



Improving The Diagnosis And Risk Stratification Of Prostate Cancer

Lina Maria Carmona Echeverria

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy
Division of Surgery and Interventional Science
University College London
June 2020

Acknowledgements

I would like to thank Prostate Cancer UK, Centre of Excellence for funding this work.

I also wish to show my appreciation to the patients that donated their samples that allowed me to complete this project.

I would like to express my gratitude to my supervisors who guided me throughout this project. Dr Hayley Whitaker, for supporting me and helping me believe in my capabilities as a researcher. Prof. Mark Emberton, whose insight and knowledge steered me through this project. Prof. Hashim Ahmed for his support, encouragement and for gently pointing out when I was making no sense. I would also like to thank Dr Hayley Pye; there is not enough space to express how grateful I am for your support and guidance. Thank you to Prof. Moore, for being a role model and always having time for a chat.

Many thanks to my collaborators; Dr Avi Rosenfeld I am extremely grateful for our friendly chats and for sharing your passion for research. Dr Haider, for teaching me pathology is beautiful. Dr Freeman, for saying yes to this time-consuming project and your guidance. Dr Yipeng Hu, for taking the time to listen to my ideas, and the patience to understand what I meant. To my colleagues in the Geography department, Prof. Longley thank you for taking that meeting two years ago to talk about prostate cancer, and to Dr Lansley and Dr Lan for translating ideas into a reality.

I would also like to thank my colleagues who supported me and offered deep insight throughout this project, particularly Urszula Stopka-Farooqui, who taught me so much about planning experiments and troubleshooting. Dr Heavey who for reminding me this is a journey not a destination, Dr Luxton for teaching this clinician how to function in a lab. Mr Simpson, Dr Stavrinides and Dr Norris, thank you for sharing this experience with me. To my Colombian mentor Dr Gomez for allowing me to dip my toes in research all those years ago.

I cannot begin to express my thanks to my family for their support during this PhD. Thank you for celebrating my successes along the way and for always believing I had what it takes. David, thank you for understanding why I had to work late nights while writing this book, and thank you for making me smile at the end of the day. Michal, thank you for being by my side and for the many FaceTime breakfasts we shared.

To my parents, thank you for giving me the tools to fly alone, both of you showed me I could be anything I wanted if I worked hard enough. I am deeply grateful for the sacrifices you made along the way. Jacquie, thank you for being a role model. Capeto and Pipe thank you for teaching me resilience like only an older brother can, and Diego for reminding me coffee is an ally. To Coco for brightening thesis writing during a pandemic.

To my friends, Luz Helena and Stephanie for being close even though we are so far away. And, to Diana for your unconditional support.

This thesis is dedicated to my grandma Clementina, who I lost so many years ago but remains in my memory.

I, Lina Maria Carmona Echeverria, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Signed

Abstract

The current diagnostic and stratification pathway for prostate cancer has led to over-diagnosis and over-treatment. This thesis aims to improve the prostate cancer diagnosis pathway by developing a minimally invasive blood test to inform diagnosis alongside mpMRI and to understand the true Gleason 4 burden which will help better stratify disease and guide clinicians in treatment planning.

To reduce the number of patients who have to undergo prostate biopsy after indeterminate or false positive prostate mpMRI, we aimed to develop a new panel of mRNA detectable in blood or urine that was able to improve the detection of clinical significant prostate cancer (Gleason 4+3 or ≥ 6 mm) in combination with prostate mpMRI. mRNA expression of 28 potential genes was studied in four prostate cancer cell lines and, using publicly available datasets, a new seven gene biomarker panel was developed using machine learning techniques. The signature was then validated in blood and urine samples from the ProMPT, PROMIS and INNOVATE trials.

To redefine the classification of Gleason 4 disease in prostate cancer patients, digital pathology was used to contour and accurately assess the burden and spread of Gleason 4 in a cohort of PROMIS patients compared to the gold standard manual pathology. There was a significant difference between observed and objective Gleason 4 burden that has implications in patient risk stratification and biomarker discovery.

The work presented in this thesis makes a significant step toward improving the patient diagnostic and risk classification pathways by ensuring only the right patients are biopsied when necessary, improving the current pathological reference standard.

Impact Statement

The impact of the work presented in this thesis can be divided into impact to patients, academia and outside academia. The impact to patients includes better risk stratification when diagnosed with prostate cancer, as new categories of risk may be found when objectively measuring the amount of Gleason 4 in prostate biopsies prior to any intervention. Additionally, using technology to standardize Gleason 4 reporting will remove inter-observer variability; this will allow for biomarkers to be refined based on more objective measurements, either confirming or refuting their utility in clinical practice. If proven to help diagnose patients early, these panels could be used in low income countries as a guidance prior to performing prostate biopsies.

In academia, the work presented here has the potential to encourage future collaboration with non-biological sciences, such as Geography. Multidisciplinary thinking involving experts from different backgrounds can help us study cancer from a different point of view and apply research methods we would not otherwise use. Furthermore, the work presented here invites reassessment of the current gold standard of Gleason classification by encouraging use of digital pathology and machine learning; this can streamline the histopathologist work load and improve our understanding of the behavior of prostate cancer. It remains unanswered if the gene expression measured in fluids (urine and blood) presented in this thesis truly corresponds to changes in the tissue. Further work can be done to demonstrate or refute this, using PROMIS samples as discussed in Chapter 6 of this thesis.

Outside academia, the work in this thesis can influence the use of biomarkers in clinical practice. Although the purpose of this work is not to compare the accuracy of other biomarkers compared to the panel described here, the work presented has the potential to encourage better guidelines, to formulate objectives and use of appropriate cohorts for discovering and validating biomarkers. In addition, this work encourages the development of biomarkers to aid or improve mpMRI with an appropriate biopsy sampling method as a baseline.

Finally, part of this work has been presented in local meetings, mainly at the Prostate Cancer UK Making progress meeting. It was also awarded the Faculty of Medical Sciences Dean's prize in 2019, and is currently under peer review for publication. This work has also led to grant applications and it the foundation to my future career as an academic.

Table of contents

| | |
|---|-----------|
| Abstract | 5 |
| Impact Statement | 6 |
| Table of contents | 7 |
| List of Figures | 16 |
| List of tables | 19 |
| List of Appendices | 20 |
| Abbreviations | 21 |
| Chapter 1. Introduction | 25 |
| 1.1. Summary | 25 |
| 1.2. Context/literature review | 26 |
| 1.2.1. Prostate cancer | 26 |
| 1.2.1.1. The prostate | 26 |
| 1.2.1.2. Epidemiology | 28 |
| 1.2.1.3. Androgens and the Androgen Receptor (AR) | 28 |
| 1.2.1.4. The diagnosis of prostate cancer | 29 |
| 1.2.1.4.1. Prostate Specific Antigen (PSA) | 29 |
| 1.2.1.4.2. Prostate Biopsy | 29 |
| 1.2.1.4.3. Multiparametric MRI | 33 |
| 1.2.1.5. Gleason system | 36 |
| 1.2.1.6. Prostate Grade Group | 36 |
| 1.2.1.6.1. The importance of Gleason pattern 4 in prostate cancer | 38 |
| 1.2.1.7. Prostate cancer staging | 38 |
| 1.2.1.7.1. The D'Amico classification | 38 |
| 1.2.1.7.2. The CAPRA-S score | 39 |
| 1.2.1.7.3. The NCCN risk classification | 39 |
| 1.2.2. Biomarkers | 42 |
| 1.2.2.1. Definition and types of biomarkers | 42 |
| 1.2.2.2. Biomarkers for prostate cancer | 43 |
| 1.2.2.2.1. Blood-based biomarkers | 46 |

| | |
|--|-----------|
| 1.2.2.2.2. Urine based biomarkers | 48 |
| 1.2.2.2.3. Tissue based biomarkers | 50 |
| 1.3. The problem with current biomarker panels..... | 53 |
| 1.4. Lay summary..... | 54 |
| Chapter 2. Materials and methods | 56 |
| 2.1. Materials..... | 56 |
| 2.1.1. Cell lines and associated media | 56 |
| 2.1.2. Drugs | 56 |
| 2.1.3. RNA extraction..... | 56 |
| 2.1.4. RNA quality control | 57 |
| 2.1.5. cDNA..... | 57 |
| 2.1.6. Preamplification | 57 |
| 2.1.7. PCR | 58 |
| 2.1.7.1. qPCR..... | 58 |
| 2.1.7.2. Fluidigm high throughput qdPCR | 58 |
| 2.2. Methods | 60 |
| 2.2.1. Cell culture | 60 |
| 2.2.1.1. Maintenance of cell lines..... | 60 |
| 2.2.1.2. Cell harvesting..... | 60 |
| 2.2.1.3. Cryopreservation of cell lines | 60 |
| 2.2.1.4. Retrieval of frozen cell lines | 60 |
| 2.2.1.5. Cell counting..... | 60 |
| 2.2.1.6. Androgen time course | 61 |
| 2.2.1.7. Androgen stimulated cells exposed to urine to simulate RNA degradation caused by urine..... | 61 |
| 2.2.2. RNA extraction..... | 61 |
| 2.2.2.1. Cell line RNA extraction | 62 |
| 2.2.2.2. PAXgene blood RNA extraction | 62 |
| 2.2.2.3. Urine RNA extraction | 62 |
| 2.2.3. RNA quality control | 62 |

| | |
|--|----|
| 2.2.3.1. Nanodrop..... | 62 |
| 2.2.3.2. Bioanalyzer..... | 62 |
| 2.2.3.3. Qubit..... | 63 |
| 2.2.4. Gene based assays | 63 |
| 2.2.4.1. cDNA synthesis..... | 63 |
| 2.2.4.2. qPCR..... | 63 |
| 2.2.4.3. Preamplification..... | 63 |
| 2.2.4.4. Fluidigm qdPCR | 64 |
| 2.2.4.4.1. cDNA | 64 |
| 2.2.4.4.2. Preamplification | 64 |
| 2.2.4.4.3. Fluidigm planning and loading..... | 66 |
| 2.2.5. Digital pathology for measurement of Maximum Cancer Core Length (MCCL) and Gleason 4 burden..... | 66 |
| 2.2.5.1. Scanning of H&E slides..... | 66 |
| 2.2.5.2. Systematic assessment of digital slides..... | 66 |
| 2.2.5.2.1. MCCL measurement in digital images | 66 |
| 2.2.5.2.2. Gleason 4 contouring | 67 |
| 2.2.5.2.3. Calculation of Gleason 4 percentage | 67 |
| 2.2.5.2.4. Data collection | 67 |
| 2.2.5.3. Applying geographical parameters, evenness and clustering to Gleason 4 patterns of distribution. | 69 |
| 2.2.5.3.1. Analysis of location and zone in relation to the evenness and clustering scores. . | 71 |
| 2.2.6. Data analysis | 73 |
| 2.2.6.1. ΔCt and $\Delta\Delta\text{Ct}$ qPCR and qdPCR data analysis | 73 |
| 2.2.6.2. Statistical analysis..... | 73 |
| 2.2.6.2.1. Assessment of normality of data distribution..... | 73 |
| 2.2.6.2.2. Assessment equality of variances. | 73 |
| 2.2.6.2.3. Comparison of means of two groups..... | 74 |
| 2.2.6.2.4. Comparison of means of more than two groups | 74 |
| 2.2.6.2.5. Time course analysis..... | 74 |
| 2.2.6.2.6. Bland-Altman test | 74 |

| | |
|--|-----------|
| 2.2.6.3. Missing data analysis | 75 |
| 2.2.6.3.1. Multiple imputation for missing values..... | 75 |
| 2.2.6.4. Data mining | 76 |
| 2.2.6.5. Confusion matrix and calculation of accuracy parameters. | 76 |
| 2.2.6.5.1. Sensitivity..... | 76 |
| 2.2.6.5.2. Specificity..... | 76 |
| 2.2.6.5.3. Positive Predictive Value (PPV) | 76 |
| 2.2.6.5.4. Negative Predictive Value (NPV) | 76 |
| 2.2.6.5.5. False Negative Rate (FNR) | 77 |
| 2.2.6.5.6. False Positive Rate (FPR) | 77 |
| 2.2.6.5.7. False Discovery Rate (FDR)..... | 77 |
| 2.2.6.5.8. False Omission Rate (FOR) | 77 |
| 2.2.6.5.9. Accuracy | 77 |
| 2.2.6.5.10. Matthews Correlation Coefficient (MCC) | 77 |
| 2.2.7. Cohorts | 79 |
| 2.2.7.1. Prostate cancer: Mechanisms of progression and treatment (ProMPT study) | 79 |
| 2.2.7.2. Prostate MR imaging study (PROMIS) | 79 |
| 2.2.7.3. Combining advaNces in imaging with biOMarkers for improved diagnosis of Aggressive prostate cancEr (INNOVATE) | 79 |
| Chapter 3. Characterisation of a novel gene panel for diagnosis and risk stratification of prostate cancer. | 80 |
| 3.1. Introduction | 80 |
| 3.1.1. The gene panel..... | 83 |
| 3.1.2. Sharma <i>et al.</i> aggressive prostate cancer panel. | 85 |
| 3.1.3. Whitaker lab gene panel. | 85 |
| 3.2. Objectives and hypothesis | 86 |
| 3.2.1. Hypothesis 1 | 86 |
| 3.2.1.1. Aim 1: Characterise the baseline expression of the primary panel in prostate cancer cell lines. | 86 |
| 3.2.1.2. Aim 2: Study the varying degree of androgen regulation in the primary gene panel in androgen sensitive prostate cancer cell lines. | 86 |

| | |
|--|-----|
| 3.2.2. Hypothesis 2 | 86 |
| 3.2.2.1. Aim 3: Characterise the RNA degradation following exposure to urine..... | 86 |
| 3.2.3. Hypothesis 3 | 86 |
| 3.2.3.1. Aim 4: To investigate the expression of the primary gene panel in publicly available datasets using data mining. | 86 |
| 3.3. Results | 87 |
| 3.3.1. Cell line model | 87 |
| 3.3.1.1. Cell lines and their characteristics | 87 |
| 3.3.1.2. Baseline expression | 87 |
| 3.3.1.3. Androgen regulation..... | 90 |
| 3.3.2. Gene expression in degraded samples | 93 |
| 3.3.2.1. Simulation of degraded RNA..... | 93 |
| 3.3.2.2. Strategies to overcome degraded RNA. | 96 |
| 3.3.2.2.1. Increasing cDNA concentration..... | 96 |
| 3.3.2.2.2. Increasing number of cycles..... | 96 |
| 3.3.2.2.3. Preamplification of degraded samples | 96 |
| 3.3.3. Gene expression in human samples | 98 |
| 3.3.3.1. Data mining and cohort introduction | 98 |
| 3.3.3.2. Gene expression of the panel in three separate cohorts | 100 |
| 3.3.3.2.1. Differentiation between benign and cancer in three independent cohorts using a refined diagnostic gene panel..... | 100 |
| 3.3.3.2.2. Differentiation between prognostic grade groups (PGG) using gene expression data in three cohorts. | 102 |
| 3.4. Discussion | 104 |
| 3.4.1. Summary of findings | 104 |
| 3.4.1.1. The refined diagnostic gene panel. | 104 |
| 3.4.1.2. The refined prognostic gene panel | 105 |
| 3.4.2. Methodological limitations..... | 105 |
| 3.4.3. Clinical utility | 107 |
| 3.5. Lay summary..... | 108 |
| 3.5.1. Objectives | 108 |

| | |
|---|------------|
| 3.5.2. Findings | 108 |
| 3.5.3. Limitations | 109 |
| Chapter 4. Measurement of mRNA gene panel expression in blood and urine samples in independent cohorts..... | 115 |
| 4.1. Introduction | 115 |
| 4.1.1. Cohorts for biomarker panel discovery and validation | 117 |
| 4.1.1.1. Cambridge ProMPT cohort | 117 |
| 4.1.1.1.1. Urine samples cohort characteristics | 117 |
| 4.1.1.1.2. Blood samples cohort characteristics | 117 |
| 4.1.1.2. Prostate MR imaging study (PROMIS) cohort at UCLH | 119 |
| 4.1.1.3. INNOVATE UCL cohort..... | 119 |
| 4.2. Objectives | 121 |
| 4.2.1. Hypothesis 1: | 121 |
| 4.2.1.1. Aim 1: Evaluate whether RNA from urine cell pellets can be effectively used to measure the refined diagnostic and prognostic gene panels in a cohort of prostate cancer patients. | 121 |
| 4.2.2. Hypothesis 2: | 121 |
| 4.2.2.1. Aim 2: Measure the RNA levels of the seven-gene panel in PAXgene blood RNA from prostate cancer patient samples. | 121 |
| 4.2.2.2. Aim 3: Study the impact of the gene panel as a diagnostic and prognostic test. | 121 |
| 4.3. Results | 122 |
| 4.3.1. Urine as a fluid for biomarker discovery | 122 |
| 4.3.1.1. Quality and control | 122 |
| 4.3.1.2. Preamplification..... | 122 |
| 4.3.1.3. qPCR of degraded urinary RNA..... | 124 |
| 4.3.1.4. Missing data | 124 |
| 4.3.2. Studying gene expression using PAXgene whole blood | 130 |
| 4.3.2.1. Gene expression in PAXgene RNA in patients from the ProMPT study | 130 |
| 4.3.2.2. Gene expression in PAXgene RNA in patients from the PROMIS and INNOVATE studies | 132 |
| 4.3.2.2.1. Quality and control..... | 132 |

| | |
|--|------------|
| 4.3.2.2.2. Measuring the gene expression of all genes in the panel in PAXgene RNA using Fluidigm | 134 |
| 4.3.3. Study of the accuracy of the diagnostic and prognostic panels. | 137 |
| 4.3.3.1. Assessment of accuracy of biomarker panels for diagnosis and risk stratification of prostate cancer..... | 137 |
| 4.3.3.2. Assessment of diagnostic accuracy of the six-gene refined diagnostic gene panel in the PROMIS cohort for model development. | 137 |
| 4.3.3.3. Assessment of accuracy of the refined prognostic three-gene panel for risk stratification of prostate cancer..... | 140 |
| 4.3.4. Panel validation in the INNOVATE cohort | 142 |
| 4.3.5. Defining the best performing qualifiers for diagnosis and risk stratification of prostate cancer. | 145 |
| 4.4. Discussion | 147 |
| 4.4.1. Summary of findings | 147 |
| 4.4.2. Methodological limitations..... | 148 |
| 4.4.3. Clinical application | 149 |
| 4.5. Lay summary..... | 150 |
| 4.5.1. Objectives | 150 |
| 4.5.2. Findings | 150 |
| 4.5.3. Limitations..... | 151 |
| Chapter 5. Critical Evaluation Of Visual Proportion of Gleason 4 and Maximum Cancer Core Length | 156 |
| 5.1. Introduction | 156 |
| 5.2. Aims and objectives | 159 |
| 5.2.1. Hypothesis 1: | 159 |
| 5.2.1.1. Aim 1: Measure the MCCL using digital images of H&E prostate core biopsies and compare these to the measurement obtained using the microscope ruler. | 159 |
| 5.2.2. Hypothesis 2: | 159 |
| 5.2.2.1. Aim 2: Objectively quantify the burden of Gleason 4 disease in a cohort of 30 men from the PROMIS study..... | 159 |
| 5.2.3. Hypothesis 2: | 159 |

| | |
|---|------------|
| 5.2.3.1. Aim 3: Understand the clustering of Gleason 4 disease by using geographical parameters. | 159 |
| 5.3. Results | 160 |
| 5.3.1. Comparison of visual and digital MCCL measurement reveals discrepancy in Gleason 4 measurement. | 160 |
| 5.3.1.1. Selection of cohort | 160 |
| 5.3.1.2. Statistical analysis | 160 |
| 5.3.1.3. Digital and visual MCCL measurement are comparable. | 162 |
| 5.3.1.4. Visual Gleason 4 overestimates Gleason 4 burden..... | 165 |
| 5.3.1.5. Index block %G4 overestimation leads to patient reclassification | 168 |
| 5.3.2. Geographical parameters can be used to study the spatial distribution of Gleason 4 in prostate biopsies..... | 172 |
| 5.3.2.1. There is no correlation between cancer area, clustering or evenness indexes and location in the prostate. | 176 |
| 5.3.2.2. Clustering index is raised in high Likert mpMRI index lesion..... | 178 |
| 5.3.2.3. The clustering index correlates to the visual and digital estimation of %G4..... | 181 |
| 5.4. Discussion | 183 |
| 5.4.1. Summary of findings | 183 |
| 5.4.2. Methodological limitations..... | 184 |
| 5.4.3. Clinical application | 185 |
| 5.5. Lay summary..... | 187 |
| 5.5.1. Objectives | 187 |
| 5.5.2. Findings | 187 |
| 5.5.3. Limitations | 188 |
| 5.5.4. Future work | 188 |
| Chapter 6. General discussion | 197 |
| 6.1. Main findings presented in this thesis | 197 |
| 6.1.1. New biomarker panel is able to detect men with Gleason 4+3 \geq 6 mm but is not better than PSA density | 197 |
| 6.1.2. Gleason 4 percentage is overestimated in the visual assessment..... | 197 |
| 6.1.3. Gleason 4 is clustered in all regions of the prostate and correlated to mpMRI visibility ... | 197 |

| | |
|--|------------|
| 6.2. Methodological limitations..... | 198 |
| 6.2.1. Biomarker panel discovery | 198 |
| 6.2.2. Urine as a fluidic biomarker | 199 |
| 6.2.3. Obtaining meaningful outcomes from cohorts with limited follow up..... | 199 |
| 6.2.4. Aim of the biomarker panel..... | 199 |
| 6.2.5. Exclusion of clinical information in the biomarker panel performance review | 200 |
| 6.3. Research implications | 201 |
| 6.3.1. Biomarker discovery | 201 |
| 6.3.2. Investigating the utility of PSA density..... | 201 |
| 6.3.3. Encouraging collaboration with physical sciences to advance our understanding of cancer | 201 |
| 6.4. Future research..... | 203 |
| 6.4.1. Biomarkers development using in the PROMIS cohort | 203 |
| 6.4.2. Study of Gleason 4 percentage estimation and impact on patient outcome | 203 |
| Bibliography | 204 |

List of Figures

| | |
|---|-----|
| Figure 1-1 Anatomy of the prostate. Schematic representation of the zonal anatomy of the prostate..... | 27 |
| Figure 1-2 Non-targeted biopsies miss significant cancer. Schematic representation of cancer detection and misses by non-targeted transrectal prostate biopsy. | 31 |
| Figure 1-3 Simplified summary of prostate biopsies techniques and approaches. Schematic representation of the possible techniques that can be used when sampling the prostate. | 32 |
| Figure 1-4 Current and NICE recommended prostate cancer diagnostic pathway. | 35 |
| Figure 1-5 Benign and cancerous prostatic tissue. H&E Eosin digitally scanned sections of prostate tissue. | 37 |
| Figure 1-6 Prostate cancer TNM classification. Schematic representation of the prostate TNM, visible tumour represented in green. | 40 |
| Figure 1-7 Biomarker use in the diagnosis and treatment planning of localised prostate cancer. | 44 |
| Figure 2-1 Gleason 4 contouring and MCCL measurement. | 68 |
| Figure 2-2 Calculation of clustering and evenness scores from H&E scanned slides. | 70 |
| Figure 2-3 Barzell zones and one-hot encoding of variables. | 72 |
| Figure 2-4 Confusion matrix and performance analysis of a test. | 78 |
| Figure 3-1 Fluidic RNA gene panel selection roadmap. | 82 |
| Figure 3-2 Taqman qPCR baseline expression in 4 commonly used prostate cancer cell lines, LNCaP, C4-2, C4-2B and PC3. | 88 |
| Figure 3-3 Difference of gene expression in 4 prostate cancer cell classifies genes into three groups. | 89 |
| Figure 3-4 Taqman qPCR gene expression of three genes from the panel in androgen stimulated prostate cancer cell lines (LNCaP, C4-2 and C42-B). | 91 |
| Figure 3-5 Gene expression in LNCaPs, C4-2 and C4-2B cell lines following androgen stimulation. . | 92 |
| Figure 3-6 Degraded RNA will affect primer effectivity in qPCR reactions. | 94 |
| Figure 3-7 Exposure to urine reduces the detection rate of targets. | 95 |
| Figure 3-8 cDNA concentration and qPCR protocols do not affect Ct values, whereas Preamplication improves detection of target mRNA. | 97 |
| Figure 3-9 Stability based selection of genes as markers for differentiation of benign and cancer reveals a seven gene signature in tissue (refined diagnostic gene panel). | 101 |
| Figure 3-10. Stability based selection of genes as markers for differentiation between PGG groups, finds a three-gene signature in tissue (refined prognostic panel). | 103 |

| | |
|---|-----|
| Figure 4-1 Low quality urinary mRNA in ProMPT patient samples are not suitable for unbiased preamplification prior to qPCR. | 123 |
| Figure 4-2 Missing data following qPCR of unamplified urinary RNA. | 125 |
| Figure 4-3 Multiple random imputation reduces the number of missing data points. | 126 |
| Figure 4-4 Relative gene expression of genes in urine RNA reveals 6 upregulated genes in cancer. | 129 |
| Figure 4-5 RPLP2 Ct values vary between disease stages and it is not an appropriate housekeeper gene for normalisation of PCR data. | 131 |
| Figure 4-6 PROMIS and INNOVATE blood RNA samples quality control. | 133 |
| Figure 4-7 Missing data analysis of blood extracted mRNA from PAXgene blood. | 135 |
| Figure 4-8 Relative gene expression of genes in PAXgene blood extracted RNA in diagnostic and prognostic signature. | 136 |
| Figure 4-9 ROC curves reveal similar performance of the refined diagnostic six-gene panel compared to PSA density, but not superior. | 138 |
| Figure 4-10 Refined prognostic three-gene panel fails to classify patients by PGG. | 141 |
| Figure 4-11 ROC curves reveal similar performance of the refined diagnostic six-gene panel in the INNOVATE cohort, but is not superior to PSAd. | 143 |
| Figure 4-12 PSA density (PSAd) and Likert have high sensitivity and specificity for the diagnosis of significant prostate cancer. | 146 |
| Figure 5-1 Visual representation of the measurements of spatial segregation. | 158 |
| Figure 5-2 Gleason 4 PROMIS cohort selection process and demographics. | 161 |
| Figure 5-3 Objective measurement of MCCL and shows a discrepancy with visual measurement and pathologist estimation. | 163 |
| Figure 5-4 No significant bias is seen between visual and digital MCCL measurement. | 164 |
| Figure 5-5 Visual Gleason 4 appraisal overestimates burden of disease. | 166 |
| Figure 5-6 Visual estimation of Gleason 4 % overestimates Gleason 4 burden. | 167 |
| Figure 5-7 Objective measurement of Gleason 4 burden shows a discrepancy between visual measurement and the digital measurement for the index block. | 169 |
| Figure 5-8 Digital assessment reclassifies patients when using the UCLH traffic light system. | 170 |
| Figure 5-9 Clustering and evenness indexes example in one patient. | 173 |
| Figure 5-10 Evenness index above 1 were found on cores where a single Gleason 4 contour was drawn. | 174 |
| Figure 5-11 Clustering and evenness distribution in twelve patients. | 175 |

| | |
|--|-----|
| Figure 5-12 The clustering index increases with the Likert score, but not the evenness. | 179 |
| Figure 5-13 Clustering index correlates to visual, digital estimation of Gleason 4 % and the difference between the two measurements. | 182 |

List of tables

| | |
|---|-----|
| Table 1.1. CAPRA score and variables. | 41 |
| Table 1.2 Summary of available biomarkers, with substrate, accuracy and cost-effectiveness evidence, if available. | 45 |
| Table 2.1 Gene name and Taqman assays..... | 59 |
| Table 2.2 Taqman assays used for pooling prior to preamplification. | 65 |
| Table 3.1. 28 gene panel GO term category classification by Categorizer. | 84 |
| Table 3.2 Data mining cohort summary and missing gene data..... | 99 |
| Table 4.1 ProMPT patient reclassification following data review..... | 118 |
| Table 4.2 Cohorts disease status, clinical, pathological and imaging data. | 120 |
| Table 5.1 Formal measures of segregation as described by Massey and Denton..... | 157 |
| Table 5.2 Comparison of age and clinical characteristics between the 85 patient cohort and the final 30 patient cohort. | 160 |
| Table 5.3 Correlation coefficients between Barzell zone and cancer area, clustering and evenness indexes. | 177 |
| Table 5.4 Correlation coefficients between the ordinal coordinate system and cancer area, clustering and evenness indexes. | 177 |
| Table 5.5 Summary of mpMRI lesions in relation to the evenness and clustering indexes and digitally obtained cancer areas and Gleason 4. | 180 |

