	62.65	69.82	13.49	11.58	23.90	12.71	48.41	50.45	−7.18
18	$CD56^+NKp46^-$	14.02	4.32	62.97	55.68	38.40	30.87	13.58	11.64	24.89	13.44	48.95	51.36	7.53
DNAM-1 %														
19	$CD56^+DNAM-1^+$	63.69	88.56	99.18	99.60	95.35	96.46	6.81	2.59	3.37	3.49	35.49	11.04	−1.11
20	$CD56^+DNAM-1^-$	0.86	0.42	37.29	11.66	4.74	3.59	6.96	2.61	3.45	3.54	36.43	11.24	1.14
NKG2D %														
21	$CD56^+NKG2D^+$	85.17	80.79	98.77	98.96	93.49	94.07	4.45	4.87	6.81	3.83	13.60	18.17	−0.58
22	$CD56^+NKG2D^-$	1.22	1.03	14.76	19.12	6.44	5.84	4.36	4.76	6.80	3.96	13.54	18.09	0.60
	PSA	4.70	4.70	19.00	19.00	8.26	8.34	3.31	3.28	3.30	4.08	14.30	14.30	−0.08
NKp44 %														
23	$CD56^+NKp44^+$	0.43	0.28	3.71	6.77	1.16	1.34	0.82	1.20	0.78	1.25	3.28	6.49	−0.18
24	$CD56^+NKp44^-$	96.10	93.70	99.53	99.70	98.82	98.64	0.83	1.13	0.80	1.25	3.43	6.00	0.18
CD85j %														
25	$CD56^+CD85j^+$	19.53	14.21	84.73	91.59	53.37	55.10	19.04	18.34	30.49	20.23	65.20	77.38	−1.74
26	$CD56^+CD85j^-$	14.93	8.50	81.54	86.08	46.94	45.24	19.21	18.43	30.28	21.48	66.61	77.58	1.69
LAIR-1 %														
27	$CD56^+LAIR-1^+$	94.97	21.43	99.90	99.89	99.07	97.47	1.07	12.19	0.49	0.47	4.93	78.46	1.60
28	$CD56^+LAIR-1^-$	0.02	0.05	5.24	78.20	0.76	2.40	1.02	12.15	0.42	0.43	5.22	78.15	−1.65
NKG2A %														
29	$CD56^+NKG2A^+$	20.43	19.01	77.57	73.01	46.14	44.24	17.41	13.73	30.82	17.47	57.14	54.00	1.90
30	$CD56^+NKG2A^-$	22.62	27.11	79.40	80.85	54.01	55.99	17.39	13.67	30.48	17.90	56.78	53.74	−1.98

Table 1 continued on next page

Table 1 continued

		Min.		Max.		Mean		Std.		IQR		Range		Diff.
		Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	Beni.	Canc.	
2B4 %														
31	$CD56^+2B4^+$	98.41	97.06	99.99	99.96	99.53	99.50	0.39	0.59	0.32	0.33	1.58	2.90	0.02
32	$CD56^+2B4^-$	0.01	0.05	1.59	2.95	0.48	0.50	0.39	0.59	0.31	0.34	1.58	2.90	−0.02

Min. is the minimum value, Max. is maximum value, Mean is the mean or average value, and Std. is Standard Deviation. Range is the difference between the minimum and maximum values. The Interquartile range (IQR) is a measure of data variability and was derived by computing the distance between the Upper Quartile (i.e. top) and Lower Quartile (i.e. bottom) of the boxes illustrated in **Figure 1**. Difference is computed as diff = mean(Benign)-mean (Cancer).

Distinguishing between benign prostate disease and prostate cancer: GA

The GA was used to identify a subset of features that, as a combination, provide an NK cell-based immunophenotypic 'fingerprint' which can determine if an asymptomatic individual with PSA levels below 20 ng ml^{-1} has benign prostate disease or prostate cancer. This fingerprint, or feature set, would then be used to construct a diagnostic/prediction model. Given that GAs stochastically select multiple individuals (i.e. features) from the current population (based on their 'fitness'), each run can return different results. A common approach to identifying the best solution(s) is, therefore, to run the algorithm several times to obtain the frequency of the solution(s). Since the aim herein is to identify the most commonly occurring subset of NK cell phenotypic predictors, the GA was applied to the dataset and the most frequent subset of features returned was considered as being the best and most promising.

Let f_c denote the number of times (frequency) a combination was returned during the n number of runs, then the relative frequency of a combination (R_{fc}) can be calculated using formula (**Equation 1**),

$$R_{fc} = \frac{f_c}{n} \tag{1}$$

Table 4 shows the most frequent feature combinations returned at the end of each of the 30 runs when setting λ to different values. In **Table 4**, λ is the number of features in a combination. *No. different comb* is the number of unique combinations returned during the n number of runs (i.e. n = 30) for a given λ; *Comb. with highest freq* is the combination which was returned most frequently during the n number of runs; *Freq of Comb.* is the frequency of the most common combination found in the previous column; *Relative Freq.* (%) is computed using formula (**Equation 1**) converted to a percentage.

As the optimum number of features is not known, the GA was run by setting $\lambda = 2, 3, \ldots, n$ where n is the total number of features in the dataset. **Table 4** shows the results for the first 10 combinations. The results indicate that the combination comprising four features is the most promising in terms of its ability to discriminate between benign prostate disease and prostate cancer on NK cell phenotypic data alone. Features 2, 20, 27, 28, were returned in all 30 runs when searching for the best combination comprising of four features. Furthermore, features 20, 27, 28 were returned together in all combinations comprising more than three features (see feature ID's in combinations $\lambda = 4$ to $\lambda = 10$ in **Table 4**). These results strongly suggest that these are good predictors when grouped. The fact that the same combination was returned in 30 iterations is a strong indicator that these four features are the most reliable for distinguishing between the presence of benign prostate disease and prostate cancer. Although the statistical analysis presented in **Table 3** determined that features: ID14: $CD56^{bright}CD8^+$, ID15: $CD56^+NKp30^+$, ID16: $CD56^+NKp30^-$, and ID17: $CD56^+NKp46^+$ were the only ones with values which were significantly different in the two groups at $\alpha = 0.05$, and for which p values were therefore less than 0.05, none of the features selected by the statistical analysis were returned by the GA when searching for the best combination of features for discriminating between the presence of benign prostate disease and prostate cancer. The features selected by the GA were: ID2: $CD56^{dim}CD16^{high}$, ID20: $CD56^+DNAM - 1^-$, ID27: $CD56^+LAIR - 1^+$, and
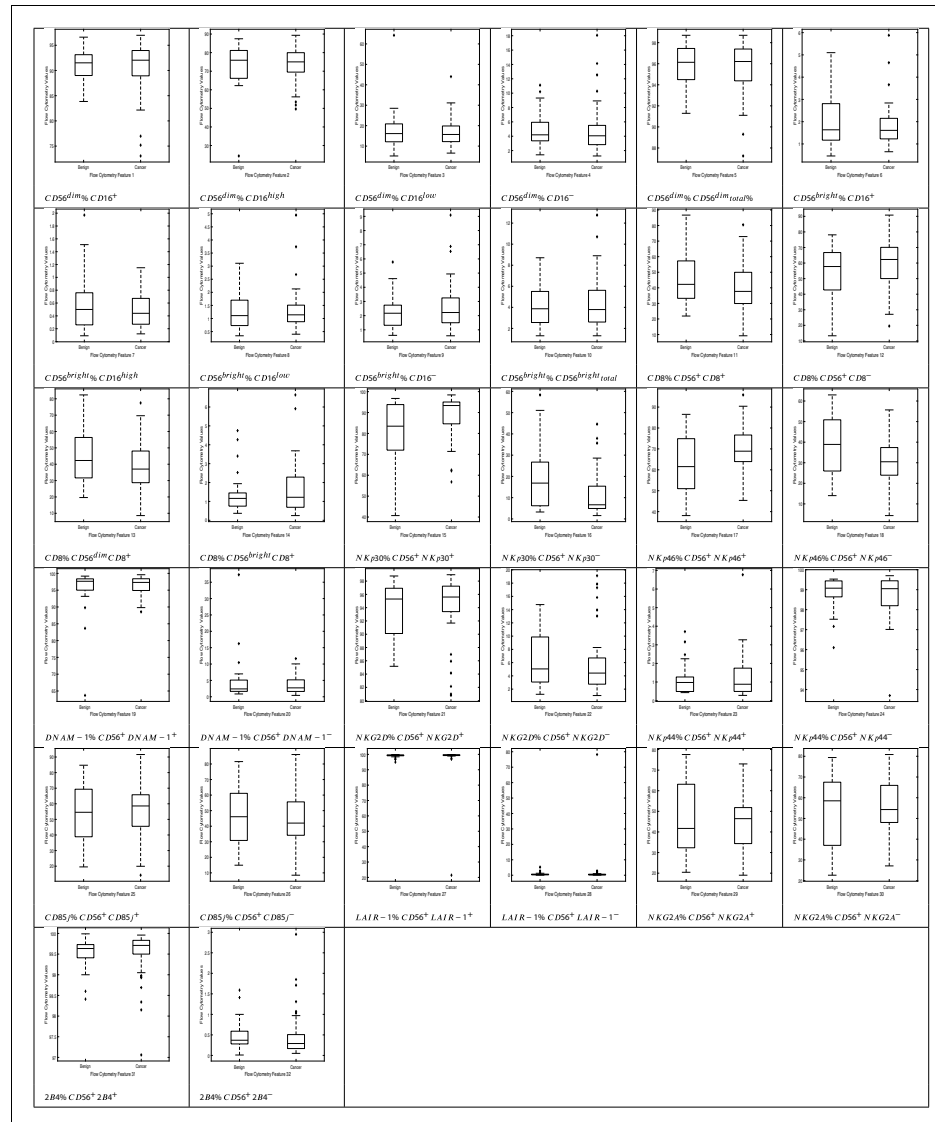
Figure 1. NK cell phenotypic features in men with benign prostate disease and patients with prostate cancer. Boxplots represent the flow cytometry values of each feature for patients with benign disease and with prostate cancer.
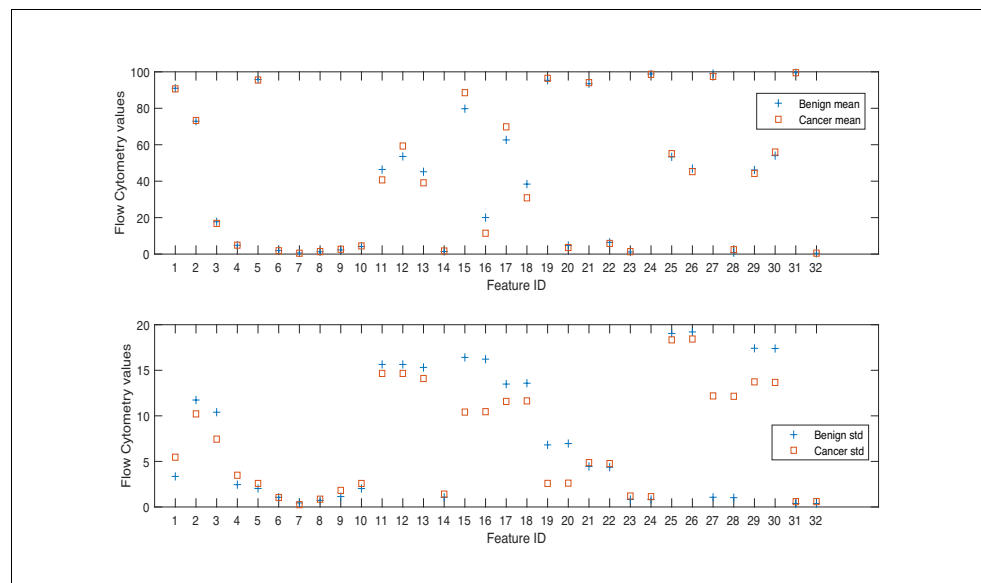
Figure 2. Mean and standard deviation values of flow cytometry features.

ID28: $CD56^+LAIR - 1^-$. Referring back to *Figure 3* and the correlation values between the selected features 2, 20, 27, 28, 14, 15, 16, 17, it is shown that these features do not have a strong positive correlation. There is a strong negative correlation between features 27 and 28, but we decided to keep both features since these were selected by the feature selection method.

The next step in the analysis involves evaluating the predictive performance of the feature subsets returned by the statistical test and by the GA. The features identified from the statistical and GA approaches were input into the proposed Ensemble Subspace kNN classifier to determine whether it can learn these features and discriminate between the presence of benign prostate disease and prostate cancer. For transparency of the machine learning model, it was important to keep the predictor selection and machine learning processes separate. The feature selection algorithm identified a set of novel NK cell phenotypic features for diagnosing the presence of prostate cancer which will be used to construct a transparent prediction tool.

Distinguishing between benign prostate disease and prostate cancer: machine learning

This section describes the outcome of experiments that were performed to determine the predictive performance of various feature subsets using the Ensemble Subspace kNN model, which was designed for the task. Machine learning classifiers that are constructed using small training sets have a large variance which means that the estimate of the target function will change if different training data are used (*Skurichina and Duin, 2002*). It is therefore expected, and normal, that classifiers will exhibit some variance. This means that small changes in input variable values can result in very different classification rules. To ensure that the proposed approach does not suffer from low variance, we evaluated the performance of the classifier using the 10-fold cross-validation approach which was repeated 30 times, for which the average and standard deviation of each run were recorded. Multiple runs of 10-fold cross-validation are performed using different partitions (i.e. folds), and the validation results are averaged over the runs to estimate a final predictive model. Each run of the cross-
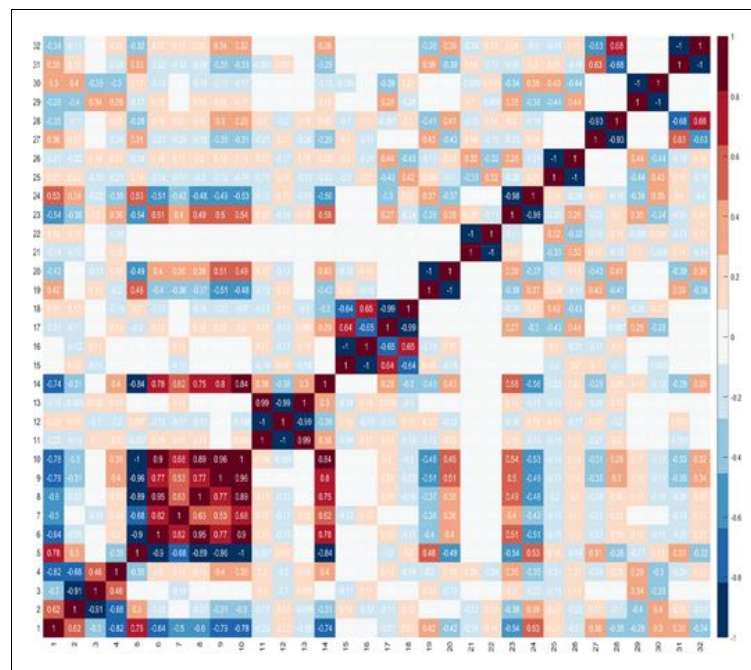
Figure 3. Correlations between features.

validation involves randomly partitioning a sample of data into complementary subsets, for which one subset is used as the training set, and the other is used as the validation subset. Cross validation randomly partitions the dataset into training and validation sets to limit overfitting problems, and to provide an insight into how the model will generalise to an independent dataset which was not previously seen by the model. A random seed generator was used to generate a different sequence of values each time the k-fold was run, and this was reseeded using a seed that was created using the current time. It is normal that a classifier returns a different validation accuracy in each fold and run, since it is training and validating on different samples. The aim is to create a low variance classifier, meaning that the results of each validation test are close together. The closer the results of each validation test, the more robust the classifier. To evaluate the predictive performance of various feature subsets derived from the computational and statistical feature selection approaches, each of these feature subsets was input into an Ensemble Subspace kNN classifier. Applying 10-fold validation resulted in 10 different partitions of the dataset of approximately 64 randomly selected samples for training and 7 randomly selected samples for validation in each partition (1 dataset comprising 63 training cases and 8 validation cases; and 9 datasets comprising 64 validation cases and 7 validation cases). All samples went through validation at some point during the evaluations. We consider 10-fold cross validation to be suitable given the small size of the dataset and the fact that sufficient samples are needed during the training process.

Table 5 shows the results of the comparison when running the 10-fold validation 30 times using six sets of features: 1) the four features selected by the GA; 2) the four features which were returned by the Kruskal-Wallis statistical test (STAT); 3) combined features selected by the GA and the statistical test (GA+STAT); 4) PSA values combined with features selected by the GA and the statistical test

Table 2. Tests of normality results.

Tests of normality

		NK cell values	Kolmogorov-Smirnova			Shapiro-Wilk		
			Statistic	df	Sig.	Statistic	df	Sig.
1	$CD56^{dim}$	$CD16^+$	0.15	71.00	0.00	0.85	71.00	0.00
2	$CD56^{dim}$	$CD16^{high}$	0.11	71.00	0.03	0.89	71.00	0.00
3	$CD56^{dim}$	$CD16^{low}$	0.17	71.00	0.00	0.79	71.00	0.00
4	$CD56^{dim}$	$CD16^-$	0.19	71.00	0.00	0.82	71.00	0.00
5	$CD56^{dim}$	$CD56^{dim}total\%$	0.15	71.00	0.00	0.91	71.00	0.00
6	$CD56^{bright}$	$CD16^+$	0.13	71.00	0.00	0.88	71.00	0.00
7	$CD56^{bright}$	$CD16^{high}$	0.15	71.00	0.00	0.87	71.00	0.00
8	$CD56^{bright}$	$CD16^{low}$	0.14	71.00	0.00	0.85	71.00	0.00
9	$CD56^{bright}$	$CD16^-$	0.16	71.00	0.00	0.86	71.00	0.00
10	$CD56^{bright}$	$CD56^{bright}total$	0.15	71.00	0.00	0.91	71.00	0.00
11	$CD8$	$CD56^+CD8^+$	0.10	71.00	0.06	0.98	71.00	0.17
12	$CD8$	$CD56^+CD8^-$	0.10	71.00	0.06	0.98	71.00	0.17
13	$CD8$	$CD56^{dim}CD8^+$	0.09	71.00	0.20*	0.98	71.00	0.24
14	$CD8$	$CD56^{bright}CD8^+$	0.19	71.00	0.00	0.82	71.00	0.00
15	$NKp30$	$CD56^+NKp30^+$	0.21	71.00	0.00	0.81	71.00	0.00
16	$NKp30$	$CD56^+NKp30^-$	0.21	71.00	0.00	0.81	71.00	0.00
17	$NKp46$	$CD56^+NKp46^+$	0.08	71.00	0.20*	0.98	71.00	0.52
18	$NKp46$	$CD56^+NKp46^-$	0.07	71.00	0.20*	0.99	71.00	0.57
19	$DNAM-1$	$CD56^+DNAM-1^+$	0.23	71.00	0.00	0.56	71.00	0.00
20	$DNAM-1$	$CD56^+DNAM-1^-$	0.23	71.00	0.00	0.55	71.00	0.00
21	$NKG2D$	$CD56^+NKG2D^+$	0.19	71.00	0.00	0.84	71.00	0.00
22	$NKG2D$	$CD56^+NKG2D^-$	0.18	71.00	0.00	0.85	71.00	0.00
23	$NKp44$	$CD56^+NKp44^+$	0.18	71.00	0.00	0.76	71.00	0.00
24	$NKp44$	$CD56^+NKp44^-$	0.17	71.00	0.00	0.78	71.00	0.00
25	$CD85j$	$CD56^+CD85j^+$	0.11	71.00	0.05	0.96	71.00	0.02
26	$CD85j$	$CD56^+CD85j^-$	0.10	71.00	0.07	0.96	71.00	0.02
27	$LAIR-1$	$CD56^+LAIR-1^+$	0.43	71.00	0.00	0.14	71.00	0.00
28	$LAIR-1$	$CD56^+LAIR-1^-$	0.43	71.00	0.00	0.14	71.00	0.00
29	$NKG2A$	$CD56^+NKG2A^+$	0.09	71.00	0.20*	0.97	71.00	0.11
30	$NKG2A$	$CD56^+NKG2A^-$	0.08	71.00	0.20*	0.97	71.00	0.10
31	$2B4$	$CD56^+2B4^+$	0.23	71.00	0.00	0.75	71.00	0.00
32	$2B4$	$CD56^+2B4^-$	0.23	71.00	0.00	0.75	71.00	0.00

*. This is a lower bound of the true significance.

Those values in bold are of those features whose data is normally distributed.

If the $p > 0.05$, we can accept the null hypothesis, that there is no statistically significant difference between the data and the normal distribution, hence we can presume that the data of those features are normally distributed.

If the $p < 0.05$, we can reject the null hypothesis because there is a statistically significant difference between the data and the normal distribution, hence we can presume that the data of those features are not normally distributed.

(PSA+GA+STAT); 5) PSA values alone as a predictor (PSA); and 6) using all 32 features (All features). The averages of the Area Under the Curve (AUC), Optimal ROC Point (ORP) False Positive Rate (FPR) of the AUC, ORP True Positive Rate (TPR) of the AUC, and Accuracy (ACC) of each fold are provided. The last column of *Table 5* shows the Rank of each model, where 1 is the best model and
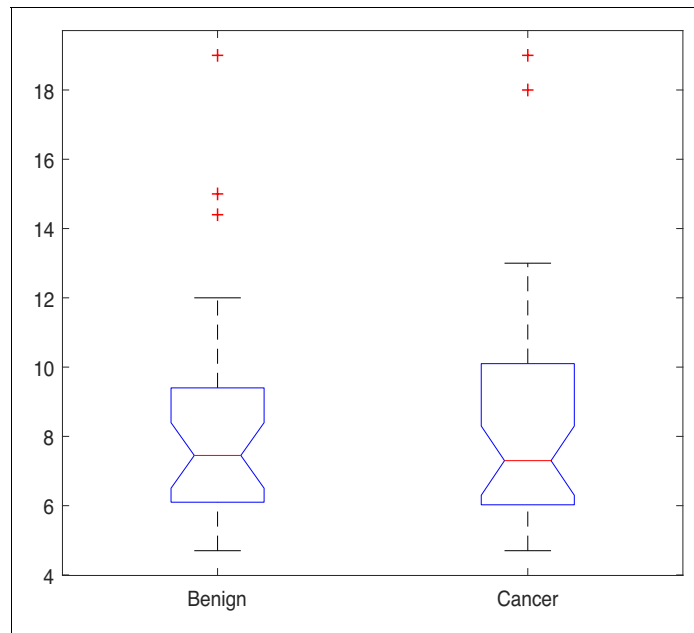
Figure 4. PSA values by group.