List of Appendices

| | |
|---|-----|
| Appendix 1. Gene expression in LNCaPs, C4-2 and C4-2B cell lines following androgen stimulation. | 111 |
| Appendix 2. Gene expression of the seven genes in the panel correlates with clinical diagnosis in three different cohorts. | 112 |
| Appendix 3. Boxplots of the relative gene expression of genes in the Ross cohort. | 113 |
| Appendix 4. Boxplots of the relative gene expression of genes in the Taylor cohort. | 114 |
| Appendix 5. Urine PCR relative gene expression data is not normally distributed and has unequal variance distribution. | 152 |
| Appendix 6. Missing data analysis of PAXgene blood extracted mRNA PCR in Fluidigm. | 153 |
| Appendix 7. Blood PCR relative gene expression data is not normally distributed and has unequal variance distribution. | 154 |
| Appendix 8. Confusion matrices for the six-gene panel for the diagnosis of prostate cancer and PROMIS definitions of significant disease in the PROMIS and INNOVATE cohorts. | 155 |
| Appendix 9 There is a very weak correlation between the clustering and evenness indexes. | 189 |
| Appendix 10 There is correlation between the clustering index and the visual estimation but not with evenness. | 191 |
| Appendix 11 There is correlation between the clustering index and the difference between the visual and digital estimation but not with evenness. | 193 |
| Appendix 12 There is correlation between the clustering index and the digital %G4 but not with evenness. | 195 |

Abbreviations

%G4 - Percentage Gleason 4

4-K score - 4 kallikrein score

ADC - Apparent Diffusion Coefficient

ANOVA - Analysis of variance

AR - Androgen Receptor

ASAP - Atypical Small Acinar Proliferation

ASCO - American Society of Clinical Oncology

AUC - Area under the Curve

BEST - Biomarker, EndpointS, and other Tools

CAPRA-S score - UCSF Cancer of the Prostate Risk Assessment Score

CCP - Cell Cycle Progression score

cDNA - Complementary DNA

CI - Confidence Interval

Ct -Threshold cycle

DNA - Deoxyribonucleic acid

DRE - Digital Rectal Examination

EAU - European Association of Urology

ERSPC - European Randomised Study of Screening for Prostate Cancer

FDA - US Food and Drug Administration

FDR - False Discovery Rate

FFPE - Formalin Fixed Paraffin Embedded

FN - False Negative

FNR - False Negative Rate

FOR - False Omission Rate

FP - False Positive

FPR - False Positive Rate

GC - Genomic Classifier

GPS - Genomic Prostate Score

H&E - Haematoxylin and Eosin

HGPIN - High Grade Prostatic Intraepithelial Neoplasia

hk2 - Human kallikrein 2

HR - Hazard Ratio

INNOVATE study - Combining advances in imaging with biomarkers for improved diagnosis of Aggressive prostate cancer study

IQR - Interquartile Range

ISUP - International Society of Urological Pathology

MAR - Missing At Random

MCAR - Missing Completely At Random

MCC - Matthews Correlation Coefficient

MCCL - Maximum Cancer Core Length

MiPS - Mi Prostate Score

MNAR - Missing Not At Random

mpMRI - Multiparametric MRI

MRI - Magnetic Resonance Imaging

mRNA - Messenger RNA

NCCN risk classification - National Comprehensive Cancer Network risk classification

NHS - National Health Service

NICE - National Institute for Health and Care Excellence

NPV - Negative Predictive Value

OR - Odds Ratio

PCA3 - Prostate Cancer Antigen

PCR - Polymerase Chain Reaction

PGG- Prostate Grade Group

PHI - Prostate Health Index

PIRADS - Prostate Imaging-Reporting and Data System

PLCO - Prostate, Lung, Colorectal and Ovarian Cancer

PPV - Positive Predictive Value

PROMIS study - Prostate MR imaging study

ProMPT study - Prostate cancer: Mechanisms of Progression and Treatment study

PSA- Prostate Specific Antigen

PSAd - PSA density

pvol - Prostate volume

QALY - Quality-Adjusted Life-Year

qdPCR - Quantitative digital Polymerase Chain Reaction

qPCR - Quantitative Polymerase Chain Reaction

Q-Q plots - Quantile-Quantile plots

RIN - RNA Integrity Number

RNA - Ribonucleic Acid

RNAse - Ribonuclease

ROC - Receiver Operating Curve

SD - Standard Deviation

STHLM3 - Stockholm-3

T2 - T2 weighted imaging

TN - True Negative

TNM staging - Tumour, Lymph Node and Metastasis staging

TP - True Positive

TRUS biopsy - Transrectal Ultrasound biopsy

UCLH - University College London Hospital

VACURG - Veterans Administration Cooperative Urological Research Group

WHO - World Health Organisation

WTA - Whole Transcriptome Amplification

Chapter 1. Introduction

1.1. Summary

Prostate cancer is the most commonly diagnosed malignancy in men (1). The incidence of prostate cancer varies from 1 in 9 in high-income countries to 1 in 52 in low-income countries (2). The diagnostic pathway for prostate cancer has dramatically changed in the last few years with imaging biomarkers included as a first-line investigation for people with suspected clinically localised prostate cancer (3). Despite these advances, over-diagnosis and over-treatment remain a challenge (4). Tissue and fluid biomarkers are currently being developed to fill the gaps in the diagnostic pathway, mainly reduce number of biopsies, find aggressive prostate cancer without the need for a biopsy and predict recurrence and death (5,6). Ideally such biomarkers would have analytic validity, clinical validity and clinical utility (7), across different populations, whilst complementing the clinical data in a reproducible manner.

Like all cancers, prostate cancer has a spectrum of aggressiveness. The Gleason score was adapted into the Prostate Grading Group (PGG) to take into account the growing evidence that the proportion of Gleason 4 impacts patient outcomes (8–13). The current Gleason and PGG classifications rely on visual percentage estimation, with the inherent subjectivity and inter-observer variability from human readers (14,15). Digital pathology has allowed novel image analysis that has been successfully linked to outcome in breast cancer (16). Since the objective for some biomarker panels is to find aggressive disease, efforts must be made to quantify aggressive disease as accurately as possible. Use of machine learning to aid Gleason scoring have already been published (17–21).

This thesis focuses on describing the discovery of a new biomarker panel to complement multiparametric MRI (mpMRI) in deciding which patients benefit from a biopsy. Chapter 3 describes the panel in detail, followed by relative gene expression of the panel in cell lines. Finally, diagnostic accuracy in three publicly available datasets is described. In Chapter 4, the relative gene expression of the new panel in urine and blood, in three different cohorts is presented, followed by performance analysis of the panel. Finally, in Chapter 5, an in-depth analysis and description of Gleason 4 burden was performed using digital pathology and geographical quantitative research in a cohort of 30 men. The following literature review will describe in depth the concepts briefly defined in here.

1.2.Context/literature review

1.2.1. Prostate cancer

1.2.1.1. The prostate

The prostate is a pelvic organ located just above the bladder neck (Figure 1-1).It is composed of a mix of glandular (70%) and fibromuscular tissue (30%),and it is surrounded by a capsule that fuses with the stroma. The stroma is composed of collagen and smooth muscle that contracts during ejaculation. The prostate can be divided into distinct zones:

1. Central zone: Located at the base of the bladder, it is conical in shape and surrounds the ejaculatory ducts. This zone contains 25% of the glandular tissue.
2. Transitional zone: Surrounds the urethra and constitutes 5-10% of the glandular tissue. Benign prostatic enlargement usually originates from this area.
3. Peripheral zone: This zone forms most of the glandular tissue (70%); it covers the posterior and lateral aspects of the gland. Traditionally, most cancers develop in this area (22).
4. Fibromuscular stroma: Extends from the bladder neck to the urethral sphincter (striated). It can take up to one-third of the prostatic mass.

A schematic representation is presented in Figure 1-1.

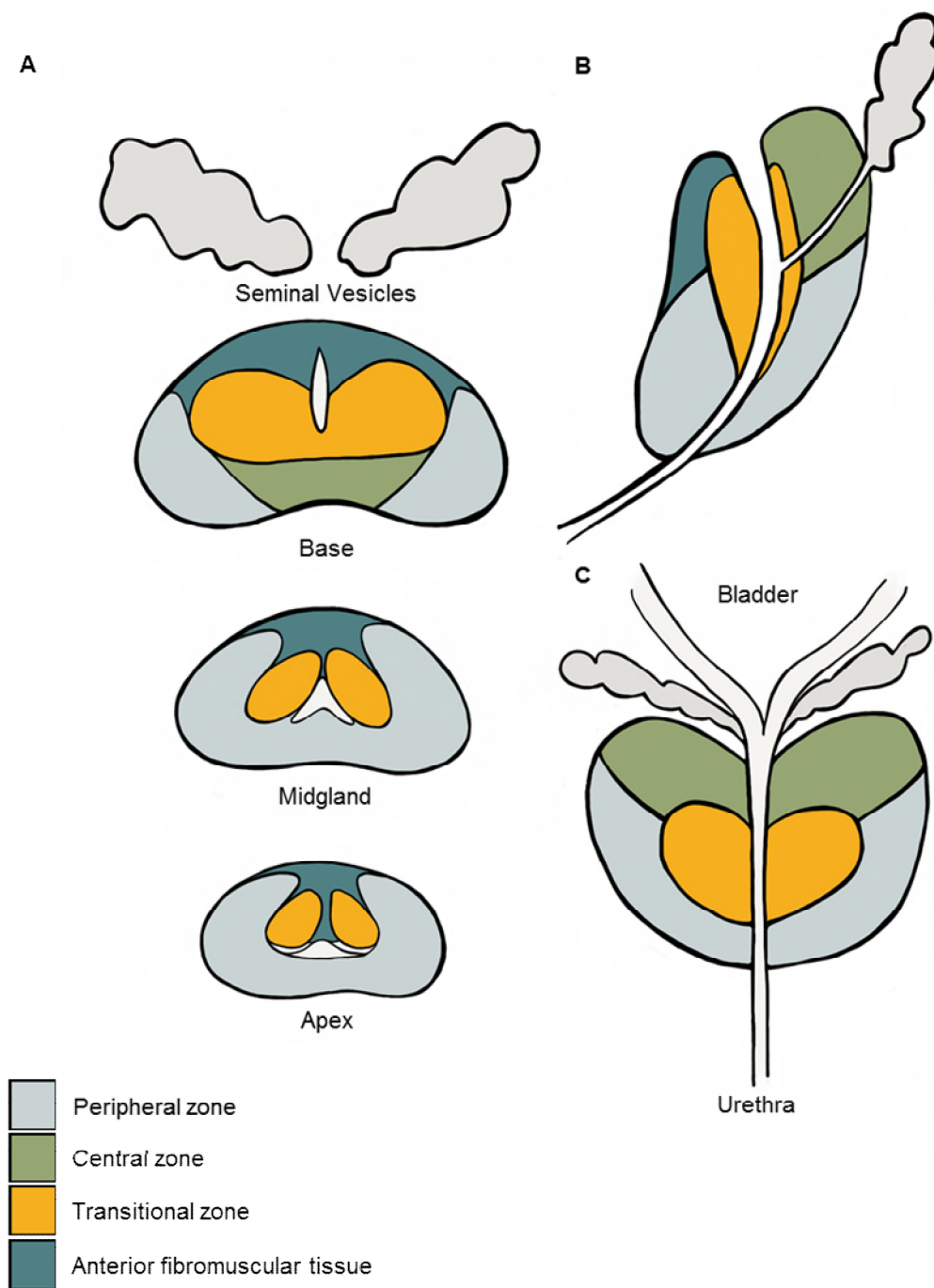


Figure 1-1 Anatomy of the prostate. Schematic representation of the zonal anatomy of the prostate.

A. Axial sections of the prostate from posterior to anterior, with the seminal vesicles as the most posterior structure, followed by the base, midgland and apex of the prostate. **B.** Sagittal section of the prostate, with the urethra traversing the prostate, the seminal vesicles just posterior to the prostate, and the ejaculatory duct connecting to the urethra. **C.** Coronal view of the prostate, the bladder located superiorly contiguous to the urethra. The seminal vesicles located just below the bladder. The prostate zones are coloured as follows: peripheral zone = light blue, central zone = green, transitional zone = yellow, anterior fibromuscular zone = dark blue. Adapted from: "Synopsis of the PI-RADS v2 Guidelines for Multiparametric Prostate Magnetic Resonance Imaging and Recommendations for Use" by Barentz et al. (23).

1.2.1.2. Epidemiology

Globally, prostate cancer is the most common cancer in men. In 2017, the incidence of prostate cancer was 1.3 million (95% CI, 1.2-1.7 million). Despite advances in treatment strategies and diagnosis, it remains the third cause of cancer-related death in men (1,2). Following the introduction of Prostate Specific Antigen (PSA) testing, the number of men with advanced disease at presentation has been reduced (24). However, PSA screening has led to over-diagnosis and over-treatment (4).

There are three well-known risk factors for developing prostate cancer. First, older men have a higher risk of developing the disease; 85% of diagnosed men are >65 years old (25). Second, men with a family history have a higher risk of developing the disease (26). Third, the incidence and mortality due to prostate cancer in men of African descent are higher compared to Caucasians. The cause for this is considered to be multifactorial. Including socioeconomic, cultural and genetic predisposition (27,28). Whereas the incidence is the lowest in Asian men (28). Additional risk factors include obesity (29,30), metabolic syndrome (29,31,32), and hypertension (33).

1.2.1.3. Androgens and the Androgen Receptor (AR)

In 1941 Huggins and Hodges published their findings on the effects of androgen in metastatic prostate cancer in dogs (34). Improved understanding of androgens and the androgen receptor in relation to prostate cancer led to the development of hormone-blocking therapies (surgical and chemical castration). Androgens are steroid hormones that influence the development, maturation and maintenance of the male reproductive system (35). They ligate to the androgen receptor (AR), a nuclear transcription factor that mediates androgen activity by initiating the transcription of androgen-responsive genes (36,37). Understanding the role of the AR in prostate cancer, the gene activation/inactivation and mechanisms of androgen insensitivity have allowed for novel therapies and biomarkers. For instance, abiraterone is an androgen biosynthesis inhibitor that has shown to reduce mortality in men with metastatic prostate cancer (38). Additionally, the presence of the AR splice variant 7 (AR-V7) can be used for clinical decision making (39,40). Antonarakis *et al.* found that men with castrate resistant prostate cancer and AR-V7 have associated resistance to enzalutamide and abiraterone (41).

In this thesis, the relative changes in gene expression of a panel of genes following androgen stimulation are presented.

1.2.1.4. The diagnosis of prostate cancer

1.2.1.4.1. Prostate Specific Antigen (PSA)

Prior to the discovery of the prostate specific antigen (PSA) by Wang *et al.*, and subsequent use as a screening tool, men presented with a plethora of symptoms, due to advanced prostate cancer (42). Symptoms include, amongst others, haematuria, obstructive lower urinary tract symptoms, and bone pain due to metastatic deposits. Before the widespread use of PSA, there was no such thing as early detection in prostate cancer.

In 1987, Stamey *et al.*, measured the PSA in 127 newly diagnosed patients and found that increased PSA levels were associated with advanced disease. Additionally, the levels of PSA dropped following radical prostatectomy, confirming its use as a monitoring tool following surgery (42). This was followed by evidence that the use of PSA as a screening tool improved the detection of prostate cancer, in combination with clinical parameters (digital rectal examination and biopsy) (43). The use of PSA as a screening tool allowed physicians to diagnose patients earlier, this combined with better surgical techniques and therapies, have improved the outcomes of patients (defined as time to recurrence and cancer-related death) (44). While PSA is recognised as a relatively useful biomarker, its widespread use has been linked to over-diagnosis and over-treatment (4). Further background into PSA and its properties as a biomarker will be discussed in section 1.2.2.2.1.1.

1.2.1.4.2. Prostate Biopsy

Until recently, the diagnostic pathway consisted of screening of men with PSA in primary care, followed by digital rectal examination (DRE) and transrectal ultrasound random biopsies, if indicated. This diagnostic pathway relied on large palpable tumours to be detected by DRE, and on the concept that most tumours were located peripherally. Additionally, and unlike other solid tumours, the random transrectal systematic biopsy consists of blind biopsies roughly targeted to the peripheral zone. This method carries medical risks (infection, sepsis, death) and the risk of entirely missing the lesion, especially if the cancer is anterior (45,46) (See Figure 1-2). More importantly, 26% of men in the PROMIS trial had cancer missed by transrectal biopsies that were detected by a targeted sampling method (47).

Reports of 6.9% risk of hospitalisation following transrectal biopsy highlight the need for careful patient selection (48). Patient reported outcomes following transrectal biopsies in the European Randomized Study of Screening for Prostate Cancer (Rotterdam section), reported 4.2% of patients developing fever (392 of 9241 questionnaires), and of those hospitalised following a biopsy 81% had an infection (48). The risk of sepsis following a transrectal biopsy has been reported between 0.5-3.1% (49).

Alternative sampling methods have been used including an MRI targeted (to a lesion), systematic (every 5 mm), sector biopsy (1-2 samples per sector), modified according to prostate size and combinations of the above, mainly targeted and systematic (See Figure 1-3) (50–54). Additionally, avoiding the transrectal route has increasingly become popular due to the lower risk of infection (49,55). Although the transperineal approach requires an operating theatre and general anaesthetic in some instances,

development of the technique and invention of novel devices may make this procedure a routine clinic/outpatient procedure (56,57).

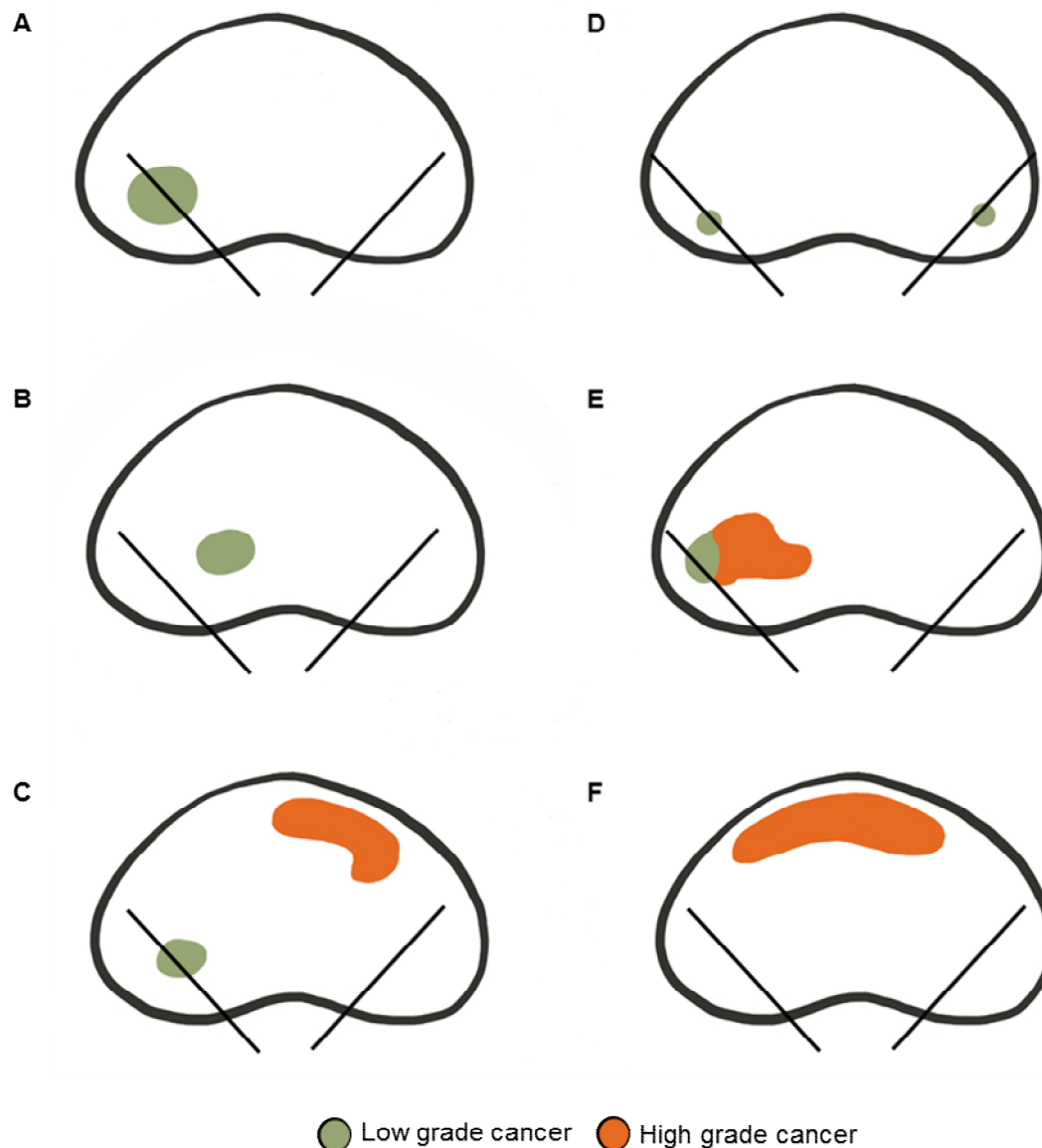


Figure 1-2 Non-targeted biopsies miss significant cancer. Schematic representation of cancer detection and misses by non-targeted transrectal prostate biopsy.

For simplicity two-needle cores represented with a black line, low-grade cancer in green, high-grade cancer in orange. **A.** Random biopsy detects low-grade cancer on the left side (true positive). **B.** Random biopsy fails to identify low-grade cancer on the left side of the prostate (false negative). **C.** Detection of low-grade but non-detection of high-grade anterior tumour. **D.** Detection of two tiny foci of low-grade disease. **E.** Sampling of a tumour area of low-grade disease, while high-grade area remains un-sampled. **F.** Peripheral zone sampling completely misses large anterior tumour.

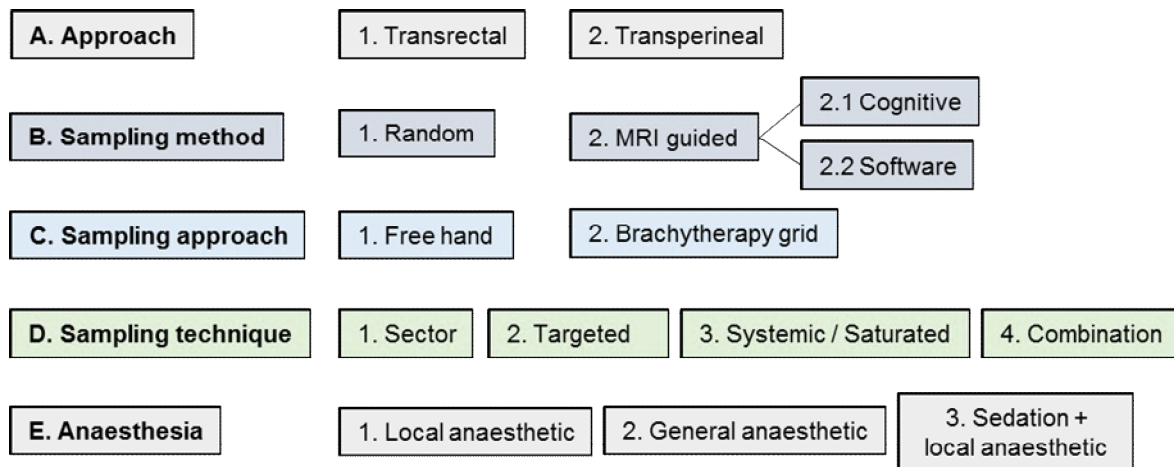


Figure 1-3 Simplified summary of prostate biopsies techniques and approaches. Schematic representation of the possible techniques that can be used when sampling the prostate.

From top to bottom **A.** Approach refers to the route through which the needle is inserted for the sample, for simplicity abdominal and other routes are not plotted here. **B.** Sampling method refers to either random sampling (1) or targeted to a lesion using mpMRI (2), this can be either cognitive (2.1), using previously contoured lesions to guide the needle, or with the assistance of software (2.2). **C.** Approach refers to either free-hand sampling, or with the assistance of a grid. **D.** Sampling technique refers to how the biopsies are taken: by pre-determined sectors, targeted to a lesion, saturated every 5 mm or combination of the previously mentioned techniques. **E.** Anaesthesia can be local pudendal block, general or a combination of sedation and local anaesthetic.

1.2.1.4.3. Multiparametric MRI

Until the emergence of multiparametric MRI (mpMRI), visualising prostate cancer was relegated to large visible tumours on ultrasound. The incidence of non-palpable tumours or T1c increased from 6% in 1989 to 47% in 1999 (58). This may be linked to an increase in diagnosis of insignificant cancers with the widespread use of PSA, or due to the difficulty in visualising prostate cancer on ultrasound. MpMRI can be used as an imaging biomarker for diagnosis; if no lesion is detected there is a high probability of not having high-risk cancer (47,59–61). The PROMIS trial enrolled 576 men to study the diagnostic accuracy of mpMRI. Using mpMRI as a triage, 27% of patients can safely avoid a biopsy, while lowering the number of men diagnosed with insignificant cancer (5%) (47). The value of mpMRI to aid the diagnosis of prostate cancer, while reducing over-diagnosis led to the addition to mpMRI as the first-line investigation for men with suspected prostate cancer (3).

For instance, the detection of transitional zone tumours can be difficult as they are not palpable on digital rectal exam. Patients with anterior tumours usually present following several negative transrectal biopsies and a persistently elevated PSA (62). Since mpMRI can improve detection of anterior tumours that would otherwise remain unseen; the incidence of anterior tumours may be higher than historically accepted.

Additionally, mpMRI allows for targeting of lesions, whether cognitively or using specialised software, allowing for improved diagnostic accuracy compared to traditional transrectal biopsy (63–66). A 2019 Cochrane review concluded that adding mpMRI to the diagnostic pathway improved the diagnostic accuracy of clinically significant cancer, while reducing the diagnosis of insignificant cancer (61). The NICE diagnostic pathway has been modified to include mpMRI before biopsy (new diagnostic pathway on Figure 1-4), and sets out the minimum requirements for appropriate imaging technique (67). Although this is the recommended pathway, access to mpMRI is not yet accessible to all patients. Data from Prostate Cancer UK's Freedom of Information Act requests in 2016 and 2018 showed an increase of use of mpMRI in the UK (68). Additionally, 68% of the centres using MRI before biopsy use this information to avoid a biopsy if the MRI is negative. While encouraging, centres are limited by several factors including radiologist expertise, time in scanner and age of the scanner available (68).

1.2.1.4.3.1. PIRADS v2

The Prostate Imaging-Reporting and Data System (PIRADS) v2, was created by the European Society of Urogenital Radiology (ESUR) to improve image acquisition, quality, and standardize mpMRI reporting (69–72). Compared to the first version, more emphasis is placed on T2 weighted imaging for transitional zone tumours, improving detection amongst benign nodules that mimic cancer (73–75). The PIRADS v2 categorizes the mpMRI parameters to obtain a final score from 1 to 5, as follows:

- PI-RADS 1 – Very low (clinically significant cancer is highly unlikely to be present).
- PI-RADS 2 – Low (clinically significant cancer is unlikely to be present).
- PI-RADS 3 – Intermediate (the presence of clinically significant cancer is equivocal).
- PI-RADS 4 – High (clinically significant cancer is likely to be present).

- PI-RADS 5 – Very high (clinically significant cancer is highly likely to be present).

1.2.1.4.3.2. Likert score

The Likert score is an mpMRI reporting tool that classifies lesions in a scale of 1 to 5. Unlike PIRADS, it considers clinical information (age, PSA, gland volume, PSA density, family history). Additionally, it does not rely on the findings in a dominant sequence. Similarly to PIRADS v2, each score relates to a degree of certainty of cancer being present in a lesion (67,76,77). However, the expertise and confidence of the reader and the prostate as a whole is taken into account in the Likert score.

The Likert score is as follows:

- Likert 1, presence of cancer is very unlikely.
- Likert 2, presence of cancer is unlikely.
- Likert 3, presence of cancer is indeterminate.
- Likert 4, presence of cancer is likely.
- Likert 5, presence of cancer is highly likely.

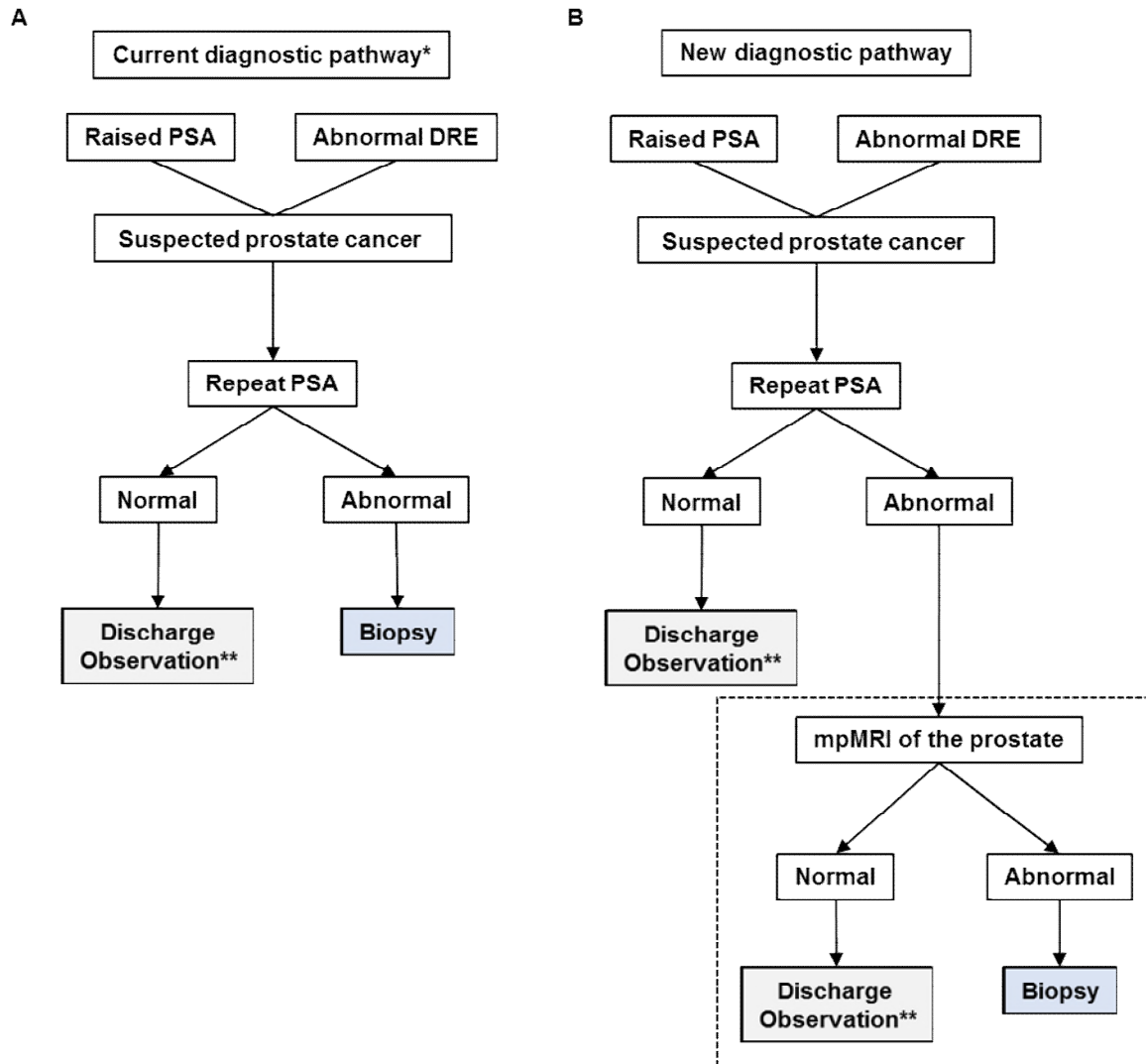


Figure 1-4 Current and NICE recommended prostate cancer diagnostic pathway.

A. Current diagnostic pathway marked with* as not all centres have access to mpMRI and may continue to follow this pathway. **B.** The NICE diagnostic pathway includes mpMRI of the prostate following a confirmed abnormal PSA. Biopsy highlighted in blue. ** Discharge or observation depending on clinical parameters such as age, PSA level, prostate size, comorbidities and other prostatic diseases, such as prostatitis.

1.2.1.5. Gleason system

In 1966, Dr Donald F. Gleason and colleagues at the Veterans Administration Cooperative Urological Research Group (VACURG) proposed a classification system for prostate cancer (78). This system compares the cancerous tissue to benign prostatic tissue in Haematoxylin & Eosin stained slides (H&E). The original classification included patterns 1 to 5 depending on four factors: tumour shape and borders, stromal invasion, tumour cell arrangement and gland size (78). Patterns 1 and 2 were later recommended not be reported and were practically removed from the classification in 2005. Currently, only patterns 3, 4 and 5 are reported (79). The Gleason score is assigned, taking into account the two most common Gleason patterns present in the specimen by visual inspection. The first number in the score corresponds to the most prevalent grade seen, and the second number to the second most common grade. This way a Gleason 4+5 consists predominantly of Gleason 4 (majority), with some Gleason 5 (minority). Whenever a tertiary pattern exists, if it corresponds to Gleason 4 and 5, it should be reported as it has an impact on pathological stage and progression (80).

Gleason 3 consists of irregular and discrete glands with variation in size and shape. Gleason 4 includes fused, poorly formed and cribriform glands that are poorly defined with irregular borders. Gleason 5 consists of solid sheets of tissue without glandular differentiation, forming chords or single “escaped” cells (See Figure 1-5 A-D). As mentioned above, the distinction between the most common and second most common Gleason pattern is based on visual estimation, and this has interobserver variability (81–84). For instance, in 2011 Netto *et al.* measured the agreement between local and central pathologists reading radical prostatectomy specimens of men that took part in the TAX 3501 trial, a randomised multinational trial comparing outcomes in patients with high-risk prostate cancer, following radical prostatectomy with or without adjuvant docetaxel. They found significant interobserver variations, with disagreement in 30% of cases, changing the progression-free survival estimates of miss-classified patients (83). Additionally, when the interobserver variability is measured for Gleason 4 only, fused and poorly fused glands seem to reduce the agreement between pathologists (85,86). This has led to the search for automated Gleason 4 grading systems to reduce variability (18,19,87–89).

1.2.1.6. Prostate Grade Group

In 2014, the International Society of Urological Pathology (ISUP) agreed on a new grading system for prostate cancer (Prognostic Grade Group or PGG), based on modifications to the original Gleason system in 2005 (90,91). The new grading system aims to simplify the Gleason score and taking into account the difference in prognosis of the different groups (Figure 1-5 E). Previously, men with Gleason 7 were grouped as a single category, regardless of the amount of Gleason 4 present. This is problematic, as men with higher amounts of Gleason 4 have been shown to have worse outcome (8–13), probably as Gleason 4 is molecularly distinct from Gleason 3 (92–94).

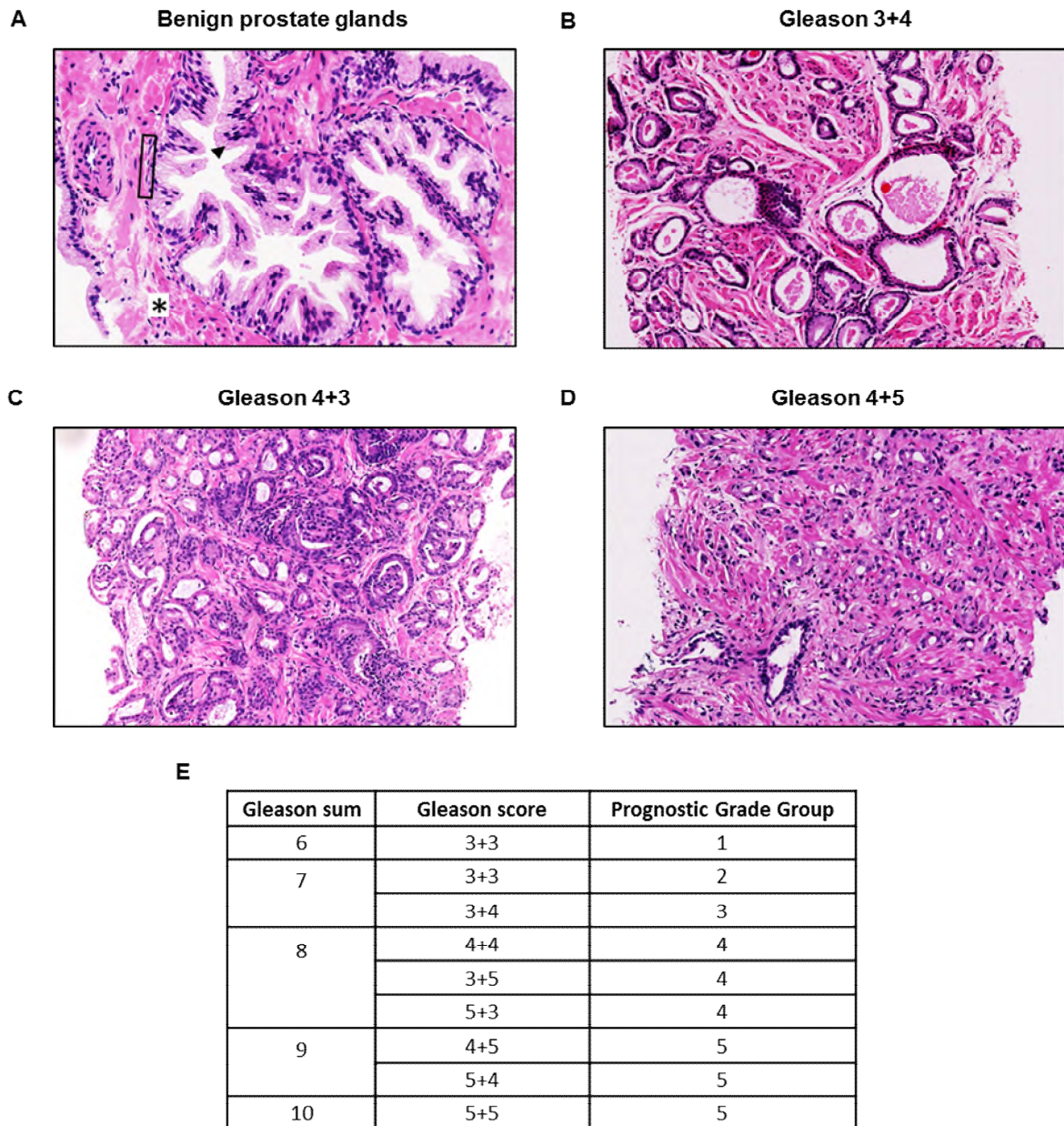


Figure 1-5 Benign and cancerous prostatic tissue. H&E Eosin digitally scanned sections of prostate tissue.

Images obtained from PROMIS scanned slides, nuclei in blue and other structures in pink. **A.** Benign prostatic tissue with well-formed glands (black arrow), surrounded by basal cells (black square) and stroma (*). **B.** Gleason 3+4 cancer, Gleason 3 retains the glandular structure remains, although both size and shape are irregular. **C.** Gleason 4+3, in Gleason 4 the glands are fused and cannot be separated from each other, the glandular structure is poorly formed. **D.** Gleason 4+5, Gleason 5 lacks glandular differentiation and is characterised by solid sheets of tissue. **E.** Table containing the Gleason sum value (left column), corresponding Gleason scores (middle column) and new Prognostic Grade Group (PGG).

1.2.1.6.1. The importance of Gleason pattern 4 in prostate cancer

When compared to Gleason 6, men with Gleason 7 have a higher risk of death from prostate cancer (95). When examining historical radical prostatectomy specimens in men with localised disease, Sweat *et al.* found that at 15 years, the rate of death was 18% to 30% for Gleason 6 and 42% to 70% for Gleason 7 (95). When comparing the outcome of Gleason 3+4 versus 4+3, men with primary Gleason 4 have been found to be more likely to have seminal vesicle involvement (34% versus 18%, $p = 0.006$), a higher pathological stage (pT3 55% versus 42%, N+ 13% versus 3%, 0.001) and extra-prostatic extension (58% versus 38%, 0.001) (96). In a separate study by Chan *et al.* tumours with primary Gleason 4 had a higher rate of extra-prostatic extension at radical prostatectomy (52.7% versus 38.2% for 3+4, $p = 0.008$), higher risk of progression at five years (15% 3+4, 40% 4+3) and earlier time to progression (3.2 years versus 4.4 years) (97).

Changes in the classification of Gleason 4 such as including cribriform pattern as Gleason 4 rather than 3, have led to increased understanding and classification of patient category risk. Patients with <5% of Gleason 4 can be candidates for active surveillance, with similar survival to men with Gleason 6 disease (13,98). When examining 12,150 prostatectomy specimens, Sauter *et al.* found a distinct and significant difference in time to biochemical recurrence when subdividing patients into 3+4 low (1-24% Gleason 4), 3+4 high (25-49% Gleason 4), 4+3 low (50-74% Gleason 4) and 4+3 high (75-94% Gleason 4), $p < 0.0001$ each (9). In 2014 and 2019, the ISUP Consensus meeting recommended Gleason 4 percentage should be reported to aid management planning (91,99). The percentage of Gleason 4 involvement from the worst core (with highest Gleason 4) and overall Gleason 4% provide accurate information as predictors of prostate cancer death (13).

1.2.1.7. Prostate cancer staging

The Tumour, Node, Metastasis (TNM) staging in prostate cancer is schematically represented in Figure 1-6. The prostate cancer staging is taken into account in risk prediction nomograms and groups. Since MRI's adoption in clinical practice, the local staging of prostate cancer has been facilitated, allowing for surgical planning (100–102). Other imaging techniques, such as molecular imaging have a role in detecting distant lymph node metastasis and recurrence following treatment (103).

1.2.1.7.1. The D'Amico classification

In 1998, Dr Anthony D'Amico proposed a model classifying patients into three risk categories depending on the TNM stage, Gleason score and preoperative PSA. The groups act as a predictor to the risk of biochemical recurrence after surgery. This classification continues to have clinical relevance but may require modernisation to adapt to current clinical practices (104).

The D'Amico classification is presented below:

1. Low risk: Clinical T1c-T2a, PSA level ≤ 10 ng/mL and Gleason score ≤ 6 .
2. Intermediate risk: Clinical T2b, PSA level of 10.1-20 ng/mL, or Gleason score 7.

3. High risk: Clinical T2c or PSA level greater than 20 ng/mL or Gleason score 8-10.

1.2.1.7.2. The CAPRA-S score

The UCSF Cancer of the Prostate Risk Assessment (CAPRA) score, is a multivariate prognostic tool that aims to predict the risk of recurrence following treatment for localised prostate cancer. It takes into account the serum PSA, Gleason score (grouping 3+4 and 4+3), T stage, percentage of positive biopsies and age to give a score from 0 to 10 (105). The CAPRA score is then tiered into three groups: 0–2 = low risk; 3–5 = intermediate risk; 6–10 = high risk. Many studies have researched the correlation between the CAPRA score and biochemical recurrence, with reported concordance scores between 0.66–0.81 (106).

1.2.1.7.3. The NCCN risk classification

The National Comprehensive Cancer Network risk classification divides patients into five distinct categories. Very low, low, intermediate, high and very high risk taking into account PSA, pathological stage, PGG group and, number of positive cores. As this classification uses PGG instead of Gleason sum, it subdivides 3+4 and 4+3 into two different categories. Additionally, this risk classification recommends active surveillance to men with low secondary Gleason 4 (107).

1. Very low risk: T1c and, PGG1 and PSA <10 ng/mL, <3 positive cores, PSA density < 0.15 ng/mg/mL.
2. Low risk: T1-T2a and PGG1 and PSA <10 ng/mL.
3. Intermediate risk: T2b-T2c, PGG 2 or 3, PSA 10-20 ng/mL and no high risk features.
 - a. Favourable intermediate: PGG 1 or 2, <50% cores positive.
 - b. Unfavourable intermediate: PGG 3, ≥50% of positive cores.
4. High risk: T3a or PGG4-5 or PSA >20 ng/mL.
5. Very high risk: T3b-T4 or primary Gleason 5 or >4 cores with grade group 4/5.

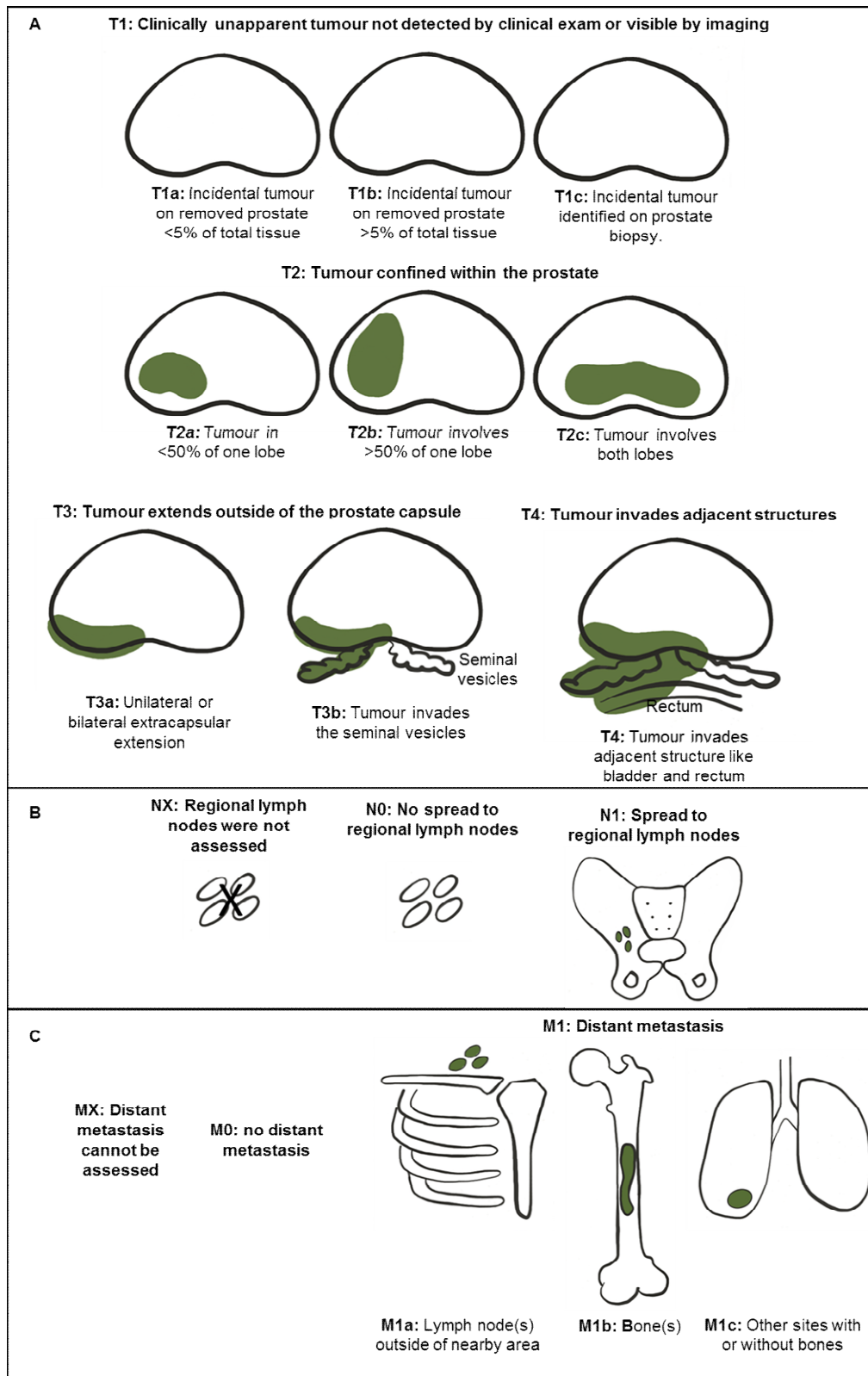


Figure 1-6 Prostate cancer TNM classification. Schematic representation of the prostate TNM, visible tumour represented in green.

A. Tumour stage is divided in T1, T2, T3 and T4. As T1 stage represents not palpable/visible tumour, no lesion was drawn. **B.** Node stage, from not assessed (NX), no regional spread (represented by white lymph nodes) and regional lymph node compromise (N1). **C.** Metastasis stage, not assessed (MX), no distant metastasis (M0) and distant metastasis (M1). M1a, distant lymph nodes (supraclavicular), M1b, bone metastasis and M1c, lung metastasis (as an example).

| Variable | Value | CAPRA-S points |
|---|--------------------------|----------------|
| PSA level (ng/mL) | 0.00 -6.00 | 0 |
| | 6.01-10.00 | 1 |
| | 10.01-20.00 | 2 |
| | 20.1-30 | 3 |
| | ≥ 30 | 4 |
| Age at diagnosis | Under 50 | 0 |
| | 50 or older | 1 |
| Gleason score | No pattern 4 or 5 | 0 |
| | Secondary pattern 4 or 5 | 1 |
| | Primary pattern 4 or 5 | 2 |
| Clinical stage | T1 or T2 | 0 |
| | T3a | 1 |
| Percentage of biopsy cores involved with cancer | Less than 34 percent | 0 |
| | 34 percent or more | 1 |

Table 1.1. CAPRA score and variables.

Table containing the variables included in the calculation of the CAPRA score (first column), followed by values (middle column) and points given depending on the values (right column).

1.2.2. Biomarkers

1.2.2.1. Definition and types of biomarkers

A biomarker or biological marker is a quantifiable indicator of a specific disease or the severity of a disease. The National Institutes of Health Biomarkers Definitions Working Group defines a biomarker as: “A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” (108). A biomarker can serve different purposes; The BEST (Biomarker, EndpointS, and other Tools) resource subdivides them into seven groups and for the purpose of this thesis, only biomarkers used for cancer will be defined here. The definitions have been modified from the BEST guidelines (109).

1. Diagnostic biomarker: A diagnostic biomarker helps determine if the patient has cancer or a subtype of cancer. A perfect diagnostic biomarker would detect all patients with the disease of interest and no healthy patients (100% sensitivity and specificity). For example, PSA is used as a diagnostic biomarker; further information on this biomarker is discussed in section 1.2.2.2.1.1.
2. Monitoring biomarker: A biomarker measured successively in order to assess the status of a medical condition, or as evidence of exposure or effect due to a medical product or environmental agent. Again, PSA can be used as an example when used to monitor patients following radical prostatectomy or radiotherapy to ensure disease control (110,111).
3. Predictive biomarker: A biomarker that helps identify patients with a higher or lower risk of good or bad outcome following exposure to a medical product or environmental agent. For instance, the presence or absence of AR-v7 can help predict which patients will benefit from taxane-based chemotherapy (39,40).
4. Prognostic biomarker: A biomarker used to predict the probability of a clinical event, disease progression or recurrence. In prostate cancer, both PSA and Gleason score have been used as predictive biomarkers to determine the risk of recurrence (112,113).
5. Susceptibility/Risk biomarker: A biomarker that indicates the likelihood of developing a disease in a patient that does not currently have the disease. The presence of BRCA2 mutations have been linked to a higher risk of prostate cancer (114).

1.2.2.2. Biomarkers for prostate cancer

This section will focus on selected biomarkers currently used for the diagnosis of prostate cancer, and well as detection of Gleason 4 disease. This does not aim to be an extensive review or meta-analysis of the available evidence, but rather a summary of biomarkers currently used in clinical practice, and promising and emerging biomarkers. A summary of available biomarkers and use can be seen in Figure 1-7 and Table 1.2.

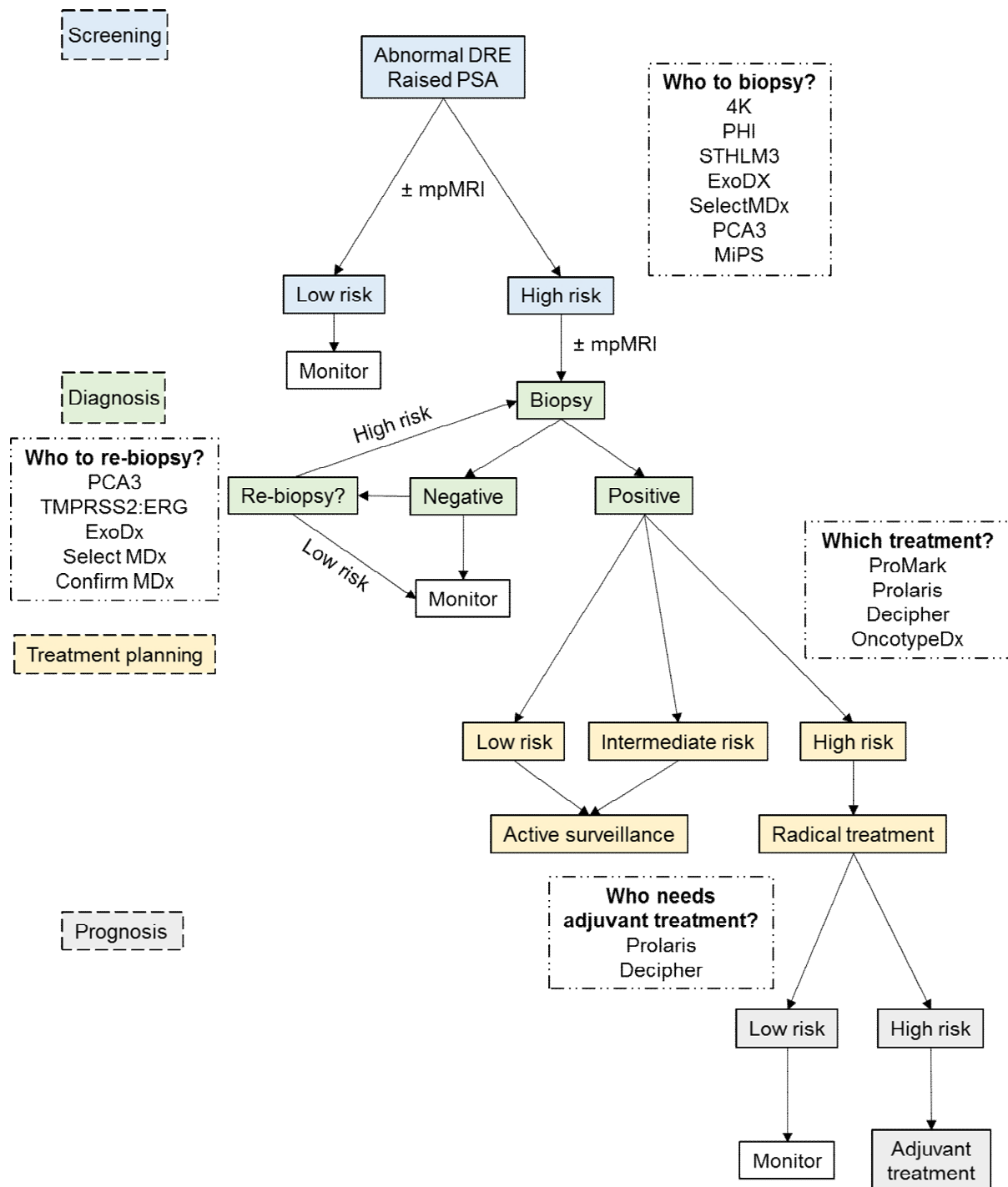


Figure 1-7 Biomarker use in the diagnosis and treatment planning of localised prostate cancer. Schematic representation of a biomarker-based diagnostic pathway including screening (blue), diagnosis (green), treatment planning (yellow) and prognosis/need for adjuvant therapy (grey). Biomarkers in dashed-dotted squares.

| Clinical pathway | Test | Company | Tissue | Marker | Accuracy | Cost-effectiveness |
|--|--------------------|-------------------------------------|--------|--|---|---|
| Diagnosis Pre-biopsy | 4K score | OPKO lab, USA | Blood | Total PSA, free PSA, intact PSA, hK2 and clinical factors | Sensitivity of 0.74 (95% CI, 0.726-0.76), specificity of 0.60 (95% CI, 0.59-0.61), AUC any cancer 0.69, 0.72 Gleason \geq 7 | |
| Diagnosis Pre-biopsy | PHI | Beckham Coulter Inc., USA | Blood | Total PSA, free PSA, [-2]proPSA | Sensitivity 0.93 (95% CI, 0.90-0.95), specificity 0.26 (95% CI, 0.25-0.28), AUC 0.72 for Gleason \geq 4+3 | Not cost-effective (115) |
| Diagnosis Pre-biopsy | STHLM3 | N/A | Blood | PSA, free PSA, intact PSA, hK2, MSMB, MIC1 and 254 SNPs and clinical | AUC 0.74 (95% CI, 0.72-0.75) For Gleason \geq 7. Reduction of the overall number of biopsies by 32% | |
| Re-biopsy | PCA3 | Hologic, USA | Urine | PCA3 | Sensitivity 47-75% , specificity 71-87%, AUC 0.68-0.87 | Not cost-effective (115) |
| Diagnosis Pre-biopsy | ExoDX Intelliscore | | Urine | PCA3, ERG, SPDEF and clinical parameters | AUC 0.77 (95% CI, 0.71-0.83) to discriminate Gleason 7, cut-off value of 15.6 avoids 26% of biopsies | Highest QALY when compared to standard of care, PHI and 4K test (116) |
| Diagnosis Pre-biopsy | Select MDx | MDxHealth, USA | Urine | DXL1, HOXC6, KLK3 | AUC of 0.90 (95% CI 0.87-0.93) with clinical variables. Cut-off of -2.8 (sensitivity: 96%, NPV 98% for Gleason \geq 7), results in 42% reduction of total number of biopsies. | Potential savings of €128 and a gain of 0.025 QALY per patient compared to the standard of care (117). |
| Diagnosis Pre-biopsy | MiPS | University of Michigan ML labs, USA | Urine | PCA3, TMPRSS-ERG, ERSPC risk parameters | 0.84 (95% CI: 0.80-0.87) for diagnosis of prostate cancer. AUC 0.84 (95% CI: 0.78-0.85) if TMPRSS2:ERG \geq 10 for diagnosis of Gleason \geq 7. | Projected savings of \$1200-2100 per patient if used as a triage prior to biopsy (118). |
| Re-biopsy? | Confirm MDx | MDxHealth, USA | Tissue | DNA hyper-methylation of GSTP1, APC and RASSF1 | AUC 0.76 and NPV of 96% for Gleason \geq 7 using Intelliscore | Savings of \$588 per patient management (119). |
| Treatment planning Adjuvant therapy | Prolaris | Myriad genetics, USA | Tissue | 31 cell cycle progression genes normalised to 15 housekeepers | Hazard ratio 1.88 per 1 point increase for biochemical recurrence. High score associated with risk of biochemical recurrence, metastasis and death (HR: 1.77) | Small savings compared to its cost, not enough evidence to recommend (120). |
| Treatment planning | ProMark | Metamark, USA | Tissue | 8 protein marker assay | PPV of score \leq 0.33 of 83.6% (95% CI 71.9-91.8%), specificity of 90%. | 0.04 more QALY and \$700 less in costs compared with usual care (121). |
| Treatment planning | OncotypeDx | Genomic Health, USA | Tissue | 12 prostate cancer-related genes and 5 reference genes | 20 point increase in the score associated with 2.3 fold increase of risk of high-grade disease (95% CI 1.5-3.7) and 1.9 fold increase of risk of \geq pT3 disease | Average saving of \$2286 per patient as more are recommended to have active surveillance over surgery (122). |
| Treatment planning Adjuvant therapy | Decipher | Genome Dx, Canada | Tissue | 22 RNA biomarkers | Prediction of metastasis (HR: 1.30, 95% CI 1.14-1.47, p <0.001), accumulated HR of 1.52 (95% CI 1.39-1.67) per 0.1 unite increase. AUC of 0.81 for 10 year metastasis when combined with clinical information | Decipher based care was less costly and resulted in greater QALYs than 100% adjuvant therapy. 100% adjuvant therapy had improved clinical outcomes at 5 and 10 years (123). |

Table 1.2 Summary of available biomarkers, with substrate, accuracy and cost-effectiveness evidence, if available.