6 is the worst. The results of each k-fold were averaged, and these average values are plotted in the box plot shown in *Figure 5*. As shown in *Table 5*, combining the features selected by the GA ID2: $CD56^{dim}CD16^{high}$, ID20: $CD56^+DNAM-1^-$, ID27: $CD56^+LAIR-1^+$, ID28: $CD56^+LAIR-1^-$; with the four features which were returned by the Kruskal-Wallis statistical test as features with values which were statistically significant between individuals with benign prostate disease and patients with prostate cancer, ID14: $CD56^{bright}CD8^+$, ID15: $CD56^+NKp30^+$, ID16: $CD56^+NKp30^-$, ID17: $CD56^+NKp46^+$ yielded the highest classification accuracy, with AUC = 0.818, ORP FPR = 0.201, ORP TPR = 0.836 and Accuracy = 0.821. PSA values input into the classifier resulted in weak classification performance, AUC = 0.698, ORP FPR = 0.217, ORP TPR = 0.609, and Accuracy = 0.692. Although PSA is used as a screening test in clinical practice for identifying prostate cancer in men, it is the weakest of all the predictors. Importantly, predictive accuracy improved when PSA is combined with GA+STAT flow cytometry features (PSA+GA+STAT): AUC = 0.812, ORP FPR = 0.208, ORP TPR = 0.832, and ACC = 0.815. Combining PSA with the NK cell phenotypic fingerprint increased accuracy by +0.123 points when compared to using PSA alone.

The closer the standard deviation value is to 0 the less spread out are the results across the 30 runs, and hence the classifier variability is low (see *Table 5*). This results in a low variance classifier. A low standard deviation indicates that the data points tend to be close to the mean (also called the expected value) of the set, whereas a high standard deviation indicates that the data points are spread out over a wider range of values. Observing the data shown in *Table 5* and *Figure 5* for each evaluation measure (i.e. AUC, ORP TPR, ORP FPR, Accuracy (ACC)), the aim is to have a high AUC and low Std.; low ORP FPR and low Std.; high ORP TPR and low Std.; and high Accuracy and low Std. The results show that the classifier achieved the best performance when using the GA+STAT input and the results using k-fold across the 30 runs returned the lowest mean standard deviation

Table 3. Results of the Kruskal-Wallis test.

			Chi-Sq.(χ^2)	Asy. sig. p value
	PSA		0	0.949
	NK cells			
1	$CD56^{dim}$	$CD16^+$	0.001	0.981
2	$CD56^{dim}$	$CD16^{high}$	0.069	0.793
3	$CD56^{dim}$	$CD16^{low}$	0.555	0.456
4	$CD56^{dim}$	$CD16^-$	0.033	0.857
5	$CD56^{dim}$	$CD56^{dim}total\%$	0.063	0.802
6	$CD56^{bright}$	$CD16^+$	0.836	0.361
7	$CD56^{bright}$	$CD16^{high}$	0.201	0.654
8	$CD56^{bright}$	$CD16^{low}$	0.106	0.744
9	$CD56^{bright}$	$CD16^-$	0.030	0.861
10	$CD56^{bright}$	$CD56^{bright}total$	2.415	0.120
11	$CD8$	$CD56^+CD8^+$	2.415	0.120
12	$CD8$	$CD56^+CD8^-$	2.849	0.091
13	$CD8$	$CD56^{dim}CD8^+$	0.417	0.518
14	$CD8$	$CD56^{bright}CD8^+$	7.230	**0.007**
15	$NKp30$	$CD56^+NKp30^+$	7.106	**0.008**
16	$NKp30$	$CD56^+NKp30^-$	4.638	**0.031**
17	$NKp46$	$CD56^+NKp46^+$	5.179	**0.023**
18	$NKp46$	$CD56^+NKp46^-$	0.001	0.981
19	$DNAM-1$	$CD56^+DNAM-1^+$	0.001	0.972
20	$DNAM-1$	$CD56^+DNAM-1^-$	0.293	0.588
21	$NKG2D$	$CD56^+NKG2D^+$	0.325	0.568
22	$NKG2D$	$CD56^+NKG2D^-$	0.033	0.857
23	$NKp44$	$CD56^+NKp44^+$	0.072	0.789
24	$NKp44$	$CD56^+NKp44^-$	0.049	0.825
25	$CD85j$	$CD56^+CD85j^+$	0.072	0.789
26	$CD85j$	$CD56^+CD85j^-$	2.135	0.144
27	$LAIR-1$	$CD56^+LAIR-1^+$	1.343	0.247
28	$LAIR-1$	$CD56^+LAIR-1^-$	0.060	0.807
29	$NKG2A$	$CD56^+NKG2A^+$	0.072	0.789
30	$NKG2A$	$CD56^+NKG2A^-$	0.879	0.348
31	$2B4$	$CD56^+2B4^+$	0.890	0.346
32	$2B4$	$CD56^+2B4^-$	0.890	0.346

and hence the least variability in the results. The results reveal that using the GA+STAT predictors delivers a more reliable classification model with regards to training and validation on new data which will be generated in the future using the prediction model.

Importance of findings

The GA+STAT prediction model achieved the best performance, in that the ORP FPR was the lowest, and the AUC, ORP TPR, and Accuracy (ACC) were the highest compared to the other prediction models. The experimental results are promising and the proposed prediction model is expected to achieve even higher classification accuracy in identifying the presence of prostate cancer in asymptomatic individuals with PSA levels < 20 ng ml^{-1} based on peripheral blood NK cell phenotypic

Table 4. Results of the Genetic Algorithm when searching for the best subset of features.

λ	No. different comb	Comb. with highest freq.	Freq. of comb.	Relative freq. (%)
2	3	17,28	16	53.3
3	2	17,27,29	23	76.7
4	1	2,20,27,28	30	100.0
5	2	3,20,27,28,32	29	96.7
6	2	3,7,20,27,28,32	26	86.7
7	3	3,7,20,23,27,28,32	24	80.0
8	4	3,7,20,22,23,27,28,32	19	63.3
9	3	3,7,19,20,22,23,27,28,32	24	80.0
10	3	2,3,7,19,20,22,23,27,28,32	21	70.0

profiles as more data become available in the future. *Table 5* shows the performance of the classifier when using various feature subsets. When using the GA+STAT features, the AUC is higher, and FPR is lower (this is an important distinction) than when using all features or the other alternative feature subsets. The most important aspect is that better performance was achieved using a much smaller set of biomarkers (features), which indicates that we have identified a fingerprint for detecting the presence of prostate cancer in asymptomatic men with PSA levels < 20 ng ml^{-1} which is indeed significant from a clinical perspective. Feature selection is important, as the fundamental aim of this project is to develop a subset of phenotypic biomarkers that is smaller than the original set of biomarkers (i.e. 32 biomarkers in total) which can confidently identify the presence of prostate cancer. Ultimately, the approach will be embedded into a software application to be used by clinicians, and the aim is to create an interface that requires the clinician to input a few values (features), that is 8 instead of 32. Importantly, identifying a small subset of 8 features which is needed for detecting the presence of prostate cancer, results in the construction of an explainable disease detection and categorization model. Working with a small set of the most promising biomarkers provides a better understanding of the disease and allows cancer immunobiologists and clinicians to focus on performing further laboratory evaluations using the specific subset of biomarkers, in a more cost effective and less time-consuming manner.

Table 5. Naming of the models includes the feature selection method (GA) combined with the proposed Ensemble Subspace kNN classifier.
Validation results are presented at k = 10 fold cross validation.

		AUC	ORP FPR	ORP TPR	ACC	Mean std.	Rank
		Results of 10-fold cross validation over 30 runs					
GA	Mean	0.776	0.296	0.833	0.781		4
	Std.	0.024	0.065	0.026	0.023	0.035	
STAT	Mean	0.769	0.303	0.828	0.774		5
	Std.	0.022	0.057	0.023	0.021	0.031	
GA+STAT	Mean	0.818	0.201	0.836	0.821		1
	Std.	0.021	0.027	0.021	0.020	0.022	
PSA+GA+STAT	Mean	0.812	0.208	0.832	0.815		2
	Std.	0.020	0.031	0.018	0.019	0.022	
PSA	Mean	0.698	0.217	0.609	0.692		6
	Std.	0.022	0.025	0.043	0.020	0.028	
All features	Mean	0.812	0.213	0.836	0.815		3
	Std.	0.022	0.035	0.021	0.021	0.025	
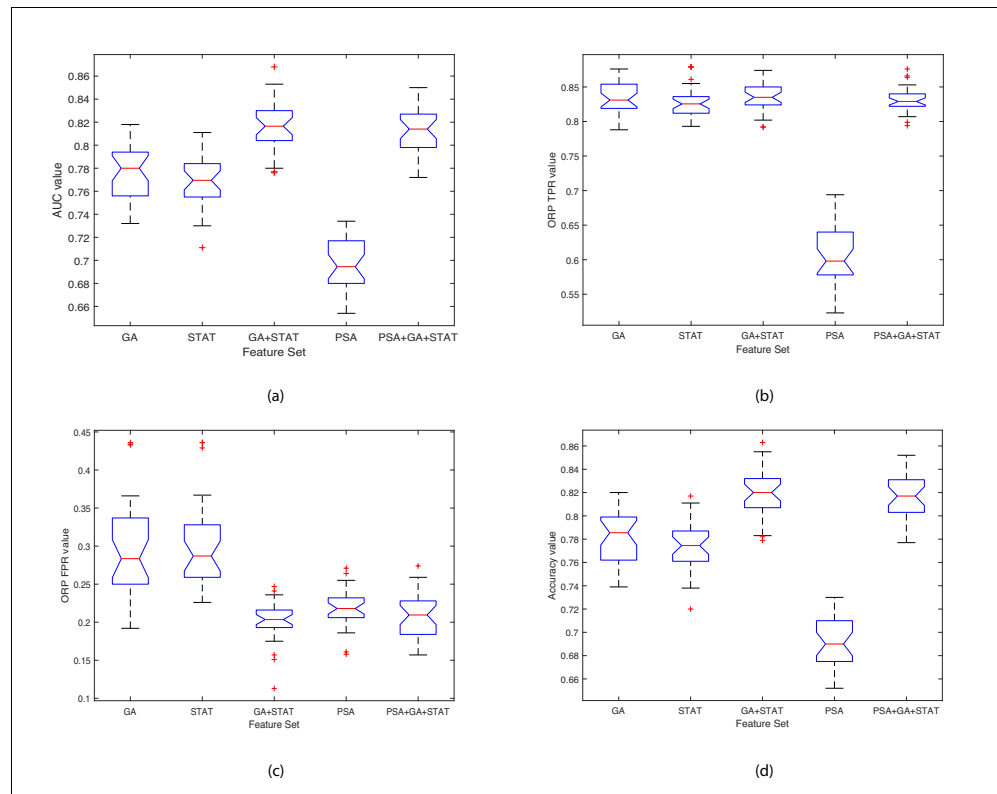
Figure 5. Boxplots illustrating the performance of the proposed model using various feature sets. (**a**) Average AUC values, (**b**) Average Optimal ROC points (TPRs), (**c**) Average Optimal ROC points (FPRs), (**d**) Average Accuracy values. Each box plot contains 30 points, where each point is the average performance evaluation value (i.e. AUC, ORP TPR, ORP FPR, Accuracy) from one 10-fold run using the various feature sets.

Comparing the performance of the proposed ensemble subspace kNN classifier with alternative classifiers

The experiments discussed thus far utilised a machine learning model comprised of an Ensemble of kNN learners (see Section 'Proposed Ensemble Learning Classifier for the task of Predicting Prostate Cancer'). We then undertook experiments to determine the impact of using the proposed Ensemble method over conventional machine learning classifiers: simple kNN; Support Vector Machine; and Naive Bayes models. The last column of *Table 6* shows the difference in the performance of the methods. The proposed method, denoted as EkNN, returned better performance than all other alternative classifiers. EkNN also returned the lowest Standard Deviation values and these are an indicator of a more stable and reliable model since the average values are clustered closely around the mean. SVM-linear returned the highest ORP TPR; however, the higher ORP FPR, higher Std. values, the low AUC, and low Accuracy values suggest that this model is worse than the proposed EkNN. Naive Bayes was the least efficient classifier, and although it returned the lowest ORP FPR, it

Table 7. Ad hoc test results.

	Ad hoc test					
	Group 1	Group 2	LI 95%	Diff. betw.means	UI 95%	p
1	GA	STAT	−12.658	1.317	15.292	1.000
2	GA	GA+STAT	−22.208	−8.233	5.742	0.525
3	GA	PSA	−4.992	8.983	22.958	0.344
4	GA	PSA+GA+STAT	−20.792	−6.817	7.158	1.000
5	STAT	GA+STAT	−23.525	−9.550	4.425	0.245
6	STAT	PSA	−6.308	7.667	21.642	0.710
7	STAT	PSA+GA+STAT	−22.108	−8.133	5.842	0.555
8	GA+STAT	PSA	3.242	17.217	31.192	0.001
9	GA+STAT	PSA+GA+STAT	−12.558	1.417	15.392	1.000
10	PSA	PSA+GA+STAT	−29.775	−15.800	−1.825	0.002

The first two columns show the groups that are compared. The third and fifth columns show the lower and upper limits for 95% confidence intervals for the true mean difference. The fourth column shows the difference between the estimated group means. The sixth column contains the p-value for testing a hypothesis that the corresponding mean difference is equal to zero.

values. There were statistically significant differences between group 8 (GA+STAT vs. GA) and 10 (PSA vs. PSA+GA+STAT) (p=0.001). We can conclude that GA+STAT returned a significantly higher AUC than PSA, and the difference between their mean ranks is diff = 17.217. PSA returned a significantly lower AUC than PSA+GA+STAT (p=0.002), and the difference between their mean ranks is diff=-15.800.

Comparing the best prediction models over 30 runs

With regard to constructing a model which has the potential to be used in clinical practice, it is necessary to finalise an initial prediction model, since the last experiment returned 30 different variations of each prediction model when using different training and validation data partitions. Those experiments were crucial in determining whether the prediction models (five models, a different one for each feature subset) suffer from low variance. We then observed the classification performance of each model for each run, to identify the highest performance achieved using a single 10-fold cross validation in any of the runs. This provides a way of comparing the performance of each prediction model as it would be used in the clinical setting. *Table 8* provides the results of the highest performing model, and the performance of the models is ranked (with 1 being the best model and 5 the worst model).

Predicting low-/intermediate risk cancer vs. high-risk cancer

The continuing, significant clinical challenge resides in distinguishing men with low- or intermediate-risk prostate cancer which is unlikely to progress (for both of which 'active surveillance' is the most appropriate approach), from men with intermediate disease which is likely to progress and men with

Table 8. Results of the best prediction models created during the 30 runs.
Validation results are presented at k = 10 fold cross validation.

	Best prediction model results				
	AUC	ORP FPR	ORP TPR	Accuracy	Rank
GA	0.818	0.192	0.829	0.820	3
GA+STAT	0.853	0.157	0.862	0.855	1
PSA	0.734	0.218	0.685	0.730	5
PSA+GA+STAT	0.844	0.175	0.864	0.848	2
STAT	0.811	0.227	0.85	0.817	4

Table 6. Comparing the performance of the proposed Ensemble Subspace kNN model against conventional machine learning models when using the GA+STAT feature set.
Results of 10-fold cross validation over 30 runs.

Proposed ensemble subspace kNN (EkNN) model
(No. of learners (NL): 30; Subspace Dimension (SD): 16)

Parameters		AUC	ORP FPR	ORP TPR	ACC	
NL: 30, SD:16	Mean	0.818	0.201	0.836	0.821	
	Std.	0.021	0.027	0.021	0.020	
Simple kNN model (Distance: Euclidean)						
k		AUC	ORP FPR	ORP TPR	ACC	Acc. Diff. (EkNN vs. kNN)
2	Mean	0.768	0.241	0.730	0.751	+0.070
	Std.	0.119	0.160	0.393	0.128	−0.108
5	Mean	0.778	0.300	0.833	0.783	+0.038
	Std.	0.107	0.265	0.103	0.103	−0.083
10	Mean	0.753	0.371	0.845	0.758	+0.063
	Std.	0.137	0.350	0.120	0.131	−0.111
Support Vector Machine models						
Kernel		AUC	ORP FPR	ORP TPR	ACC	Acc. Diff. (EkNN vs. SVM)
Linear	Mean	0.782	0.342	0.860	0.784	+0.037
	Std.	0.126	0.352	0.110	0.120	−0.100
Gaussian	Mean.	0.808	0.353	0.876	0.799	+0.022
	Std.	0.112	0.416	0.107	0.111	−0.091
Naïve Bayes model						
Predictor distributions		AUC	ORP FPR	ORP TPR	ACC	Acc. Diff. (EkNN vs. Naïve Bayes)
Normal	Mean.	0.695	0.132	0.455	0.662	+0.159
	Std.	0.169	0.163	0.493	0.181	−0.161

also returned the lowest ORP TPR, lowest AUC and Accuracy values; and its Std. values were also higher than those of the EkNN model.

Statistically significant differences in predictive performance when using various feature subsets

The next step in the analysis is to determine whether statistically significant differences exist between the average AUC performance values of the classifier when using the various feature subsets, for which Friedman's two-way Analysis of Variance (ANOVA) test was used. It was also important to observe whether including the PSA test values significantly strengthens the diagnostic accuracy and capacity. The average k-fold values across the 30 runs for each feature set were computed. A matrix C was derived which holds the results of the classifier when using one of five feature subsets. Friedman's chi-square statistic compares the mean values of the columns of matrix C. The test returned a statistically significant difference in the AUC predictive performance depending on which type of feature subset was input into the classifier, $\chi^2(4) = 106.55$, $p = 3.968E - 22$. This suggests that the mean AUC ranks of at least one feature subset are significantly different than the others. The mean ranks were as follows: GA = 12.050, STAT = 10.733, GA+STAT = 20.283, PSA = 3.067, PSA+GA+STAT = 18.867. A post hoc test was run alongside the Friedman test to pinpoint which feature subsets differ from each other. Post hoc analysis using a Bonferroni correction was used to reduce the likelihood of erroneously declaring a statistically significant due to multiple comparisons (a Type I error). *Table 7* shows the results of multiple comparisons and adjusted p

high-risk prostate cancer (both of which require treatment). The diagnosis of men with low-risk or small volume intermediate-risk prostate cancer as having prostate cancer is unhelpful as these men will very rarely require treatment. The inappropriate assignment of men to potentially life-threatening invasive procedures and life-long surveillance for prostate cancer has significant psychological, quality of life, financial, and societal consequences. Furthermore, the definitive diagnosis of prostate cancer currently requires painful invasive biopsies with which is associated a risk of potentially life-threatening urosepsis in 5% of individuals. We, therefore, undertook experiments to train the proposed Ensemble Subspace kNN model to predict the D'Amico Risk Classification for those patients with prostate cancer (see subsection 'The cancer patients dataset used for building the risk prediction modelin Methods), in terms of Low/Intermediate (L/I) risk and High (H) risk disease using NK cell phenotypic data alone.

The Ensemble model was modified to take as input all 32 features (described in *Table 1*), and was trained to classify the disease in patients with prostate cancer as being L/I or H risk disease (see Figure 9 in Materials and methods). Hence, given a new patient record, which comprises of 32 inputs, the model predicts whether the patient is D'Amico L/I risk (not clinically significant) or H (clinically significant) risk. The flow charts in *Figure 6* illustrate the process to detect the presence and risk of prostate cancer and patient outcomes. Of those 54 patient records, a total of 10 randomly selected records (5 from the L/I group and 5 from the H group) were extracted from the dataset such that they can be used at the testing (mini clinical trial) stage. To ensure thorough experiments, a rigorous methodology was adopted. More specifically, a 10-fold cross validation method was adopted, and the experiments were run in 30 iterations, for which each iteration provided an average validation result across 10 folds. Each iteration consists of 10 different 'train and validation' data arrangements (hence 300 tests were carried out using a different mix of train and validation records). The 10 test records were input into each trained model (i.e. iteration) to predict their accuracy, and
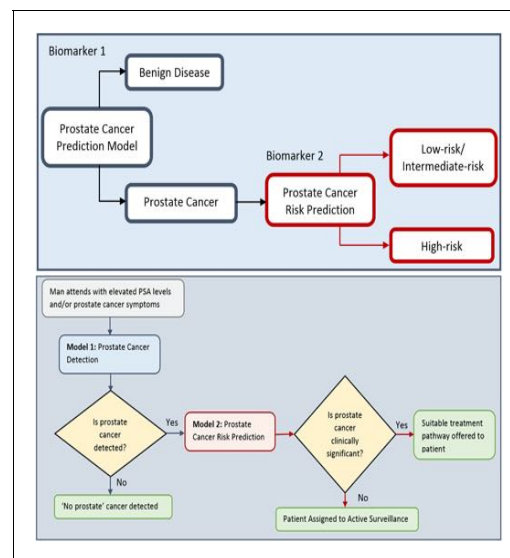
Figure 6. Flow charts illustrating the process to detect the presence and risk of prostate cancer and patient outcomes. Model 1: Distinguishes between men with benign prostate disease and prostate cancer; Model 2: predicts risk (in terms of clinical significance) in men identified as having prostate cancer in Stage 1. Note that Model 1 can detect prostate cancer in men with PSA < 20 ng ml^{-1}.

to evaluate the model when it is trained and validated using different variations of patient data. The model can highly accurately differentiate between L/I risk group and H risk group patients. The k-fold validation results across 30 iterations were AUC: 0.98(±0.03); FPR: 0.03(±0.05), TPR: 0.99 (±0.01), Accuracy: 0.99(±0.02); and results using the test set were AUC: 0.98(±0.03); FPR: 0.03 (±0.05), TPR: 0.99(±0.01), Accuracy: 0.97(±0.02). Accuracy has been near perfect in all iterations (i.e. using different train and validation data cases in each iteration). *Figure 7* illustrates the performance of the model obtained across the 30 runs during the k-fold cross validation and independent testing using the 10 patient samples. The results demonstrate that the proposed model predicts with near-perfect accuracy, the result of the D'Amico Risk Classification (L/I vs High) using NK cell phenotypic data alone, and without requiring the PSA, Gleason, and tumor stage data.