1.2.2.2.1. Blood-based biomarkers

1.2.2.2.1.1. PSA

The performance of a diagnostic test relies on several indicators, mainly the ability of the test to find disease and not to diagnose healthy patients (sensitivity and specificity). The level of PSA to trigger investigations varies depending on age; for instance, a PSA of ≥ 3 ng/mL triggers a referral to a urologist in a 50 year old man. An 85-year-old man would need to have a much higher PSA value and higher suspicion of cancer to be referred. This is based on two large randomised PSA-screening trials, the European Randomised Study of Screening for Prostate Cancer (ERSPC) and the Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) trials (124,125). The widespread use of PSA has led to controversy, as some believe the overuse of PSA has not necessarily reduced prostate cancer mortality. Therefore, different organisations recommend for or against the use of PSA as a screening tool (126–128). In the UK, PSA is not used as a screening tool, instead it is requested on patient per patient basis, as part of an informed choice programme.

The Prostate Cancer Prevention Trial performed prostate biopsy routinely in asymptomatic men with PSA values below, to investigate the prevalence of cancer in this group. 449 (15.2%) of patients had cancer, of which 67 (14.9%) were high grade (defined as Gleason 7 or greater) (129). Thompson *et al.* calculated the PSA Receiver Operating Curve (ROC) in the placebo group of the same trial, and found no optimal cut-off point with high sensitivity and specificity. For instance, the ROC for any cancer vs no cancer was 0.678 (95% CI, 0.666-0.689). For Gleason ≥ 7 vs no or lower Gleason the ROC was 0.782 (95% CI, 0.748-0.816), and Gleason ≥ 8 versus no or lower Gleason, the ROC was 0.827 (95% CI, 0.761-0.893). Lower cut-off values of 1.1 ng/mL yielded a sensitivity of 83.4% and specificity of 38.9%, compared to 20.5% sensitivity and 93.8% specificity of a cut-off of 4.1 ng/mL (130). Furthermore, a meta-analysis calculated a pooled sensitivity, specificity, and positive predictive value for PSA of 72.1%, 93.2% and 25.1%, respectively (131). Although PSA is not the perfect diagnostic and prognostic biomarker, the use of PSA derivatives have shown some promise as tools to improve the diagnostic pathway and risk stratification of patients.

1.2.2.2.1.2. PSA density

The PSA density (PSAd) value is obtained by dividing the PSA level by the prostate volume. This takes into account the fact that prostate cancer releases more PSA per unit of volume than benign prostatic enlargement (132,133). Although some authors have reported that PSAd is only useful in a subset of patients (134) or it does not improve biparametric-MRI detection of prostate cancer (135), others have reported utility in combining PSAd with mpMRI information to aid biopsy decision making (136–138). For example, Washino *et al.* found that PSAd acted as an independent predictor for prostate cancer and significant prostate cancer (Gleason $\geq 3 + 4$ and/or an MCCL ≥ 4 mm). The AUC for PSAd was 0.815 (95% CI: 0.767–0.863), compared to PSA alone (0.642, 95%CI: 0.578–0.706) when combining with mpMRI information (PIRADSv2), and applying PSA density ranges of <0.15 , $0.15-0.29$, ≥ 0.3 ; the detection rate was highest when MRI suspicion was high (PIRADS 4-5), and PSA density was high

(76% if 0.15-0.29, and 97% if PSA_d ≥0.3). Additionally, in men with non-conclusive mpMRI (PIRADS 3), a PSA_d ≥0.3 had a detection rate of 86% (139).

1.2.2.2.1.3. 4K score

The 4K score (4 kallikrein) combines total PSA, free PSA, intact PSA and human kallikrein 2 (hk2) with factors such as age, digital rectal exam (DRE) findings and previous biopsy status (140–142). The 4K panel was initially developed on 740 men that underwent biopsy as part of the ERSPC trial. Here, Vickers *et al.* found that compared to PSA alone, the 4K panel improved the AUC from 0.68 to 0.83 when including age and PSA in the model, and from 0.72 to 0.84 when including PSA and digital rectal exam (142). Based on these findings, a predictive model was created and validated in the French cohort of the ERSPC, resulting in a calculated reduction of unnecessary biopsies (573 per 1000 men with raised PSA), while missing a small number of men with cancer (42 per 1000 men) (141). This was then, further validated in the Rotterdam subset of the ERSPC cohort, with similar findings (140). The 4K score may aid clinician's biopsy decision making, helping avoid between 41-71% of biopsies (143). In a meta-analysis by Russo *et al.* found a pooled sensitivity of 0.74 (95% CI, 0.726-0.76), pooled specificity of 0.60 (95% CI, 0.59-0.61) and a diagnostic odds ratio to detect high grade prostate cancer of 10.15 (95% CI, 8.06-12.79) (144). Currently, the European Association of Urology (EAU) guidelines do not recommend the routine use of 4K as an aid for prostate biopsy decision making (145).

1.2.2.2.1.4. Prostate Health Index (PHI)

This industry developed test, calculates an index based on total PSA, free PSA and [-2]proPSA ([-2] isoform of proenzyme prostate specific antigen) (146). The score is calculated as follows: $([-2]proPSA / \text{free PSA value}) \times \sqrt{\text{total PSA value}}$. Catalona *et al.* found that increasing PHI was related to a higher risk of prostate cancer (4.7 fold) and 1.61-fold increase of Gleason ≥7. When comparing PHI and percentage free PSA, PHI was better at discerning between Gleason ≥4+3 versus lower grade or negative biopsies. The authors concluded that PHI might be useful in reducing biopsies in men with a PSA between 2 and 10 ng/mL and a negative digital rectal exam (146). PHI has been approved by the US Food and Drug Administration (FDA) for use in the “grey” area when men have a PSA between 4-10 ng/mL. Russo *et al.* calculated a pooled sensitivity and specificity of 0.93 (95% CI, 0.90-0.95) and 0.26 (95% CI, 0.25-0.28), respectively (144). The same meta-analysis noted that there is some heterogeneity on cut-off values used in the different studies, to the advantage of sensitivity (144). For both 4K score and PHI, the index test is transrectal biopsy, and significant cancer is defined as Gleason 7 without specifying Gleason 4 component or MCCL. When used in conjunction with mpMRI, PHI has been shown to enhance the diagnostic ability (147).

1.2.2.2.1.5. Stockholm-3 (STHLM3)

In 2015, Grönberg *et al.* published their findings on a prospective population-based diagnostic study in Stockholm, Sweden. Patients aged 50-69 years old were invited randomly to take part in the study. A new model called STHLM3 was developed and consists of plasma biomarkers (PSA, free PSA, intact PSA, hK2, MSMB and MIC1), genetic markers (254 SNPs) and clinical characteristics (DRE and prostate volume).

The STHLM3 model aimed to identify men with Gleason ≥ 7 , and was developed in a cohort diagnosed with transrectal ultrasound. The model helped reduce the overall number of biopsies by 32%, 17% in Gleason 6 and 44% of benign patients. When comparing the ability to detect MCCL 0-10 mm versus >10 mm, there was a reduction in detection of Gleason 3+3 tumours compared to a PSA threshold of 3ng/mL (761 versus 608, $p < 0.001$), but not for Gleason >7 ($p = 0.82$). In other words, using a 3ng/mL threshold detects more insignificant cancer compared to STHLM3 (0-10 mm), but detects a similar amount of high-grade tumours regardless of the MCCL length (148).

The performance of this test in conjunction with mpMRI was tested in 532 men from a population-based study. Grönberg *et al.* found that using STHLM3 prior to targeted and systematic biopsy in men with a positive STHLM3 score, increased detection of Gleason ≥ 7 tumours by 10% ($n = 178$ vs 162 ; 95% CI 1.03–1.18). While sparing 38% of performed biopsy procedures. Compared to performing systematic and targeted biopsies in all men, this approach shows a lower detection of Gleason ≥ 7 cancer (Relative sensitivity 0.92; 95% CI 0.88–0.95) (148).

1.2.2.2.2. Urine based biomarkers

1.2.2.2.2.1. PROGENSA Prostate Cancer Antigen 3 (PCA3)®

In 1999 by Bussemakers *et al.* reported the discovery of a novel prostate-specific gene with high expression in prostate cancer (149). The transcripts of PCA3 mRNA can be detected in urine; this discovery led to the development of a commercially available quantitative test, measure following a digital rectal massage (150,151). A PCA3 score is calculated by dividing the PCA3 mRNA value by the PSA mRNA and the resulting value multiplied by 1000. This test is approved by the Food and Drug Administration (FDA) in men with a persistently raised PSA and negative biopsy. It has been reported to reduce the number of biopsies between 37 and 77.1% (59). Vlaeminck-Guillem *et al.* performed an analysis of 11 studies evaluating the diagnostic accuracy of PCA3 and found a sensitivity of 53-84% and specificity of 71-80% in biopsy naïve men. In men with a previous negative biopsy, the sensitivity was between 47-75%, and the specificity was between 71-87% (152). The ability of PCA3 to detect high-grade disease has been inconsistent, with some authors reporting a correlation between PCA3 score and T stage, tumour volume or Gleason score (153–155), others have reported that PCA3 cannot predict either (156–160). When PCA3 and clinical information is added to mpMRI, the AUC improved from 0.76 (95% CI 0.70-0.82) to 0.80 (95% CI 0.742-0.874). Whereas PCA3 alone has AUC values of 0.64 (95% CI, 0.58-0.70) with a threshold of ≥ 27 , 0.59 (95% CI, 0.51-0.66) with a threshold ≥ 35 and 0.67 (95% CI, 0.60-0.74) for a threshold of ≥ 44 (161).

1.2.2.2.2.2. ExoDX/ Intelliscore

ExoDX measures the expression of three genes (PCA3, ERG and SPDEF) from exosomes in urine (162). Exosomes are small vesicles secreted by normal and cancer cells that contain mainly mRNAs and microRNAs, making these vesicles enriched of material that may be low in cancer cells (162). It does not require a prior DRE or preservatives, improving its potential to apply to clinical practice (163). A score from 0-100 is obtained after applying an algorithm to the normalised PCA3/ERG RNA levels. McKiernan *et al.* examined the utility of the ExoDx and standard of care (PSA, age, race and family

history) versus standard of care alone to discriminate Gleason 7 and 6 and no cancer. A cut-off value of 15.6 avoided 26% of biopsies, while missing 5% of patients with dominant Gleason 4 patients. The calculated AUC was 0.77 (95% CI, 0.71-0.83) vs 0.66 (95% CI, 0.58-0.72) ($p < 0.001$) for standard of care alone (164). When comparing the cost-effectiveness of PHI, 4Kscore, Select MDX score, ExoDX and standard of care, the ExoDX provided the highest quality-adjusted life years (QALY) with an incremental cost-effectiveness ratio of \$58,404 per QALY, whilst avoiding 34.17% of unnecessary biopsies and reducing over-diagnosis by 20%, when compared to standard of care (116).

1.2.2.2.2.3. Select MDx™

The Select MDx™ measures mRNA relative expression of two genes: DLX1 and HOXC6. KLK3 (PSA) is used as an internal reference. First described by Leyten *et al.* in 2015. When compared to PCA3 and PSA, Select MDx had an improved AUC of 0.77 (95% CI, 0.71–0.83) compared to PCA3 AUC of 0.68 (95% CI, 0.62–0.75). This was further improved when serum PSA was added to obtain an AUC of 0.81 (95% CI, 0.75–0.86) (165). The test was validated in two prospective cohorts, aiming to develop a model combining molecular profiling and clinical risk factors to identify Gleason ≥ 7 (166). The AUC for each marker was PCA3: 0.65, HOXC6: 0.73 DLX1: 0.65 and TDRD1 0.69. The combination of HOXC6 and DLX1 yielded the highest AUC of 0.74. A combined logistic regression model was created by adding clinical information including age, PSA, PSA density, family history DRE and history of prostate biopsy reaching an AUC of 0.90 (95% CI 0.87-0.93). When a cut-off of -2.8 was applied (sensitivity: 96%, NPV 98% for Gleason ≥ 7), resulted in 53% decrease of unnecessary biopsies and a reduction of the total number of biopsies if 42% (166).

Using SelectMDx could reduce over-diagnosis and over-treatment in men with PSA levels of $>3\text{ng/mL}$, with potential savings of €128 and a gain of 0.025 QALY per patient compared to the standard of care (117). Combination of SelectMDx with mpMRI has shown a statistically significant difference between PI-RADS 3 and 4 ($p = < 0.01$) and between PI-RADS 4 and 5 ($p = < 0.01$) (167). Systematic transperineal biopsies combined with mpMRI fusion biopsies outperformed SelectMDx in a cohort of men in active surveillance, sensitivity: 66.6 versus 55.6%, specificity: 87.7 versus 65.8%, and diagnostic accuracy 84.9 versus 70.3%, respectively (168).

1.2.2.2.2.4. Mi Prostate Score (MiPS)™

The MiPS consists of is a post digital rectal exam urine test that measures RNA levels of the TMPRSS2:ERG fusion and PCA3 in addition to serum PSA levels (169). This test has been reported to improve the detection of prostate cancer and Gleason >6 disease with AUC of 0.77 (169). Other groups have measured the same biomarkers; for instance, Leyten *et al.* found that TMPRSS2:ERG provided additional predictive value to PCA3 and the European Randomised Study of Screening for Prostate Cancer (ERSPC) risk calculator parameters (serum PSA, volume, abnormal digital rectal exam and abnormal transrectal ultrasound). In the diagnosis of prostate cancer, the ERSPC parameters had an AUC of 0.79 (95% CI: 0.75-0.84), the addition of PCA3 score ≥ 25 increased the AUC to 0.83 (95% CI: 0.79-0.87), this was slightly increased to 0.84 (95% CI: 0.80-0.87) when TMPRSS:ERG was added. For detection of Gleason score ≥ 7 the values were 0.80 (95% CI: 0.73-0.86) for ERSPC parameters,

increased to 0.84 (95% CI: 0.78-0.85) if TMPRSS2:ERG ≥ 10 (170). This test is considered investigational by the EAU guidelines (145). When examining the cost-effectiveness of the MiPS test, 42% of biopsies could be avoided and projected savings of \$1200-2100 per patient if used as a triage prior to biopsy (118). The prevalence of the TMPRSS2:ERG fusion in different races was found to be higher in men of European descent (49%), followed by men of Asian descent (27%) and the lowest on men of African descent (25%). This may limit the use of this test in minorities (171).

1.2.2.2.3. Tissue based biomarkers

1.2.2.2.3.1. ConfirmMDx®

The ConfirmMDx® test evaluates DNA hyper-methylation of CpG islands of the promoter region of three genes: GSTP1 (DNA detoxification), APC (apoptosis), and RASSF1 (cycle cell regulation). This test studies negative formalin fixed paraffin embedded (FFPE) prostate biopsy tissue, in order to determine the risk of a nearby undetected tumour (172). ConfirmMDx® has been validated in a European cohort (United Kingdom and Belgium) and the United States of America (173–175). Stewart *et al.* tested the panel on bio-banked FFPE tissue from 498 patients with an initial negative biopsy, followed by repeat biopsy at 30 months. The NPV was 90% (95% CI 87-93). Following a multivariate model correcting for age, PSA, digital rectal exam and characteristics of the first biopsy (HGPIN, inflammatory atrophy and adenoids), the epigenetic panel was found to be an independent predictor of a positive subsequent biopsy (OR 3.17, 95% CI 1.81-5.53) (175). A separate study showed an NPV of 88% (95% CI 85-91), and an OR of 2.69 (95% CI 1.60-4.51) when the same multivariate model was applied (173). When validated on 211 African-American men, Waterhouse *et al.* found a similar diagnostic accuracy, with NPV of 78.8% (95% CI 71.5-84.6) for any prostate cancer and Gleason ≥ 7 94.2% (95% CI 88.7-97.1), this study had 81 men with cancer, of which 66% had Gleason 6 disease and included men with low PSA values (lowest value of 0.62 ng/mL); additionally, most men with cancer had incidental disease (T1c in 69%) (174).

In 2016 Van Neste *et al.* published Episcore, a methylation intensity algorithm that was then combined with clinical characteristics resulting in an NPV of 96% for Gleason ≥ 7 (166). All these studies lack long-term follow up and have small high grade (Gleason ≥ 7) cohorts. No comparison with or along mpMRI is available on the literature.

1.2.2.2.3.2. Prolaris®

Also referred to as the Cell Cycle Progression score (CCP), this test is an RNA expression panel of 31 genes relative to 15 housekeeper genes, for a total of 46 genes. The test can be performed on prostate biopsy and radical prostatectomy tissue, to predict several outcomes including decision making between active surveillance and radical treatment, prediction of biochemical recurrence following surgery and mortality (176–178). A systematic review and meta-analysis by Sommarivva *et al.* calculated a pooled hazard ratio (HR) for biochemical recurrence per 1 point increase in the score of 1.88 on univariate analysis, in other words for each increase in the score the risk of biochemical recurrence almost doubles (179). The multivariate model, including clinical and pathological variables (extent of disease, age clinical stage and use of hormones) obtained a combined HR of 1.63. The ASCO

guidelines recommend this test can be offered only when the result may change management (180). When examining the correlation between CCP score and mpMRI features, Wibmer *et al.* found that men with mpMRI signs of extracapsular extension had a significantly higher CCP score. There was no significant association between the CCP score and any of the PIRADSV2 score, or other mpMRI features (tumour volume and length on T2 and ADC, tumour ADC value or restricted diffusion (181). A systematic review failed to find enough evidence to support the cost-effectiveness of the test due to low quality of studies; the authors found that the test appeared to affect clinical decision in 64.9% of cases. Additionally, Prolaris® offered small savings compared to its cost (120).

1.2.2.2.3.3. ProMark ®

In 2014, Shipitsin *et al.* published a new proteomic biomarker panel for the prediction of prostate cancer aggressively and lethality. The panel was developed using two tissue microarrays made up of tissue from radical prostatectomy, 37 samples from Gleason 3+3 and 35 with Gleason 4+3 or 3+3/3+4 with T3b disease. From each tumour area, a low and high-grade area were selected, to model for random biopsy sampling missing the index biopsy. 120 biomarkers were selected from publicly available data based on four biological criteria, followed by technical criteria based on tissue staining of 62 candidates. 39 of these were then used to stain the tissue microarrays, to obtain a final 12 marker selection (182). The panel was further refined in 380 patients with an average 4 year follow up after radical prostatectomy, resulting in an 8 biomarker assay panel identifying favourable (Gleason \leq 3+4, \leq T2) versus unfavourable pathology (Gleason \geq 4+3, \geq T3/N/M). A risk score from 0 to 1 is calculated, the positive predictive value of a score \leq 0.33 was calculated at 83.6% (95% CI 71.9-91.8%), with specificity of 90%. Only 23.1% of patients with a risk score $>$ 0.80 had favourable disease (95% CI 11.1-39.3%) (183). The authors proposed a potential use for active surveillance, but no clinical trials to evaluate the performance of this test have been done so far. As with Prolaris®, the use of this panel is recommended by ASCO only if clinical management is likely to be changed based on the result (180).

1.2.2.2.3.4. Oncotype Dx Genomic Prostate Score ®

This test was designed to be used on prostate biopsy FFPE tissue; it consists of five reference genes and 12 cancer genes with a known role in prostate cancer, including androgen pathway (AZGP1, KLK2, SRD5A2 and FAM13C), cellular organisation (FLNC, GSN, TMP2 and GSTM2), proliferation (TPX2), and stromal response (BGN, COL1A1 and SFRP4) (184). Normalised gene expression is calculated and used to generate individual group scores that are then algorithmically combined to obtain and unscaled Genomic Prostate Score (GPS), ranging from 1 to 100 (184). The score aims to predict adverse pathology in order to risk assess patients before active surveillance. Multivariate analysis by Klein *et al.* calculated that a 20 point increase in the score was associated with a 2.3 fold increase of risk of high-grade disease (95% CI 1.5-3.7) and 1.9 fold increase of risk of \geq pT3 disease (185). All cohorts described by Klein *et al.* had a high grade of upgrade at radical prostatectomy, for example; the validation cohort at biopsy had 94 men with Gleason 7 disease, this number increased to 207 (52%) with Gleason \geq 7. This reflects the random biopsy approach with a relatively high proportion of patients upgraded at the time of surgery.

In an active surveillance retrospective study by Kornberg *et al.* men with a higher GPS score had a higher risk of biopsy upgrading (HR 1.28, 95% CI 1.19–1.39, $p < 0.01$). The same was found comparing PI-RADSv2 scores 1-2 versus 4 (HR 2.62, 95% CI 1.45–4.76, $p < 0.01$), and 1-2 versus 5 (HR 4.38, 95% CI 2.36–8.16, $p < 0.01$) (186). Although not discussed by the authors, the GPS score could be useful in men with PIRADS scores 1-2 to decide suitability to defer biopsy.

1.2.2.2.3.5. Decipher®

The Decipher® test or Genomic classifier (GC) was developed as an early prediction tool to determine the risk of metastasis after surgery. 545 patients from the Mayo Clinic Radical Prostatectomy Tumor Registry were included, all patients had a radical prostatectomy between 1987 and 2001 and were divided into three groups: no evidence of disease, PSA recurrence (matched controls) and clinical metastasis (cases). A 22 marker panel was generated from differential RNA expression using high-density expression arrays, obtaining a score from 0 to 1. The markers biological function include cell differentiation and proliferation, adhesion and motility, immune response, cell cycle progression/mitosis and androgen receptor signalling (187). The performance on the validation cohort of the GC alone, clinical characteristics alone and in combination resulted in AUC of 0.75, 0.74 and 0.69, respectively. A meta-analysis of five studies including data for 855 patients found that the Decipher test was able to predict metastasis (HR: 1.30, 95% CI 1.14-1.47, $p < 0.001$), with an accumulated HR of 1.52 (95% CI 1.39-1.67) per 0.1 unite increase. The clinical model alone had a c-index of 0.76, increased to 0.81 when the GC was added (188).

The decipher test has been reported to predict metastasis at 10 years at point of biopsy with HR of 1.34 to 1.75 per 10% increase in the score (185,189,190). In another study, the Decipher test had an AUC of 0.82, 95% CI 0.76-0.86, $p=0.0003$) when predicting metastatic prostate cancer following biochemical recurrence. Gleason score alone had an AUC of 0.63 (95% CI 0.58-0.70) (191). Higher GC has also been found to correlate to Gleason 4 percentage, expansive cribriform, both histopathological signs of tumour aggressiveness (192,193). As with Oncotype Dx, Prolaris and ProMark, the use of this test is only recommended by ASCO if management is likely to change when considered with clinical parameters (180).

1.3. The problem with current biomarker panels

As described above, multiple biomarkers have been developed aimed at answering different questions in the clinical pathway (Figure 1-7). Most of these were aiming at reducing the number of unnecessary biopsies or decide whom to re-biopsy and were developed in an era when prostate cancer was diagnosed with blind and random biopsies. With the emergence of multiparametric MRI as a tool to visualise prostate cancer and aid targeting of samples, many of these biomarkers utility in the diagnostic pathway seems lost. For instance, PHI and PCA3 have been examined in conjunction with mpMRI, and were found not to be a cost-effective strategy for the NHS (115). Although there is some evidence that some of these biomarkers enhance the diagnostic ability of mpMRI (147,161,194), studies looking at mpMRI in conjunction with biomarkers are rare. Development and examination of biomarkers in conjunction with a technique that has been shown to reduce the number of biopsies and over-diagnosis should be a priority (47). Novel biomarker panels should aim to improve mpMRI, especially when the results are inconclusive.

Additionally, the panels presented here were designed to detect “aggressive” disease, usually defined as Gleason ≥ 7 . As discussed on 1.2.1.6.1, Gleason 3+4 and 4+3 are distinct entities, especially when the percentage of Gleason 4 is taken into account. The Gleason 7 classification has left to be forgotten and has been replaced by the distinction between 3+4 and 4+3 as two separate categories. Therefore, biomarker panels should aim to detect primary Gleason 4, or ideally percentage of Gleason 4 with limits based on long term follow up with death outcomes. Furthermore, the detection of Gleason 3+4/4+3 on most panels was based on transrectal blinded ultrasound; for instance, Confirm MDx aims to detect missed cancer following a negative biopsy, this can be reduced by using mpMRI prior to biopsy (195,196).

Lastly, the estimation of Gleason 4 disease is performed visually by a trained histopathologist, as the exact percentage of Gleason 4 is of primordial importance both for risk stratification and as a target for biomarker development. Attempts should be made to streamline this process, possibly taking advantage of digital pathology and machine learning.

In this thesis, the development and validation of a novel biomarker panel that can be detected in fluidic samples is presented, taking advantage of cohorts with mpMRI before transperineal prostate biopsy. The biomarker panel would act as an extra piece of information and will be evaluated on its performance alone and in conjunction with mpMRI. Second, Gleason 4 was examined using quantitative methods and to compare digital pathology estimations to the standard Gleason 4 percentage estimation.

1.4. Lay summary

In the UK, one in six men will be diagnosed with prostate cancer in their lifetime. The prostate gland is located in the pelvis just below the bladder, and it depends on androgens to function. Like normal tissue, prostate tumours are dependent on androgens to grow and proliferate. This mechanism has been used to treat prostate cancer and has helped scientists understand how genes are turned on and off by increasing or decreasing the levels of androgens.

Problem 1: The diagnostic pathway

Currently, upon suspicion of prostate cancer, men will have a blood test called Prostate Specific Antigen (PSA), followed by samples of the prostate (biopsy), or in some centres, an image of the prostate using Magnetic Resonance Imaging (MRI) to decide if a sample needs to be taken. Although these tests have improved early diagnosis and lowered death rate, in some cases the tests miss men who need biopsies (false negative) or are inaccurate and advice a biopsy in a man who has no cancer (false positive). Additionally, the current pathway has led to significant over-diagnosis, ranging from 1.7 to 67% of cases.

Most centres take these biopsies through the rectum, carrying a risk of serious infections that require hospitalisation. Samples can also be taken through the perineum (the skin between the testicles and the rectum), but this also has some side effects, including pain and bleeding. As such, taking biopsies of the prostate should be performed after careful consideration.

Therefore, developing new tests with better accuracy, that reduce the number of unnecessary biopsies and over-diagnosis are acutely needed. Many companies have tried to develop such tests. Unfortunately, the available evidence does not support using these in clinical practice, and their use is minimal.

Problem 2: Prostate cancer biopsy-based classification can be subjective

If a man has a sample taken from their prostate, a pathologist will examine the tissue, and given a score called the Gleason score. Briefly, the pathologist compares the cancer to normal tissue. The Gleason score goes from 3 to 5, where 3 is very similar to normal tissue, 5 looks nothing like normal tissue, and 4 is something in between. The two most common scores in the sample are used to give the final Gleason sum = most common + second most common. Let us imagine a cancer with only Gleason 3 and 4; if there is <50% of Gleason 3 then the final score would be 3+4, if there is ≥50% Gleason 4, the final score would be 4+3.

Since greater amounts of Gleason 4 have been linked to increased mortality, many of the tests trying to see how aggressive the cancer in an attempt to predict when a patient has a lot of Gleason 4. This percentage calculation is performed visually by a trained histopathologist and can be influenced by experience. It can also change depending on who is examining the biopsy (subjective). New techniques like digital pathology, where the images are converted to digital images, could make this estimation

more accurate. For example, artificial intelligence algorithms have been used in pathology as a fast alternative that can be used to read the images and make the readings more objective.