The dataset that was utilized to identify the biomarker (that comprised eight features) for detecting the presence of prostate cancer (i.e. benign prostate disease vs prostate cancer) in 71 men, and thus it was large enough to perform the combinatorial feature selection task for finding the best subset of features. The GA that was used for the combinatorial feature selection task is described in Section Computational Methods. Given that detecting the presence of prostate cancer and its risk if present are two different tasks, it is expected that the biomarkers for those tasks will be different since a different target is given to the GA (i.e. the target for the prostate cancer detection model comprises 0 (benign prostate disease) and 1 (prostate cancer) values; the target for the prostate cancer risk prediction model comprises 0 (L/I risk) and 1 (High risk) values). For the L/I vs H risk task, the dataset was small (n = 54 men (L/I = 36, H = 16)), and we could not perform the combinatorial feature selection task with confidence. Hence, it was decided to use the entire feature set for the risk prediction task. The results obtained from the risk prediction model were very promising as shown experimentally, and this provided the confidence to report these preliminary results. The combinatorial feature selection task to identify the best subset of features for the risk prediction task will be performed once a larger dataset is available.

Herein, we demonstrate that all 32 phenotypic features are required to distinguish between low/intermediate risk cancer (L/I) and high risk (H) cancer. However, we expect to be able to identify smaller subset(s) of these features as the datasets increase and the prediction model is retrained on the larger dataset. As indicated above, the generation and delivery of additional datasets is beyond the scope of this paper.

Discussion

The clinical challenge in prostate cancer diagnosis resides in distinguishing men with low- or small volume intermediate-risk prostate cancer which is unlikely to progress (both require 'active surveillance') from men with intermediate disease which is likely to progress or high-risk disease (both of
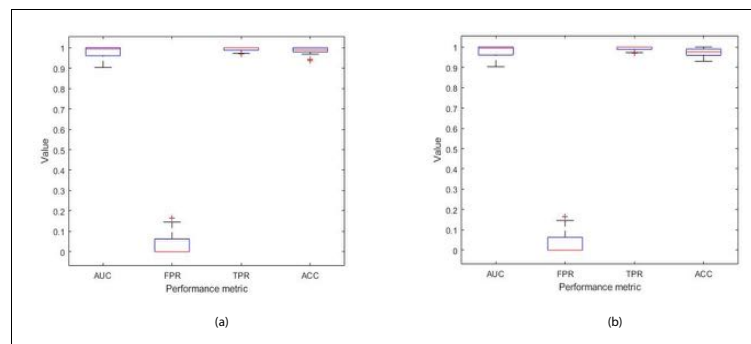
Figure 7. Each box plot contains 30 points, where each point is the average performance evaluation value (i.e. AUC, FPR, TPR, Accuracy (ACC)) from one 10-fold run during (**a**) k-fold validation results, and (**b**) independent testing results (i.e. using 10 patient records).

which require treatment). It is essential that men with low-risk prostate abnormalities are not diagnosed as having prostate cancer, as those with low-risk/grade disease do not require active treatment. Furthermore, unnecessarily labeling men as having prostate cancer can assign these men to life-long surveillance and have significant psychological, quality of life, financial and societal consequences. Recent findings from a decade-long study involving 415,000 British men (The Cluster Randomized Trial of PSA Testing for Prostate Cancer (CAP) Randomized Clinical Trial) have not supported single PSA testing for population-based screening and suggest that asymptomatic men should not be routinely tested to avoid unnecessary anxiety and treatment. It is therefore essential that new approaches for enabling more definitive, early detection of prostate cancer are developed. The reliable diagnosis of prostate cancer based on PSA levels alone is not possible and confirmation using invasive biopsies or other approaches such as MRI and biopsy are currently required. Although interest in the potential diagnostic capabilities of MRI scanning is developing, MRI cannot currently be used as a sole diagnostic as a positive MRI can be incorrect in approximately 25% of cases and a negative MRI can be incorrect in approximately 20% of cases *Ahmed et al., 2017*. Although the findings from the CAP study do not support using the PSA test as an approach for population-based screening, combining PSA measurements with other approaches that either identify individuals for additional testing or strengthen the capacity to diagnose prostate cancer have significant merit, and it is based on this concept that the current study has been performed. The studies presented herein have focused on asymptomatic men with a PSA < 20 ng/ml, as men with a PSA level > 20 ng/ml are more likely to harbour prostate cancer and are thereby less likely to pose a clinical diagnostic quandary. In contrast, men with a PSA < 20 ng/ml pose a major problem because although only 30–40% of these men will have prostate cancer, all currently undergo potentially unnecessary invasive prostate biopsies to determine who has the disease. It is, therefore, this group of men for which the development of new and more accurate approaches for the early detection of cancer is a clear unmet clinical need, and for whom the benefits of such an approach will be most relevant and significant.

Comparing results to the previous study

We have previously shown that incorporating peripheral blood immune phenotyping-based features into a computation-based prediction tool enables the better detection of prostate cancer and, furthermore, strengthens the accuracy of the PSA test in asymptomatic individuals having PSA levels < 20 ng/ml (*Cosma et al., 2017*). The phenotypic feature set which was shown to be discriminatory between benign disease and prostate cancer comprised $CD8^+CD45RA^-CD27^-CD28^-$ ($CD8^+$ Effector Memory cells), $CD4^+CD45RA^-CD27^-CD28^-$ ($CD4^+$ Effector Memory Cells), $CD4^+CD45RA^+CD27^-CD28^-$ ($CD4^+$ Terminally Differentiated Effector Memory Cells re-expressing CD45RA), $CD3^-CD19^+$ (B cells), $CD3^+CD56^+CD8^+CD4^+$ (NKT cells).

Using samples from the same cohort of asymptomatic individuals, herein we have further investigated the phenotype and function of NK cell subsets. Using a combination of statistical and computational feature selection approaches, we have identified a subset of eight phenotypic features $CD56^{dim}CD16^{high}$, $CD56^+DNAM-1^-$, $CD56^+LAIR-1^+$, $CD56^+LAIR-1^-$, $CD56^{bright}CD8^+$, $CD56^+NKp30^+$, $CD56^+NKp30^-$, $CD56^+NKp46^+$ which distinguish between the presence of benign prostate disease and prostate cancer. These features were used to implement a prediction model. The kNN machine learning approach developed in our previous study (*Cosma et al., 2017*) has been extended to an Ensemble of kNN learners to improve performance in identifying patterns in even more complex data. As was observed in our previous study, flow cytometry predictors significantly outperform the PSA test. The findings presented herein significantly reinforce our previous finding (*Cosma et al., 2017*) that complementing the PSA prediction model with a subset of flow cytometry-based phenotypic predictors can significantly increase the accuracy of the initial prostate cancer test and reduce misclassification. The performance of the prediction model which was built using the phenotypic 'signature' presented in our previous study $-CD8^+CD45RA^-CD27^-CD28^-$, $CD4^+CD45RA^-CD27^-CD28^-$, $CD4^+CD45RA^+CD27^-CD28^-$, $CD3^-CD19^+$, $CD3^+CD56^+CD8^+CD4^+$ (*Cosma et al., 2017*), is similar to the model which was built using the NK cell-based phenotypic signature presented herein, $CD56^{dim}CD16^{high}$, $CD56^+DNAM-1^-$, $CD56^+LAIR-1^+$, $CD56^+LAIR-1^-$, $CD56^{bright}CD8^+$, $CD56^+NKp30^+$, $CD56^+NKp30^-$, $CD56^+NKp46^+$. Specifically, the prediction model using the five flow cytometry features identified in *Cosma et al.,*

Cancer Biology | Computational and Systems Biology

2017 achieved Accuracy: 83.33% , AUC: 83.40%, ORP TPR: 82.93%, FPR: 16.13%, whereas the prediction model presented herein achieved AUC: 85.3%, ORP FPR: 15.7%, ORP TPR: 86.2%, Accuracy: 85.5%. Across the 30 runs the average performance of the prediction model presented herein is AUC: 81.8%, ORP TPR: 83.6%, FPR: 20.1%, Accuracy: 82.1%.

The difference in the performance of the model presented in the first study (*Cosma et al., 2017*) and the study described herein is a consequence of different data and prediction models being used in each study. Given that the phenotypic features that were used to create the prediction models were different, the studies resulted in different prediction models. In particular, the model presented previously (*Cosma et al., 2017*) was based on a kNN classifier, and herein the kNN classifier was extended to construct an Ensemble Subspace kNN method which comprised several kNN classifiers (see Figure 9). The dataset used herein was more complex, and it was therefore necessary to create a more complex classifier. At this point in the studies, it is not possible to determine which set of phenotypic features is better at identifying prostate cancer. However, it is evident that both approaches have significant promise. Since the publication of our previous study (*Cosma et al., 2017*), the model developed for that study was used to predict the outcomes of a further 20 new patients which were previously unseen by the prediction model. The model correctly identified the presence of prostate cancer in 19 of the 20 patients (data not shown).

Encouragingly, the prediction models generated in the study reported upon herein selected phenotypic features that are associated with the expression of activating receptors NKp30, NKp46, and DNAM-1 by NK cells. *Pasero et al., 2015* demonstrated that these activating receptors, in addition to NKG2D, are involved in the recognition of prostate cancer cell lines. Furthermore, they identified that the intensity of NKp30 and NKp46 expression on the surface of NK cells isolated from the peripheral blood of patients with metastatic prostate cancer was predictive of time to hormone (castration) resistance and overall survival. This suggests that our computational analysis is selecting phenotypic features that are of biological/clinical relevance. Thus far, our identification of disease predictive phenotypic immune features has been limited to effector immune populations (T, B, and NK cells). The responsiveness of these cells is known to be influenced by the presence of innate immune cell populations that can be polarized by the tumor toward an immunosuppressive state (*Vitale et al., 2014*; *Anderson et al., 2017*). Therefore, future studies will investigate the identification and inclusion of phenotypic features from innate immune subpopulations such as monocytes and neutrophils into prediction models to assess whether their inclusion enhances predictive capability and enables a better assessment of patient prognosis in line with the D'Amico Risk Classification.

The proposed machine learning model was adapted to predict the D'Amico Risk Classification of patients with prostate cancer using NK cell phenotypic data alone. Experiments with data from 54 patients revealed the significant potential of using the proposed machine learning model for determining if men with prostate cancer are in the low-/intermediate- or high-risk groups, without the need for additional clinical data (i.e. PSA, Gleason, clinical stage data). One limitation of the current study is that the small patient numbers required for low- and intermediate-risk patients to be grouped. Future work, for which additional sample collections are required, will train the model to separately predict low-, intermediate- and high-risk cancer. Future work involves collecting more patient samples to conduct further testing of the proposed machine learning models. In terms of future work from a computational perspective, once we have a larger patient dataset we plan to design deep learning models and compare their performance to the conventional machine learning model which was proposed in this paper.

Potential impact

Currently available screening methods and tests for prostate cancer lack accuracy and reliability, the consequence of which is that many men unnecessarily undergo invasive tests such as biopsy and/or are misdiagnosed as having the disease. Furthermore, a biopsy involves removing samples of tissue from the prostate and it is an extremely uncomfortable procedure which also puts men at risk of developing life-threatening infections. As biopsy results are not definitive, there is a significant potential for misdiagnosis and over- and under- treatment. It is therefore essential that new non-invasive approaches such as blood tests that are more accurate than the Prostate Specific Antigen (PSA) test are developed to reduce misdiagnosis and unnecessary procedures. Misdiagnosis unnecessarily subjects many men to lifelong monitoring for prostate cancer which can have undesirable psychological and quality of life side-effects, as well as place a significant financial burden on the

NHS and other healthcare systems. This paper proposes a computerised model, which detects the presence of prostate cancer in men by analyzing immune system cells in the blood. The model uses the data from the blood tests and artificial intelligence-based computing (machine learning) to more accurately detect the presence of prostate cancer. A preliminary model has also been presented to detect the clinical risk that any prostate cancer which is present poses. The tool has two elements, the first detects whether a man has prostate cancer. If prostate cancer is detected, the second element will detect the clinical risk of the disease (low, intermediate, high) and thereby enable the clinician to decide whether the patient requires no further investigation/treatment ('watch and wait') or whether further investigation and treatment are required.

To our knowledge, these are the first studies to employ computational modeling of peripheral blood NK cell phenotyping data for the early detection of cancer and its clinical significance. They also illustrate the potential for this approach to decipher clinically relevant immune features that can distinguish between benign prostate disease and prostate cancer in asymptomatic individuals for whom the management and treatment strategy is unclear. Of translational importance is that our prediction models are interpretable, can be explained to patients and clinicians and can be continually refined and improved as data are collected.

The novelty of this approach is that it interrogates the immunological response to the tumour, not the tumour itself and that it requires a simple blood test (liquid biopsy). Based on current practice, we expect that this approach could avoid up to 70% of prostate biopsies, thereby sparing men with benign prostate disease or low-risk prostate cancer from unnecessary invasive procedures with which are associated significant side-effects. Furthermore, more accurate diagnosis would reduce the demands of healthcare provision and resources associated with treatment and continual surveillance, thereby reducing costs and improving healthcare. We envisage that, in the future, men with a mildly elevated PSA will also undergo an immune status test and those with a suspicion for significant prostate cancer will then undergo an MRI. Although the current study focuses on prostate cancer, its fundamental principles and approaches are highly likely to be applicable across many, if not all, cancer entities.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Biological Sample	Hyclone fetal bovine serum (FBS)	GE Healthcare Life Sciences	SV30180.03	
Antibody	Monoclonal mouse IgG1 kappa anti human DNAM-1 (CD226) (clone 11A8); FITC	BioLegend	338304	5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human NKG2D (CD314) (clone 1D11); PE	eBioscience	12-5878-42	5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human CD56 (clone N901); ECD (PE-Texas Red)	Beckman Coulter	A82943	2.5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human CD16 (clone 3G8); PerCP-Cy5.5	BioLegend	302028	5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human NKp46 (CD335) (clone 9E2); PE-Cy7	BioLegend	331916	5 µl per tube / 10^6cells

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	Monoclonal mouse IgG1 kappa anti human NKp30 (CD337) (clone P30-15); Alexa Fluor 647	BioLegend	325212	5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human CD3 (clone UCHT1); Alexa Fluor 700	BioLegend	300424	2 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human CD19 (clone HIB19); Alexa Fluor 700	BioLegend	302226	1 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human CD8 (clone SK1); APC-Cy7	BioLegend	344714	2.5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG2b anti human CD85j (ILT2) (clone GHI/75); FITC	Miltenyi Biotec	130-098-437	10 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human LAIR-1 (CD305) (clone DX26); PE	BD Biosciences	550811	20 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG2b anti human NKG2A (CD159a) (clone Z199); PE-Cy7(PC7)	Beckman Coulter	B10246	20 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human NKp44 (CD336) (clone P44-8); Alexa Fluor 647	BioLegend	325112	5 µl per tube / 10^6cells
Antibody	Monoclonal mouse IgG1 kappa anti human 2B4 (CD244.2) (clone C1.7); FITC	BioLegend	329506	5 µl per tube / 10^6cells
Chemical Compound	LIVE/DEAD Fixable Violet Dead Stain	Thermo Fisher Scientific	L34955	1 µl in 1 µl
Chemical Compound	Novagen Benzonase Nuclease	Merck Millipore	70664	
Chemical Compound	CTL Wash Solution	Cellular Technology Limited	CTLW-010	
Chemical Compound	Trypan Blue viability stain	Santa Cruz	sc-216028	
Chemical Compound	Dimethyl sulfoxide (DMSO)	Santa Cruz	sc-202581	
Chemical Compound	Calbiochem bovine serum albumin (BSA)	Merck Millipore	2905-OP	
Chemical Compound	Sigma-Aldrich sodium azide	Merck Millipore	S8032	
Chemical Compound	Sigma-Aldrich lithium heparin	Merck Millipore	H0878	
Chemical Compound	Ficoll-Paque	GE Healthcare Life Sciences	17-1440-03	
Chemical Compound	Isoton II isotonic buffered saline solution	Beckman Coulter	844 80 11	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Chemical Compound	RPMI medium	Lonza	12-167Q	
Chemical Compound	Phosphate Buffered Saline (PBS)	Lonza	17-517Q	
Other	Leucosep tubes	Greiner Bio-One International	227290	
Software	Kaluza v1.3	Beckman Coulter		

Data collection

Peripheral blood samples were obtained from individuals suspected of having prostate cancer that attended the Urology Clinic at Leicester General Hospital (Leicester, UK) between 24th October 2012 and 15th August 2014. Only patients who had provided informed consent and met the criteria of being biopsy naive, a benign feeling Digital Rectal Examination (DRE) with a PSA level of < 20 ng ml^{-1} and agreeing to undergo a simultaneous 12 core TRUS biopsy and a 36 core transperineal template prostate biopsy (TPTPB) were included in the study. Further details regarding the TPTPB technique are provided in *Nafie et al., 2014b*. A total of 71 males (30 patients diagnosed with benign disease and 41 patients diagnosed with cancer, as confirmed by pathological examination of TPTPB biopsies) met the criteria. Of the 30 patients diagnosed with benign disease; 9 patients were diagnosed with High Grade Prostatic Intraepithelial Neoplasia (PIN), 10 patients were diagnosed with Atypia and 2 patients were diagnosed with Atypical Small Acinar Proliferation. The remainder were diagnosed with benign disease. Of the men diagnosed with prostate cancer, 16 had Gleason 6 disease, 23 had Gleason 7 disease and 2 had Gleason 9 disease on biopsy-based evidence. The clinical features of individuals with benign disease and patients with prostate cancer are provided in *Table 9*.

The cancer patients dataset used for building the risk prediction model

Data derived from the 41 individuals with prostate cancer were extracted from the dataset shown in *Table 9*. All 41 patients had PSA < 20 ng ml^{-1}. However, three of the 41 patients who had a High D'Amico risk were removed because their clinical profiles were very different from those of other high risk patients. They were patients with either a Gleason score 3+3 or had a benign biopsy. In the future, we aim to collect more data from such infrequent patient groups to train the algorithms on patients with such clinical profiles. The remaining 38 patients had PSA levels < 20 ng ml^{-1} and belonged to the D'Amico L/I risk group.

Data were collected from an additional 16 patients with prostate cancer who were diagnosed as having a D'Amico High risk profile (see *Table 10*). Thus, the new cancer patient dataset comprised 54 patients with prostate cancer, of which 38 patients belonged to the D'Amico L/I risk group and all had PSA<20 ng ml^{-1}, and 16 patients belong to the D'Amico H risk group and have PSA 4.3 ng

Table 9. Patient clinical features.

Patient group	Gleason score	Number of patients	Age range (years)	PSA range (ng/ml)
Benign	Benign	9	64-71	5.3–15
Benign	HGPIN	9	54–70	5.1–12
Benign	Atypia	10	50–76	4.7–19
Benign	ASAP	2	59–60	5.3–7.8
Cancer	Gleason 6	16	55–80	4.7–11
Cancer	Gleason 7	23	53–77	4.7–19
Cancer	Gleason 9	2	65–75	6.3–18

Table 10. Dataset used for differentiating between patients with L/I and H cancer.

Patient group	Count	%
L/I	38	70.37
H	16	29.63

ml^{-1} $\leq$ PSA $\leq$ 2617 ng ml^{-1}. The 16 patients were diagnosed with Gleason scores of: 4+4 = 8 (n = 2), 5+4 = 9 (n = 2), and 4+5 = 9 (n = 11), and 1 patient was diagnosed with small cell cancer. The combined dataset (i.e. 38+16 = 54) comprised 15 patients with Gleason 6 (3+3), 18 patients with Gleason 7 (3+4), 5 patients with Gleason 7 (4+3), 2 patients with Gleason 8 (4+4), 11 patients with Gleason 9 (4+5), 2 patients with Gleason 9 (5+4), and 1 patient with small cell cancer.

Since 11 of those 16 patients had a PSA > 20 ng ml^{-1}, their data could only be utilised for building the prostate cancer risk prediction model, as the detection model focuses on detecting prostate cancer in asymptomatic men with PSA< 20 ng ml^{-1}.

Flow cytometric analysis

Peripheral blood (60 ml) was collected from all patients using standard clinical procedures. Aliquots (30 ml) were transferred into two sterile 50 ml polypropylene (Falcon) tubes containing 300 μl of sterilized Sigma Aldrich Lithium Heparin (1000 U/ml, Merck Millipore). Anti-coagulated samples were transferred to the John van Geest Cancer Research Centre at Nottingham Trent University (Nottingham, UK) and processed immediately upon receipt (always within 3 hr of collection). Peripheral blood (60 ml) was mixed with Phosphate Buffered Saline (PBS, 30 ml, Lonza) and layered over Ficoll-Paque (GE Healthcare Life Sciences) in Leucosep tubes (20 ml blood per tube) and then centrifuged at 800 g for 20 min. The peripheral blood mononuclear cell (PBMC) fraction was harvested and washed twice with PBS before being re-suspended in Hyclone fetal bovine serum (FBS, GE Healthcare Life Sciences). Viable cells were counted using trypan blue (0.1 % v/v trypan blue, Santa Cruz) and a haemocytometer. Cells were frozen in 90% v/v FBS, 10% v/v DMSO (Santa Cruz) in aliquots of 10 $\times$ 10^6 PBMC/vial and stored in liquid nitrogen until phenotypic analysis. At the time of analysis, one vial from each patient was thawed by mixing with 10 ml 'thaw' solution (90% v/v RPMI (Lonza)), 10% v/v CTL wash solution (Cellular Technology Limited) and 10 μl of Novagen Benzonase (Merck Millipore) at room temperature.

PBMCs were centrifuged at 400 g for 5 min followed by resuspension in 1 ml of RPMI (supplemented with 10% v/v FBS, 1% v/v L-glutamine (Lonza)). Cells were rested for 1 hr at 37, after which viable cells were counted using trypan blue dye (Santa Cruz) exclusion. For each monoclonal antibody (mAb) panel shown in *Table 11*, 1 $\times$ 10^6 cells were washed and incubated in 100 μl of Wash Buffer (PBS +2% w/v Calbiochem bovine serum albumin (BSA, Merck Millipore) +0.02% w/v sodium azide (Sigma)) containing the relevant mAb cocktail for 15 min, after which cells were washed with 1 ml PBS and then incubated in 1 ml LIVE/DEAD Fixable Violet dead stain (Thermo Fisher Scientific) for 30 min. All incubations were performed at 4 protected from light. The cells were washed with PBS and then re-suspended in Beckman Coulter Isoton isotonic buffered saline solution.

Data (on viable cells) were acquired within 1 hr using a 10-color/3-laser Beckman Coulter Gallios flow cytometer and analyzed using Beckman Coulter Kaluza v1.3 data acquisition and analysis software. Controls used a Fluorescence Minus One (FMO) approach. A typical gating strategy for the analyses is presented in *Figure 8*.

Computational methods

Initially, the GA by *Ludwig and Nunes, 2010* was adopted to identify the best subset of features (i.e. predictors), and thereafter a prediction model was constructed using the Ensemble classifier. This section also explains the metrics adopted for evaluating the performance of the prediction model.

GA for selecting the best subset of features

The GA is a metaheuristic, commonly used to generate solutions to optimization and search problems. Given the large number of combinations, the process of selecting the best subset of flow
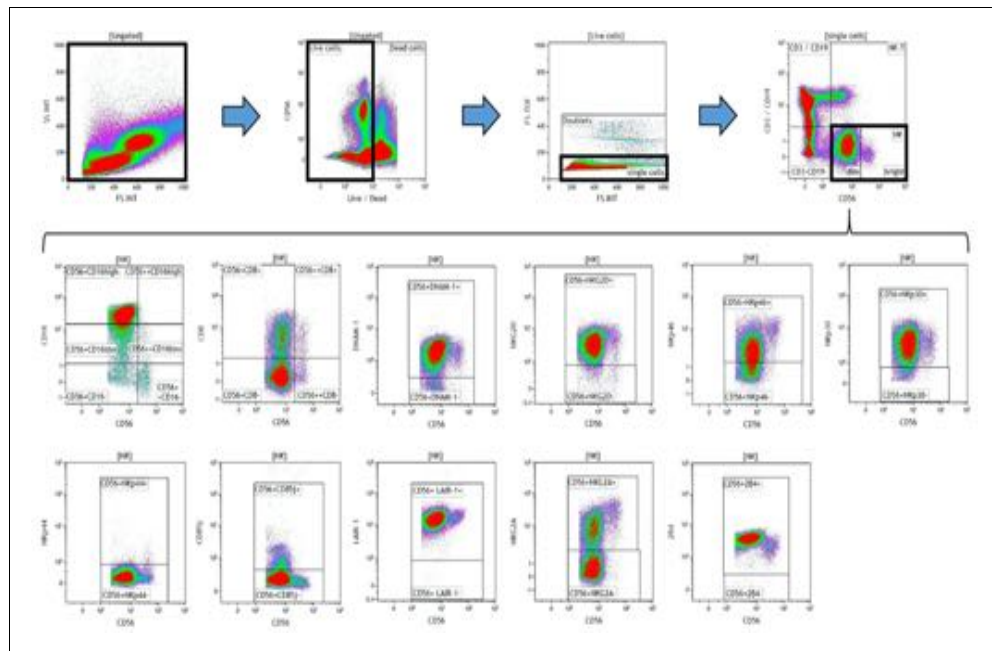
Figure 8. Representative gating strategy for analyzing the expression of activating and inhibitory receptors on peripheral blood natural killer (NK) cells. Using density plots, the NK cell phenotypic profiles were determined by first gating on 'live cells' in the forward scatter (FSc) linear vs side scatter (SSc) linear density plot and then gating on single cells (determined by FSc Linear vs FS time of flight). The expression of activating and inhibitory receptors was determined by gating on $CD3^- CD19^- CD56^+$ cells using fluorescence minus one (FMO) controls. The expression of each NK cell receptor was measured using the 'Logical' setting.

0 for benign). Hence, each patient x is mapped to a diagnosis y. The GA takes three inputs: 1) the feature-by-patient matrix X; 2) the vector y which holds the corresponding labels for each patient record; and 3) the desired number of features, λ. The GA returns the IDs of the best subset of features, where the subset has size λ. GAs stochastically select multiple features from the current population and thus each run of the GA can return different results. Consequently, we proposed an approach to identify the best subset of features by running the algorithm several times and then obtaining the frequency of the subsets.

Proposed ensemble learning classifier for identifying the presence of prostate cancer

This section discusses the machine learning classifier which was developed for the task of identifying the presence of benign prostate disease or prostate cancer using the identified subset of phenotypic features. The challenging task is that a suitable and reliable classifier must be developed using only 72 patient records. A limitation is that classifiers that have been trained on small sample size data are likely to be unstable because small changes in the training set cause large changes in the classifier. It was for this reason that the Ensemble machine learning classifier was preferred as an approach for developing a more stable and reliable classifier. Ensemble classifiers achieve stability and reliability by constructing many 'weak' classifiers instead of a single classifier and then combine the weak classifiers (i.e. weak learners) to create a more powerful decision rule than that constructed

Table 11. Antibody panels for measuring the phenotype of Natural Killer cells.

Antibody	Fluorochrome	Clone no.	Supplier
Panel 1			
DNAM-1 (CD226)	FITC	11A8	BioLegend
NKG2D (CD314)	PE	1D11	eBioscience
CD56	ECD (PE-Texas Red)	N901	Beckman Coulter
CD16	PerCP-Cy5.5	3G8	BioLegend
NKp46 (CD335)	PE-Cy7	9E2	BioLegend
NKp30 (CD337)	Alexa Fluor 647	P30-15	BioLegend
CD3	Alexa Fluor 700	UCHT1	BioLegend
CD19	Alexa Fluor 700	HIB19	BioLegend
CD8	APC-Cy7	SK1	BioLegend
Live/Dead	Dye (violet)		Thermo Fisher Scientific
Panel 2			
CD85j (ILT2)	FITC	GHI/75	Miltenyi Biotec
LAIR-1 (CD305)	PE	DX26	BD Biosciences
CD56	ECD (PE-Texas Red)	N901	Beckman Coulter
CD16	PerCP-Cy5.5	3G8	BioLegend
NKG2A (CD159a)	PC7 (PE-Cy7)	Z199	Beckman Coulter
NKp44 (CD336)	Alexa Fluor 647	P44-8	BioLegend
CD3	Alexa Fluor 700	UCHT1	BioLegend
CD19	Alexa Fluor 700	HIB19	BioLegend
CD8	APC-Cy7	SK1	BioLegend
LIVE/DEAD	Dye (violet)		Thermo Fisher Scientific
Panel 3			
2B4 (CD244.2)	FITC	C1.7	BioLegend
CD56	ECD (PE-Texas Red)	N901	Beckman Coulter
CD16	PerCp-Cy5.5	3G8	BioLegend
CD3	Alexa Fluor 700	UCHT1	BioLegend
CD19	Alexa Fluor 700	HIB19	BioLegend
CD8	APC-Cy7	SK1	BioLegend
LIVE/DEAD	Dye (violet)		Thermo Fisher Scientific

cytometry features for creating the prediction algorithm is performed using a GA. The GA adopted in the experiments was developed by *Ludwig and Nunes, 2010*. The particular GA performs combinatorial optimization to identify a subset of features that comprises the optimum feature set, in which the order of features has no relation with their importance. The algorithm works by maximising the mutual information between the target y (where y can have a value 1 for cancer or 0 for benign) and the input features (i.e. these are the 32 features listed in *Table 1*). Mutual information is the measure of the mutual dependence between the two variables, i.e. an input feature and the target. Adopting a GA eliminates the computational effort which is necessary to evaluate all the possible combinations of features. The fitness function of the GA (*Ludwig and Nunes, 2010*) is based on the principle of max-relevance and min-redundancy (mRMR), for which the objective is that the outputs of the selected features present discriminant power, thereby avoiding redundancy. The principle of max-relevance and min-redundancy corresponds to searching the set of feature indexes that are mutually exclusive and correlated to the target output. Let $m \times n$ be a feature-by-patient matrix, $X = [x_{ij}]$ with m features and n patients. Thus, the matrix element x_{ij} is the flow cytometry value i of patient j. Let y be a vector of size $1 \times n$ which holds the diagnosis of each patient (1 for cancer and

when using a single classifier. In clinical applications, it is important to construct prediction models which have a low bias, meaning that the classifier suggests fewer assumptions about the form of the target function. Because Ensemble learning makes fewer assumptions about the form of the target function, it was considered to be a suitable classifier for the task. Several techniques for combining the classifiers of an Ensemble model exist and these include Boosting, Bagging, and Random Subspace Dimension.

In the proposed method, the Random Subspace Dimension approach was utilised as a strategy for combining the kNN classifiers, to create the Ensemble of kNN classifiers. In machine learning, the Random Subspace Method (*Ho, 1998*), also called attribute bagging (*Bryll et al., 2003*) or feature bagging, is an Ensemble learning method which attempts to reduce the correlation between estimators in an Ensemble by training them on random samples of features instead of the entire feature set. In the Random Subspace method, classifiers are constructed in random subspaces of the data feature space. These classifiers were combined by simple majority voting in the final decision rule, and we used the k Nearest Neighbor method (see *Figure 9*). In particular, we used the Random Subspace ensemble-aggregation method coupled with k Nearest Neighbours weak learners to produce an Ensemble of classifiers, and this resulted to a better classification rule. Thus, the Random Space modifies the training data set, builds classifiers on these modified training sets, and then combines them into a final decision rule by simple or weighted majority voting.

Figure 9 provides an overview of the architecture of the proposed kNN Ensemble learning, and the description that follows explains the architecture in more detail. Let m be the number of dimensions (variables) to sample in each learner minus 1. Let d be the number of dimensions in the data, which is the number of predictors in the data matrix X. Let n be the number of learners in the ensemble. The basic random subspace algorithm performs the following steps using the above-mentioned parameters:
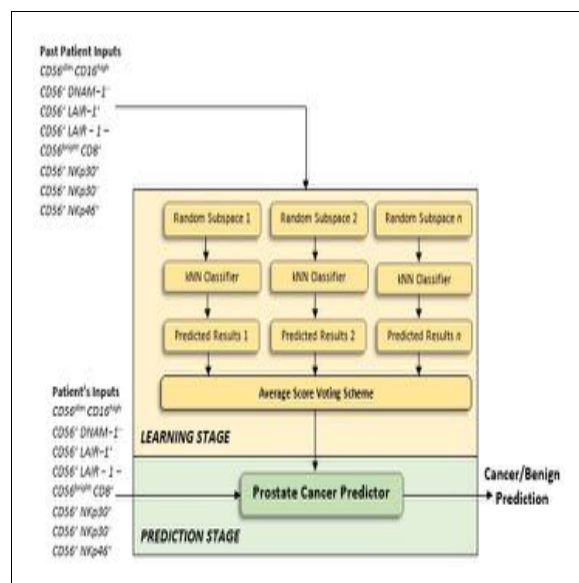
Figure 9. Proposed Ensemble Subspace kNN model. Ensembles combine predictions from different models to generate a final prediction. Because Ensemble approaches combine baseline predictions, they perform at least as well as the best baseline model.

1. Choose without replacement a random set of m predictors from the d possible values.
2. Train a weak learner using just the m chosen predictors.
3. Repeat steps 1 and 2 until there are n weak learners.
4. Predict by taking an average of the score prediction of the weak learners and classify the category with the highest average score.

Performance evaluation measures

A variety of relevant evaluation metrics were adopted for the task of evaluating the performance of the machine learning prostate cancer presence and risk prediction models.

Prostate cancer presence prediction models: Let $|TP|$ be the total number of patients with cancer who were correctly classified as having cancer; $|TN|$ be total the number of individuals with benign disease who were correctly classified as having benign disease; $|FP|$ be the total number of individuals with benign disease who were incorrectly classified as having cancer; $|FN|$ be the total number of patients with cancer who were incorrectly classified as having benign disease; $|P|$ be the total number of patients with cancer that exist in the dataset, where $|P| = |TP| + |FN|$; and $|N|$ be the total number of individuals with benign disease that exist in the dataset, where $|N| = |FP| + |TN|$. The following commonly used evaluation measures can be defined.

$$Accuracy = \frac{|TP| + |TN|}{|TP| + |FP| + |FN| + |TN|}, \in [0,1]. \tag{2}$$

$$TPR = \frac{|TP|}{|TP| + |FN|}, \in [0,1]. \tag{3}$$

$$TNR = \frac{|TN|}{|TN| + |FP|}, \in [0,1]. \tag{4}$$

$$FNR = \frac{|FN|}{|TP| + |FN|} = 1 - Sensitivity, \in [0,1]. \tag{5}$$

$$FPR = \frac{|FP|}{|FP| + |TN|} = 1 - Specificity, \in [0,1]. \tag{6}$$

The closer the values of Accuracy, True Positive Rate (i.e. TPR, Sensitivity) and True Negative Rate (i.e. TNR, Specificity) are to 1.0, then the better the classification performance of a system.

The Receiver Operating Characteristic (ROC) is an effective measure for evaluating the quality of a prediction model's performance. The ROC curve has an optimal ROC point which comprises two values: the False Positive Rate (FPR) and the True Positive Rate (TPR) values. The optimal ROC point is computed by function (*Equation 7*) for finding the slope, S.

$$S = \frac{Cost(P|N) - Cost(N|N)}{Cost(N|P) - Cost(P|P)} \times \frac{N}{P}, \tag{7}$$

where $Cost(N|P)$ is the cost of misclassifying a positive class (i.e. cancer) as a negative class (i.e. benign); $Cost(P|N)$ is the cost of misclassifying a negative class, as a positive class; P, and N, are the total instance counts in the cancer and benign class, respectively. The optimal ROC point is identified by moving the straight line with slope S from the upper left corner of the ROC plot ($FPR = 0$, $TPR = 1$) down and to the right, until it intersects the ROC curve.

The Area Under the ROC Curve (AUC) is another important performance evaluation metric which reflects the capacity of a model capacity to discriminate between the data obtained from individuals with benign disease and patients with cancer. The larger the AUC, the better the overall capacity of the classification system to correctly identify benign disease and cancer.

Prostate cancer risk prediction models: When applying the above-mentioned measures to evaluate the performance of the risk prediction models, the Positive class, P, was changed to be the High-risk group and the Negative class, N, was changed to be the L/I group.

eLife Research article

Acknowledgements

Additional information

Competing interests

Funding

Author contributions

Simon P Hood, Gemma A Foulds, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - review and editing; Georgina Cosma, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Catherine Johnson, Stephen Reeder, Investigation, Methodology; Stéphanie E McArdle, Formal analysis, Investigation, Methodology, Writing - review and editing; Masood A Khan, Resources, Data curation, Validation, Investigation, Writing - review and editing; A Graham Pockley, Conceptualization, Supervision, Funding acquisition, Investigation, Methodology, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Georgina Cosma https://orcid.org/0000-0002-4663-6907
A Graham Pockley https://orcid.org/0000-0001-9593-6431

Ethics

Human subjects: Research Protocols were registered and approved by the National Research Ethics Service (NRES) Committee East Midlands and by the Research and Development Department in the University Hospitals of Leicester NHS Trust. All participants were given information sheets explaining the nature of the study and all provided informed consent. All samples were collected by suitably qualified individuals using standard procedures. Ethical approval for the collection and use of samples from the TPTPB cohort (Project Title: Defining the role of Transperineal Template-guided prostate biopsy) was given by NRES Committee East Midlands- Derby 1 (NREC Reference number: 11/EM/3012; UHL11068).

Decision letter and Author response
Decision letter https://doi.org/10.7554/eLife.50936.sa1
Author response https://doi.org/10.7554/eLife.50936.sa2

Additional files

Supplementary files
• Source data 1. Prostate Cancer Dataset.
• Transparent reporting form

Data availability

A spreadsheet of the immune cell phenotypic data has been provided as Source Data File 1.

References

Aganovic D, Prcic A, Kulovac B, Hadziosmanovic O. 2011. Prostate Cancer detection rate and the importance of premalignant lesion in rebiopsy. *Medicinski Arhiv* **65**:109–112. PMID: 21585187

Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, Collaco-Moraes Y, Ward K, Hindley RG, Freeman A, Kirkham AP, Oldroyd R, Parker C, Emberton M. 2017. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate Cancer (PROMIS): a paired validating confirmatory study. *The Lancet* **389**: 815–822. DOI: https://doi.org/10.1016/S0140-6736(16)32401-1

Anderson KG, Stromnes IM, Greenberg PD. 2017. Obstacles posed by the tumor microenvironment to T cell Activity: A Case for Synergistic Therapies. *Cancer Cell* **31**:311–325. DOI: https://doi.org/10.1016/j.ccell.2017.02.008, PMID: 28292435

Bryll R, Gutierrez-Osuna R, Quek F. 2003. Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets. *Pattern Recognition* **36**:1291–1302. DOI: https://doi.org/10.1016/S0031-3203(02)00121-8

Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. 1991. Measurement of prostate-specific antigen in serum as a screening test for prostate Cancer. *New England Journal of Medicine* **324**:1156–1161. DOI: https://doi.org/10.1056/NEJM199104253241702, PMID: 1707140

Chang DT, Challacombe B, Lawrentschuk N. 2013. Transperineal biopsy of the prostate–is this the future? *Nature Reviews Urology* **10**:690–702. DOI: https://doi.org/10.1038/nrurol.2013.195, PMID: 24061531

Cosma G, McArdle SE, Reeder S, Foulds GA, Hood S, Khan M, Pockley AG. 2017. Identifying the presence of prostate Cancer in individuals with PSA levels <20 ng ml^{-1} Using Computational Data Extraction Analysis of High Dimensional Peripheral Blood Flow Cytometric Phenotyping Data. *Frontiers in Immunology* **8**:1771. DOI: https://doi.org/10.3389/fimmu.2017.01771, PMID: 29326690

Dimmen M, Vlatkovic L, Hole KH, Nesland JM, Brennhovd B, Axcrona K. 2012. Transperineal prostate biopsy detects significant Cancer in patients with elevated prostate-specific antigen (PSA) levels and previous negative transrectal biopsies. *BJU International* **110**:E69–E75. DOI: https://doi.org/10.1111/j.1464-410X.2011.10759.x, PMID: 22093091

Hankey BF, Feuer EJ, Clegg LX, Hayes RB, Legler JM, Prorok PC, Ries LA, Merrill RM, Kaplan RS. 1999. Cancer surveillance series: interpreting trends in prostate Cancer–part I: evidence of the effects of screening in recent prostate Cancer incidence, mortality, and survival rates. *JNCI Journal of the National Cancer Institute* **91**:1017–1024. DOI: https://doi.org/10.1093/jnci/91.12.1017, PMID: 10379964

Ho TK. 1998. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**:832–844. DOI: https://doi.org/10.1109/34.709601

Ludwig O, Nunes U. 2010. Novel maximum-margin training algorithms for supervised neural networks. *IEEE Transactions on Neural Networks* **21**:972–984. DOI: https://doi.org/10.1109/TNN.2010.2046423, PMID: 20409990

Nafie S, Mellon JK, Dormer JP, Khan MA. 2014a. The role of transperineal template prostate biopsies in prostate Cancer diagnosis in biopsy naïve men with PSA less than 20 ng ml(-1.). *Prostate Cancer and Prostatic Diseases* **17**:170–173. DOI: https://doi.org/10.1038/pcan.2014.4, PMID: 24590360

Nafie S, Pal RP, Dormer JP, Khan MA. 2014b. Transperineal template prostate biopsies in men with raised PSA despite two previous sets of negative TRUS-guided prostate biopsies. *World Journal of Urology* **32**:971–975. DOI: https://doi.org/10.1007/s00345-013-1225-x, PMID: 24337167

Naughton CK, Miller DC, Mager DE, Ornstein DK, Catalona WJ. 2000. A prospective randomized trial comparing 6 versus 12 prostate biopsy cores: impact on Cancer detection. *Journal of Urology* **164**:388–392. DOI: https://doi.org/10.1016/S0022-5347(05)67367-3, PMID: 10893592

NICE. 2014. *Costing Statement: Prostate Cancer:diagnosis and Treatment. Implementing the Nice Guideline on Prostate Cancer (Cg175)*: NICE. https://www.nice.org.uk/guidance/cg175.

---

eLife Research article

Pal RP, Elmussareh M, Chanawani M, Khan MA. 2012. The role of a standardized 36 core template-assisted transperineal prostate biopsy technique in patients with previously negative transrectal ultrasonography-guided prostate biopsies. *BJU International* **109**:367–371. DOI: https://doi.org/10.1111/j.1464-410X.2011.10355.x, PMID: 21883818

Pasero C, Gravis G, Granjeaud S, Guerin M, Thomassin-Piana J, Rocchi P, Salem N, Walz J, Moretta A, Olive D. 2015. Highly effective NK cells are associated with good prognosis in patients with metastatic prostate Cancer. *Oncotarget* **6**:14360–14373. DOI: https://doi.org/10.18632/oncotarget.3965, PMID: 25961317

Quandt D, Dieter Zucht H, Amann A, Wulf-Goldenberg A, Borrebaeck C, Cannarile M, Lambrechts D, Oberacher H, Garrett J, Nayak T, Kazinski M, Massie C, Schwarzenbach H, Maio M, Prins R, Wendik B, Hockett R, Enderle D, Noerholm M, Hendriks H, et al. 2017. Implementing liquid biopsies into clinical decision making for Cancer immunotherapy. *Oncotarget* **8**:48507–48520. DOI: https://doi.org/10.18632/oncotarget.17397, PMID: 28501851

Skurichina M, Duin RPW. 2002. Bagging, boosting and the random subspace method for linear classifiers. *Pattern Analysis & Applications* **5**:121–135. DOI: https://doi.org/10.1007/s100440200011

Vitale M, Cantoni C, Pietra G, Mingari MC, Moretta L. 2014. Effect of tumor cells and tumor microenvironment on NK-cell function. *European Journal of Immunology* **44**:1582–1592. DOI: https://doi.org/10.1002/eji.201344272, PMID: 24777896

Yuasa T, Tsuchiya N, Kumazawa T, Inoue T, Narita S, Saito M, Horikawa Y, Satoh S, Habuchi T. 2008. Characterization of prostate Cancer detected at repeat biopsy. *BMC Urology* **8**:14. DOI: https://doi.org/10.1186/1471-2490-8-14, PMID: 19000320

Phenotype and Function of Activated Natural Killer Cells From Patients With Prostate Cancer: Patient-Dependent Responses to Priming and IL-2 Activation

Simon P. Hood[1][*][†], Gemma A. Foulds[1], Heather Imrie[2], Stephen Reeder[1], Stéphanie E. B. McArdle[1], Masood Khan[3] and Alan Graham Pockley[1][*]

[1] John van Geest Cancer Research Centre, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom, [2] School of Animal Rural and Environmental Sciences, Nottingham Trent University, Nottingham, United Kingdom, [3] Department of Urology, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

Background: Although immunotherapy has emerged as the "next generation" of cancer treatments, it has not yet been shown to be successful in the treatment of patients with prostate cancer, for whom therapeutic options remain limited to radiotherapy and androgen (hormone) deprivation therapy. Previous studies have shown that priming natural killer (NK) cells isolated from healthy individuals via co-incubation with CTV-1 cells derived from an acute lymphoblastic leukemia (ALL) enhances their cytotoxicity against human DU145 (metastatic) prostate cancer cells, but it remains unknown to what extent NK cells from patients with prostate cancer can be triggered to kill. Herein, we explore the phenotype of peripheral blood NK cells in patients with prostate cancer and compare the capacity of CTV-1 cell-mediated priming and IL-2 stimulation to trigger NK cell-mediated killing of the human PC3 (metastatic) prostate cancer cell line.