This thesis has two main aims:

1. To describe the process of discovering and validating a new blood or urine test that measures to determine which patients would benefit from a biopsy (Chapter 3 and 4).
2. Study the amount of Gleason 4 and cancer length (in a biopsy) using digital pathology and compare this to what the pathologist is reporting (Chapter 5).

Chapter 2. Materials and methods

2.1. Materials

Unless otherwise stated all chemicals and reagents were supplied by Sigma Aldrich.

2.1.1. Cell lines and associated media

LNCaP (ATCC, CRL-1740), C4-2 (ATCC, CRL-3314), C4-2B (ATCC, CRL-3315) and PC3 (ATCC, CRL-1435)

Cultured in RPMI cell media supplemented with 10% FBS (Life technologies, Thermo Fisher Scientific, Massachusetts, US) stored at 4°C.

LNCaP, C4-2, C4-2B and PC3 starvation media

Phenol red free RPMI (Gibco™ through Thermo Fisher Scientific, Massachusetts, US) supplemented with 10% charcoal stripped FBS (Thermo Fisher Scientific, Massachusetts, US) stored at 4°C.

2.1.2. Drugs

Synthetic Androgen (R1881)

R1881 (CHEMOS GmbH) was diluted in 100% ethanol (Thermo Fisher Scientific, Massachusetts, US) to a working concentration of 1 nM.

2.1.3. RNA extraction

Cell line RNA extraction

miRNeasy mini kit (Qiagen, Manchester, UK) was used to extract RNA.

Blood mRNA extraction

PAXgene blood RNA kit (Qiagen, Manchester, UK) was used for extraction of RNA from blood stabilized in PAXgene blood RNA tubes (Qiagen, Manchester, UK).

Urine RNA extraction

QIAzol lysis reagent (Qiagen, Manchester, UK)

Chloroform:Isoamyl alcohol 24:1

Ethanol 100 % (Thermo Fisher Scientific, Massachusetts, US)

Nuclease-free water not DEPC-Treated (Ambion - Thermo Fisher Scientific, Massachusetts, US)

2.1.4. RNA quality control

Bioanalyzer automated electrophoresis

Agilent RNA 6000 Pico and RNA 6000 Nano reagents were used depending on the final expected RNA concentration (Agilent, California, US).

Qubit

Qubit RNA HS Assay (Thermo Fisher Scientific, Massachusetts, US).

2.1.5. cDNA

2720 Thermal cycler (Life technologies - Thermo Fisher Scientific, Massachusetts, US)

Cell line cDNA

High-Capacity cDNA reverse transcription kit (Thermo Fisher Scientific, Massachusetts, US).

Blood cDNA

Fluidigm reverse transcription master mix (Fluidigm, San Francisco, US).

2.1.6. Preamplification

Urine

Complete whole transcriptome amplification WTA kit (Sigma-Aldrich, Missouri, US)

GenElute PCR clean-up kit (Sigma-Aldrich, Missouri, US)

PAXgene whole blood

Preamp Master Mix (Fluidigm, San Francisco, US).

2.1.7. PCR

2.1.7.1. qPCR

Taqman gene expression assays (Thermo Fisher Scientific, Massachusetts, US) used in this thesis are described in Table 2.1.

TaqMan fast advanced master mix (Thermo Fisher Scientific, Massachusetts, US) was used in all reactions.

96 well hard-shell plate (BIO-RAD)

CFX-Connect Real-Time System (BIO-RAD)

2.1.7.2. Fluidigm high throughput qdPCR

Taqman primers and fast advanced master mix were used in all Fluidigm chips.

Control Line Fluid syringes for 192.24 and 96.96 IFC (Fluidigm, San Francisco, US).

96.96 and 192.24 IFC Fluidigm chips (Fluidigm, San Francisco, US)

96.96 and 192.24 IFC Fluidigm reagents (Fluidigm, San Francisco, US).

Juno System and Biomark HD System (Fluidigm, San Francisco, US) access courtesy of Dr Emma Ashton at Great Ormond Street Hospital for Children NHS Foundation Trust.

| Gene | Taqman probe |
|---------------------------|---------------------|
| <i>ABHD12</i> | Hs01018047 |
| <i>AGR2</i> | Hs00356521 |
| <i>EIF2B5</i> | Hs00384949 |
| <i>GFM1</i> | Hs00227997 |
| <i>MAD1L1</i> | Hs00269119 |
| <i>NDUFB11</i> | Hs00372638 |
| <i>NT5DC3</i> | Hs00213132 |
| <i>PECI (ECI2)</i> | Hs00196146 |
| <i>SEC61A1</i> | Hs01037684 |
| <i>SLC26A2</i> | Hs00164423 |
| <i>STIL</i> | Hs00161700 |
| <i>TM4SF1</i> | Hs00371997 |
| <i>TNFSF10</i> | Hs00921974 |
| <i>TRMT12</i> | Hs00535286 |
| <i>TSPAN13</i> | Hs00917717 |
| <i>XRCC3</i> | Hs00193725 |
| <i>AR</i> | Hs00907244 |
| <i>CBLN2</i> | Hs00699345 |
| <i>COL6A3</i> | Hs00915125 |
| <i>GDF15</i> | Hs00171132 |
| <i>GTF2H4</i> | Hs00231008 |
| <i>MSMB</i> | Hs00159303 |
| <i>NAALADL2</i> | Hs00822484 |
| <i>NT5C2L1</i> | Hs00261330 |
| <i>KLK3</i> | Hs02576345 |
| <i>SLC25A36</i> | Hs00216785 |
| <i>TMPRSS2</i> | Hs00237175 |
| <i>VPS28</i> | Hs00211938 |
| <i>PCA3</i> | Hs01371939 |
| <i>RPLP2</i> | Hs01115128 |
| <i>UBC</i> | Hs01871556 |
| <i>GAPDH</i> | HS02758991 |

Table 2.1 Gene name and Taqman assays.

List of all Taqman assays used for qPCR and qdPCR for cell lines and human sample analysis. Column on the left denotes gene ID name, column on the right indicates Taqman probe ID. All probes were purchased from Thermofisher.

2.2.Methods

Unless indicated otherwise all experiments were repeated independently on at least three occasions.

2.2.1. Cell culture

2.2.1.1. Maintenance of cell lines

Cell culture was performed in a class II laminar flow hood using aseptic technique. Cells were grown up to 60-70% confluence, washed with room temperature PBS and treated with trypsin-EDTA. Cells were divided 1:3 (LNCaP), 1:5 to 1:8 (PC3, C4-2 and C4-2B) twice a week. Flasks were labelled to include date, cell line and passage number.

2.2.1.2. Cell harvesting

Cells were grown up to 60-70% confluence. Following removal of media, the flasks were placed on ice. Two washes with cold PBS cells were scraped on to 1ml of PBS and pipetted into a labelled eppendorf tube. The tube was centrifuged at 5000 RPM for 4 minutes. PBS was removed prior cell pellet processing.

2.2.1.3. Cryopreservation of cell lines

Cells were grown to 60-70% confluence, washed twice with PBS and treated with trypsin-EDTA. Cells well pelleted by centrifuging at 1100 RPM for 3 minutes. Pellets were re-suspended in 1 ml of cell media with 10% DMSO and transferred to a labelled cryovial. Cells were slowly frozen using Nalgene® Mr Frosty (Sigma-Aldrich, Missouri, US) at -80°C before transferring to long-term storage in liquid nitrogen.

2.2.1.4. Retrieval of frozen cell lines

Cryovials were removed from liquid nitrogen and placed in a water bath at 37°C until thawed. Cells were transferred onto a T75 flask with 15 ml of pre-warmed culture media ensuring cell separation. Cells were passaged at least twice before use in any experiments.

2.2.1.5. Cell counting

Cells were harvested at specific densities depending on experimental set up. Cells were washed twice with PBS and scrapped on to PBS to create a cell suspension. 0.5 ml of such suspension was pipetted onto an eppendorf tube and trypan blue (0.4% final concentration) was added to give a 1 in 4 dilution. 100 µL of trypan blue-treated cell suspension was placed on the haemocytometer allowing the cell suspension to be drawn out by capillary action. The cells in the four large corner squares were counted

and an average cell number obtained. Using the following formula the number of cells per ml was calculated.

$$\text{Viable cells per ml} = \frac{\text{Average of counted live cells}}{\text{Number of squares (4)}} \times \text{Dilution} \times 10,000$$

2.2.1.6. Androgen time course

Cells at 50-70% confluence were serum starved using phenol red-free media and charcoal stripped FBS for 24 hours. 1 nM of R1881 or 0.01% ethanol was added prior cell harvesting and RNA extraction at 4, 8, 16, 24 and 48 hours.

2.2.1.7. Androgen stimulated cells exposed to urine to simulate RNA degradation caused by urine.

Donated urine from female volunteers was pooled and kept on ice. 160 ml of urine was taken and divided into four conical tubes. Two tubes were centrifuged at 2400 g for 20 minutes at 4°C to remove urine sediment (spun urine). Two labelled T75 flasks were filled with 40ml of unspun urine, two with spun urine and one with 40 ml of PBS.

LNCaP cells that had been androgen stimulated for 24 hour were harvested on to an eppendorf tube and mixed with 4.5 ml of cold PBS and counted. The cells were divided equally into three T75 flasks (PBS, unspun and spun urine) and one eppendorf for immediate pelleting and freezing at -80°C. The five T75 flasks were incubated for one hour at 37°C.

The contents of the T75 flasks were transferred to labelled 50 ml conical tubes and centrifuged at 4° C at 2400 g for 20 minutes. The supernatant was discarded and 1 ml of cold PBS added and mixed to re-suspend the pellet. 2µl of tRNA (9mg/mL) was added to each sample and the sample were stored at -80 °C for 30 minutes or until sample was frozen. tRNA was added to aid precipitation of low occurring RNA in the urine.

2.2.2. RNA extraction

Following extraction all RNA samples were aliquoted and diluted depending on experimental setting following extraction. Aliquots were frozen at -20 °C for short term storage and -80 °C for long term storage.

2.2.2.1. Cell line RNA extraction

Total RNA from cell lines was extracted using miRNeasy mini kit (Qiagen) following the manufacturer's instructions and omitting DNase digestion step. RNA was eluted in 40 µL of RNase free water. RNA yield and quality was measured using a Nanodrop spectrophotometer.

2.2.2.2. PAXgene blood RNA extraction

RNA was extracted using PAXgene blood RNA kit (RNA extraction was assisted by Ula Stopka-Farooqui); this tube contains a preservative that protects RNA from degradation by RNases and other molecules and triggers cell lysis of cells to reduce ex-vivo changes (197). An extra BR4 was buffer wash prior to elution was added to the protocol, following optimization performed by Ula Stopka-Farooqui. Following extraction RNA yield and quality was measured using a Nanodrop spectrophotometer, Qubit and Bioanalyzer as required for Fluidigm experiment. All samples were aliquoted as follows: 5uL for quality and control, 5uL for Fluidigm experiment, 10 and 60uL for future experiments.

2.2.2.3. Urine RNA extraction

Urine samples were collected as per local guidelines and RNA extracted by Dr Hayley Luxton and Dr Jonathan Kay with carrier RNA (PROMPT samples only). Cell pellets were lysed cells using QIAzol followed by a 5 minute incubation at RT. 200 µL of chloroform was added and samples repeatedly inverted for 15 seconds followed by a 2 minute incubation at RT. Samples were centrifuged at 12,000 g for 15 minutes at 4 °C. The colourless upper aqueous phase was removed to a fresh eppendorf tube and RNA precipitated by adding 500 µL 100 % isopropyl alcohol followed by centrifugation at 12,000 g for 10 minutes at 4 °C. Supernatant was removed and 1 ml of 75% ethanol added before vortexing and centrifuging at 7,500 g for 5 minutes at 4° C. The pellet was briefly air dried and re-suspended in 40 µL of RNase free water.

2.2.3. RNA quality control

2.2.3.1. Nanodrop

After zeroing with RNase free water, 1µl of each sample was used to measure RNA concentration. The amplitude of the curves and A260/230 and A260/280 ratios recorded to determine the concentration, purity or contamination of the samples prior to aliquoting.

2.2.3.2. Bioanalyzer

Agilent RNA Nano and Agilent RNA Pico kits were used following manufacturer's instructions. The resulting RNA integrity number (RIN) was recorded.

2.2.3.3. Qubit

RNA concentration was measured using Qubit RNA HS assay kit and 1 µl of sample. Prior to each use the machine was calibrated following manufacturer instructions.

2.2.4. Gene based assays

Unless otherwise specified all gene based assays and preparation was made on ice.

2.2.4.1. cDNA synthesis

Following normalisation of RNA concentration to 2µg (cell lines) or to the lowest concentration across all samples (urine), RNA was diluted to give a total volume of 10 µl. Following manufacturer's instructions for the high-capacity cDNA reverse transcription kit cDNA, samples were placed in a thermal cycler for conversion to complementary DNA using the following parameters: 25 °C for 25 minutes, 37 °C for 120 minutes, 85 °C for 5 minutes and 4°C indefinitely. 80 µl of RNase-free water was added to the final cDNA to dilute it.

2.2.4.2. qPCR

For all qPCR one to three housekeepers and a negative and positive control were included where available.

2 µl of cDNA were loaded on to a 96 well hardshell PCR plate with 18 µl of mastermix (10 µl Taqman fast advanced mastermix, 1µl of Taqman assay and 7µl of RNase free water). qPCR was performed using a CFX-Connect Real-Time System with a 2 minute UNG hold at 50 °C, followed by polymerase activation at 95 °C for 20 seconds and 40 PCR cycles (1 second at 95 °C to denature and 20 seconds at 60 °C to extend/anneal). For data analysis see 2.2.6.1.

2.2.4.3. Preamplification

RNA was amplified using the complete whole transcriptome amplification WTA kit following manufacturer's instructions. The resulting cDNA was purified GenElute PCR clean up kit following manufacturer's instructions.

2.2.4.4. Fluidigm qdPCR

2.2.4.4.1. cDNA

All samples were normalised to 80 ng. 4µl of reverse transcription pre-mix was added to 1 µl of RNA and placed on a thermal cycler for reverse transcription following the manufacturer's protocol. Samples were immediately pre-amplified.

2.2.4.4.2. Preamplification

Following the gene expression pre-amplification with Fluidigm PreAmp mastermix and Taqman assays protocol all samples underwent 16 cycles of Preamplification. A total of 25 assays were included in the Taqman assay pool (see

Table 2.2). Products were diluted on 20 µl of TE Buffer (10mM Tris-HCL, 1.0 mM EDTA) and stored at -20 °C.

| Chip Type | | | | |
|-------------------|----------|-----|---------|-----|
| 96*96 | | | 192*24 | |
| Chip number | | | | |
| One | Two | and | Four | and |
| | three | | five | |
| Taqman assays | | | | |
| NT5DC1 | NT5DC1 | | | |
| TRMT12 | TRMT12 | | | |
| ECI2 | ECI2 | | ECI2 | |
| NAALADL2 | NAALADL2 | | | |
| TM4SF1 | TM4SF1 | | TM4SF1 | |
| TSPAN13 | TSPAN13 | | TSPAN13 | |
| AR | AR | | | |
| GFM1 | GFM1 | | | |
| NDUFB11 | NDUFB11 | | | |
| STIL | STIL | | STIL | |
| SLC25A36 | SLC25A36 | | | |
| SLC26A2 | SLC26A2 | | SLC26A2 | |
| COL6A3 | COL6A3 | | | |
| EIF2B5 | EIF2B5 | | EIF2B5 | |
| TNFSF10 | TNFSF10 | | TNFSF10 | |
| TNFSF11 | AGR2 | | AGR2 | |
| GDF15 | GDF15 | | GDF15 | |
| ABHD12 | ABHD12 | | ABHD12 | |
| NT5DC3 | NT5DC3 | | | |
| GTF2H4 | GTF2H4 | | | |
| MAD1L1 | MAD1L1 | | MAD1L1 | |
| Housekeeper genes | | | | |
| RPLP2 | RPLP2 | | RLPL2 | |
| UBC | UBC | | UBC | |
| GAPDH | GAPDH | | GAPDH | |

Table 2.2 Taqman assays used for pooling prior to preamplification.

A total of five chips were used for three sample and assay replicates. Genes used in each chip are denoted in left column (chip one), middle column (chips 2 and 3) and right column (chips 4 and 5).

2.2.4.4.3. Fluidigm planning and loading

Depending on the Fluidigm chip to be used (96 x 96 or 192 x 24 wells) assays and sample position was planned using Excel. Each Fluidigm chip was primed on the Juno system and loaded following manufacturer's instructions. Gene expression was measured using the Biomark module and data downloaded for analysis using the Fluidigm real time PCR analysis software.

2.2.5. Digital pathology for measurement of Maximum Cancer Core Length (MCCL) and Gleason 4 burden

Pathologists were blinded to the PROMIS Gleason score; each patient was given an ID from 1 to 30, blinded from hospital number, PROMIS ID and pathology number. Using a random number generator (Excel) a patient was selected and assessed by two experienced uropathologists (Alex Freeman/Aiman Haider) using NDP.View 2 software.

2.2.5.1. Scanning of H&E slides

Haematoxylin and eosin (H&E)-stained 4 µm sections were retrieved from the UCL/UCLH Biobank for Studying Health & Disease. Slides were scanned using a Hamamatsu scanner (model C10730-12, Hamamatsu Photonics, Japan), uploaded to NDPscan software (Hamamatsu Photonics, Japan) at 454nm/pixel resolution, source lens 40x, with capability to zoom up to 80x. All scanning performed by Dominic Patel (UCL/UCLH Biobank for Studying Health & Disease).

2.2.5.2. Systematic assessment of digital slides

Each slide was methodically assessed as follows: 1. Each core was numbered from left to right. 2. Length of cancer was measured. 3. Areas containing any cancer were contoured. 4. Areas containing Gleason 4 were contoured.

2.2.5.2.1. MCCL measurement in digital images

The MCCL was reported prospectively by the pathologists during the PROMIS trial using the integrated ruler in the microscope; this measurement was assigned as 'visual' MCCL. In PROMIS, the MCCL was reported by taking into account intervening benign glands (ISUP) and measuring cancer only (non-ISUP). For the purposes of this thesis, the ISUP measurement was used. The 'digital' MCCL was derived as follows: If a core was straight, a single measurement was performed. If there was any curvature, manual sequential measurements were performed along the core axes and combined to give the final measurement (Figure 2-1 A).

2.2.5.2.2. Gleason 4 contouring

All cancer was contoured in yellow (Figure 2-1B), followed by contouring of areas with Gleason 4 (Figure 2-1B). Each Gleason 4 contour was given an ID based on its location on the core (Pathology number–block number, core number, contour number).

2.2.5.2.3. Calculation of Gleason 4 percentage

Percentage Gleason 4 (%G4) was not collected as part of the original trial, pathologists retrospectively visually estimated the %G4 per patient to the closest 10% using the annotated images. This was assigned as 'visual' %G4. For digital %G4, the software performs instant area measurements based on the contours of all cancer and individual areas of pattern 4 previously delineated. The resulting area (for each yellow and black contours) was prospectively recorded, and an objective percentage of G4 was calculated as shown in equation 1 (Figure 2-1C). This total was assigned as 'digital' %G4. A separate analysis of the index block was performed separately. The index block was defined as the block with the highest Gleason score and MCCL in combination with concordance with the index lesion on mpMRI.

2.2.5.2.4. Data collection

Study data were collected and managed using REDCap electronic data capture tools hosted at University College London (198,199). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing an interface for validated data capture, audit trails for tracking data manipulation, automated export procedures; and procedures for data integration with external sources.

The initial data dictionary was created with help of Feargus Hosking-Jervis and modified to fit the purpose of the study. Access to the database was extended to Urszula Stopka-Farooqui and Cristina Cardona Barrena for data input. The data collected includes: Pathology ID, PROMIS ID, pathologist scoring (AH) and revision of scoring (AF). For each patient, volume of each contour (cancer and Gleason 4) was recorded and added to obtain a volume per core. This was then added to obtain a total volume per patient.

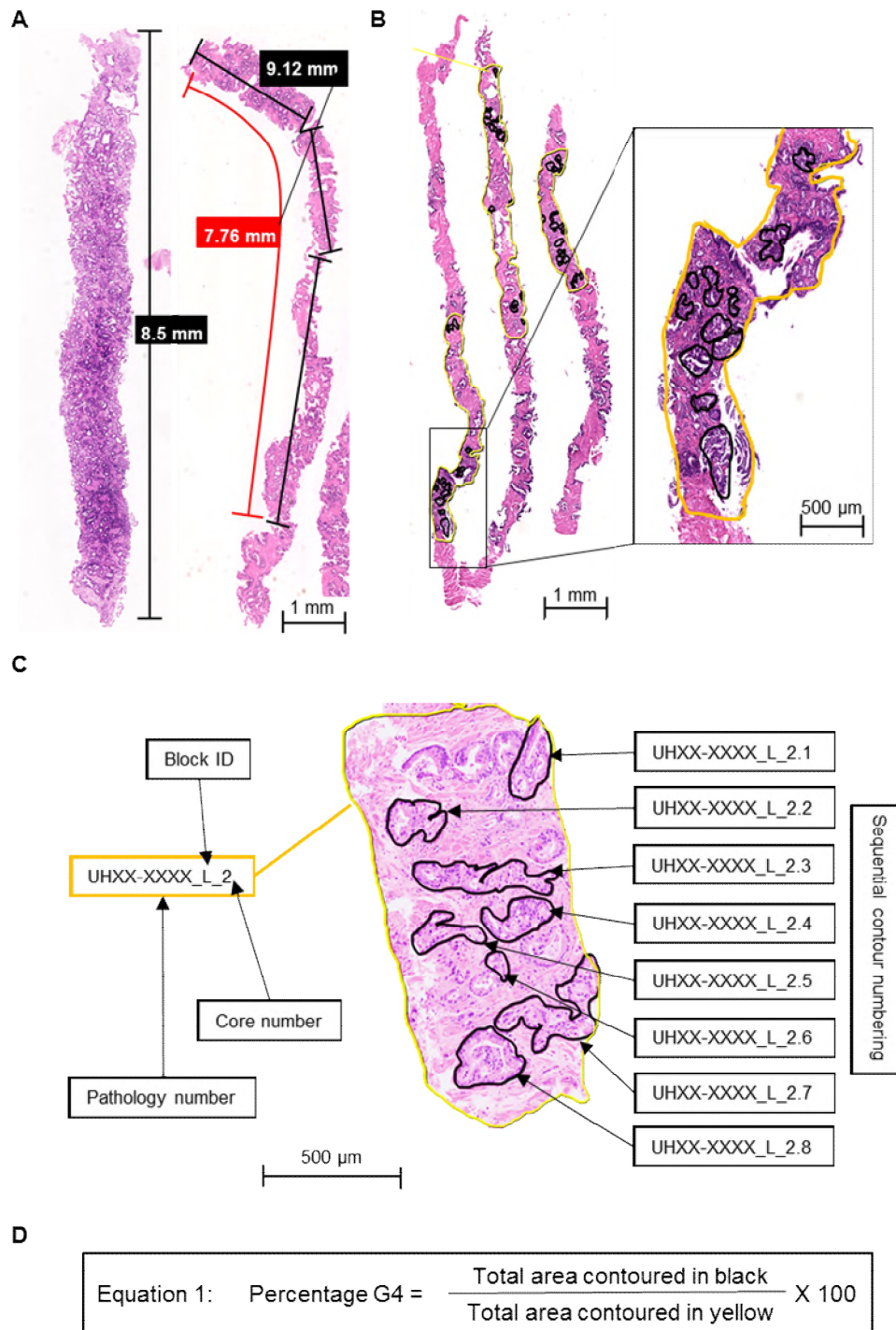


Figure 2-1 Gleason 4 contouring and MCCL measurement.

NDPview2 image of scanned H&E slide of prostate cores from transperineal biopsies, where nuclei are shown in blue, and other structures in pink. **A.** From left to right, MCCL measurement in a straight core of 8.5 mm. Approximate visual pathologist measurement marked with a red line (7.76 mm). Following the axis of the core, three measurements in black of 2.53 mm, 2.11 mm and 4.48 mm for a total of 9.12 mm for the digital measurement. **B.** Three prostate cores, areas with cancer were contoured in yellow; areas with Gleason 4 were contoured in black. Close up of contours shown in black box. Non-contoured areas correspond to benign prostatic tissue. **C.** Block, core and contour numbering for data analysis. Each core was identified using the pathology ID starting with UH- number, followed by block ID (depending on location of biopsy) from A to T, and core number. Each black contour was given an identifier from top to bottom as X.1, X.2, etc. **D.** Equation used to derive percentage Gleason 4.

2.2.5.3. Applying geographical parameters, evenness and clustering to Gleason 4 patterns of distribution.

In collaboration with Tian Lan from the UCL Geography department an in-depth analysis of the Gleason 4 distribution was performed. The work presented here is Tian Lan's contribution to this study. Annotations were extracted from NDP.View files onto .csv files. Analogy was drawn to the study of residential segregation among black and white ethnic groups in social science (200–202). Here, Gleason 4 was considered as the “minority” population surrounded by the majority population (Gleason 3).

Automated calculation of evenness and clustering scores was performed using Python 2.7.12 (64bit) and the following packages: Geojson 2.3.0, Fiona 1.7.11.post11, shapely 1.6.4.post1 and rtree 0.8.3 (203–207). Annotations were extracted into GIS format file Geojson for processing, this was done core by core (Figure 2-2 A). To model the Gleason 3 and 4 distribution, the contoured areas were sampled into grid points and it was assumed that cells were distributed evenly.

Each G3 or G4 point has a conceptualised “neighbourhood” (see Figure 2-2 E), which is a circular area around the point with pre-defined bandwidth or radius. Measurement H (Equation 1) and SP (Equation 4) respectively quantify the average entropy and average pairwise distance of two group members in these neighbourhoods, weighted by population.

The evenness and clustering are measured respectively by the spatial entropy index (Theil's information theory index) H and spatial proximity index SP, these are explained in depth in Chapter 5 (Section 5.3.2). For each core, an evenness and clustering score was obtained.

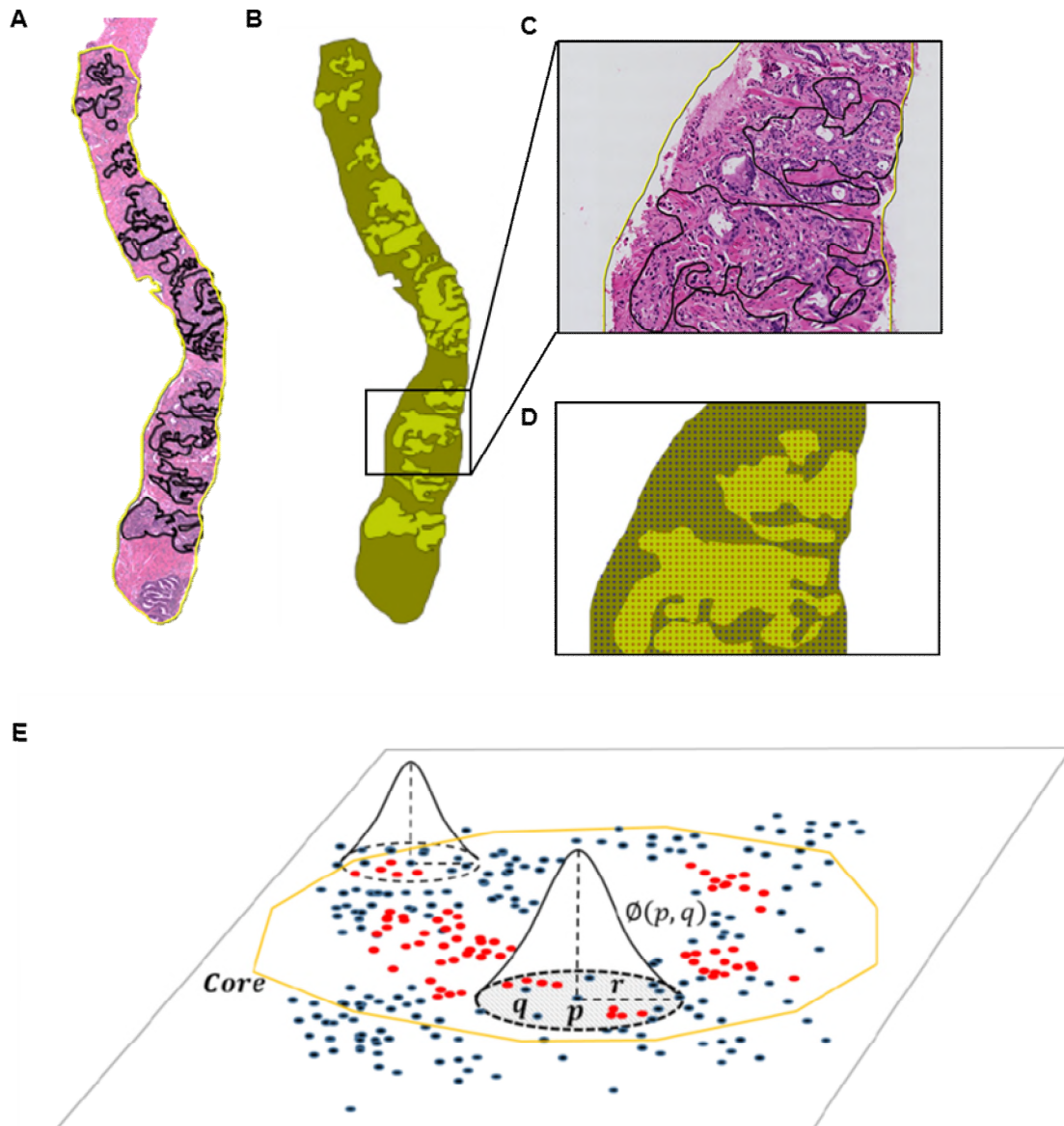


Figure 2-2 Calculation of clustering and evenness scores from H&E scanned slides.