Methods: The phenotype of resting, primed (co-incubation with CTV-1 cells for 17 h) and IL-2 activated (100 IU/ml IL-2 for 17 h) NK cells isolated from frozen-thawed peripheral blood mononuclear cell (PBMC) preparations from patients with benign disease ($n = 6$) and prostate cancer ($n = 18$) and their cytotoxicity against PC3 and K562 cells was determined by flow cytometry. Relationship(s) between NK cell phenotypic features and cytotoxic potential were interrogated using Spearman Rank correlation matrices.

Results and Conclusions: NK cell priming and IL-2 activation of patient-derived NK cells resulted in similar levels of cytotoxicity, but distinct NK cell phenotypes. Importantly, the capacity of priming and IL-2 stimulation to trigger cytotoxicity was patient-dependent and mutually exclusive, in that NK cells from ~50% of patients preferentially responded to priming whereas NK cells from the remaining patients preferentially responded to cytokine stimulation. In addition to providing more insight into the biology of primed

168 | P a g e

and cytokine-stimulated NK cells, this study supports the use of autologous NK cell-based immunotherapies for the treatment of prostate cancer. However, our findings also indicate that patients will need to be stratified according to their potential responsiveness to individual therapeutic approaches.

Keywords: natural killer (NK) cells, priming, CTV-1, cytotoxicity, phenotype, prostate, cancer, TNF receptors

INTRODUCTION

The 2012 GLOBOCAN project revealed prostate cancer to be the 4th most common cancer in the world, with 1.1 million cases reported. Prostate cancer typically occurs in men over the age of 50 yrs and is the most common male cancer in the developed world (1). In 2014, there were 45,406 new cases of prostate cancer reported and 12,082 deaths from the disease (33 per day) in the UK (2).

Patients with confirmed prostate cancer are stratified according to the D'Amico risk classification (i.e., low, intermediate, and high risk) which predicts the likelihood of a patient suffering biochemical recurrence following treatment (3, 4). Stratification is based on the patient's clinical tumor, node and stage of metastasis (TNM), serum prostate specific antigen (PSA) levels, and biopsy Gleason score (3, 4). Due to the slow growing nature of prostate cancer, patients with low risk disease are typically assigned to active surveillance, as the disease is unlikely to progress within their life time. In contrast, patients at intermediate and high risk, and young patients at low risk undergo active treatment as disease progression is more likely in these individuals (5).

The primary treatment for advanced metastatic prostate cancer is androgen (hormone) deprivation therapy (ADT), with upfront chemotherapy if medically fit and with good renal function. Although the majority of patients initially respond to ADT, as evidenced by disease regression and disease stability (5), it is inevitable that disease will progress and become hormone-resistant. At this point, second-line hormone therapy followed by further hormone manipulation therapy is considered, but will typically deliver only a very limited effect.

Immunotherapy involving stimulating the patient's own immune system to retarget their cancer is emerging as the next generation of cancer treatment (6). Currently, the only approved immunotherapy for treating castration-resistant prostate cancer is Sipuleucel-T immunotherapy which has been shown to improve the median overall survival by 4.1 months compared to a placebo group (7). Although preventing tumor-mediated immunoregulation using immune checkpoint inhibitors such as Ipilimumab has shown some success in treating immunogenic cancers such as melanoma and non-small cell lung cancer, their use in patients with prostate cancer has not been shown to improve overall survival (8). However, some evidence of beneficial effects have been observed and clinical trials testing Ipilimumab in combination with other standard prostate cancer treatments (e.g., ADT) are ongoing (9).

Natural killer (NK) cells were first identified on the basis of their natural cytotoxicity toward cancerous cells and

a number of NK cell-based immunotherapies are now in development (10–15). As reviewed by Sabry and Lowdell (16), the cytotoxic function of NK cells is controlled by the balance of signals transduced via activating and inhibitory receptors following ligation with stress ligands and MHC class I molecules, respectively (Dynamic Equilibrium Theory). Bryceson et al. demonstrated that natural cytotoxicity requires the co-engagement of multiple activating receptors (17, 18). Furthermore, work by Lowdell et al. led to the hypothesis that the natural cytotoxicity mechanism can be divided into two discrete stages; "priming" and "triggering" (16, 19, 20). For this, they hypothesized that the "priming" signal can be delivered either by the ligation of the appropriate number and combination of activating receptors with their target ligands or via an activating cytokine (e.g., IL-2). The "triggering" signal requires the ligation of at least one additional activating receptor to its target ligand that is specific to stressed cells (16).

Tumor primed NK cells (TpNK) can be generated *in vitro* by co-incubating resting NK cells with the acute lymphoblastic leukemia (ALL) cell line CTV-1 (19). Phenotypically, tumor primed NK cells appear distinct from resting NK cells in that they exhibit reduced expression of activating receptors (e.g., CD16, NKG2D, NKp46), both in terms of intensity and proportion, whereas both the proportion and intensity of expression of co-receptors (e.g., CD69 and CD25) are up-regulated (19, 20). Priming NK cells from healthy volunteers in this way has been reported to enhance their cytotoxicity against NK cell-resistant tumor cell lines such as the human metastatic prostate cancer cell line DU145 (20).

The therapeutic potential of an autologous NK cell-based therapy requires that patient-derived NK cells can be appropriately triggered. Herein, we determined whether activation of NK cells isolated from thawed peripheral blood mononuclear cell (PBMC) preparations derived from patients with prostate cancer by either co-incubation with mitomycin C treated CTV-1 cells or stimulation with IL-2 enhanced their capacity to kill the human metastatic disease-derived prostate cancer cell line PC3.

Tumor priming and IL-2 stimulation of patient-derived NK cells resulted in similar levels of cytotoxicity, but distinct NK cell phenotypes. Importantly, the capacity of priming and IL-2 stimulation to trigger cytotoxicity was patient-dependent and mutually exclusive, in that NK cells from ∼50% of patients preferentially responded to tumor priming, whereas NK cells from the remaining patients preferentially responded to IL-2 stimulation. In addition to providing more insight into the biology of tumor primed and cytokine-stimulated NK cells, this study supports the use of autologous NK cell-based

immunotherapies for the treatment of prostate cancer. However, our findings also indicate that patients will need to be stratified according to their potential responsiveness to individual therapeutic approaches.

METHODS

Patients and Ethical Approval

Ethical approval for the study cohort (Ethical Approval Number 14/ES/1014) was obtained from the East of Scotland Research Ethics Service (EoSRES). Patients suspected of having prostate cancer who attended the Urology Clinic at Leicester General Hospital (Leicester UK) between 14th August 2014 and 3rd December 2015 were given the opportunity to take part in the study and provide a peripheral blood sample. Approval for the collection of peripheral blood from healthy volunteers was obtained from the Nottingham Trent University College of Science and Technology Human Ethics Committee (Application Number 435). Healthy volunteers and patients were given information sheets detailing the nature of the study and those wishing to take part were provided the opportunity to discuss and ask questions. All participants provided informed consent and were assigned a number to maintain anonymity. Participants provided a 60 mL peripheral blood sample which was obtained by venepuncture. Of the 24 individuals who attended the Urology Clinic at Leicester General Hospital and were included in the study, 6 were diagnosed as having benign disease, and 18 patients were diagnosed with prostate cancer, as determined by TRUS biopsy. Gleason scores of the 18 cancer patients were; Gleason 6 ($n = 3$), Gleason 7 ($n = 5$), Gleason 9 ($n = 8$), and Gleason 10 ($n = 2$).

Cell Culture

The CTV-1 cell line was purchased from the Leibniz-Institut DSMZ—Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Braunschweig, Germany) and maintained in RPMI 1640 (LONZA) supplemented with 10% v/v fetal bovine serum (FBS) (Hyclone) and 1% v/v L-Glutamine (LONZA). The PC3 cell line was purchased from ATCC and maintained in Hams F-12K (Kaighn's) medium (GIBCO™) supplemented with 10% v/v FBS and passaged using 1X trypsin—versene (LONZA). The K562 cell line was purchased from ATCC and maintained in iscove's modified dulbecco's medium (IMDM) (LONZA) supplemented with 10% v/v FBS.

Peripheral Blood Mononuclear (PBMC) Isolation

Peripheral blood (60 ml) was collected using standard procedures and aliquoted (30 ml) into two sterile 50 ml polypropylene Falcon™ tubes containing 300 µl of heparin (1,000 IU/ml, Sigma). Samples were immediately transferred to the John van Geest Cancer Research Centre at Nottingham Trent University (Nottingham, UK) and processed immediately upon receipt—all samples were processed within 2 h. Blood was diluted 1 in 3 with phosphate buffered saline (PBS, LONZA) and layered over Ficoll Paque (GE Healthcare Life Sciences) in LeucoSep® tubes (20 ml per tube). The tubes were subsequently centrifuged at 800 g for 20 min. The PBMC layer was collected and washed twice with

PBS before being counted using trypan blue dye exclusion (Santa Cruz Biotechnology). PBMCs were frozen down in 90% v/v FBS, 10% v/v Dimethyl sulfoxide (DMSO) at 10^6 cells per vial and stored in liquid nitrogen.

NK Cell Isolation and Activation

PBMCs were defrosted, washed and rested for 30 min in complete medium (RPMI 1640 supplemented with 10% v/v FBS and 1% v/v L-Glutamine). Natural killer (NK) cells were isolated by magnetic beads using the human NK cell negative selection isolation kit (Miltenyi Biotec) according to the manufacturer's instructions and counted using trypan blue with the cells exhibiting a viability of >98%. Flow cytometry analysis revealed that following isolation, the purity of the NK cells averaged 82%. CTV-1 cells were pelleted, counted by trypan blue exclusion, re-suspended to a concentration of 7×10^6 viable cells/ml and then treated with 33 µg/ml Mitomycin C (Sigma) for 2 h at 37°C, 5% v/v CO_2. Following treatment, the cells were washed three times with PBS and then re-suspended in complete medium and recounted. For each patient, NK cells were activated by (1) co-incubation with mitomycin C treated CTV-1 cells at a 1:2 (NK:CTV-1) ratio in complete medium or (2) incubation in complete medium containing 100 IU/ml of IL-2 (PeproTech). As a control, resting NK cells were incubated at a concentration of 2×10^6 cells/ml in complete medium for 17 h overnight.

Cytotoxicity Assay

The non-adherent K562 cells were pelleted at 400 g for 5 min, whereas the adherent PC3 cells were harvested using 1X Trypsin-Versene and then pelleted at 300 g for 5 min. Both cell lines were counted using trypan blue dye exclusion (>95% viability) and were used as target cells. Target cells were stained with 200 nM MitoTracker™ Green FM (ThermoFisher Scientific) in RPMI 1640 alone at a concentration of 2×10^6 viable cells/ml for 20 min at 37°C, 5% v/v CO_2 in the dark. Following incubation, the target cells were washed three times with PBS and then re-suspended in IMDM (K562) or Hams F-12K (Kaighns) medium (PC3). The cells were then recounted using trypan blue dye exclusion (>95% viability) and diluted to a concentration of 1.0×10^6 viable cells/ml.

NK cells activated overnight for 17 h were pelleted, re-suspended in fresh complete RPMI medium, counted (NK cells are smaller in size compared to CTV-1 cells) and diluted to a concentration of 1.0×10^6 NK viable cells/ml. NK cells (150 µL, 150,000 cells) were co-incubated with target cells (30 µL PC3 or K562, 30,000 cells) plus an additional 120 µL complete medium in 12 × 75 mm polycarbonate tubes for 3 h at 37°C, 5% v/v CO_2. As a control, target cells were also analyzed alone for background death. For this, 10 µl of Propidium Iodide (50 µM/ml) was added to each tube prior to sample analysis. Samples were acquired on a Beckman Coulter Gallios™ flow cytometer. For each tube a minimum of 5,000 events were acquired and data analyzed using the Beckman Coulter Kaluza™ v1.3 software.

Characterizing the NK Cell Phenotype

To measure the phenotype of both resting and activated NK cell populations, 1.5×10^5 NK cells were added to 12 × 75 mm tubes and washed with Wash Buffer (PBS + 0.5% w/v

TABLE 1 | Monoclonal antibody panel 1.

Antibody	Fluorochrome	Clone	Supplier
CD2	FITC	TS1/8	BioLegend
CD96	PE	NK92.39	BioLegend
CD56	ECD	N901	Beckman Coulter
CD16	PerCP-Cy5.5™	3G8	BioLegend
CD137	PE-Cy7	4B4-1	BioLegend
CD69	APC	FN50	BioLegend
CD3	Alexa Fluor™ 700	UCHT1	BioLegend
CD19	Alexa Fluor™ 700	HIB19	BioLegend
CD107a	APC-Cy7™	H4A3	BioLegend
LIVE/DEAD™	dye		ThermoFisher scientific

TABLE 2 | Monoclonal antibody panel 2.

Antibody	Fluorochrome	Clone	Supplier
DNAM-1	FITC	11A8	BioLegend
NKG2D	PE	5C6	eBioscience
CD56	ECD	N901	Beckman Coulter
CD16	PerCP-Cy5.5™	3G8	BioLegend
NKp46	PE-Cy7™	9E2	BioLegend
GITR	APC	ebioAITR	BioLegend
CD3	Alexa Fluor™ 700	Cr24.1	BioLegend
CD19	Alexa Fluor™ 700	HIB19	BioLegend
OX40	APC-Cy7™	Ber-ACT35	BioLegend
LIVE/DEAD™	Dye	Violet	ThermoFisher scientific

BSA, 0.02% w/v sodium azide). NK cells were incubated with antibody panels (detailed in **Tables 1**, **2**) for 15 min in the dark at room temperature. The cells were washed with PBS, after which they were incubated with 1 ml of LIVE/DEAD™ Fixable Violet solution (ThermoFisher Scientific) according to the manufacturer's instructions. Cells were washed with Wash Buffer and re-suspended in Isoton™ II diluent Beckman Coulter. Data were acquired on a Beckman Coulter Gallios™ flow cytometer and analyzed using Beckman Coulter Kaluza™ v1.2 software.

Statistical Analysis

Graphs were created using GraphPad Prism v7. All datasets were assessed for Gaussian distribution. Significant differences in the expression of NK cell receptors between resting NK cells, CTV-1 primed NK cells and IL-2 stimulated NK cells were assessed by repeated measures two way ANOVA with a Tukey multiple comparisons *ad-hoc* test. Correlations between receptor expression and cytotoxic killing were assessed using two tailed non-parametric Spearman Correlation tests.

RESULTS

Optimisation of the NK Cell Priming Assay and Generation of Evidence to Suggest That Down-Regulation of Activating Receptors Does not Necessarily Indicate NK Cell Dysfunction

It has previously been reported that NK cells from healthy individuals can be primed by co-incubation with CTV-1 cells,

thereby enabling them to kill human DU145 prostate cancer cells (20). During optimisation of our methods utilizing healthy cells, NK cells were primed using different NK:CTV-1 ratios (i.e., 2:1, 1:1, 1:2, and 1:4) and their cytotoxicity against PC3 cells (metastatic prostate cancer cell) assessed. As a reference, the cytotoxic capacity of CTV-1 primed NK cells against K562 (chronic myelogenous leukemia) cells and resting NK cells against both K562 and PC3 cell lines was also determined.

Resting NK cells from three healthy individuals killed 2.7, 10, and 8.8% of PC3 cells and 68.6, 83.7, and 43.8% of K562 cells, respectively (**Figures 1A–C**). Priming with CTV-1 cells enhanced the cytotoxicity of NK cells from the healthy volunteers against PC3 cells: For two healthy volunteers (1 and 3), PC3 lysis was maximal at the 1:2 ratio (lysed PC3 cells; 35.5 and 26.5%, respectively), whereas for the third (healthy volunteer 2) PC3 lysis was maximal at the 1:1 ratio (34% lysed cells) (**Figures 1A–C**). In general, increasing the proportion of CTV-1 cells within the priming ratio beyond that which induced a peak level of cytotoxicity resulted in a reduced lysis of PC3 cells. Interestingly, for the NK cells from healthy volunteers 1 and 2, priming with CTV-1 cells reduced their cytotoxicity against K562 cells compared to their resting NK cell counterparts. The reduction in K562 lysis increased further as the proportion of CTV-1 cells in the priming ratio increased (**Figures 1A,B**). For healthy volunteer 3, priming at the 2:1 ratio resulted in a 34.6% increase in the killing of K562 cells, with further increases in the proportion of CTV-1 in the priming ratio resulting in a reduced lysis of K562 cells (**Figure 1C**). Analysis of the NK cell phenotype before and after priming revealed that the reduction in K562 lysis significantly correlated with a down-regulation in the proportion of NK cells expressing the activating receptors NKG2D (rs = 0.780, $P = 0.0015$) and NKp46 (rs = 0.643, $P = 0.0153$) (**Figures 1D,E**). For subsequent experiments, NK cells were primed at the 1:2 NK:CTV-1 ratio.

Influence of CTV-1 Priming or IL-2 Activation on the Cytotoxicity of NK Cells From Patients With Prostate Cancer Against K562 and PC3 Cells

Next, we wished to assess the influence of CTV-1 priming and IL-2 (100 U/ml) activation on the cytotoxicity of NK cells isolated from thawed PBMC preparations derived from patients with prostate cancer against PC3 and K562 cells. Sufficient NK cells to perform both the priming and IL-2 stimulation assays were only obtained from 21 of 24 patients (5 benign, 16 cancer). Sufficient NK cells to test the priming assays, but not the IL-2 stimulation assays were obtained from one of the individuals with benign disease. Only the phenotype and function of resting NK cells was possible for the remaining two patients with prostate cancer.

As shown in **Figure 2A**, resting NK cells from 7 of 22 patients killed between 56 and 87% of K562 cells, whereas the NK cells from the other 15 patients only killed between 4.8 and 23.1% of K562 cells. Priming at a 1:2 NK:CTV-1 ratio for 17 h decreased NK cell-mediated killing of K562 cells (median −40.7%, range −11.4% to −58.8%) for 9 out of 22 patients, whereas a small increase in killing of K562 cells was observed for the other 13 patients (median +7%, range +0.6% to +33.4%). Stimulation of

FIGURE 1 | Optimisation of the NK:CTV-1 priming ratio using NK cells from healthy volunteers. NK cells were isolated from thawed PBMCs that had previously been generated from the peripheral blood of healthy individuals and co-incubated with mitomycin C treated CTV-1 cells at ratios 2:1, 1:1, 1:2, 1:4 for 17 h at 37°C. NK cells incubated in isolation were used as controls. The phenotype and cytotoxic function of the primed NK cells were then assessed by flow cytometry. For each priming ratio, the ability of primed NK cells to lyse K562 cells and PC3 cells at a 5:1 effector to target ratio were measured. **(A–C)** Cytotoxicity against K562 and PC3 cells for three healthy volunteers. **(D,E)** Significant correlation between K562 lysis and expression of activating receptors NKG2D and NKp46 on primed NK cells, respectively.

the NK cells with 100 IU of IL-2 over 17 h resulted in increased killing of K562 cells (median $+17.4\%$, range $+6.4\%$ to $+46.9\%$) for all but one individual, for whom killing decreased by 1.6% (**Figure 2B**). IL-2 stimulation was more effective at enhancing the NK cytotoxicity toward K562 cells, whereas priming often resulted in a reduction in lysis (**Figure 2C**).

Although resting NK cells from the majority of patients exhibited low cytotoxicity against PC3 cells (median 3.6%, range 0% to 32.1%), cytotoxic potential was enhanced following priming with CTV-1 (median 14.4%, range 6.2% to 44.2%) (**Figure 2D**). This increased ability to kill PC3 cells was comparable to that which was induced by IL-2 stimulation (median 15.7%, range 2.8% to 61.9%) (**Figure 2E**). Due to the large range of cytotoxic responses against PC3 cells that were observed for activated NK cells (by either method), the patients from whom these were isolated were divided into two groups; (1) those that functionally responded better to NK cell priming than IL-2 stimulation (10 out of 21 patients) and (2) those that functionally responded better to IL-2 stimulation than NK cell

priming (11 out of 21 patients). As shown in **Figure 2F**, those patients that responded better to NK cell priming exhibited a median increase in killing of PC3 cells of $+19\%$ (range of $+6\%$ to $+36.5\%$) compared to a median increase of $+5.3\%$ (range -3.9% to $+19.1\%$) when NK cells were stimulated with IL-2. For the patients that responded better to IL-2 stimulation, the median increase in lysis of PC3 cells was $+26.5\%$ (range $+8.4\%$ to $+40.3\%$) compared to a median increase of $+4.9\%$ (range -20.8% to $+15.2\%$) when the NK cells were primed (**Figure 2G**).

Patients were then grouped according to whether priming increased or decreased K562 lysis (**Figure 2H**). For each group, the increase in cytotoxic response toward PC3 cells against that achieved when the NK cells were stimulated with IL-2 were compared (**Figures 2I–K**). NK cells from all but one of the patients that exhibited a reduced cytotoxicity against K562 cells after priming with CTV-1 cells ($n = 9$) exhibited a better cytotoxic response to PC3 cells when stimulated with IL-2 (**Figures 2J,K**). In contrast, NK cells from 9 of 11 patients that exhibited an increased cytotoxicity

FIGURE 2 | Influence of CTV-1 priming and IL-2 activation on the cytotoxicity of patient-derived NK cells against K562 and PC3 target cells. NK cells were isolated from thawed PBMC samples derived from peripheral blood of patients with benign prostate disease and patients with prostate cancer. For each patient, the NK cells were either primed with mitomycin C treated CTV-1 cells at a 1:2 ratio or activated with 100 IU IL-2, after which their ability to lyse K562 cells and PC3 cells at a 5:1 effector to target ratio was measured using flow cytometry. Comparison of cytotoxic responses toward K562 cells; **(A)** resting NK cells vs. primed NK cells, **(B)** resting NK cells vs. IL-2 activated NK cells, **(C)** change in lysis of K562 cells effected by priming compared to IL-2 activation. Assessment of cytotoxic responses toward PC3 cells; **(D)** resting NK cells vs. primed NK cells **(E)** resting NK cells vs. IL-2 activated NK cells, **(F)** patient-derived NK cells that lysed more PC3 cells following priming compared to following IL-2 activation **(G)** patient-derived NK cells that lysed more PC3 cells following IL-2 activation than following priming. **(H–K)** Alternative analysis of data presented in **(A, D, F, G)** by grouping patients according to increased or decreased lysis of K562 cells following priming.

against K562 cells after priming exhibited a better cytotoxic response to PC3 cells compared to that which was induced following stimulation with IL-2 (**Figures 2J,K**). Overall, these findings demonstrate that CTV-1 priming can enhance cytotoxic responses of patient NK cells toward a metastatic prostate cancer cell line, that is comparable to that of IL-2 stimulation, and that this enhancement appears to be irrespective of disease severity. Furthermore, these results also demonstrate that the responsiveness to NK cell CTV-1 priming and IL-2 activation is patient-dependent, thereby suggesting that treatment plans will need to be tailored to the patient in order to achieve an appropriate level of efficacy.

Correlation Between the Phenotype and Cytotoxic Function of Resting NK Cells and NK Cells Following CTV-1 Priming or IL-2 Activation

In parallel with the assessment of cytotoxic function, the phenotype of the patient-derived NK cells before and after activation was interrogated using the antibody panels described in **Tables 1**, **2**. In order to associate NK cell phenotype with cytotoxic function, we performed a correlation matrix with the percentage of resting NK cells expressing each receptor and the percentage of K562 and PC3 lysis (**Supplementary Figure 1**). We observed significant positive correlations between the percentage

of K562 lysis and the percentage of resting NK cells expressing NKp46 (activating receptor), CD69 (early activation marker), CD137 and GITR (TNF receptors). Furthermore, there was a significant positive correlation between the percentage of NK cells expressing CD69 and both TNF receptors. Graphical representations of a selection of these correlations are shown in **Figure 3**. Although the percentage of K562 lysis correlated with the percentage of NK cells expressing NKp46 (rs = 0.516, $P = 0.0099$), thereby reflecting what was observed in our experiments with healthy volunteers (**Figure 3A**), the most prominent correlation appeared to be between the percentage of K562 lysis and the proportion of NK cells expressing CD69 (rs = 0.688, $P = 0.0002$) (**Figures 3B**). As shown in **Figure 3B**, NK cell preparations that preferentially responded to IL-2 stimulation (green), as previously measured by a greater cytotoxic response to PC3 cells, generally exhibited a greater proportion of CD69$^+$ NK cells compared to NK cell preparations that preferentially responded to priming (blue).

It is plausible that thawing PBMCs up-regulated CD69 on a greater proportion of the NK cells recovered from the IL-2 responders and that further activation of these NK cells by priming reduced their cytotoxic potential. Resting NK cells have been reported not to express CD137 and to only express low levels of GITR. However, both receptors are up-regulated following NK cell activation (21, 22). This may account for the significant positive correlations between CD69 expression and both CD137 and GITR expression (rs = 0.638, $P = 0.0008$ and rs = 0.713, $P < 0.0001$, respectively) reported herein (**Figures 3C,D**).

There was no correlation between the percentage of K562 lysis and the percentage of PC3 lysis exhibited by CTV-1 primed NK cells (**Figure 4A**). However, there was a significant positive correlation when comparing the percentage change in lysis of K562 cells against the percentage change in lysis of PC3 cells induced by priming (rs = 0.687, $P = 0.0004$) (**Figure 4B**). In contrast, as shown in **Figures 4C,D**, only the percentage of K562 lysis and the percentage of PC3 lysis exhibited by IL-2 stimulated NK cells were positively correlated (rs = 0.672, $P = 0.008$). Taken together, it appears that CTV-1 priming and IL-2 activation differentially regulate NK cell function. Since NK cell function is controlled by signals delivered via activating and inhibitory receptors expressed on the NK cell surface, it is likely that the disparities in NK cell function are due to the differences in the way the activation methods influence the expression of relevant NK cell receptors.

Influence of CTV-1 Priming and IL-2 Activation on CD16, NKG2D, NKp46 and CD69 Expression

Having observed differences in NK cell-mediated cytotoxic responses following tumor priming and IL-2 stimulation, we next wanted to determine whether these two approaches differentially regulate the expression of NK cell activating receptors. Our initial analysis on the total NK cell population revealed that CTV-1 priming down-regulated the percentage of NK cells expressing the activating receptors CD16, NKG2D, and NKp46 in conjunction with an up-regulation in the percentage of NK cells

expressing CD69. These results supported the observations made by Lowdell et al. (19, 20).

The analysis was then focussed on the CD56^{dim} NK cell subset which can be subdivided into three subpopulations based on CD16 expression; CD56^{dim}CD16^{high}, CD56^{dim}CD16^{low}, and CD56^{dim}CD16^{neg}. The down-regulation of CD16 expression by CTV-1 priming altered the proportions of these three CD56^{dim} subpopulations. Representative density plots and gating strategies are shown in **Figures 5A,B**. A decrease in the median percentage of CD56^{dim}CD16^{high} NK cells (resting 72.6 vs. primed 48.7%) was observed, whereas there was an increase in the median percentage of CD56^{dim}CD16^{low} NK cells (resting 17.7 vs. primed 32.7%), and CD56^{dim}CD16^{neg} NK cells (resting 3.2 vs. primed 11%). In comparison, the proportion of the CD56^{dim}CD16$^{+/-}$ subpopulations was not significantly altered following stimulation with IL-2 and remained significantly different to that of the CTV-1 primed NK cell subpopulations (**Figure 5C**).

Having observed alterations in the proportions of CD56^{dim}CD16$^{+/-}$ subpopulations following NK cell priming, the expression of NKG2D, NKp46, and CD69 on these subpopulations was analyzed in order to determine whether their expression differed between the three subpopulations. As a comparison, the phenotype of IL-2 stimulated NK cells was also determined. As shown in **Figures 5D,E**, compared to resting NK cells, priming significantly decreased the proportion of NK cells within the three CD56^{dim}CD16$^{+/-}$ subpopulations expressing NKG2D and NKp46. In contrast, only the proportion of the CD56^{dim}CD16^{neg} NK cells expressing NKp46 was decreased following IL-2 stimulation, whereas the proportion of all three CD56^{dim}CD16$^{+/-}$ subpopulations expressing NKG2D was increased by IL-2 stimulation compared to resting NK cells. When resting, the median proportion of the three CD56^{dim}CD16$^{+/-}$ NK cell subpopulations expressing CD69 ranged between 73.6 and 79.7% (**Figure 5F**). IL-2 stimulation significantly increased the proportion of NK cells within the three subpopulations expressing CD69 (median 94.4% to 97.1%). Interestingly, priming significantly upregulated the expression of CD69 on only a proportion of NK cells within the CD56^{dim}CD16^{low} and CD56^{dim}CD16^{neg} NK cell subpopulations. The median expression for the two subpopulations was 84.9 and 89.7%, respectively (**Figure 5F**). Despite observing alterations to the proportion of CD56^{dim}CD16$^{+/-}$ subpopulations expressing CD69 following activation by either method, only the IL-2 stimulated NK cell subpopulations exhibited a significant increase in intensity of CD69 expression (**Supplementary Figure 2**).

Influence of CTV-1 Priming and IL-2 Activation on the Expression of TNF Receptors and the CD107a Degranulation Receptor and Their Differential Effects on the Expression of DNAM-1 and CD96

In addition to determining the influence of CTV-1 priming on the expression of NK cell activating receptors and co-receptors that have previously been described by Lowdell et al. (19, 20),

FIGURE 3 | Correlations between phenotype and cytotoxic function of resting NK cells. NK cells were isolated from thawed PBMCs that had previously been generated from the peripheral blood of patients with benign prostate disease and patients with prostate cancer and rested over 17 h. NK cell phenotype and cytotoxic function against K562 cells were subsequently analyzed using flow cytometry. Correlations between phenotype and function were evaluated using non-parametric Spearman Rank correlation matrix. Graphical representations of selected significant correlations are shown; **(A,B)** % of K562 lysis vs. % of NK cells expressing NKp46 and CD69, respectively, **(C,D)** % of NK cells expressing CD69 vs. those expressing CD137 and GITR, respectively. Green dots indicate patients whose NK cells responded better to IL-2 stimulation. Blue dots indicate patients whose NK cells responded better to being primed. Black dots indicate patients for which only priming experiments were undertaken due to low NK cell numbers.

their influence on the expression of additional receptors was examined. Lowdell et al. have previously shown that the priming of NK cells using CTV-1 cells requires cell-to-cell contact (19). We therefore decided to measure the expression of CD107a on the surface of the primed NK cells. Since the primed NK cells used for phenotyping came from the same pool of cells as those used for the functional assays, we did not include monensin during the 17-h co-incubation between patient-derived NK cells and CTV-1 cells.

IL-2 stimulation and CTV-1 priming significantly increased the proportion of all three $CD56^{dim}CD16^{+/-}$ NK cell subpopulations expressing CD107a. However, the proportion of primed $CD56^{dim}CD16^{low}$ and $CD56^{dim}CD16^{neg}$ NK cell subpopulations expressing CD107a was significantly higher than their IL-2 stimulated counterparts (**Figure 6A**). A significant increase in the median fluorescence intensity (MFI) of CD107a expression was also only observed for primed $CD56^{dim}CD16^{low}$ and $CD56^{dim}CD16^{neg}$ subpopulations. IL-2 stimulation did not significantly increase their intensity of CD107a expression (**Supplementary Figure 3A**). Interestingly, it was observed that proportionally, $CD56^{dim}CD16^{neg}$ NK cells were more able to

up-regulate the CD107a receptor, whereas the $CD56^{dim}CD16^{high}$ subpopulation were the least able to up-regulate the CD107a receptor (**Figure 6A**). This was true for both primed and IL-2 stimulated NK cells.

Tumor necrosis factor (TNF) receptors such as OX40, CD137, and GITR are up-regulated on the surface of activated NK cells. Although well studied in T cell biology, little is known about the role these receptors play in NK cell immunity (23). We wanted to observe whether these receptors are up-regulated on CTV-1 primed NK cells and play a role in the triggering of NK cell cytotoxic responses. Resting patient-derived NK cells expressed very little CD137 (median 1.29%, range 0.23% to 6.91%) and OX40 (median 1.05%, range 0.11% to 5.68%) in contrast to expression of GITR (median 26% range 3.63% to 95%). Priming significantly increased the proportion of $CD56^{dim}CD16^{low}$ and $CD56^{dim}CD16^{neg}$ subpopulations expressing the three TNF receptors. This contrasts with IL-2 stimulation which significantly upregulated expression of the three TNF receptors on each $CD56^{dim}CD16^{+/-}$ subset (**Figures 6B–D**). Similar to the up-regulation of CD107a, the $CD56^{dim}CD16^{neg}$ subpopulation exhibited the greatest proportion of cells expressing the TNF

FIGURE 4 | Comparing the capacity of primed and IL-2 activated NK cells to lyse K562 cells and PC3 cells. NK cells were isolated from thawed PBMCs generated from the peripheral blood of patients with benign prostate disease and patients with prostate cancer. For each patient, the NK cells were either primed with mitomycin C treated CTV-1 cells at a 1:2 ratio or activated with 100 IU IL-2, after which their ability to lyse K562 cells and PC3 cells at a 5:1 effector to target ratio was measured using flow cytometry. As assessed by Spearman Rank correlations, the ability of activated NK cells (primed or IL-2 stimulated) to lyse K562 cells compared to PC3 cells are shown **(A)** direct comparison between % of lysed K562 cells vs. % of lysed PC3 cells by primed NK cells **(B)** change in lysis of K562 cells vs. change in lysis of PC3 cells by NK cells upon priming, **(C)** direct comparison between % of lysed K562 cells vs. % of lysed PC3 cells by IL-2 (100 IU) activated NK cells, **(D)** change in lysis of K562 cells vs. change in lysis of PC3 cells by NK cells after IL-2 activation. Green dots indicate patients whose NK cells responded better to IL-2 activation. Blue dots indicate patients whose NK cells responded better to being primed. Black dots indicate patients for which only priming experiments were done due to low NK cell numbers.

receptors following NK cell activation by either method. Again, proportionally the $CD56^{dim}CD16^{high}$ subpopulation was the least able to up-regulate the TNF receptors. Furthermore, in the context of primed NK cells, that on average exhibited the greatest up-regulation of both CD137 and CD107a, co-expression of these two markers on the same primed NK cell was rarely observed. However, it should be noted that the majority of primed NK cells did not express either marker (**Supplementary Figure 4**). Due to the antibody panel design it was not possible to directly assess the co-expression of CD107a with OX40 and GITR.

Expression of the adhesion receptors DNAM-1 and CD96 (TACTILE) was also determined. It has been proposed that DNAM-1 is involved in the recognition stage of forming an immunological synapse and ligates with the nectin-like protein CD155 on the surface of the target cell (24, 25). CD96 is thought

to compete with DNAM-1, along with the inhibitory receptor TIGIT, for ligation with CD155 (25). CTV-1 priming upregulated DNAM-1 expression by the $CD56^{dim}CD16^{neg}$ subpopulation, both in terms of the proportion of NK cells expressing the receptor (**Figure 6E**) and the intensity of expression (**Supplementary Figure 2B**). Priming also significantly increased the intensity of DNAM-1 expression on $CD56^{dim}CD16^{low}$ NK cells (**Supplementary Figure 2B**). In comparison, IL-2 stimulation only increased the intensity of CD96 expression on the $CD56^{dim}CD16^{high}$ and $CD56^{dim}CD16^{low}$ subpopulations (**Supplementary Figure 2B**). The change in expression of CD96 following CTV-1 priming was different to that of DNAM-1. The proportion of NK cells expressing CD96 within all three $CD56^{dim}CD16^{+/-}$ subpopulations decreased after priming, with the intensity of CD96 expression only decreasing

FIGURE 5 | Influence of priming and IL-2 activation on CD16 expression by CD56^{dim} NK cells and the expression of NKG2D, NKp46, and CD69 expression by CD56^{dim} NK cell subpopulations expressing different levels of CD16. For each patient, isolated NK cells were either primed with mitomycin C treated CTV-1 cells at a 1:2 ratio or activated with 100 IU IL-2, after which their expression of CD16, NKG2D, NKp46, and CD69 was measured by flow cytometry and compared to that of resting NK cells. **(A,B)** Strategies for the gating of live, single cell, NK subpopulations based on CD56, and CD16 expression for both resting and primed NK cells, respectively. **(C)** Box and whisker plot comparing the proportion of CD56^{dim}CD16^{high}, CD56^{dim}CD16^{low}, CD56^{dim}CD16^{neg} populations before and after priming, and IL-2 activation. For each CD56^{dim} subpopulation before and after NK cell activation, box and whisker plots analyzing the proportion of NK cells expressing **(D)** NKG2D, **(E)** NKp46, and **(F)** CD69. Statistical analysis was performed using two way ANOVA repeated measure tests combined with Tukey's multiple comparisons test and 95% confidence intervals. *< 0.05, **< 0.01, ***< 0.001, ****< 0.0001.

on the CD56^{dim}CD16^{neg} subpopulation (**Figure 6F** and **Supplementary Figure 3E**). In contrast, IL-2 stimulation only significantly upregulated the intensity of CD96 expression on the CD56^{dim}CD16^{low} and CD56^{dim}CD16^{neg} NK cell subpopulations (**Supplementary Figure 3E**).

Correlations Between Changes in NK Cell Phenotype and Changes in Cytotoxic Function Following CTV-1 Priming and IL-2 Activation

Lowdell et al. have proposed that the NK cell cytotoxic mechanism can be split into two stages; priming and triggering

(16, 19). It was suggested that the CD69 receptor acts as a triggering receptor, but that more triggering receptors exist (19). In an attempt to assign potential "priming" and "triggering" attributes to the NK cell receptors that were measured in this study, for each patient we calculated (1) the change in expression of all receptors measured and (2) the change in lysis of K562 and PC3 cells, following NK cell activation and performed a series of non-parametric correlation matrices using data on receptor expression on total NK cells (**Supplementary Figure 5**) and data on receptor expression by the three CD56^{dim}CD16$^{+/-}$ subpopulations (**Supplementary Figure 6**). Key significant correlations from these matrices were selected and XY scatter plots created in

FIGURE 6 | Influence of priming and IL-2 activation on the phenotype of CD56^{dim} NK cell subpopulations. For each patient, isolated NK cells were either primed with mitomycin C treated CTV-1 cells at a 1:2 ratio or activated with 100 IU IL-2, after which the expression of CD107a, CD137, OX40, GITR, DNAM-1, and CD96 by CD56^{dim}CD16^{high}, CD56^{dim}CD16^{low}, CD56^{dim}CD16^{neg} NK subpopulations was measured by flow cytometry and compared to that of resting NK cells. Box and whisker plots show the proportion of each NK cell subpopulation expressing **(A)** CD107a, **(B)** CD137, **(C)** OX40, **(D)** GITR, **(E)** DNAM-1, **(F)** CD96. Statistical analysis was performed using two way ANOVA repeated measure tests combined with Tukey's multiple comparisons test and 95% confidence intervals. $*< 0.05$, $**< 0.01$, $***< 0.001$, $****< 0.0001$.

order to better observe and interpret these correlations from a biological point of view. We color coded the points to identify patients who preferentially responded to IL-2 stimulation for the lysis of PC3 cells (green) and patients who preferentially responded to CTV-1 priming (blue). Data from patients from whom too few NK cells were recovered to perform both the priming and IL-2 stimulation experiments are colored black.

Only the expression of the activating receptors NKp46 and DNAM-1 correlated with the lysis of K562 and PC3 cells (**Figure 7**). A decrease in the proportion of NK cells expressing NKp46, as result of CTV-1 priming, positively correlated with a decrease in the ability of the NK cells to kill both K562 cells (rs = 0.487, P = 0.0252) and PC3 cells (rs = 0.457, P = 0.0372) (**Figures 7A,B**). Despite

all three $CD56^{dim}CD16^{+/-}$ subpopulations down-regulating NKp46 (**Figure 5E**) following NK cell priming, only changes in the phenotype of the $CD56^{dim}CD16^{low}$ and $CD56^{dim}CD16^{neg}$ populations significantly correlated with the decrease in lysis of both target cells (**Figures 7C–F**). The data revealed a trend that those patients whose NK cells preferentially responded to IL-2 stimulation, thereby enhancing their ability to lyse PC3 cells, tended to down-regulate more NKp46 than the NK cells from patients that preferentially responded to tumor priming (**Figures 7D,F**). Changes in the proportion of NK cells expressing DNAM-1 in the three $CD56^{dim}CD16^{+/-}$ subpopulations as a result of priming were small (<6%) (**Figures 7G,H**). An increase in the proportion of $CD56^{dim}CD16^{low}$ NK cells expressing DNAM-1 positively correlated with the lysis of K562 (rs = 0.567, $P = 0.0073$). In contrast, a decrease in the proportion of $CD56^{dim}CD16^{high}$ NK cells expressing DNAM-1 negatively correlated with an increase in the lysis of PC3 cells (rs = 0.661, $P = 0.0011$).

In contrast to primed NK cells, lysis of K562 and PC3 cells by IL-2 stimulated NK cells did not correlate with the change in expression of any activating receptors (**Figure 8**). Instead, the lysis of K562 cells only positively correlated with an increased proportion of CD69 positive cells (rs = 0.558, $P = 0.0085$) (**Figure 8A**). However, this increase in the proportion of CD69 positive cells negatively correlated with the ability of the IL-2 stimulated NK cells to lyse PC3 cells (rs = −0.506, $P = 0.0193$) (**Figure 8B**). Interestingly, an increase in the proportion of NK cells up-regulating the TNF receptors CD137 and OX40 following stimulation with IL-2 positively correlated with an increase in the ability of the NK cells to kill PC3 cells (rs = 0.484, $P = 0.0261$ and rs = 0.461, $P = 0.0353$, respectively) (**Figures 8C,D**).

NKG2D has been commonly reported to be an important activating receptor for the recognition and killing of cancer cells, including prostate cancer (26–29). However, changes in the proportion of NK cells expressing NKG2D after NK cell priming did not correlate with the ability of the primed NK cells to lyse K562 and PC3 cells in the current study. Instead, proportionally, NKG2D expression positively and significantly correlated with the expression of CD96 and the three TNF receptors (**Figures 9A–D**). The results showed that a priming-induced reduction in the proportion of NK cells expressing NKG2D was associated with a reduction in the proportion of NK cells expressing CD96 (rs = 0.459, $P = 0.0362$) (**Figure 9A**). As previously shown in **Figures 6B–D**, priming up-regulated CD137, OX40, and GITR expression. The correlations in **Figures 9B–D** show that the up-regulation of these TNF receptors on primed NK cells was associated with retaining NKG2D on the surface of primed NK cells. The greater the proportion of primed NK cells down-regulating NKG2D, then the smaller the proportion of NK cells up-regulating CD137 (rs = 0.466, $P = 0.0331$), OX40 (rs = 0.529, $P = 0.0136$), and GITR (rs = 0.529, $P = 0.0136$). Interestingly the expression of CD69 also correlates with the expression of CD96 and the expression of the three TNF receptors. The greater the proportion of primed NK cells down-regulating CD96 expression the lower the proportion of primed NK cells up-regulating CD69 expression (rs = 0.551, P

= 0.0097) (**Figure 9E**). In general, an increased proportion of NK cells up-regulating CD69 expression upon CTV-1 priming was associated with an increased proportion of NK cells expressing CD137 (rs = 0.662, $P = 0.0011$), OX40 (rs = 0.603, $P = 0.0038$), and GITR (rs = 0.902, $P < 0.0001$) (**Figures 9F–H**). However, it was noted that an up-regulation in the expression of all three receptors also occurred in the absence of an up-regulation in the expression of CD69. Considering that the up-regulation of the three TNF receptors on NK cells correlates with CTV-1 priming and the expression of CD69, it was not surprising that expression of the three TNF receptors positively, and significantly correlated with each other (**Figures 10A–C**). However, only the expression of GITR positively correlated with the expression of CD96 (**Figure 10D**). In general, the greater the proportion of NK cells down-regulating CD96, then the lower the proportion of NK cells up-regulating GITR (rs = 0.455, $P = 0.0384$).