A. NDPview2 image of scanned H&E slide of prostate core, where nuclei are shown in blue, and other structures in pink. Cancer area contoured in yellow, Gleason 4 areas contoured in black. **B.** Converted core into Geojson format, where light green corresponds to Gleason 4 areas and dark green to Gleason 3 areas. Grid distribution of Gleason pattern 4 in red and Gleason 3 in blue. Close up of a core region as H&E **C.** and as a Geojson image. **D.** Dots are evenly spaced for simplicity of the model. **E.** Neighbourhood conceptualization by Gleason 3 or 4, each point is surrounded by a circular area with a pre-defined bandwidth. Each dot corresponds to Gleason 4 (red) or Gleason 3 (blue), representing uneven distribution of the two Gleason grades.

2.2.5.3.1. Analysis of location and zone in relation to the evenness and clustering scores.

In order to analyse the data, correlation analysis was performed by Saheer Saeed under supervision of Yipeng Hu (University College London). Four variables were used in this analysis:

1. Core location using the 20 modified Barzell zones (Figure 2-3).
2. Area of cancer obtained from cancer contours described on section 2.2.5.2.
3. Clustering score.
4. Evenness score.

2.2.5.3.1.1. One-hot encoding for location information.

In order to analyse categorical data, such as location in the prostate, data was one-hot encoded. This method involves creating dummy variables for each Barzell zone. Each dummy variable has a binary variable or either 1 if the sample comes from that zone, or 0 if it doesn't (Figure 2-3 B, C). A total of 20 dummy variables were created (A-T or 1-20)

2.2.5.3.1.2. Ordinal encoding using a discrete 2D coordinate system.

A coordinate system was created to obtain coordinates of zone location. A grid was overlaid onto the Barzell zone scheme, encoding location as follows: -3 to +3 from left to right and -2 to +2 from top to bottom. Each block/zone has a coordinate, for example: block 2 (1,-1), block 3 (-1,2), etc. (Figure 2-3D). All blocks were encoded using this coordinate system, creating two ordinal variables, corresponding to the location on the x-axis and y-axis (Figure 2-3 E).

2.2.5.3.1.3. One-hot encoding for central and peripheral zone

Two binary dummy variables were created for the central and peripheral zone. If a sample originates from the central zone, the dummy variable for central zone takes on a value of 1 and the dummy variable for peripheral zone takes a value of 0.

2.2.5.3.1.4. Point bi-serial correlation analysis

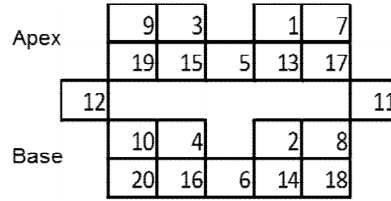
The point bi-serial test splits the population in two: a population where the binary variable is 1 and one where the binary variable is 0. Using the dummy variables, a correlation coefficient was obtained by doing a bi-serial test between the binary dummy variable for the Barzell zone and the area, evenness and clustering values. Similarly, a bi-serial test was performed for the zonal binary dummy variable and area, evenness and clustering.

2.2.5.3.1.5. Spearman's rank correlation

The Spearman's rank correlation is a nonparametric test that measures the association between two ranked variables. This test was applied to the ordinal data obtained from the grid system.

A

Modified Barzell zones



| | | | |
|----|----------------------------------|----|-----------------------------------|
| 1 | Left parasagittal anterior apex | 11 | Left lateral |
| 2 | Left parasagittal anterior base | 12 | Right lateral |
| 3 | Right parasagittal anterior apex | 13 | Left parasagittal posterior apex |
| 4 | Right parasagittal anterior base | 14 | Left parasagittal posterior base |
| 5 | Midline apex | 15 | Right parasagittal posterior apex |
| 6 | Midline base | 16 | Right parasagittal posterior base |
| 7 | Left medial anterior apex | 17 | Left medial posterior apex |
| 8 | Left medial anterior base | 18 | Left medial posterior base |
| 9 | Right medial anterior apex | 19 | Right medial posterior apex |
| 10 | Right medial anterior base | 20 | Right medial posterior base |

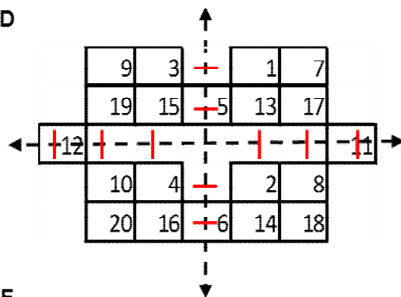
B

| Sample | Barzell zone |
|-------------|--------------|
| UHXX-XXXX_2 | 1 |
| UHXX-XXXX_3 | 3 |
| UHXX-XXXX_4 | 2 |
| UHXX-XXXX_5 | 1 |

C

| Dummy variable corresponding to Barzell zone | | | |
|--|---|---|---|
| Sample | 1 | 2 | 3 |
| UHXX-XXXX_2 | 1 | 0 | 0 |
| UHXX-XXXX_3 | 0 | 0 | 1 |
| UHXX-XXXX_4 | 0 | 1 | 0 |
| UHXX-XXXX_5 | 1 | 0 | 0 |

D



E

| Barzell zone | x-axis value | y-axis value |
|--------------|--------------|--------------|
| 10 | -2 | -1 |
| 10 | -2 | -1 |
| 13 | 1 | 1 |
| 20 | -2 | -2 |
| 2 | 1 | -1 |

F

| Dummy variable corresponding to Barzell zone | | |
|--|---------|------------|
| Zone of sample | Central | Peripheral |
| Central | 1 | 0 |
| Central | 1 | 0 |
| Peripheral | 0 | 1 |

Figure 2-3 Barzell zones and one-hot encoding of variables.

A. Graphical representation of the anatomical position of each modified Barzell zone, from 1 to 20. Where zone 1 corresponds to block A, zone 2 to block B, and so on. Table contains exact anatomical position of each zone. **B.** Example of sample ID and corresponding Barzell zone. **C.** Hot-one encoded samples, where dummy variables are created for each Barzell zone, presence or absence of that core in X zone is encoded as 1 if present and 0 if absent. **D.** Grid overlaid onto Barzell zone map, red dashes correspond to discrete numbers from -3 to +3 from left to right, and -2 and +1 from bottom to top. **E.** Location coordinate of sample Barzell zones using the grid presented in D. **F.** One-hot encoding for central and peripheral zones.

2.2.6. Data analysis

2.2.6.1. Δ Ct and $\Delta\Delta$ Ct qPCR and qdPCR data analysis

To obtain the Δ Ct values the genes of interest were normalised to between one and three housekeeper genes depending on experimental setting, in the case of human samples only Δ Ct was attainable, as no standard exists to obtain $\Delta\Delta$ Ct. The $\Delta\Delta$ Ct was calculated by subtracting the sample of interest minus a standard or reference sample (208). The resulting number was then unlogged by subtracting $2^{\Delta\Delta\text{Ct}}$ from the $\Delta\Delta$ Ct (See equation below).

$$\begin{aligned}\Delta\text{Ct} &= (\text{Average Ct gene of interest}) - (\text{Average Ct housekeeper gene}) \\ \Delta\Delta\text{Ct} &= \text{Experimental } \Delta\text{Ct} - \text{Control } \Delta\text{Ct} \\ \text{Fold change} &= 2^{-\Delta\Delta\text{Ct}}\end{aligned}$$

For data visualisation GraphPad Prism 7 or R programming environment (<http://www.R-project.org/>, version 3.5.1) and the ggpubr package were used (209).

2.2.6.2. Statistical analysis

2.2.6.2.1. Assessment of normality of data distribution

Normality, or lack of, was analysed using a combination of visual methods and normality tests as follows:

2.2.6.2.1.1. Quantile-quantile plots (Q-Q plots)

Q-Q plots provide a visual representation of the correlation between given samples and the normal distribution. A perfect 45-degree reference line is added. A perfectly correlated dataset will plot on the $x=y$ line. Plots were created using base R (210,211).

2.2.6.2.1.2. Density plots

The density plots provide another visual representation of the data distribution, and how it resembles a normal distribution bell shape. Plots were created using ggpubr package in R (209).

2.2.6.2.1.3. Shapiro-Wilk's test

The Shapiro-Wilk tests the null hypothesis that a sample came from a normally distributed population. If the p value is below 0.05, the null hypothesis is rejected and data is considered not normally distributed. Analysis was performed using base R (210,211).

2.2.6.2.2. Assessment equality of variances.

Parametric tests, such as t-test and ANOVA, rely on equal variance distribution between the group/groups being compared. Equality of variances in this thesis was analysed using Bartlett's/F test or Levene's test if data was normally or not normally distributed, respectively.

2.2.6.2.2.1. Bartlett's test

Bartlett's test checks homogeneity of variances in normally distributed data (homoscedasticity). The null hypothesis is that there is no difference in variances between the groups. When the resulting p value is >0.05 there is homoscedasticity. Analysis done with base R (210,211).

2.2.6.2.2.2. F- tests

The F-test also measures the homogeneity of variances; this test is extremely sensitive to non-normality. Analysis performed using base R (210,211).

2.2.6.2.2.3. Levene's test

This is an alternative to Bartlett's test, if data are not normally distributed. Analysis performed using car package in R (210,211).

2.2.6.2.3. Comparison of means of two groups

Student's t-test was performed when samples were normally distributed and had homoscedasticity. If these conditions were not met, two alternatives are possible: Data with equality of variances but not normally distributed, here a Wilcoxon signed-rank test was performed. The second possibility is data that normally distributed with heteroscedasticity, here a Welch-Satterthwaite t-test was performed instead. Analysis done with base R (210,211).

2.2.6.2.4. Comparison of means of more than two groups

ANOVA test was performed when assumptions of normality and variance were met. Tukey post-hoc test was performed, comparing the mean of all groups against each other. The Kruskal-Wallis rank sum test, is the non-parametric alternative to the ANOVA test, when assumptions of normality and variance are not met. Post hoc test for the Kruskal-Wallis applied was Nemenyi test using the PMCMR package (212).

2.2.6.2.5. Time course analysis

Time course results were plotted using GraphPad Prism 7. A two tailed paired t-test was performed to compare each treated time point to the control using base R (210,211).

2.2.6.2.6. Bland-Altman test

The Bland-Altman method measures the correlation between two measurements measuring the same outcome. The resulting plot describe the agreement between the two measurements by constructing limits of agreement between the two tests. The difference between the two measurements is calculated and plotted against the mean of the two measurements. Ideally the plotted values should lie between $\pm 2s$ of the mean difference (213).

Bland–Altman plots allow identification of any systematic difference between the measurements (i.e., fixed bias) or possible outliers. The mean difference is the estimated bias, and the SD of the differences measures the random fluctuations around this mean. If the mean value of the difference differs significantly from zero based on a 1-sample t-test, this indicates the presence of fixed bias. If there is a consistent bias, it can be adjusted for by subtracting the mean difference from the new method. It is common to compute 95% limits of agreement for each comparison (average difference ± 1.96 standard deviation of the difference), which tells us how far apart measurements by two methods were more likely to be for most individuals. If the differences within mean ± 1.96 SD are not clinically important, the two methods may be used interchangeably. The 95% limits of agreement can be unreliable estimates of the population parameters especially for small sample sizes so, when comparing methods or assessing repeatability, it is important to calculate confidence intervals for 95% limits of agreement.

2.2.6.3. Missing data analysis

Since statistical inferences are drawn from data analysis, ensuring that potential biases are taken into account can allow for a more accurate interpretation of the data. As a result, when missing data is simply deleted, sample size is reduced, biases are introduced as follows:

1. The final population has changed, for example, if most missing values belong to one group.
2. Effect estimates are biased, for example when comparing means in a t-test (as above).
3. Sample analysis may no longer belong to a random population, by selecting samples that express a gene, data may be biased towards more aggressive tumours.

In general, there are three missingness mechanisms: MCAR, MAR and MNAR (missing completely at random, missing at random and missing not at random, respectively) (214). Data is MCAR when there are no systematic differences between observed and not-observed data. Data is considered MAR, when there are systematic differences between missing and observed values and these can be explained by differences in the observed data. For instance, non-detectable PSA values in men who have had radical prostatectomy when comparing to men who had hormone therapy, the absence of a prostate could explain why PSA is not detectable in one group. Finally, data is MNAR, when there are systematic differences between observed and not-observed data but these cannot be explained by observed data. For example, when erectile function questionnaires are not filled by patients due to unwillingness to share personal data. In this case, erectile function cannot be predicted using other variables.

Pattern analysis was performed using IBM SPSS Statistics for Windows, version 24.0 and UpSet package in R (210,211). Little's multivariate MCAR test was performed, the null hypothesis being that values are missing independently from the data values (214,215).

2.2.6.3.1. Multiple imputation for missing values

A multiple random imputation method was used following confirmation that the data was missing completely at random. Thousands of logistical regressions based on existing data were made by using

IBM SPSS Statistics for Windows, Version 24.0. A random combination of imputed data was included in the existing database. Instances where no values or two out of three values were missing were not included in the simulation. The resulting data was included in the final analysis.

2.2.6.4. Data mining

Data from there previously published datasets was downloaded from PubMed Gene Expression Omnibus (GEO) by Benjamin Simpson using the following IDs: Taylor *et al.* (GSE21034), Varambally *et al.* (GSM74875) and Ross-Adams *et al.* (GSE70770) each .rar file was converted to csv file. Gene probe conversion was done using g:Profiler g:Convert functionality. When multiple probes were identified for one gene in Affymetrix, Jetset package was used. This package provides a score based on specificity, coverage, and degradation resistance (216). When more than one gene was found on Illumina, the most variable probe was selected (as measure by IQR), followed by sequence BLAST of the selected probe to confirm specificity to the gene of interest.

2.2.6.5. Confusion matrix and calculation of accuracy parameters.

Confusion matrices or 2 x2 contingency tables were obtained by Avi Rosenfeld (Jerusalem Institute of Tehcnology). Analysis of performance parameters were calculated using the formulas depicted on Figure 2-4. Confusion matrices allow visualisation of true positive (test positive, disease present), false positive (test positive, no disease), true negative (test negative, no disease) and false negative (test negative, disease present). From these parameters, the performance of a test can be calculated, obtaining the following measurements:

2.2.6.5.1. Sensitivity

Measures the proportion of positive cases that are correctly found by the test. It is also called true positive rate. It ranges from 0-100% (or 0-1).

2.2.6.5.2. Specificity

Measures the proportion of negative cases that do not have a positive test, in other words, measures the true negative rate. It ranges from 0-100% (or 0-1).

2.2.6.5.3. Positive Predictive Value (PPV)

The PPV takes into account the sensitivity, specificity and prevalence. It measures the probability that positive tests mean the disease is actually present. A result of 1 means all the positive tests have the disease.

2.2.6.5.4. Negative Predictive Value (NPV)

As with the PPV (2.2.5.2.9.2), it takes the sensitivity, specificity and prevalence into account, in this case to obtain the probability that a negative test is actually free of disease. A result of 1 means the a negative test has a 100% probability of not having the disease.

2.2.6.5.5. False Negative Rate (FNR)

Measures the probability of a negative test to have the disease, or incorrectly identified as not having the disease. It is represented by the symbol α , where $1-\alpha$ is equal to the specificity of the test.

2.2.6.5.6. False Positive Rate (FPR)

Is the proportion of positive tests that did not have the disease, or incorrectly identified as having the disease. It is represented by the symbol β , where $1-\beta$ equals the sensitivity of the test.

2.2.6.5.7. False Discovery Rate (FDR)

The FDR represents the proportion of type I errors (incorrectly rejecting the null hypothesis). It is complementary of the PPV, as a PPV of 70% would correspond to a FDR of 30%. The FDR can be used to adjust the test to detect more or less false positives, depending on the objective of the test.

2.2.6.5.8. False Omission Rate (FOR)

The FOR measures the proportion of having the disease when testing negative. It is complementary of the NPV.

2.2.6.5.9. Accuracy

Represents the number of correctly “diagnoses” cases in the entire population. In other words, how often does the test get it right?.

2.2.6.5.10. Matthews Correlation Coefficient (MCC)

The Matthews correlation coefficient (MCC) takes into account the results and performance of the test in the four categories studied in the confusion matrix: false and true positives and false and true negatives. The result is a number between -1 and 1. Where -1 indicates complete incongruity between the observation and prediction (test), 0 where there is not better than a random prediction and 1, where the test is perfect at predicting the outcome (217).

A

| | | Disease | | |
|------|---------|----------------|----------------|----------------|
| | | YES (Ca) | NO (B) | |
| Test | Pos (1) | True Positive | False Positive | Total positive |
| | Neg (0) | False Negative | True Negative | Total negative |
| | | Total diseased | Total Healthy | Total cohort |

B

| | |
|---|--|
| Sensitivity | TP/totaldiseased |
| Specificity | TN/total healthy |
| Positive Predictive Value (PPV) | TP/TP+FP |
| Negative Predictive Value (NPV) | TN/TN+FN |
| False Negative Rate (FNR) | FN/TF+TP |
| False Positive Rate (FPR) | FP/FP+TN |
| False Discovery Rate (FDR) | FP/FP+TP |
| False Ommision Rate (FOR) | FN/FN+TN |
| Accuracy | TP+TN/Total Population |
| Matthews correlation coefficient | $((TP*TN)-(FP*FN))/((TP+FP)(TP+FN)(TN+FP)(TN+FN))^{(1/2)}$ |

Figure 2-4 Confusion matrix and performance analysis of a test.

- A.** Confusion matrix where prediction by the test is plotted against the confirmed diagnosis (disease).
B. Performance calculations on the left column and formula on the right column.

2.2.7. Cohorts

2.2.7.1. Prostate cancer: Mechanisms of progression and treatment (ProMPT study)

All clinical samples were collected from Cambridge University Hospitals NHS Trust, ethical approval was granted by the local research and ethics committee (LREC number: 02/281M) and by the multicentre research and ethics committee (MREC number 01/4/061). Informed consent was obtained from all subjects (218).

2.2.7.2. Prostate MR imaging study (PROMIS)

All clinical samples were collected from University College London Hospital NHS Trust patients who had consented to the study. Ethics committee approval was granted by National Research Ethics Service Committee London (reference 11/LO/0185). Access to biobank samples was obtained (reference (EC/21.16) (47).

2.2.7.3. Combining advances in imaging with biomarkers for improved diagnosis of Aggressive prostate cancer (INNOVATE)

Ethics committee approval was granted by National Research Ethics Service Committee London (reference 15/LO/2099). Samples were collected from June 2017 to May 2018. Clinical samples were taken at University College London Hospital NHS Trust and Barts Health NHS Trust in consented patients only (219).

Chapter 3. Characterisation of a novel gene panel for diagnosis and risk stratification of prostate cancer.

3.1. Introduction

Molecular signatures in different biological matrices are one way to improve the accuracy of prostate cancer diagnosis and risk stratification (220–223). The utility of these signatures can only be measured if they are linked to meaningful key events such as biochemical recurrence, metastasis free survival or death (See section 1.2.2.1). Commercial biomarkers panels can be used for risk stratification and as decision making tools in the clinic (224). Many of these require tissue from either a prostate biopsy or radical prostatectomy. For example, the Oncotype DX® Assay, was developed by Genomic Health Inc., as a multi-gene validated assay that studies the patient's tumour unique genetic makeup to help guide treatment decisions in breast, colon and prostate cancer (225–228). The Oncotype DX® Prostate Cancer Assay was optimised on prostate biopsy tissue, and aims to estimate the probability of PGG3 (Gleason 4+3) or >pT3 disease (184,229,230). The test has been found to correlate to adverse pathologic features, metastases and time to biochemical recurrence (230).

In contrast, a liquid biopsy (blood or urine) could avoid unnecessary sampling and stratify patients to ensure biopsies are only performed if high suspicion of significant cancer exists. To create a novel fluid biomarker panel, companion to mpMRI that helps decide which patients to biopsy, I hypothesised that tissue-based biomarkers could be detectable in blood and urine.

The discovery of new biomarker panels is limited by the inherent challenges of prostate cancer diagnostics including sampling error, heterogeneity, and a lack of true normal/benign samples due to the field effect (231,232). Since many available biomarkers rely on prostate tissue, I aimed to develop a new biomarker panel that can be detected in blood or urine, avoiding unnecessary biopsies to men at low risk of prostate cancer, thus preventing the risks associated with prostate biopsies.

In 2013, Sharma *et al.* analysed AR DNA binding sites in prostate cancer tissue and identified a 16 gene panel that differentiates advanced disease (castrate-resistant). The panel consisted of genes with overexpression in castrate resistant prostate cancer and downregulation in castration xenografts. This signature was identified as possible targets of progression and response to therapy in castrate-resistant prostate cancer (233). An additional 12 genes were added to this list from a panel showing promise in the Whitaker lab (234,235,244,236–243).

It was hypothesised that the resulting 28-gene panel could act as a biomarker panel to aid in the diagnosis and risk stratification of prostate cancer in men. The primary aim of the signature was to correctly identify men that would benefit from a prostate biopsy, therefore avoiding unnecessary biopsies. Such panel would help differentiate clinically significant cancer that requires sampling from prostate cancer that poses no risk to the patient. This improved clinical pathway could reduce the number of unnecessary biopsies and streamline the prostate cancer diagnostic pathway.

Before the use of patient samples, the baseline expression of the gene panel in different cell lines was assessed. In addition, the degree of androgen regulation of selected genes was evaluated, to provide further biological information on the 28 genes and act as a positive control for future experiments. A cell line model was selected as it allows controlled experimental testing in different prostate cancer stages (from androgen-sensitive to insensitive). The four cell lines used for this work include LNCaP, C42, C42B and PC3.

The LNCaP cell line was obtained from a lymph node needle aspiration of a man with prostate cancer. LNCaP and MS fibroblast cells were co-injected to immunocompromised mice, followed by harvesting of cells four weeks post castration to produce C4 cells. C4 cells were then co-injected with MS fibroblast cells to immunosuppressed-castrated mice and harvested at 12 weeks to produce the C4-2 cell line. The resulting C4-2 cells were inoculated onto castrated mice; the resulting bone deposits were then harvested and formed the C4-2B cell line. All of these retain their androgen sensitivity; in contrast, the PC3 cell line lacks the androgen receptor and is a model of androgen resistant disease (245,246).

Unlike cell lines, RNA extracted from urine is not routinely processed to achieve preservation, since the urine is an acidic and hostile environment for RNA (247). Working under the hypothesis that RNA from prostate cancer cells would undergo degradation due to exposure to urine, a model using cancer cells exposed to urine was created. The effect of urine on potential fluid biomarkers was evaluated; here I demonstrated how the exposure to urine affects the measurement of relative gene expression; and how this can be overcome by amplifying the extracted RNA before PCR.

Selection of the final signature based on cell line data can be flawed; due to cross-contamination, genetic and epigenetic changes secondary to prolonged culture and the inherent lack of a microenvironment (248). For this reason, taking advantage of the rise of bioinformatics tools to study publicly available prostate cancer datasets, the primary gene panel (28 genes) was further studied using data from a selection of studies with benign samples to act as a control. Through collaboration with Avi Rosenfeld (Jerusalem Institute of Technology, Israel), the primary gene panel was reduced to a refined diagnostic gene panel and a refined prognostic gene panel. A roadmap of this chapter can be seen on Figure 3-1.

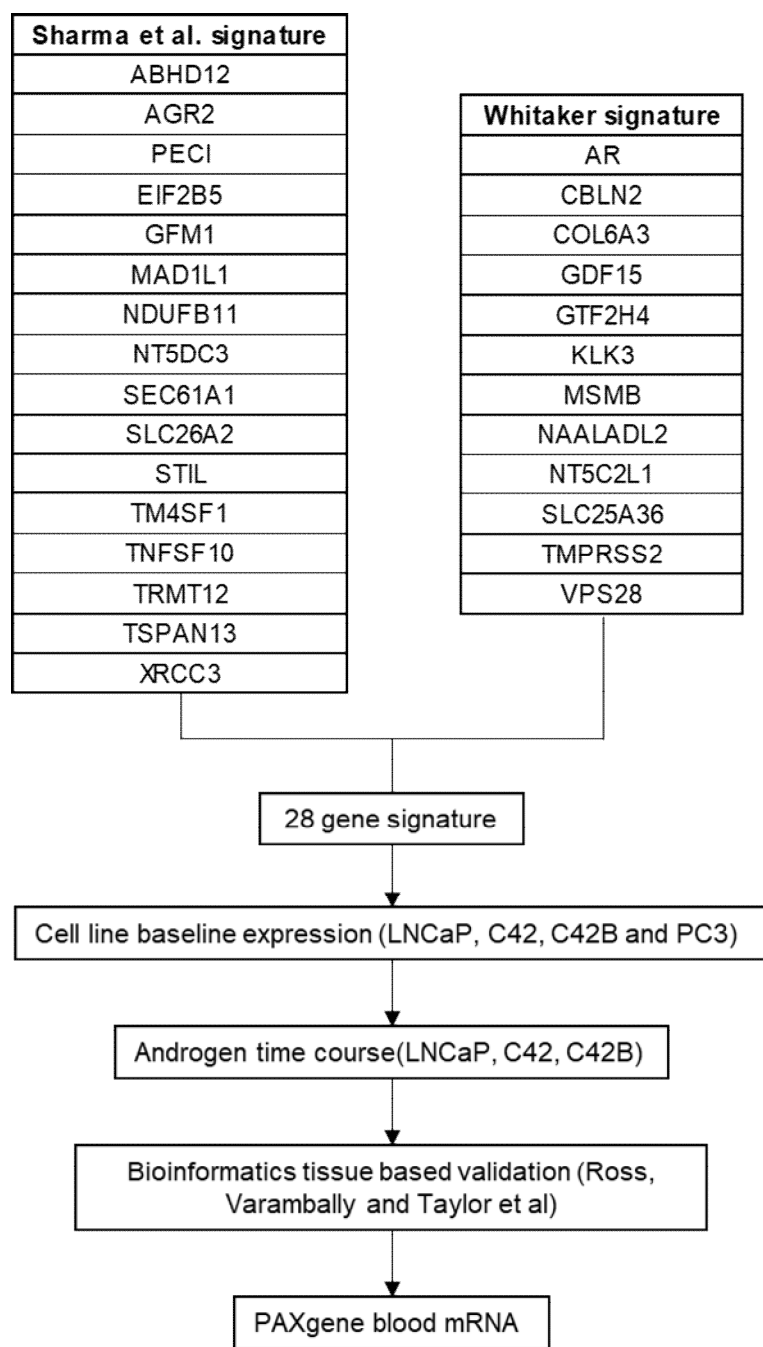


Figure 3-1 Fluidic RNA gene panel selection roadmap.

Initial gene panel is formed by 28 genes, of which 16 are part of a tissue-based signature (Sharma *et al.*), and 12 were added by the Whitaker lab based on previous and current work. Baseline cell expression was measured in 4 prostate cancer cell lines, followed by androgen time course for 72 hours. Data mining of three available datasets was used to obtain a final signature to measure in PAXgene blood

3.1.1. The gene panel

As mentioned in the section 3.1, the Sharma *et al.* panel was included in addition to 12 genes being developed at the Whitaker lab (Figure 3-1). A literature review was performed for all markers and key papers discussed below; the 28 genes were also analysed using Categorizer to identify enriched biological processes in the panel. Categorizer is a publicly available tool that categorises genes into predetermined categories and calculates p-values for the enrichment of the categories. It classifies the best-fit category for each gene by using semantic similarity measure of GO terms (249) (Table 3.1).

| Function | Gene | Panel | Evidence |
|-------------------------|----------|-------|--|
| Apoptosis | TNFSF10 | S | Selectively induces death in cancer cells Has a tumour supportive immuno modulatory role and cell migration (250–254), |
| Cytoskeleton | MAD1L1 | S | Mutation impairs the mitotic checkpoint, contribution in cell immortalization and chromosomal instability (255,256). |
| | STIL | S | Increased expression correlated with increasing grade and metastatic potential, knockdown induces apoptosis of cells. (257–259). |
| | COL6A3 | W | Decreased expression with increasing tumour stage in prostate cancer. Long isoform present in ~50% of metastatic prostate cancer (260–263) |
| DNA repair | XRCC3 | S | SNP at exon 7 increase risk of breast and prostate cancer. Correlation to high Gleason score and metastasis (264). |
| | GTF2H4 | W | SNP potentially associated with prostate cancer risk. Linked to MYC-SL transcriptional machinery with potential druggable target (238,265). |
| Golgi ER | SEC61A1 | S | No other published data (233). |
| | GDF15 | W | Knockdown may improve sensitivity to docetaxel. SNPs associated with GDF-15/MIC-1 linked to aggressive prostate cancer. Overexpression initial control of local disease but promotion of metastasis in advanced disease (237,266,267). |
| Metabolism | NT5DC1 | W | No specific data on prostate cancer. Function unknown. |
| | NT5DC3 | S | Overexpression in sub-group of bone metastatic tumours giving better prognosis and response to therapy (233). |
| | ECI2 | S | Increased expression predicts poor outcome. Reduced glucose utilization and fatty acid accumulation when knocked-down (268). |
| Mitochondria | NDUFB11 | S | Overexpressed in tumours with low relapse free survival (269). |
| | SLC25A36 | W | Up-regulation in metastatic prostate cancer compared to localised (242,270). |
| Receptors | AGR2 | S | Overexpression in high-grade prostate cancer. May have a protective role in localised cancer, enhances invasion in metastasis (271–274). |
| | AR | W | Several reported mutations predispose to prostate cancer, this includes N-terminal domain repetitions, activation of androgen dependent transcription factors and structural changes of AR. Its role in prostate cancer is well known, and has been used as a target to treat prostate cancer (234,275–279). |
| | ABHD12 | S | Copy number aberration in colorectal cancer. (Yoshida) Acts as a brake of immune system, may have a role in enhancing T-cell response in cancer and infections (280). |
| RNA processing | TRMT12 | S | Overexpression in breast cancer, could affect wybutosine (yW) biochemical pathway resulting in hypo-modified tRNA (281). |
| Signalling | NAALADL2 | W | NAALADL2 was shown to predict poor survival following radical prostatectomy (240). |
| Translation | EIF2B5 | S | Head and neck cancers increased survival of cells under hypoxic conditions. Upregulated in liver cancer (282). |
| | GFM1 | S | Upregulation in mamosphere model of cancer. Potential tumour suppressor role in head and neck cancers (283,284). |
| Transmembrane transport | TSPAN13 | S | Inversely correlated with Gleason score and presenting PSA, its expression is correlated with favourable outcome. Potential anticancer target in breast cancer. When knockdown less migratory potential in osteosarcoma model (285–287). |
| Vesicles | KLK3 | W | Extensive evidence of its use in the diagnosis of prostate cancer. However its raised in benign conditions and its not reliable in stratification of risk (239,288,289). |
| | SLC26A2 | S | Mediator of TRAIL resistance in some cancers, significant decrease in relapse-free survival (290). |
| | VPS28 | W | Low or moderate VPS28 expression lower risk of recurrence compared to high expression (244,291). |
| | TMPRSS2 | W | Associated with more aggressive prostate cancer, it is believed that its fusion (TMPRSS2:ERG) mediate invasion (243,292–295). |
| No location | CBLN2 | W | Reported frequency in prostate cancer: 20-25% Primary cancer, 33% Advanced cancer , 50% DCT or CTC (235,296). |
| | MSMB | W | Downregulation in prostate cancer, levels in serum may be useful biomarker in prognosis and diagnosis in tissue. Levels decrease when CRPC is present (241,297,298). |
| | TM4SF1 | S | Role in regulating tumour cell invasion and migration in ovarian, pancreatic and prostate cancer (299,300). |

Table 3.1. 28 gene panel GO term category classification by Categorizer.

28 genes organised by Categoriser GO term category, where more than one category was found the highest enrichment was selected. CBLN2, MSMB and TM4SF1 had no category. Panel column is either S (Sharma) or W (Whitaker) signature. Evidence column contains literature based evidence for each gene.

3.1.2. Sharma *et al.* aggressive prostate cancer panel.

ABHD12 and TNFSF10 have an immunomodulatory role in cancer, while ABHD12 may have a role in enhancing anti-tumour T-cell response, TNFSF10 has a tumour supportive role (254,280). Four genes have shown to have a potentially protective role in prostate cancer: AGR2, NDUFB11 and TSPAN13 (269,274), also AGR2 has been found to enhance invasion in metastasis (301). STIL, TM4SF1 and XRCC3 have a role in cell migration and metastasis (258,264,299). SLC26A2 and PECI overexpression are correlated with poor outcome and MAD1L1 and TRMT12 with cell immortalisation (255,268,281,290). There was no literature available on SEC61A1 relating to cancer. EIF2B5 and GFM1 are expressed in head and neck cancers, increasing survival of cancer cells in hypoxic conditions and as a tumour suppressor respectively (283,302).

3.1.3. Whitaker lab gene panel.

The 12 extra genes were carefully selected based on published and unpublished data obtained in house. AR, TMPRSS2 and MSMB were included due to the extensive evidence of their roles in prostate cancer (234,241,243,275,276,293,297). CBLN2 is downregulated in castrate resistant prostate cancer; in contrast, SLC25A36 is upregulated (270). COL6A3 expression is inversely correlated with tumour stage, expression of NAALADL2 and VPS28 has been shown to predict poor survival post radical prostatectomy (240,261,291). Finally, GTF2H4 has an SNP that was associated with prostate cancer risk (265). NT5DC1's function is unknown and was included as it is part of NT5DC family.

3.2.Objectives and hypothesis

3.2.1. Hypothesis 1

The expression of the 28 gene panel varies depending on the cell line is measured and degree of androgen independence.

3.2.1.1. Aim 1: Characterise the baseline expression of the primary panel in prostate cancer cell lines.

Objective 1: Determine the baseline gene expression of the gene panel using qPCR in four commonly used prostate cancer cell lines that mimic the states of androgen responsive to androgen independent states as a model for the progression of prostate cancer.

3.2.1.2. Aim 2: Study the varying degree of androgen regulation in the primary gene panel in androgen sensitive prostate cancer cell lines.

Objective 2: Determine the androgen dependence or independence of the primary gene panel in androgen stimulated prostate cancer cell lines.

3.2.2. Hypothesis 2

Cell line RNA is degraded following exposure to urine, decreasing its downstream utility as a fluid biomarker source.

3.2.2.1. Aim 3: Characterise the RNA degradation following exposure to urine.

Objective 3: Determine the effect in RNA yield, quality and gene expression in androgen stimulated cell lines following exposure to urine.

3.2.3. Hypothesis 3

Analysis of the expression of the primary gene signature in previously published datasets reveals a novel gene panel for the diagnosis and risk stratification of prostate cancer.

3.2.3.1. Aim 4: To investigate the expression of the primary gene panel in publicly available datasets using data mining.

Objective 4: Study the gene expression in previously published datasets and determine the ability of a refined panel of selected genes to diagnose and risk-stratify prostate cancer using machine-learning algorithms (Collaboration with Avi Rosenfeld, Jerusalem Institute of Technology, Israel).

3.3. Results

3.3.1. Cell line model

3.3.1.1. Cell lines and their characteristics

Aggressive prostate cancer follows a progression from androgen sensitivity to androgen insensitivity and a tendency to metastasise to bone. The difference in baseline expression of four well-known cancer cell lines was measured to model different stages of the disease. Individual cell line characteristics were discussed in section 3.3.1.1.

3.3.1.2. Baseline expression

Cells were grown and harvested as described in materials and methods. RNA was extracted, followed by qPCR of the primary gene panel (see Figure 3-2). Data was analysed using the $\Delta\Delta C_t$ method as detailed in section 2.2.6.1 in the materials and methods.

There was expression of all 28 genes in our primary gene panel in all the cell lines tested, with the exception of COL6A3 (Figure 3-2). Fold change between cell lines is shown for all the genes alongside example plots of a gene where expression was higher in androgen sensitive cell lines (TRMPSS2), a gene where expression was higher in androgen insensitive cell lines (TM4SF1) and a gene where expression levels did not change in the difference cell lines (XRCC3).

The genes were grouped into three categories, group 1; higher expression in androgen-regulated cell lines (LNCaP/C4-2/C4-2B > PC3), group 2: no difference in expression between the cell lines and group 3: higher expression in the androgen resistant cell line (PC3 > LNCaP/C42/C42B) (Figure 3-3).

AR, KLK3, NAALAD2 and TMPRSS are all genes previously known to be androgen-regulated, higher expression in the androgen regulated cells (group 1) is not surprising (240,277,288). Most of the genes in the signature showed no difference in expression between the four cell lines (Group 2). Six genes had higher expression in the PC3 cells/castrate resistant model (Group 3). Genes and the groups are represented in Figure 3-3.

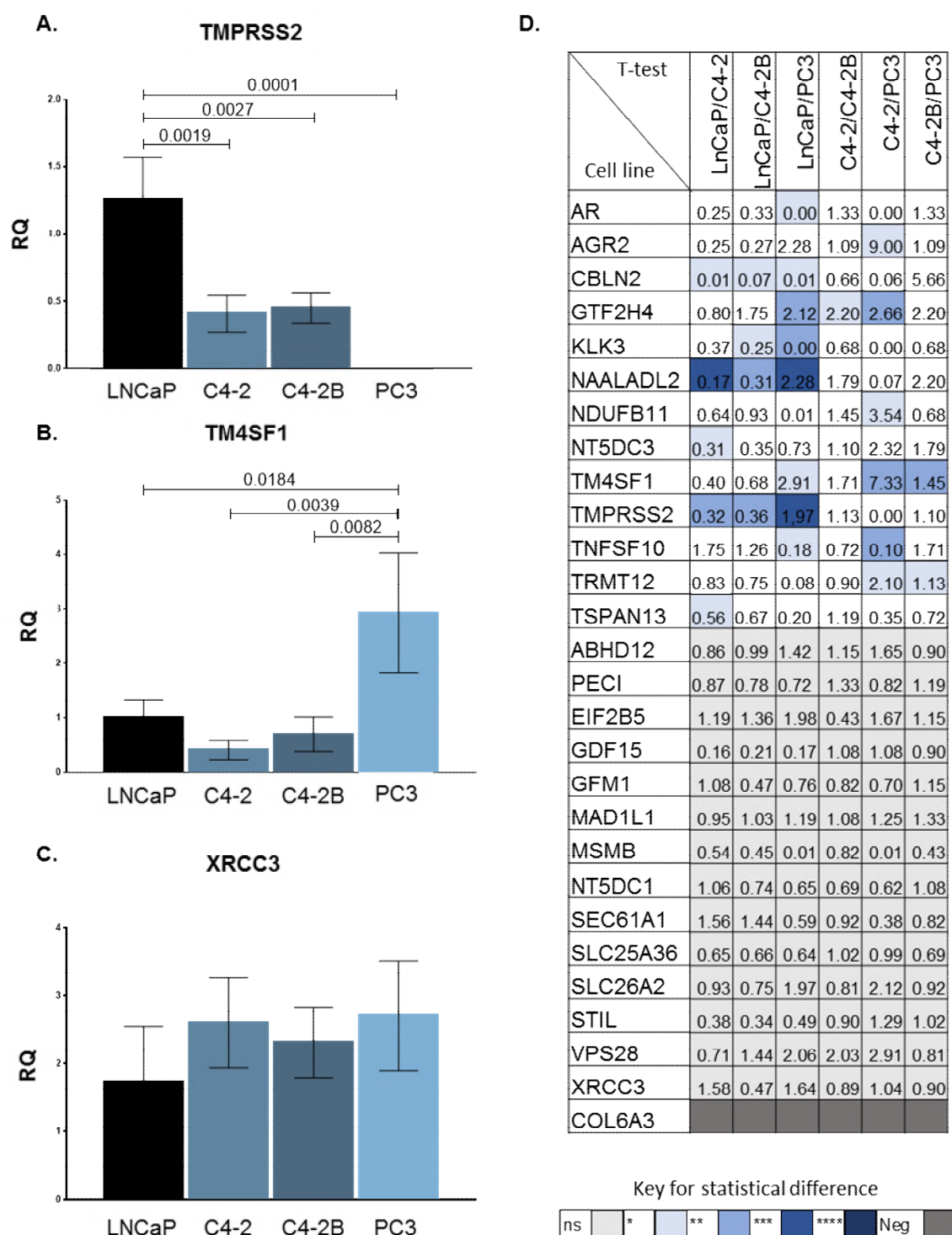


Figure 3-2 Taqman qPCR baseline expression in 4 commonly used prostate cancer cell lines, LNCaP, C4-2, C4-2B and PC3.

mRNA expression of TMPRSS2, TM4SF1 and XRCC3 in 4 prostate cancer cell lines, each cell line has n=3 replicates, and the error line represents the standard deviation. **A.** TMPRSS2 expression in LNCaPs is higher compared to C4-2, C4-2B and PC3 (One-way anova:0.0002). **B.** TM4SF1 shows the opposite trend (One-way anova:0.0035). **C.** XRCC3 has no significant difference in expression. **D.** Heatmap representing the expression of the 28 genes, the top 13 genes have significant difference represented by the shades of blue. The following 14 genes show no difference in expression and COL6A3 was not detected in any cell line. Numbers represent fold change in reference to LNCaP (first three columns), C4-2 (columns 4 and 5) and C4-2B (column 6) * p < 0.05, ** < 0.01, *** p < 0.001, **** p < 0.0001

| Group 1 | Group 2 | Group 3 |
|----------|-----------|---------|
| AR | ABHD12 | AGR2 |
| CBLN2 | PECI | GTF2H4 |
| KLK3 | EIF2B5 | NDUFB11 |
| NAALADL2 | GDF15 | SLC26A2 |
| NT5DC3 | GFM1 | TM4SF1 |
| TMPRSS2 | MAD1L1 | TRMT12 |
| TNFSF10 | MSMB* | |
| TSPAN13 | NT5DC1 | |
| | SEC61A1* | |
| | SLC25A36 | |
| | SLC26A2** | |
| | STIL | |
| | VPS28 | |
| | XRCC3 | |

Figure 3-3 Difference of gene expression in 4 prostate cancer cell classifies genes into three groups.

Group 1: Where gene expression was higher in cell lines that express androgen sensitivity. **Group 2:** No difference in expression between the cell lines, genes with an asterisk did not reach statistical significance but had higher expression in androgen sensitive cell lines* or insensitive cell line**. **Group 3:** Higher expression of these genes was seen in the PC3 cell line.

3.3.1.3. Androgen regulation

To further understand the relationship of androgen status and gene expression of the gene panel, LNCaP, C4-2 and C4-2B cells were stimulated with a synthetic androgen for up to 48 hours as described in materials and methods 2.2.1.6. Androgen exposure time courses were performed in a selection of genes from Group 1 (excluding NT5DC3), Group 2 (excluding SEC61A1, VPS28 and XRCC3) and Group 3 (excluding GTF2H4); COL6A3 was not detected on baseline measurement and was not measured in the time courses. The time course was not performed on the aforementioned genes, as they were not included in either refined gene panel. The resulting gene expression profile allowed for the identification of a positive control for future qPCR experiments in patients.

Figure 3-4 represents data from three genes to show an example of a gene with downregulation (AR), no change (GDF15) and upregulation (TMPRSS2). Figure 3-4 also includes published where Massie *et al.* performed an androgen regulation time course, extracting RNA every 30 minutes for four hours and then every hour for 24 hours, the gene expression was measured using Illumina BeadArrays (37).

A change in gene expression was seen for most genes; fold change in gene expression after androgen stimulation for all genes is shown in Figure 3-5. Genes were organised by upregulation (red), downregulation (blue) and no change (grey). Upregulation was defined as sustained fold change of > 2 , with an upward trend for ≥ 4 hours compared to the control. Downregulation was defined as sustained fold change > 2 with a downward trend from the baseline for ≥ 4 hours.

There was upregulation of six genes (most evident in TMPRSS2 and KLK3). CBLN2 showed initial downregulation followed by upregulation at 48 hours in LNCaP, whereas there was marked and sustained downregulation in C42B cells. This is possibly explained by a change in gene function in a more aggressive prostate cancer cell line in line with the baseline expression profile (Figure 3-3); where the expression of this gene in PC3 cells is extremely low if non-existent. AR was downregulated following androgen stimulation as previously demonstrated (279). There was no significant change in two genes (TSPAN13 and NDUFB11). The remaining genes are plotted in Appendix 1.

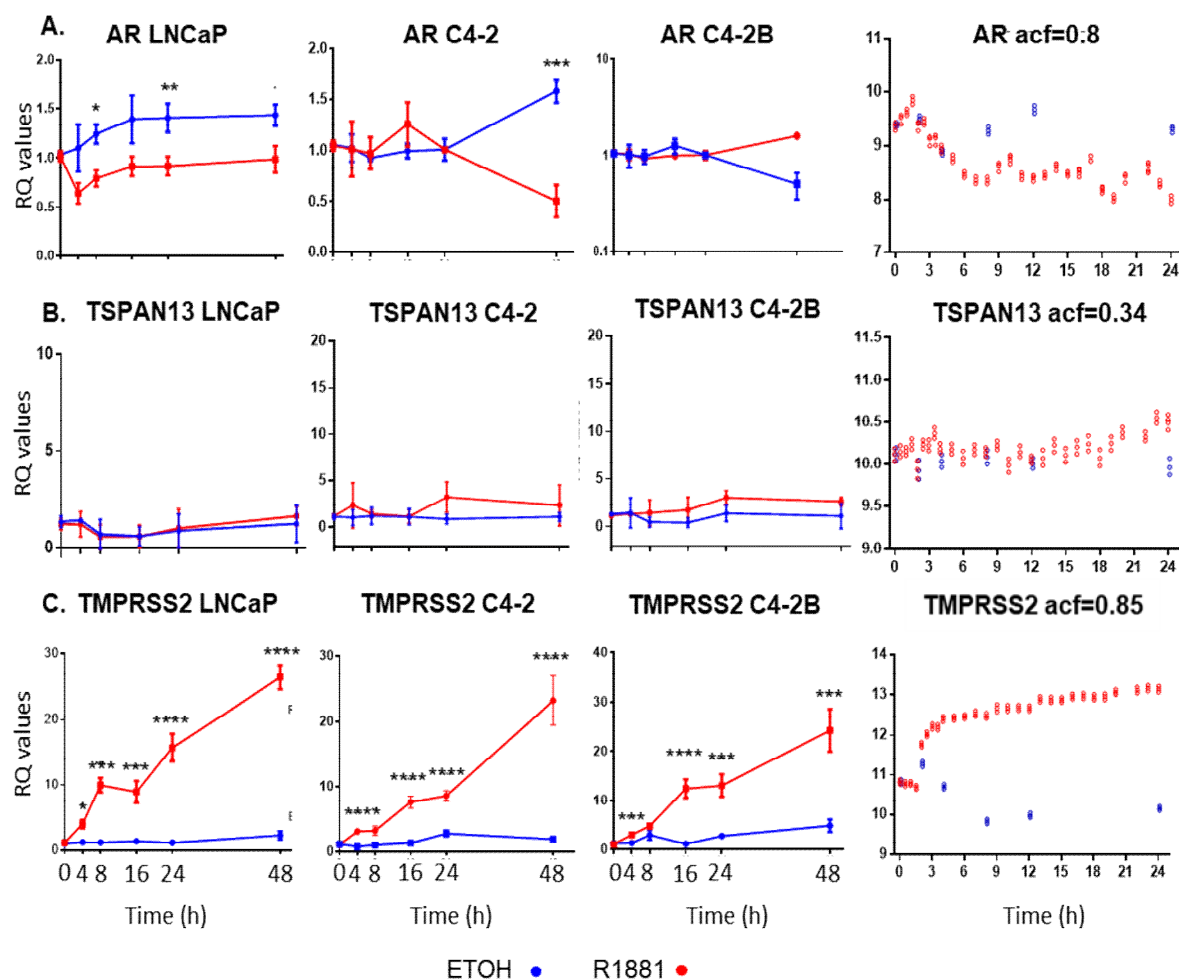


Figure 3-4 Taqman qPCR gene expression of three genes from the panel in androgen stimulated prostate cancer cell lines (LNCaP, C4-2 and C4-2B).

Cells were stimulated with synthetic androgen R1881 (red) and ethanol as a control (blue), RNA was extracted at different time points (x-axis). The RQ values ($2^{-\Delta\Delta C_t}$) are represented on the y-axis. The last column shows previously published sequencing data in LNCaP cell after stimulation with R1881 (Massie *et al.*) A. mRNA expression of AR, B. TSPAN13 and C. TMPRSS2 in LNCaP, C4-2 and C4-2B cell lines. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

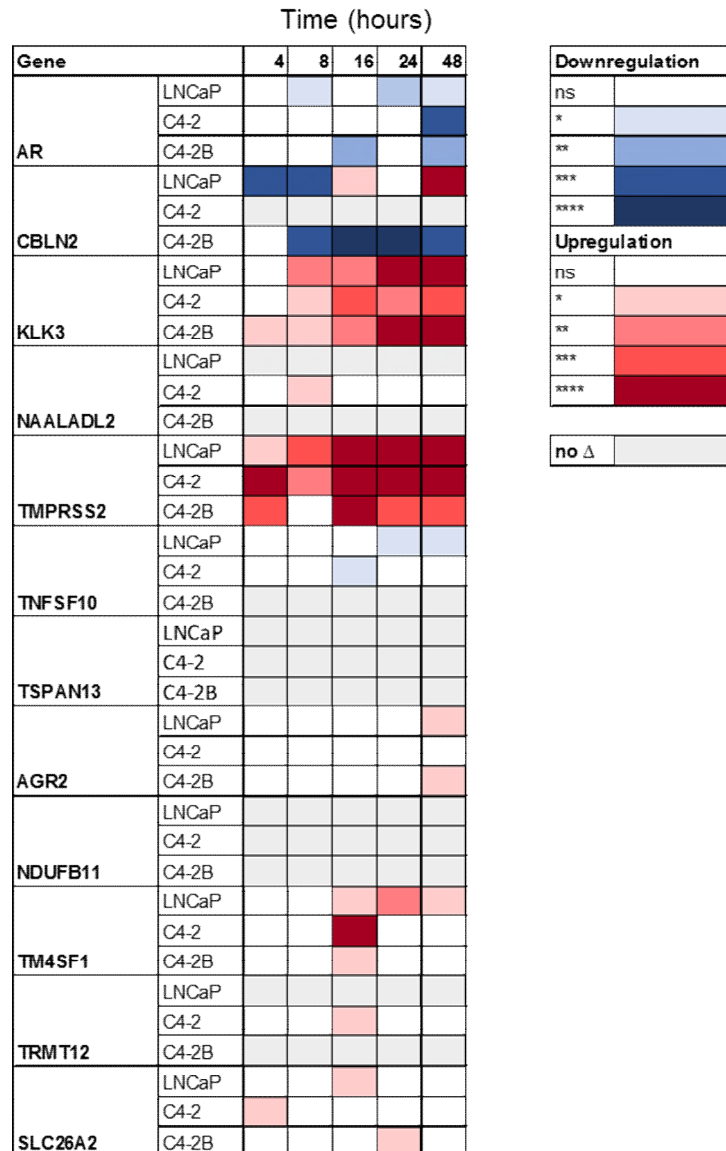


Figure 3-5 Gene expression in LNCaPs, C4-2 and C4-2B cell lines following androgen stimulation.

Heatmap representing gene expression ranked by upregulation, downregulation and no change. Gene on left column, followed by time in hours post stimulation (4, 8, 16, 24 and 48). Grey represents no androgen regulation after R1881 stimulation (no Δ). Red gradient represents statistical significance if there is upregulation (KLK3, TMPRSS2, and MSMB) and blue if there is downregulation (AR, CBLN2). If there was no statistically significant difference, the areas are in white (GDF15 in LNCaP) or in grey if not change was seen (EIF2B5).

3.3.2. Gene expression in degraded samples

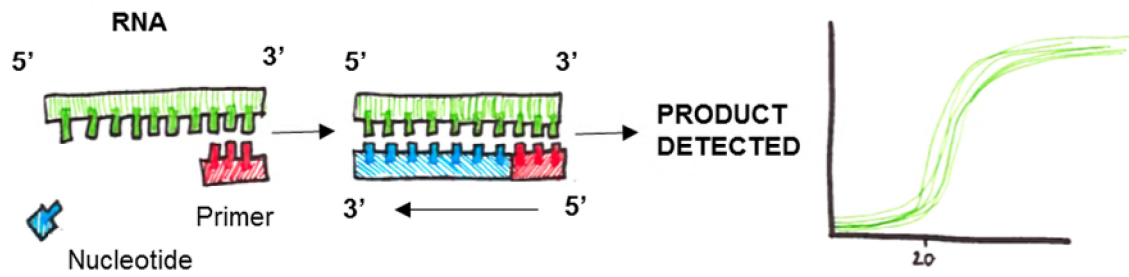
3.3.2.1. Simulation of degraded RNA

Experiments were designed to explore the effect of urine on LNCaP cells in order to investigate the suitability and stability of urinary RNA for a biomarker panel. As discussed in the section 3.1, gene expression quantification relies on un-degraded RNA for appropriate detection of the gene of interest, urine exposure and its degradation of RNA affects the PCR target amplification. As a result, PCR primers may be unable to attach to the fragmented RNA; See Figure 3-6 for a schematic representation of this effect. To quantify the impact of RNA degradation in urine and the impact on the genes of interest, an experiment was designed to expose androgen stimulated LNCaPs to urine (see materials and methods 2.2.1.7).

To determine the integrity of the RNA, quality and control was performed using the Bioanalyzer. The Bioanalyzer is a commonly used method to assess RNA degradation by providing a calculated RNA integrity number (RIN) (303). This tool uses the entire electrophoretic trace of the sample. Intact RNA will have a RIN above seven, as it is usually the case with cell line extracted RNA. A RIN of three to seven is acceptable, while a degraded sample will have a RIN of equal to or less than three. RNA extracted from LNCaP cells had RIN numbers between 8.9 and 9.3 (n=3), whereas urine exposed LNCaPs had RINs of 2.3 of 2.6 consistent with RNA degradation.

The urine exposed RNA was assessed by qPCR. Ct values are a relative measure of the concentration of target in the PCR reaction, the lower the Ct value the higher the relative concentration of the target. To assess the relative concentration, values of the genes of interest are compared to that of a housekeeper gene (a gene whose expression remains unchanged despite change in cell conditions). When cells were exposed to urine, the Ct values increased in all genes examined, there was an increase in average of 2.91 cycles when cells were exposed to PBS (range 1.98-5.53) compared to 4.72 (range 3.27-6.10) when exposed to urine (Figure 3-7). An increase in Ct values indicates a lower relative concentration of the target. In three biological replicates a gene of interest, (NAALADL2) and a housekeeper (RPLP2) were plotted. LNCaPs not exposed to urine or PBS, had a high relative concentration of RPLP2, suggesting that when exposed to urine, is due to RNA degradation increases, resulting in lower product detection. This would suggest that any RNA from prostate cancer cells that finds its way into the urine would likely be degraded; alternative post-collection measures to improve the detection of targets were investigated.

A.



B.

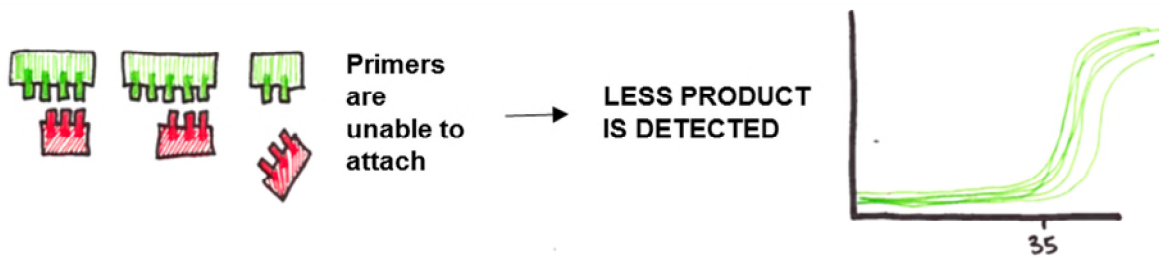


Figure 3-6 Degraded RNA will affect primer effectivity in qPCR reactions.

RNA integrity influences the PCR reaction. **A.** RNA represented in green as a continuous strand from 5' to 3', with primer attachment and consequent cDNA reverse transcription and subsequent gene of interest detection with qPCR. **B.** When RNA is degraded, the primers cannot attach correctly with incomplete formation of the cDNA and ensuing lower product detection.

A.

| | Ct values in different conditions (average values, n=3) | | | Difference in Ct values between cells and exposed cells | |
|----------|--|--------------|----------------|---|------------------------|
| Gene | LNCaP | LNCaP in PBS | LNCaP in Urine | LNCaP in PBS - LNCaP | LNCaP in Urine - LNCaP |
| AR | 26.02 | 29.06 | 30.60 | 3.04 | 4.58 |
| MSMB | 30.79 | 33.25 | 35.21 | 2.46 | 4.43 |
| PSA | 22.84 | 25.63 | 27.35 | 2.79 | 4.51 |
| GDF15 | 25.14 | 27.12 | 29.28 | 1.98 | 4.14 |
| NAALADL2 | 27.89 | 33.42 | 33.99 | 5.54 | 6.10 |
| PCA3 | 33.79 | 36.72 | 37.06 | 2.93 | 3.27 |
| CBLN2 | 31.54 | 33.76 | 37.09 | 2.22 | 5.55 |
| TMPRSS2 | 23.65 | 26.03 | 28.88 | 2.38 | 5.23 |
| RPLP2 | 23.74 | 26.53 | 28.83 | 2.79 | 5.09 |

B.