Although priming resulted in a positive correlation between CD69 expression and TNF receptor expression, IL-2 stimulation resulted in a negative correlation between the expression of CD69 and that of the TNF receptors CD137 and OX40. As shown in **Figures 10E,F**, respectively, an increase in the proportion of NK cells up-regulating CD69 negatively correlated with the proportion of NK cells up-regulating CD137 (rs = −0.710, $P < 0.0001$) and OX40 (rs = −0.565, $P = 0.0076$). Interestingly, NK cells from those patients that already included a high proportion of NK cells expressing CD69 prior to IL-2 stimulation appeared more likely to up-regulate CD137 and OX40. Similarly to primed NK cells, a highly significant positive correlation between the up-regulation of OX40 and CD137 by IL-2 stimulated NK cells was observed (rs = 0.741, $P = 0.0001$) (**Figure 10G**). In contrast to primed NK cells, no significant correlation was observed between the proportion of IL-2 stimulated NK cells expressing GITR and those expressing CD69. However, the proportion of IL-2 stimulated NK cells up-regulating NKG2D did positively correlate with the proportion of NK cells up-regulating GITR (rs = 0.603, $P = 0.0038$) (**Figure 10H**). Overall, it appears that NK cell priming and IL-2 activation differentially regulate CD137, OX40, and GITR expression.

DISCUSSION

NK cells have immunotherapeutic potential for the treatment of cancer due to their natural ability to kill cancerous cells and a number of NK cell-based immunotherapies are now in development (10–15).

One approach has previously been proposed by Lowdell et al. who discovered that CTV-1 cells (ALL cell line) could "prime" NK cells from healthy volunteers and enhance their ability to kill NK cell-resistant cancer cell lines such as the DU145 metastatic prostate cancer cell line (19, 20). Furthermore, Lowdell et al. also noted that NK cells undergo alterations in phenotype upon priming (19, 20). Most notably, primed NK cells down-regulated their expression of activating receptors NKG2D, NKp46, and NKp80 and this was associated with enhanced cytotoxic function (20). This is in contrast to the down-regulation of NK cell activating receptors and resulting inhibition of NK cell cytotoxic

FIGURE 7 | Correlation between changes in expression of NKp46 and DNAM-1 by patient-derived NK cells following priming with changes in their ability to lyse K562 and PC3 cells. Using phenotypic profiles from patients with prostate cancer, a Spearman Rank correlation matrix was performed to identify correlations between the changes in NK cell phenotype and cytotoxic potential after priming at a 1:2 NK:CTV-1 ratio. Correlation between changes in the proportion of NK cells expressing NKp46 and target cell lysis; **(A,B)** killing of K562 and PC3 cells by the total NK cell population, **(C,D)** killing of K562 and PC3 cells by the CD56^{dim}CD16^{low} NK cells, **(E,F)** killing of K562 and PC3 cells by CD56^{dim}CD16^{neg} NK cells. Correlation between changes in the proportion of CD56^{dim}CD16^{low} NK cells expressing DNAM-1
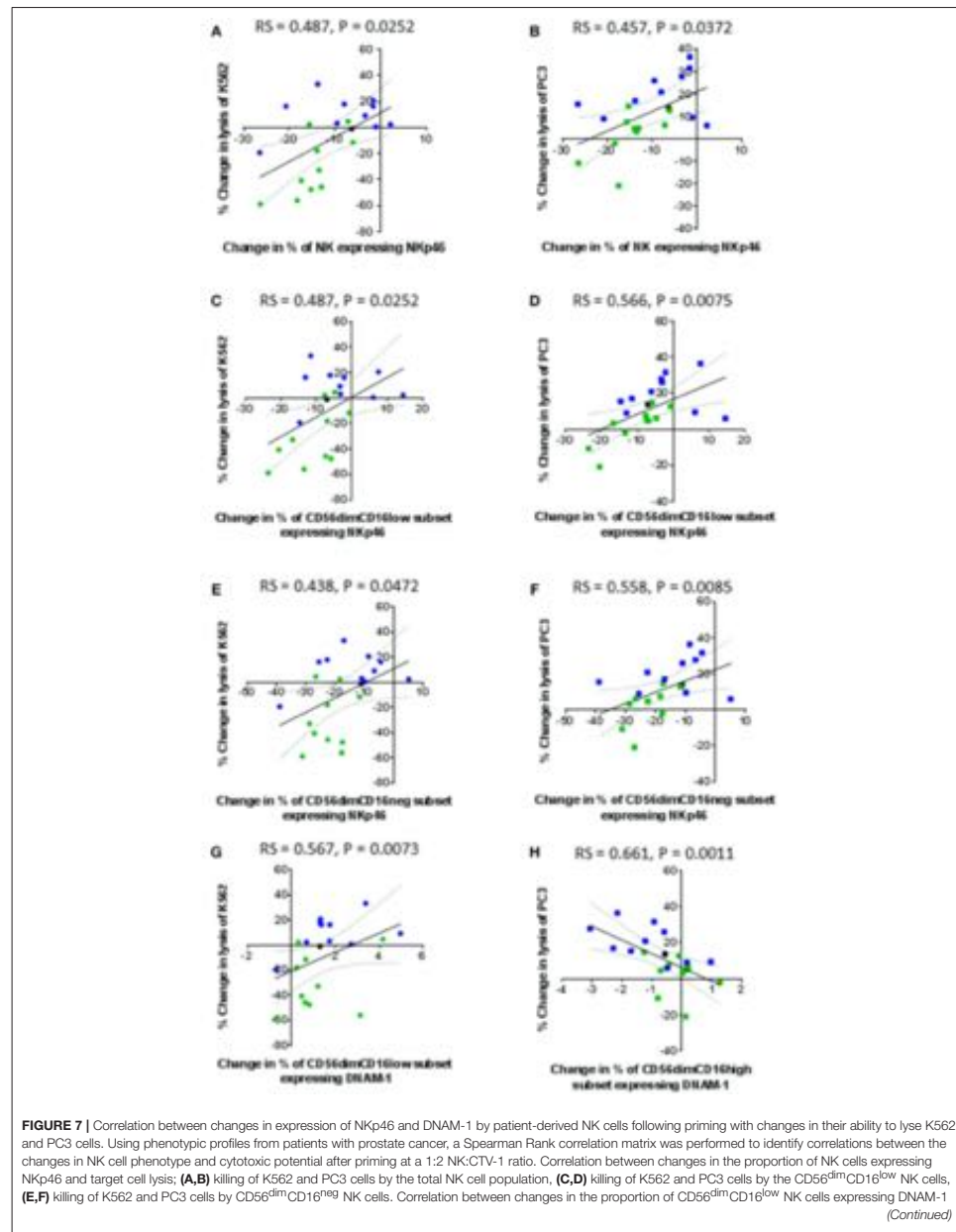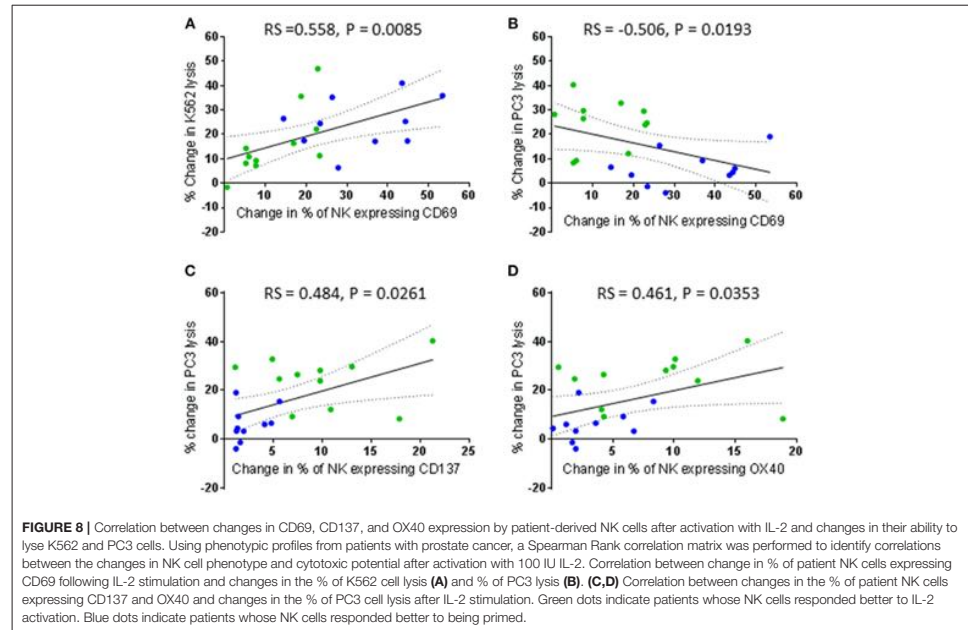*(Continued)*

FIGURE 7 | and change in the lysis of K562 cells **(G)**. Correlation between changes in the proportion of CD56^dimCD16^high NK cells expressing DNAM-1 and change in lysis of PC3 cells **(H)**. Green dots indicate patients whose NK cells responded better to IL-2 activation. Blue dots indicate patients whose NK cells responded better to being primed. Black dots indicate patients for which only priming experiments were done due to low NK cell numbers.
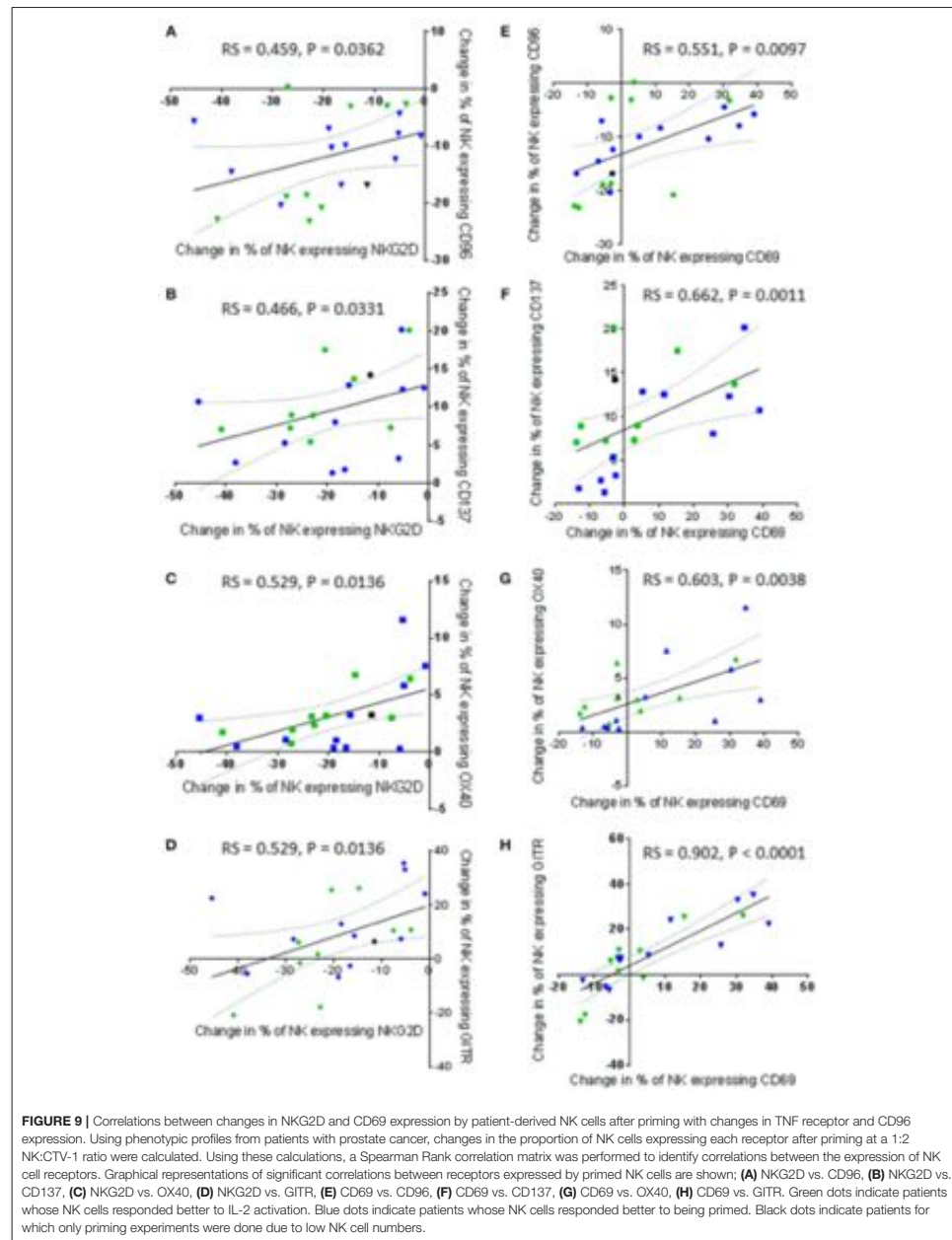

FIGURE 8 | Correlation between changes in CD69, CD137, and OX40 expression by patient-derived NK cells after activation with IL-2 and changes in their ability to lyse K562 and PC3 cells. Using phenotypic profiles from patients with prostate cancer, a Spearman Rank correlation matrix was performed to identify correlations between the changes in NK cell phenotype and cytotoxic potential after activation with 100 IU IL-2. Correlation between change in % of patient NK cells expressing CD69 following IL-2 stimulation and changes in the % of K562 cell lysis **(A)** and % of PC3 lysis **(B)**. **(C,D)** Correlation between changes in the % of patient NK cells expressing CD137 and OX40 and changes in the % of PC3 cell lysis after IL-2 stimulation. Green dots indicate patients whose NK cells responded better to IL-2 activation. Blue dots indicate patients whose NK cells responded better to being primed.

functions which is generally associated with exposure of NK cells to immunosuppressive cytokines (e.g., TGF-β) produced by tumors and suppressive immune cell populations (e.g., Tumor Associated Macrophages, TAMs) (16, 20, 30). Therefore, not only is the priming of NK cells using CTV-1 cells a method of NK cell activation with potential immunotherapeutic application, it may also serve as a model that can be used to improve our understanding of the mechanisms involved in NK cell cytotoxicity. Although the application of CTV-1 primed NK cells in a clinical setting has to date been limited, the adoptive transfer of CTV-1 primed NK cells into humans has shown that primed NK cells can promote durable complete remission in some high risk patients with acute myeloid leukaemia (AML) who were not candidates for hematopoietic cell transplantation (31).

The aim of this study was to compare and contrast the influence of CTV-1 priming and IL-2 activation of NK cells, an approach which has been shown to have immunotherapeutic potential in a number of settings (14, 15), from patients with prostate cancer on their ability to kill the NK cell-resistant cell line PC3 which is considered to represent an aggressive form of metastatic prostate cancer (32). We also

wanted to understand how changes in the phenotype of primed and IL-2 activated NK cells influence their cytotoxic function.

Irrespective of disease status (i.e., benign, low grade cancer or high grade cancer) our data showed that patient-derived NK cells respond to being primed with CTV-1 cells and that this triggered an increase in their ability to lyse PC3 prostate cancer cells (up to 37%). This increased ability to lyse PC3 cells was comparable to that of IL-2 stimulation. Importantly, the capacity of priming and IL-2 stimulation to trigger cytotoxicity was patient-dependent, in that NK cells from ∼50% of patients preferentially responded to CTV-1 priming whereas NK cells from the remaining patients preferentially responded to IL-2 stimulation. Such patient-specific responsiveness has been observed in other immunotherapeutic settings such as the use of checkpoint inhibitors (33, 34) and reiterates the need for the development of "companion diagnostics" that can identify those patients that will benefit from a defined immunotherapy (35).

In general, CTV-1 priming and IL-2 stimulation enhanced NK cell cytotoxicity against PC3 cells. However, although IL-2 stimulation always enhanced K562 lysis, priming did not. K562

FIGURE 9 | Correlations between changes in NKG2D and CD69 expression by patient-derived NK cells after priming with changes in TNF receptor and CD96 expression. Using phenotypic profiles from patients with prostate cancer, changes in the proportion of NK cells expressing each receptor after priming at a 1:2 NK:CTV-1 ratio were calculated. Using these calculations, a Spearman Rank correlation matrix was performed to identify correlations between the expression of NK cell receptors. Graphical representations of significant correlations between receptors expressed by primed NK cells are shown; **(A)** NKG2D vs. CD96, **(B)** NKG2D vs. CD137, **(C)** NKG2D vs. OX40, **(D)** NKG2D vs. GITR, **(E)** CD69 vs. CD96, **(F)** CD69 vs. CD137, **(G)** CD69 vs. OX40, **(H)** CD69 vs. GITR. Green dots indicate patients whose NK cells responded better to IL-2 activation. Blue dots indicate patients whose NK cells responded better to being primed. Black dots indicate patients for which only priming experiments were done due to low NK cell numbers.

FIGURE 10 | Correlations between expression of TNF receptors, NKG2D, CD69, and CD96 by patient-derived NK cells after priming or IL-2 activation. Using phenotypic profiles from patients with prostate cancer, changes in the proportion of NK cells expressing each receptor after priming or activation with 100 U/ml IL-2 were calculated. Using these calculations, Spearman Rank correlation matrices for primed NK cells and IL-2 stimulated NK cells were used to identify correlations in expression of NK cell receptors. Graphical representations of significant correlations between pairs of receptors expressed by primed NK cells are shown; **(A)** CD137 vs. GITR, **(B)** GITR vs. OX40, **(C)** CD137 vs, OX40, **(D)** GITR vs. CD96. Graphical representations of significant correlations between pairs of receptors expressed
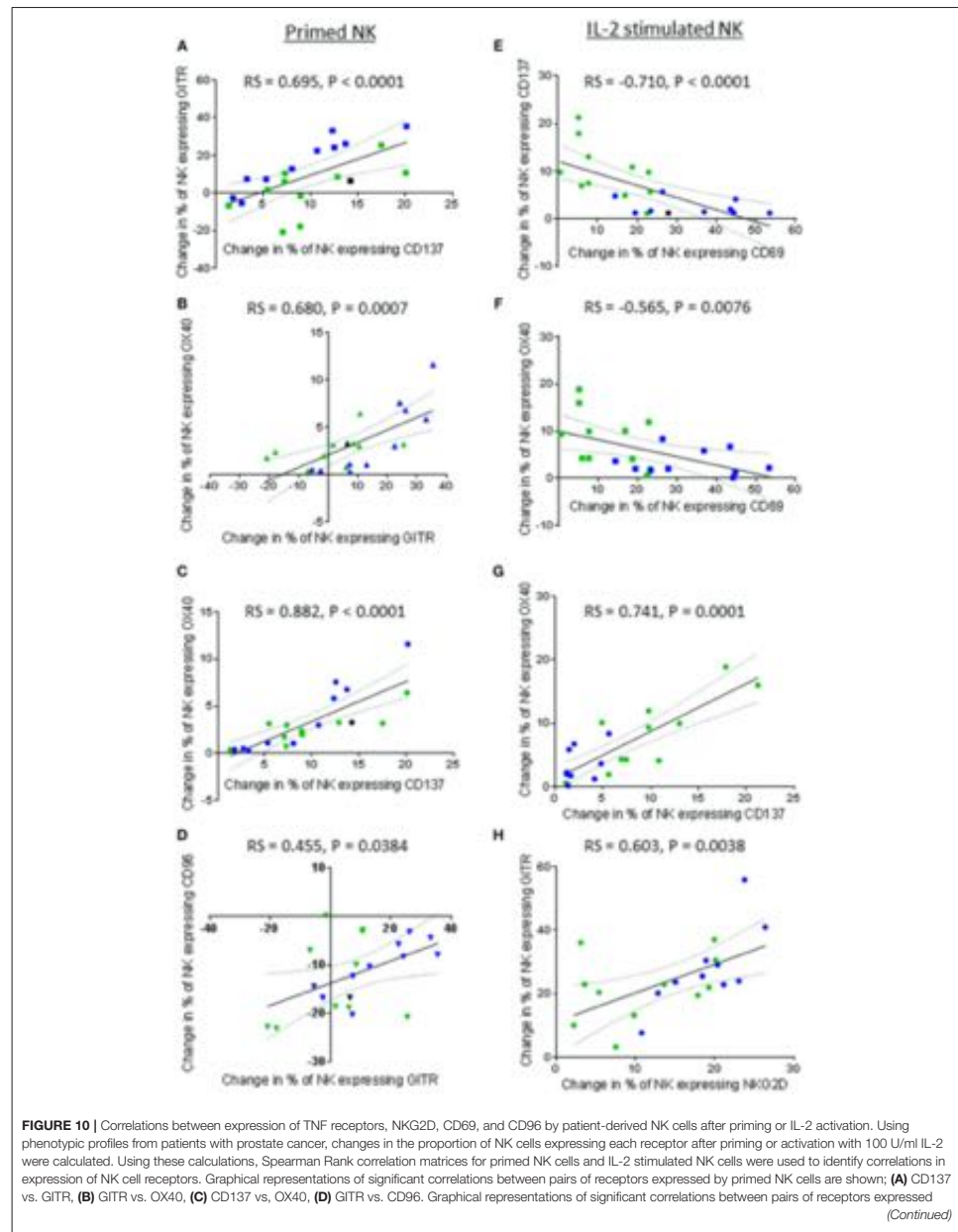*(Continued)*

FIGURE 10 | by IL-2 stimulated NK cells are shown; **(E)** CD69 vs. CD137, **(F)** CD69 vs. OX40, **(G)** CD137 vs. OX40, **(H)** NKG2D vs. GITR. Green dots indicate patients whose NK cells responded better to IL-2 activation. Blue dots indicate patients whose NK cells responded better to being primed. Black dots indicate patients for which only priming experiments were done due to low NK cell numbers.

cells are MHC class I negative cells and are thus highly susceptible to lysis by NK cells as they do not provide MHC class I-triggered inhibitory signals to the NK cell. Yet despite this, NK cells from only 7 of the 22 patients in our study were able to lyse >56% of the K562 cells within the timeframe of the cytotoxicity assay, whereas NK cells from the remaining patients lysed <24%. Although it might be assumed that NK cells from these patients were exhibiting a degree of cytotoxic dysfunction, our data suggest that this is not the case, as the NK cells from these patients responded better to being primed by CTV-1 cells, thereby enabling them to increase their lysis of PC3 cells to greater extent than when they were stimulated with IL-2. In contrast, NK cells from the seven patients that exhibited a high level of K562 lysis at rest responded poorly to being primed with CTV-1 cells.

Overall our data questions current understanding of NK cell dysfunction. It appears that NK function is determined by (1) the composition and phenotypes of the NK cell populations at rest, (2) the type of stimulus used to activate the NK cell (i.e., cytokine or a combination of membrane bound ligands), and (3) the effect this stimulus has on the NK cell phenotype which then determines their ability to lyse subsequent targets. Our data suggest that NK cells need the correct stimulation to achieve a functional response and a failure to respond to one target (or stimulus) does not necessarily indicate a failure to respond to another.

Our observations support the notion that the NK cell cytotoxic mechanism can indeed be divided into the two stages; "priming" and "triggering," first postulated by Lowdell et al. (16, 19). Our data suggest that only a proportion of the NK cells can respond to an activation stimulus and this is likely due to the vast number of different NK cell subpopulations that exist in one individual, with each subpopulation expressing a different combination of receptors, as has been reported by Horowitz et al. (36). Although the NK cell activating and inhibitory receptor repertoire appears to be large, NK cells do not appear to express all the receptors at the same time. Therefore, when NK cells are stimulated with a cytokine (e.g., IL-2) or come into contact with a target cell expressing a specific combination of ligands, only a proportion of the NK cells are capable of responding. IL-2 stimulation had no effect on the expression of CD16, but did increase the proportion of NK cells expressing NKG2D, CD69, CD107a, CD137, OX40, GITR, and also the intensity of CD96 expression. However, it should be noted that the correlation between the expression of CD69 and the two TNF receptors CD137 and OX40 was negative, which is in contrast to the positive correlation exhibited by primed NK cells. In the case of priming, down-regulation of CD16 appeared to be an important indicator of response to stimulus, although not exclusively as a small proportion of the CD56^{dim}CD16^{high} population upregulated CD107a expression which is indicative of degranulation. Similar to interactions between NK cells and K562 cells, cell-to-cell contact between NK

cells and CTV-1 cells has been reported to down-regulate or lead to the shedding of CD16 (19, 37–39). We found that the shedding of CD16 on a proportion of NK cells following priming with CTV-1 cells decreases the proportion of CD56^{dim}CD16^{high} NK cells in the NK cell population which coincided with an increase in the proportion of CD56^{dim}CD16^{low} or CD56^{dim}CD16^{neg} NK cells. This observation is analogous to that observed by Jewett et al. who observed that loss of CD16 expression by NK cells following exposure to K562 cells was only observed on those NK cells that could form conjugates with the NK cells (37). In our study, analysis of the three CD56^{dim}CD16$^{+/-}$ subpopulations following priming with CTV-1 cells revealed a link between the extent of CD16 shedding and the up-regulation of the activation marker CD69 similar to that observed by Jewett et al. (37). Additionally, we observed the up-regulation of three TNF receptors (CD137, OX40, GITR) and the up-regulation of the degranulation receptor CD107a. Compared to resting NK cells, both CD56^{dim}CD16^{low} NK cells and CD56^{dim}CD16^{neg} NK cells significantly upregulated all five receptors. Proportionally, a greater percentage of NK cells within the CD56^{dim}CD16^{neg} subpopulation expressed the CD107a and TNF receptors compared to the CD56^{dim}CD16^{low} subpopulation. In contrast, the CD56^{dim}CD16^{high} subpopulation only significantly up-regulated CD107a. The extent of TNF receptor up-regulation on primed NK cells positively correlated with the up-regulation of CD69, which itself correlated positively with the expression of CD96. Interestingly, only the CD56^{dim}CD16^{low} and CD56^{dim}CD16^{neg} subpopulations significantly up-regulated CD69 expression, therefore explaining the absence of TNF receptor up-regulation on the CD56^{dim}CD16^{high} subpopulation.

To put our observations into context with the literature, our data suggest that successful priming of NK cells by tumor cells involves the ligation of multiple NK cell activating receptors, two of which appear to be NKG2D and CD96. Interestingly the expression of NKG2D and CD96 positively correlates with each other, with the down-regulation of one being associated with the down-regulation of the other. Retention of CD96 expression at the NK cell surface is important for NK cell activation, as measured by CD69 expression. The retention of NKG2D and the up-regulation of CD69 at the cell surface of primed NK cells appears to be important for the up-regulation of the TNF receptors. It is unclear whether the down-regulation of the activating receptors that occurs with priming to a degree is the result of ligation with ligands expressed on the CTV-1 cell surface or due to exposure of immunosuppressive cytokines secreted by CTV-1 cells, despite them having been treated with mitomycin C. The data suggest that both NKG2D and CD96 are required for priming NK cells, but not for the triggering of cytotoxic granule release. In contrast, the activating receptor NKp46 showed no association with NK cell activation, but did appear to correlate with target cell lysis of both K562 and PC3

cells, thereby suggesting a role in the triggering of cytotoxic responses. Strangely, CD107a did not positively correlate with K562 and PC3 lysis. This may be due to the fact that we did not use monensin to retain CD107a expression on the surface of the primed NK cells, which is common practice in indirect cytotoxic killing assays (40). Although monensin was not used, we still observed CD107a expression on the surface of NK cells 17 h post co-incubation with CTV-1 cells, thereby suggesting that NK cells continue to kill CTV-1 cells over this length of time. Interestingly, following contact with CTV-1 cells over this 17 h period, which for some primed NK cells resulted in a cytotoxic response, the same primed NK cells could lyse a metastatic prostate cancer cell line which is typically resistant to lysis by resting NK cells. At this point we do not know whether it is the same primed NK cells lysing both CTV-1 and PC3 cells. A study by Jewett et al. revealed that NK cells dissociated from MHC class I deficient K562 cells following initial conjugation display anergy resulting in decreased cytotoxic function due to a reduced ability to form conjugates (37). Subsequent work by this group revealed that the NK cells that formed conjugates could be further subdivided into "binders" and "killers". The "binders" and "killers" displayed "split anergy" with "binder" NK cells forming conjugates for a longer period of time compared to "killers." As a result, "binders" suffered from target induced inactivation and induction of apoptosis (39). Although we did not specifically set out to measure conjugation between NK cells and CTV-1 cells, we did observe a reduction in the number of viable NK cells as a result of priming (data not shown) which suggests that conjugates between NK cells and CTV-1 cells may also promote a split anergy. However, in contrast to K562 cells, CTV-1 cells enhance the ability of NK cells to kill PC3 cells, suggesting that CTV-1 (MHC class I positive) cells do not impair the ability of primed NK cells to form conjugates with subsequent target cells.

Enhanced cytotoxic function by NK cells following initial exposure to acute lymphoblastic leukemia cells has been observed in other studies. A study by Pal et al. revealed that upon primary exposure to acute B cell precursor leukemic cell lines, healthy NK cells acquired a mode of functional memory enabling them to enhance their cytotoxic capacity against the same cell line upon secondary exposure (41). However, in contrast to the CD56^{dim}-based phenotype of the CTV-1 primed NK cells described in this study, the tumor induced memory-like NK cells described by Pal et al. were instead associated with the CD56^{bright} subpopulation. NK cells appear to be relatively "plastic" and their immune response depends on the extent and type of external stimuli they receive.

To fully realize the immunotherapeutic potential of CTV-1 primed NK cells, further characterization of their phenotype, function, and the mechanisms involved in their generation is required (42). Current immunotherapeutic strategies have shown little or no efficacy for the treatment of prostate cancer and therefore new alternative strategies need to be explored (9, 43). This study supports a role for stratified NK cell-based therapeutics in the prostate cancer setting and warrants further investigation. However, it should be noted that one limitation of this study was the use of only one allogenic

metastatic prostate cancer cell line (i.e., PC3 cells) as a target for primed and IL-2 activated NK cells *in vitro* and an inability to undertake profiling of cytokine responses following priming and activation due to limited sample availability. Further assessment of cytotoxic responses against multiple metastatic prostate cancer cell lines and/or primary prostate cancer cell lines are needed and, ideally, these would be combined with an assessment of "triggering" receptor ligand interactions in order to further interrogate the efficacy of using CTV-1 primed and IL-2 activated NK cell populations for the treatment of prostate cancer.

Our study also highlights an area for further investigation which concerns the role of TNF receptors in NK cell biology. A limitation of the current study is that the analysis was limited to the co-expression of CD137 and CD107a. The data suggested that these two receptors are rarely co-expressed and therefore CD137 does not appear to be associated with the triggering of NK cell cytotoxic responses. However, since we did not use monensin in our experiments to retain CD107a expression at the cell surface, we cannot completely rule out this possibility. TNF receptors act as co-stimulatory receptors that provide bidirectional signaling between effector cells and their targets (23). Currently, the majority of information regarding TNF receptors has been derived from T cell studies, although information regarding their function in the setting of NK cells is beginning to emerge. It has been shown that ligation of OX40 with OX40L expressed on activated T cells and activated monocytes promotes NK cell proliferation (44, 45). Ligation of GITR with GITRL secreted by tumor cells or with agonistic anti-GITR antibodies results in down-regulation of NK cell cytotoxic responses, proliferation and IFN-γ production, while promoting NK cell apoptosis (21, 46). Ligation of CD137 with CD137L expressed on AML cells from patients also reduced NK cell cytotoxicity and IFN-γ production. Blocking CD137 and CD137L interactions restored NK cell cytotoxicity, but not IFN-γ production (22). In our study, co-incubation of NK cells with CTV-1 cells induced the expression of all three TNF receptors, thus representing an opportunity not only to investigate these receptors, but also to further manipulate the primed NK cell response in a therapeutic setting using monoclonal antibodies that target these receptors. Such antibodies are currently in development or/and undergoing clinical trials (47, 48).

In summary, this study presents a comprehensive, comparative analysis of changes in NK cell phenotype and function following priming by CTV-1 cells and activation by IL-2 and new insights into the correlation between changes in phenotype and function. Although the findings demonstrate that the priming and activation of NK cells has potential as an immunotherapy for the treatment of prostate cancer, it also shows that not all patients will benefit from a particular therapeutic approach. Our interrogation of the consequences of priming and activation on NK cell biology provides an opportunity to predict and optimize therapeutic potential. Our findings also confirm that the targeting of TNF receptors and other pathways using monoclonal antibodies may further enhance the cytotoxic potential of NK cells by blocking

NK cell-directed inhibitory signals, as has been discussed previously (15).

AVAILABILITY OF DATA AND MATERIALS

All relevant data generated or analyzed during this study are included in this published article and its **Supplementary Information Files**.

AUTHOR CONTRIBUTIONS

SH, GF, MK, and AP contributed to the conception and design of the study. MK provided access to clinical samples and clinical data. SH, GF, and SM contributed to the development of methodology. SR processed the clinical samples and performed the peripheral blood mononuclear cell extractions. HI and GF performed the flow cytometry experiments and acquired the data. SH, GF, and SM analyzed and interpreted the flow cytometry data. SH, GF, MK, and AP prepared and revised the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2018.03169/full#supplementary-material

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* (2015) 136:E359–86. doi: 10.1002/ijc.29210
2. WHO. *United Kingdom Cancer Country Profile.* World Health Organisation (2014). Available online at: http://www.who.int/cancer/country-profiles/gbr_en.pdf?ua=1
3. D'Amico AV, Whittington R, Malkowicz S, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* (1998) 280:969–74. doi: 10.1001/jama.280.11.969
4. Hernandez DJ, Nielsen ME, Han M, Partin AW. Contemporary evaluation of the D'Amico risk classification of prostate cancer. *Urology* (2007) 70:931–5. doi: 10.1016/j.urology.2007.08.055
5. Kalina JL, Neilson DS, Comber AP, Rauw JM, Alexander AS, Vergidis J, et al. Immune modulation by androgen deprivation and radiation therapy: implications for prostate cancer immunotherapy. *Cancers* (2017) 9:13. doi: 10.3390/cancers9020013
6. Farkona S, Diamandis EP, Blasutig IM. Cancer immunotherapy: the beginning of the end of cancer? *BMC Med.* (2016) 14:73. doi: 10.1186/s12916-016-0623-5
7. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, et al. Sipuleucel-T Immunotherapy for castration-resistant prostate cancer. *N Engl J Med.* (2010) 363:411–22. doi: 10.1056/NEJMoa1001294
8. Kwon ED, Drake CG, Scher HI, Fizazi K, Bossi A, van den Eertwegh, AJ, et al. Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. *Lancet Oncol.* (2014) 15:700–12. doi: 10.1016/s1470-2045(14)70189-5
9. Modena A, Ciccarese C, Iacovelli R, Brunelli M, Montironi R, Fiorentino M, et al. Immune checkpoint inhibitors and prostate cancer: a new frontier? *Oncol Rev.* (2016a) 10:293. doi: 10.4081/oncol.2016.293
10. Trinchieri G. Biology of natural killer cells. *Adv Immunol.* (1989) 47:187–376
11. Miller JS. Therapeutic applications: natural killer cells in the clinic. *Hematol Am Soc Hematol Educ Program* (2013) 2013:247–53. doi: 10.1182/asheducation-2013.1.247
12. Specht HM, Ahrens N, Blankenstein C, Duell T, Fietkau R, Gaipl US, et al. Heat shock protein 70 (Hsp70) peptide activated natural killer (NK) cells for the treatment of patients with non-small cell lung cancer (NSCLC) after radiochemotherapy (RCTx) - from preclinical studies to a clinical phase II trial. *Front Immunol.* (2015) 6:162. doi: 10.3389/fimmu.2015.00162

13. Shevtsov M, Multhoff G. Heat shock protein–peptide and HSP-based immunotherapies for the treatment of cancer. *Front Immunol.* (2016a) 7:171. doi: 10.3389/fimmu.2016.00171
14. Chiossone L, Vienne M, Kerdiles YM, Vivier E. Natural killer cell immunotherapies against cancer: checkpoint inhibitors and more. *Semin Immunol.* (2017) 31:55–63. doi: 10.1016/j.smim.2017.08.003
15. Shevtsov M, Multhoff G. Immunological and translational aspects of NK cell-based antitumor immunotherapies. *Front Immunol.* (2016b) 7:492. doi: 10.3389/fimmu.2016.00492
16. Sabry M, Lowdell MW. Tumor-primed NK cells: waiting for the green light. *Front Immunol.* (2013) 4:408. doi: 10.3389/fimmu.2013.00408
17. Bryceson YT, March ME, Ljunggren HG, Long EO. Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. *Blood* (2006) 107:159–66. doi: 10.1182/blood-2005-04-1351
18. Bryceson YT, Ljunggren HG, Long EO. Minimal requirement for induction of natural cytotoxicity and intersection of activation signals by inhibitory receptors. *Blood* (2009) 114:2657–66. doi: 10.1182/blood-2009-01-201632
19. North J, Bakhsh I, Marden C, Pittman H, Addison EG, Navarrete C, et al. Tumor-primed human NK cells lyse NK resistant tumor targets evidence of a two stage process in resting NK cell activation. *Immunology* (2007) 178:85–94. doi: 10.4049/jimmunol.178.1.85
20. Sabry M, Tsirogianni M, Bakhsh IA, North J, Sivakumaran J, Giannopoulos K, et al. Leukemic priming of resting NK cells is killer Ig-like receptor independent but requires CD15-mediated CD2 ligation and natural cytotoxicity receptors. *J Immunol.* (2011) 187:6227–34. doi: 10.4049/jimmunol.1101640
21. Liu B, Li Z, Mahesh SP, Pantanelli S, Hwang FS, Siu WO, et al. Glucocorticoid-induced tumor necrosis factor receptor negatively regulates activation of human primary natural killer (NK) cells by blocking proliferative signals and increasing NK cell apoptosis. *J Biol Chem.* (2008) 283:8202–10. doi: 10.1074/jbc.M708944200
22. Baessler T, Charton JE, Schmiedel BJ, Grunebach F, Krusch M, Wacker A, et al. CD137 ligand mediates opposite effects in human and mouse NK cells and impairs NK-cell reactivity against human acute myeloid leukemia cells. *Blood* (2010) 115:3058–69. doi: 10.1182/blood-2009-06-227934
23. Melero I, Hirschhorn-Cymerman D, Morales-Kastresana A, Sanmamed MF, Wolchok JD. Agonist antibodies to TNFR molecules that costimulate T and NK cells. *Clin Cancer Res.* (2013) 19:1044–53. doi: 10.1158/1078-0432.ccr-12-2065
24. Mace EM, Dongre P, Hsu HT, Sinha P, James AM, Mann SS, et al. Cell biological steps and checkpoints in accessing NK cell cytotoxicity. *Immunol Cell Biol.* (2014) 92:245–55. doi: 10.1038/icb.2013.96
25. Martinet L, Smyth MJ. Balancing natural killer cell activation through paired receptors. *Nat Rev Immunol.* (2015) 15:243–54. doi: 10.1038/nri3799

26. El-Sherbiny YM, Meade JL, Holmes TD, McGonagle D, Mackie SL, Morgan AW, et al. The requirement for DNAM-1, NKG2D, and NKp46 in the natural killer cell-mediated killing of myeloma cells. *Cancer Res.* (2007) 67:8444–9. doi: 10.1158/0008-5472.CAN-06-4230

27. Mamessier E, Sylvain A, Thibult ML, Houvenaeghel G, Jacquemier J, Castellano R, et al. Human breast cancer cells enhance self tolerance by promoting evasion from NK cell antitumor immunity. *J Clin Invest.* (2011) 121:3609–22. doi: 10.1172/JCI45816

28. Morgado S, Sanchez-Correa B, Casado JG, Duran E, Gayoso I, Labella F, et al. NK cell recognition and killing of melanoma cells is controlled by multiple activating receptor-ligand interactions. *J Innate Immun.* (2011) 3:365–73. doi: 10.1159/000328505

29. Pasero C, Gravis G, Granjeaud S, Guerin M, Thomassin-Piana J, Rocchi P, et al. Highly effective NK cells are associated with good prognosis in patients with metastatic prostate cancer. *Oncotarget* (2015) 6:14360–73. doi: 10.18632/oncotarget.3965

30. Vitale M, Cantoni C, Pietra G, Mingari MC, Moretta L. Effect of tumor cells and tumor microenvironment on NK-cell function. *Eur J Immunol.* (2014) 44:1582–92. doi: 10.1002/eji.201344272

31. Fehniger TA, Miller JS, Stuart RK, Cooley S, Salhotra A, Curtsinger J, et al. A Phase 1 trial of CNDO-109-activated natural killer cells in patients with high-risk acute myeloid leukemia. *Biol Blood Marrow Transplant.* (2018) 24:1581–9. doi: 10.1016/j.bbmt.2018.03.019

32. Ravenna L, Principessa L, Verdina A, Salvatori L, Russo MA, Petrangeli E. Distinct phenotypes of human prostate cancer cells associate with different adaptation to hypoxia and pro-inflammatory gene expression. *PLoS ONE* (2014) 9:e96250. doi: 10.1371/journal.pone.0096250

33. Horvat TZ, Adel NG, Dang TO, Momtaz P, Postow MA, Callahan MK, et al. Immune-related adverse events, need for systemic immunosuppression, and effects on survival and time to treatment failure in patients with melanoma treated with ipilimumab at Memorial Sloan Kettering Cancer Center. *J Clin Oncol.* (2015) 33:3193–8. doi: 10.1200/jco.2015.60.8448

34. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med.* (2015) 373:1803–13. doi: 10.1056/NEJMoa1510665

35. Massari F, Santoni M, Ciccarese C, Santini D. The immunocheckpoints in modern oncology: the next 15 years. *Expert Opin Biol Ther.* (2015) 15:917–21. doi: 10.1517/14712598.2015.1035251

36. Horowitz A, Strauss-Albee DM, Leipold M, Kubo J, Nemat-Gorgani N, Dogan OC, et al. Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry. *Sci Transl Med.* (2013) 5:208ra145. doi: 10.1126/scitranslmed.3006702

37. Jewett A, Bonavida B. Target-induced anergy of natural killer cytotoxic function is restricted to the NK-target conjugate subset. *Cell Immunol.* (1995) 160:91–7

38. Grzywacz B, Kataria N, Verneris MR. CD56^{dim}CD16$^+$ NK cells downregulate CD16 following target cell induced activation of matrix metalloproteinases. *Leukemia* (2007) 21:356–9; author reply 359. doi: 10.1038/sj.leu.2404499

39. Jewett A, Bonavida B. Target-induced inactivation and cell death by apoptosis in a subset of human NK cells. *J Immunol.* (1996) 156:907–15

40. Alter G, Malenfant JM, Altfeld M. CD107a as a functional marker for the identification of natural killer cell activity. *J Immunol Methods* (2004) 294:15–22. doi: 10.1016/j.jim.2004.08.008

41. Pal M, Schwab L, Yermakova A, Mace EM, Claus R, Krahl AC, et al. Tumor-priming converts NK cells to memory-like NK cells. *Oncoimmunology* (2017) 6:e1317411. doi: 10.1080/2162402x.2017.1317411

42. Doherty E, Rouce RH. Primed to kill: CTV-1 stimulated haploidentical natural killer cells for consolidation of AML. *Biol Blood Marrow Transplant.* (2018) 24:1533–5. doi: 10.1016/j.bbmt.2018.06.019

43. Quinn DI, Shore ND, Egawa S, Gerritsen WR, Fizazi K. Immunotherapy for castration-resistant prostate cancer: progress and new paradigms. *Urol Oncol.* (2015) 33:245–60. doi: 10.1016/j.urolonc.2014.10.009

44. Pollmann J, Gotz JJ, Rupp D, Strauss O, Granzin M, Grunvogel O, et al. Hepatitis C virus-induced natural killer cell proliferation involves monocyte-derived cells and the OX40/OX40L axis. *J Hepatol.* (2018) 68:421–30. doi: 10.1016/j.jhep.2017.10.021

45. Turaj AH, Cox KL, Penfold CA, French RR, Mockridge CI, Willoughby JE, et al. Augmentation of CD134 (OX40)-dependent NK anti-tumour activity is dependent on antibody cross-linking. (2018) 8:2278. doi: 10.1038/s41598-018-20656-y

46. Baltz KM, Krusch M, Bringmann A, Brossart P, Mayer F, Kloss M, et al. Cancer immunoediting by GITR (glucocorticoid-induced TNF-related protein) ligand in humans: NK cell/tumor cell interactions. *FASEB J.* (2007) 21:2442–54. doi: 10.1096/fj.06-7724com

47. Mentlik James A, Cohen AD, Campbell KS. Combination immune therapies to enhance anti-tumor responses by NK cells. *Front Immunol.* (2013) 4:481. doi: 10.3389/fimmu.2013.00481

48. Chester C, Ambulkar S, Kohrt HE. 4-1BB agonism: adding the accelerator to cancer immunotherapy. *Cancer Immunol Immunother.* (2016) 65:1243–8. doi: 10.1007/s00262-016-1829-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

188 | P a g e