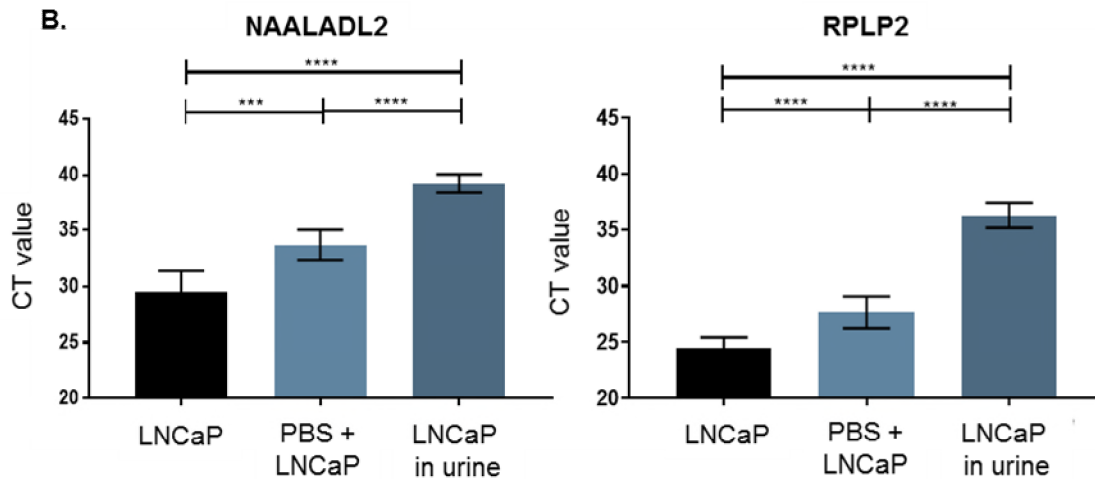


Figure 3-7 Exposure to urine reduces the detection rate of targets.

A. Averaged Ct values of nine genes following no exposure (LNCaP) and exposure to PBS and spun urine. Note the increase in Ct values following exposure to PBS and urine, more marked when exposed to urine. The effect is consistent in genes of interest and housekeepers (RPLP2). **B.** Ct values of NAALADL2 and RPLP2 (C) (housekeeper gene) in non-exposed cells (LNCaP), cells suspended in PBS and cells suspended in urine. Note the Ct value increases when cells are exposed to urine.

3.3.2.2. Strategies to overcome degraded RNA.

If RNA degradation in patient biological fluids cannot be avoided, several measures can be used to compensate during the qPCR at four levels: Increasing the amount of cDNA added to the reaction, increasing the number of PCR cycles/longer PCR protocol and sample pre-amplification.

3.3.2.2.1. Increasing cDNA concentration

The concentration of cDNA used for the PCR reaction can have an impact in the reverse transcription reaction. This is more noticeable in abundant genes, as the rate of reverse transcription inhibition is higher in targets that are more concentrated; in other words, a higher initial concentration of cDNA will decrease the resulting Ct value. In addition, the reaction will be inhibited if initial cDNA concentration is too high (304).

Since the housekeeper, genes used to normalise the reaction are the most likely affected, careful consideration of the optimal cDNA concentration was considered. The PCR reaction was done with a cDNA volume of 2 μ L and 6 μ L, two biological replicates and two technical replicates for one gene of interest (TMPRSS) and one housekeeper (RPLP2). Relative concentration (Ct) was lower at 6 μ L for the housekeeper while it improved for TMPRSS. This confirmed that the optimal cDNA volume to use in future reactions was 2 μ L to reduce the risk of RT inhibition in more abundant genes (Figure 3-8A).

3.3.2.2.2. Increasing number of cycles

The target cDNA is subjected to cycles of denaturation, annealing, and extension. The number of cycles can vary depending on the amount of cDNA and desired yield of product, and is usually up to 40 cycles. Over 45 cycles, there is nonspecific amplification and accumulation of by-products lowering PCR efficiency (304). cDNA from cell lines, was cycled up to 45 times to determine if low occurring genes would be amplified at later cycles in degraded RNA, there was no difference in detection of targets between 40 and 45 cycles. In addition, a ChiP PCR protocol was used, with more extended incubation periods and time inversion in the amplification step (15 seconds at 95°C and 60 seconds at 60°C), there was no difference in detection of the genes of interest or housekeeper (Figure 3-8 B).

3.3.2.2.3. Pre-amplification of degraded samples

Pre-amplification of degraded RNA can improve the detection of genes with low availability. However, bias can be introduced if the amplification is uneven. A whole transcriptome amplification (WTA) kit by was used to reliably amplify degraded FFPE samples without 3' bias, using quasi-random 3' primers with universal ends. Pre-amplification can introduce bias in the final analysis. For example, poor efficiency in amplification of certain targets can result in under/over representation of the gene of interest if the gene is under or over pre-amplified. Amplification of urine exposed cells showed appropriate amplification of genes tested (Figure 3-8 C).

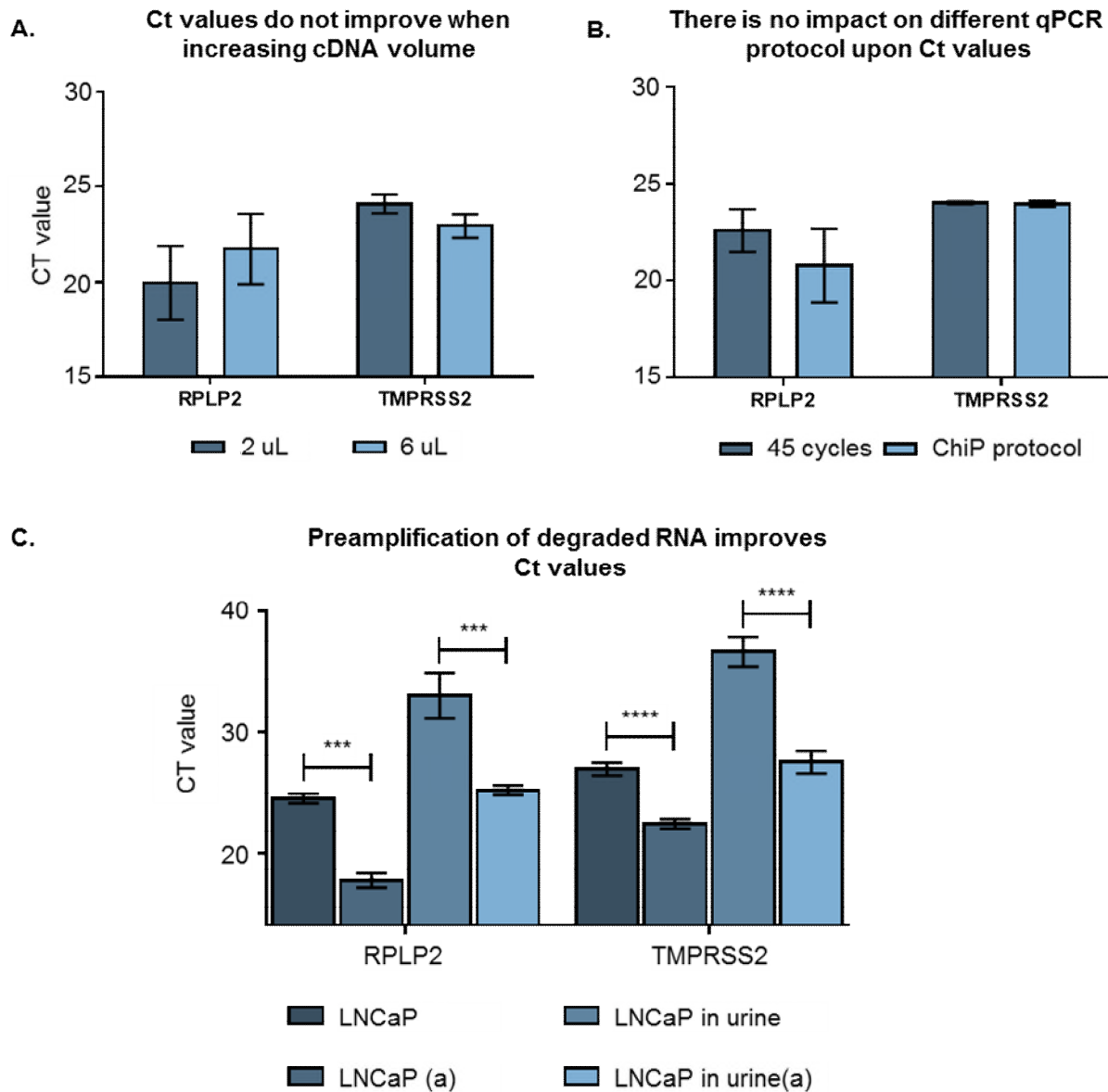


Figure 3-8 cDNA concentration and qPCR protocols do not affect Ct values, whereas Preamplification improves detection of target mRNA.

A. Comparison of Ct values of a housekeeping gene (RPLP2) and TMPRSS2, there is no significant difference between using 2 o 6 μ L of cDNA. **B.** Increasing the number of cycles or using a new PCR protocol did not improve the detection of the targets.. **C.** Preamplification of degraded samples improves the detection of target genes in an unbiased manner. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, if no value the difference was not significant (A and B).

3.3.3. Gene expression in human samples

3.3.3.1. Data mining and cohort introduction

Prior to testing in clinical samples, bio-statistical analysis of the 28 genes in the original panel was performed in three publicly available cohorts with benign sample controls: Taylor, Varambally and Ross-Adams (Table 3.2). Data was accessed using Gene Expression Omnibus (GEO) repository (see 2.2.6.4). In total, Ross had data for 27/28 genes, Taylor 18/28 and Varambally 26/28 (Table 3.2). A proportion of Taylor and Ross-Adams patients had matched benign samples that could be directly compared, and although Varambally has benign samples, the cohort was small and had limited information available. Only data from the Cambridge arm were included in the analysis for the Ross-Adams cohort. The gene expression in these cohorts was measured from tissue extracted from radical prostatectomy specimens (220,305,306).

| Study | Taylor et al 2010 | Varambally et al 2005 | Ross-Adams et al 2015 |
|---------------------------|--|-----------------------------------|--|
| Technology | Agilent 244K aCGH* microarray | Affymetrix U133 2.0 plus array | Illumina OMNI2.5M Genotype Beadchip |
| Localised | 131 | 5 | 112 |
| Metastatic | 19 | 4 | 13 |
| Benign | 29** | 4 | 74** |
| Age range (median) | 37.3-83 (58) | NA | 44-73 (61) |
| PSA range (median) | 1.15-506 (6.6) | NA | 4-23.7 (7.63) |
| Gleason 6 | 79 | NA | 9 |
| Gleason 7 | 50 | NA | 58 |
| Gleason ≥8 | 19 | NA | 6 |
| Tumour content | ≥70% | 90% | ≥20% |
| Missing gene data | PECI, GDF15, GTF2H4, MSMB, NT5DC3, TM4SF1, TNFSF10, TRMT12, VPS28, XRCC3 | AR, MSMB, | N/A |

Table 3.2 Data mining cohort summary and missing gene data.

Summary of clinical characteristics of the three cohorts. Missing gene data for Taylor cohort and Varambally. *aCGH array comparative genomic hybridization. ** Matched benign samples from cancerous prostates.

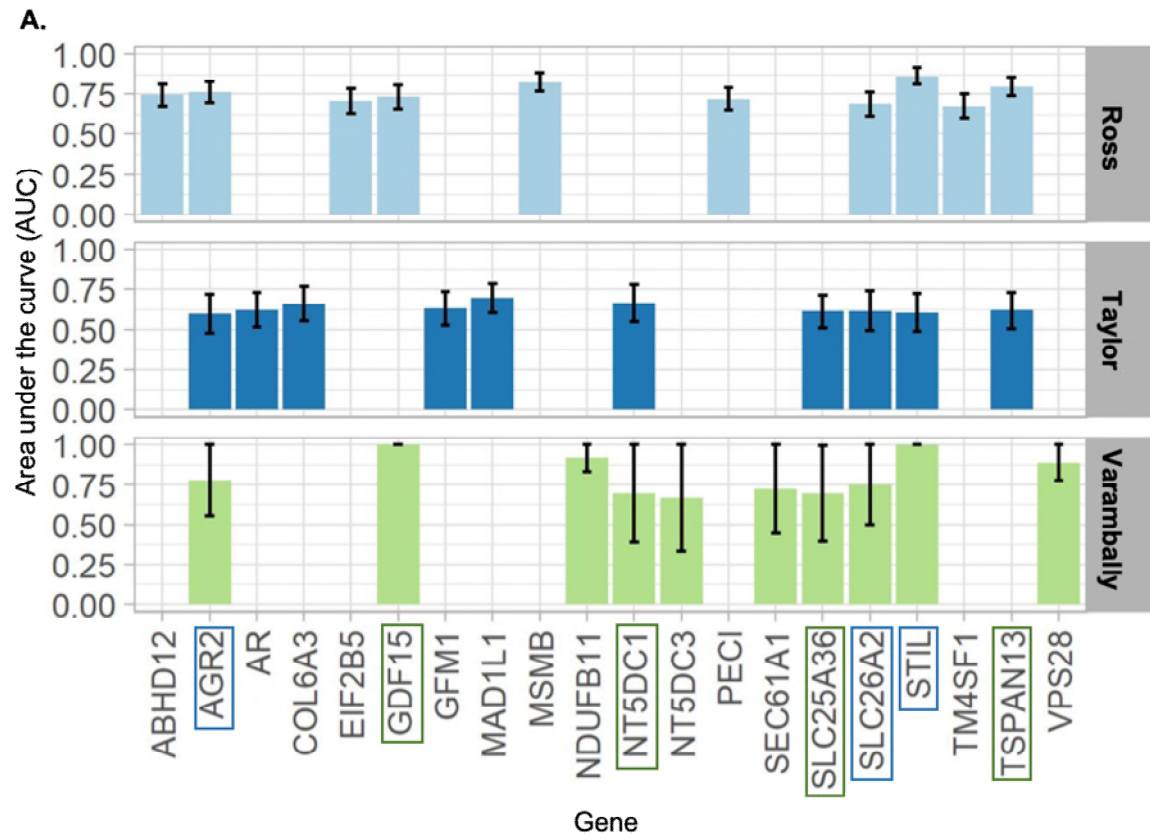
3.3.3.2. Gene expression of the panel in three separate cohorts

3.3.3.2.1. Differentiation between benign and cancer in three independent cohorts using a refined diagnostic gene panel.

In collaboration with Avi Rosenfeld- Roni Reznik (Israel Institute of Technology), the gene expression data from the three previously mentioned cohorts was analysed. The aim was to select genes that were present reliably in more than one dataset. To do this we used “stability” of a feature as a measurement of whether the likelihood that the set of markers in a signature are specific to the study, or applicable to other cohorts. i.e., to determine if an effect is real or biased due to patient selection.

Data was then analysed to generate areas under the curve (AUC) data to predict benign versus cancer and different prostate prognostic grades (PGG). The data was ranked by the AUC; using the intersection of the top ten per cohort was used to establish which markers were consistently found in the three datasets (Figure 3-9). In the diagnostic signature, three genes (STIL, SLC26A2 and AGR2) were in the top 10 of all cohorts with ROC values >0.6. GDF15, NT5DC1, TSPAN13 and SLC25A36 were in two of three cohorts, of these five markers only GDF15 was absent from one cohort (Taylor) (See Figure 3-9 A).

Relative gene expression data from the three datasets was plotted comparing benign vs cancer. Appendix 2 shows this data for the final seven chosen for the panel. There is concordance, defined as a similar trend of gene expression between benign and cancer in all three cohorts, only for STIL and TSPAN13. GDF15 is concordant between Varambally and Ross; AGR2 expression matches between Taylor and Ross, SLC26A2 and SLC25A36 match between Taylor and Varambally (see Appendix 2). This could be explained by the difference in technologies and probes used in each study, the proportion of tumour content and population variability. Concordance was not used as a rationale for removing any genes from the panel (307).



B.

| Position | Ross | Taylor | Varambally |
|----------|---------|----------|------------|
| 1 | STIL | MAD1L1 | STIL |
| 2 | MSMB | NT5DC1 | GDF15 |
| 3 | TSPAN13 | COL6A3 | NDUFB11 |
| 4 | AGR2 | GFM1 | VPS28 |
| 5 | ABHD12 | AR | AGR2 |
| 6 | GDF15 | TSPAN13 | SLC26A2 |
| 7 | PECI | SLC26A2 | SEC61A1 |
| 8 | EIF2B5 | SLC25A36 | SLC25A36 |
| 9 | SLC26A2 | STIL | NT5DC1 |
| 10 | TM4SF1 | AGR2 | NT5DC3 |

| | |
|--|---------------|
| | Three studies |
| | Two studies |

| Gene panel |
|------------|
| STIL |
| SLC26A2 |
| AGR2 |
| GDF15 |
| NT5DC1 |
| TSPAN13 |
| SLC25A36 |

Figure 3-9 Stability based selection of genes as markers for differentiation of benign and cancer reveals a seven gene signature in tissue (refined diagnostic gene panel).

A. Faceted barplot for each cohort, x-axis denotes genes where those in the top ten for three and two cohorts are highlighted with a blue rectangle, and green rectangle respectively. Y-axis contains the values for the AUC for each gene. **B.** Table representing the stable feature selection; each cohort was ranked and the top ten markers were selected. Markers for the refined diagnostic gene panel were carefully chosen if they were present in three (blue) or two out of three cohorts (green).

3.3.3.2.2. Differentiation between prognostic grade groups (PGG) using gene expression data in three cohorts.

The Gleason score expressed as the PGG, was grouped as PGG 1 and 2 (Gleason 3+3 and 3+4) vs PGG 3 (Gleason 4+3), 4 (Gleason 4+4) and 5 (any Gleason 5) to differentiate between lower and higher grade cancers ($PGG \leq 2$ and $PGG \geq 3$). Only two cohorts were included in this analysis (Ross and Taylor) as Varambally lacked PGG/Gleason score information. Three genes made up the final signature: NT5DC1, MAD1L1 and EIF2B5. These were present in both cohorts with AUC values > 0.57 (lowest for MAD1L1 in Taylor), see Figure 3-10 A-B. When plotting the gene expression data across the grade groups for each gene selected for the panel, only NT5DC1 had significant differences between groups (one way ANOVA) in both cohorts (Figure 3-10 C-D). The distribution of the relative gene expression of this gene has a different distribution, where it increases in the Taylor cohort but decreases in the Ross samples.

Gene expression data for all 28 genes across PGG can be found in Appendix 3 and Appendix 4.

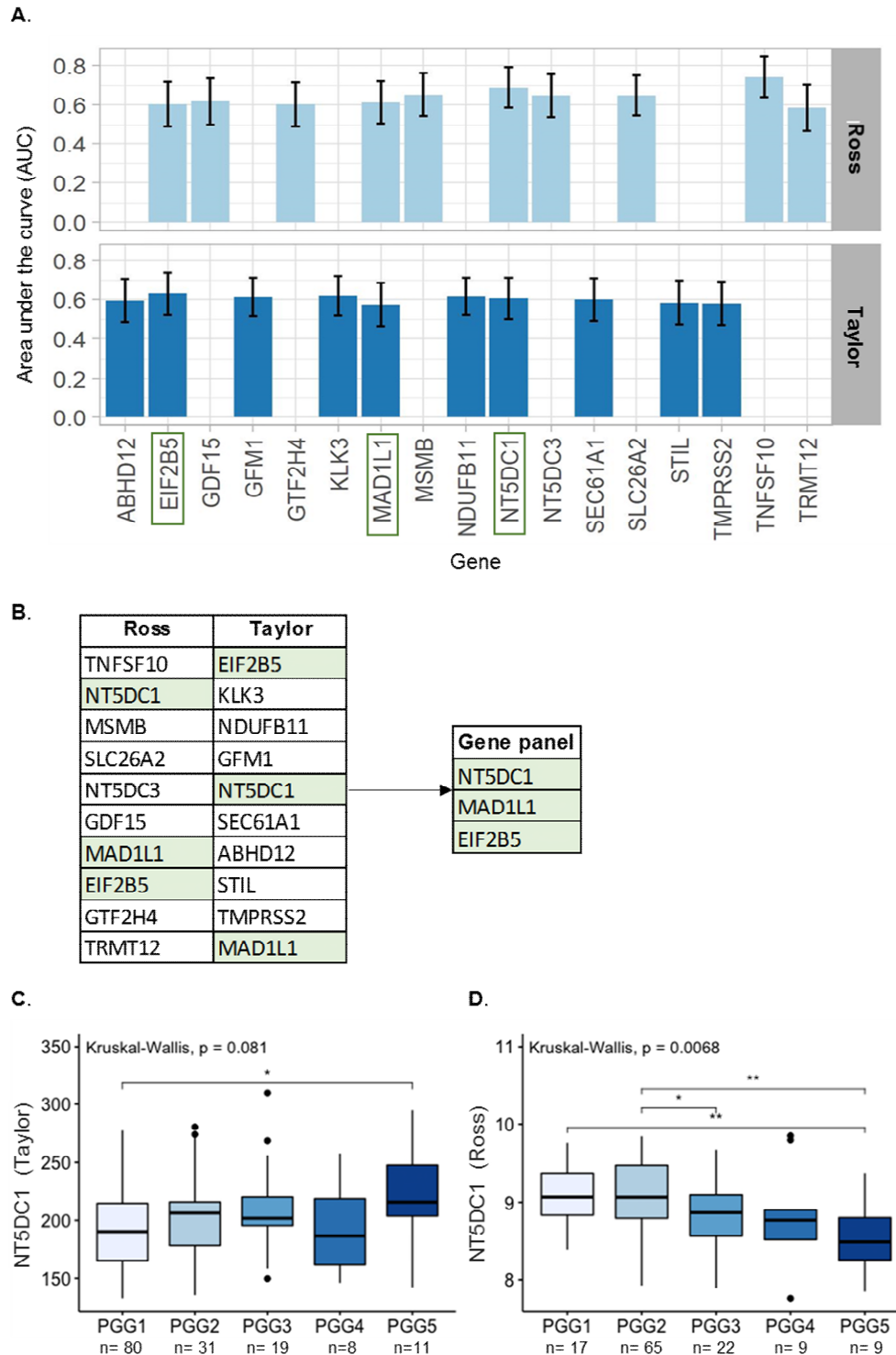


Figure 3-10. Stability based selection of genes as markers for differentiation between PGG groups, finds a three-gene signature in tissue (refined prognostic panel).

A. Faceted barplot for each cohort, the x-axis denotes genes that were ranked in the top ten for each cohort. Genes highlighted with a green rectangle if present in both cohorts. The y-axis contains the values for the AUC for each gene. **B.** Table representing the stable feature selection, each cohort was ranked, and the top ten markers were selected. Markers for the final gene panel were chosen if they were present in both cohorts. **C-D.** Boxplots of the relative gene expression of NT5DC1 in the Taylor (C) and Ross (D) cohorts. The x-axis contains the PGG groups from one to five, the y-axis the relative gene expression. Kruskal-Wallis value displayed if a significant difference on student t-test result is displayed. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

3.4.Discussion

3.4.1. Summary of findings

In this chapter, the relative gene expression of a tissue-based gene panel in four different cell lines was measured (from androgen dependent to independent); grouping the genes in three groups (See Figure 3-3). The gene expression of 22 genes following androgen stimulation also grouped the genes in three groups: upregulated, downregulated and not regulated (See Figure 3-5). Additionally, cell line RNA was significantly degraded following exposure to urine, adding a preamplification step improved the quality for downstream analysis (See Figure 3-6 and Figure 3-7, respectively). Finally, using previously published data, a novel biomarker panel was developed to aid detection of prostate cancer.

Relative baseline gene expression of all genes in the four cell lines studied, allowed grouping of the genes based on in-vitro models of prostate cancer progression. Genes in-group 3 (higher expression in PC3 to other cell lines) may represent a signature for more aggressive forms of the disease. Of these, AGR2 is associated with prostate cancer metastasis and aggressive prostate cancer, and TM4SF1 is known to contribute to cell motility and metastasis (274,299). While androgen regulation enhanced our understanding of the potential implication of these genes as possible biomarkers of disease.

3.4.1.1. The refined diagnostic gene panel.

The diagnostic refined gene panel was developed using publicly available data, consistent expression of the signature in three separate cohorts was demonstrated. There was concordant expression in at least two of the three cohorts. While concordance in all three studies would be ideal, the discrepancy between studies is expected, as there is heterogeneity in sample size, race, and methods of collection, patient selection and technologies used. While many approaches exist for scoring the relative effectiveness of markers, we chose the well-accepted AUC measure to determine the relative predictive ability of a given marker. We chose the top ten markers based on their AUC value for each of these datasets. Encouragingly, the seven genes had steadily high AUC values. Whilst a variety of ranking methods are available, we chose AUC as it consistently yielded a reliable range across all datasets. The Varambally cohort had the highest AUCs, due to the smaller cohort size. The proposed refined diagnostic signature has the potential to enhance the current clinical diagnosis pathway.

The refined diagnostic gene panel is comprised of androgen-regulated and non-androgen regulated genes. Here, these will be discussed according to the baseline group classification shown on Figure 3-3. TSPAN13 (group 1), has been shown to be overexpressed in prostate cancer and has a potential role in prostate cancer progression (285).

Five genes belonged to group 2, NT5DC1, STIL, SLC26A2, GDF15 and SLC25A36. The function of NT5DC1 is unknown. STIL, is a promotor of tumour growth that showed androgen downregulation in LNCaPs and upregulation in C42B cells, consistent with its ability to regulate pathways that influence invasion and apoptosis of prostate cancer cells (259). SLC26A2, this gene was upregulated when exposed to synthetic androgen, its exact role in prostate cancer is not entirely clear, but it has been

shown to mediate TRAIL resistance in breast cancer (290). GDF15, was also upregulated when exposed to R1881, it has been reported to have potential influence in metastatic disease (308) and as a possible biomarker for prostate cancer (309). SLC25A36 was not androgen-regulated, it is essential for the synthesis and breakdown of mitochondrial DNA and RNA, and has been shown to be upregulated in metastatic prostate cancer (270); this is one of two genes in this panel that lacks androgen regulation.

The last gene part of the refined diagnostic signature is AGR2 (group 3), this gene showed androgen upregulation, and it is known to correlate to aggressive prostate cancer (274).

3.4.1.2. The refined prognostic gene panel

Accurate risk stratification of men with prostate cancer without relying on tissue would provide patient benefit. The signature described in here is formed of three genes: NT5DC1, MAD1L1 and EIF2B5. Only NT5DC1 is present in the refined diagnostic and prognostic panel, as previously mentioned its function is unknown. Interestingly, its relative expression plots between Ross and Taylor show conflicting trends, where in Taylor, the relative expression is higher in PGG5 compared to PGG1. In Ross, PGG1 is higher than PGG5. This is the case for some of the significant genes in the Ross and Taylor cohorts (Appendix 3 and Appendix 4).

MAD1L1 is a mitotic spindle assembly checkpoint protein, it has been found to induce stem-like cell properties in nasopharyngeal carcinoma (256). Mutant MAD1L1 has been correlated to chromosomal instability and immortalization in lymphoma cell lines (255). In the Taylor cohort, there was a significant difference in expression between PGG1 and PGG3, with an upward trend (Appendix 4). However, the PGG5 expression had the lowest median value; this could be explained by the presence of the more aggressive mutant form. Interestingly MAD1L1 expression did not change upon androgen stimulation in cell lines (Appendix 1).

Finally, EIF2B5 is a translation initiation factor; it has been found to be upregulated in liver cancer and associated with the histologic and clinical stage (282). It has also been found in head and neck cancers cell, where cells adapt to hypoxia and increase their survival. This gene also had significant differences in the relative expression on the Taylor cohort. Its expression decreases with PGG grade.

3.4.2. Methodological limitations

The use of cell lines as a model for biomarker discovery offers several benefits including: availability, experiments can be scaled up rapidly, experimental settings can be rigorously controlled, therefore reducing potential confounding factors, and extensive available -omic data on public datasets. However, immortalised cell line models grown in 2D plastic plates and cannot be directly compared to a 3D cancer model in a living patient. Additionally, with increasing passage number, the risk of cross-contamination and novel mutations increases, altering the phenotype of the cells to the point it does not resemble the primary tissue (310,311). Additionally, with the exception of PC3, all cell lines used in this work share the same parental lineage. Therefore, although modelling of prostate cancer stages was attempted by using four different cell lines, the model is not a perfect simulation of clinical presentation. Alternative

prostate cancer cell lines such as DU-145 (no androgen receptor), LAPC-4 or VCaP (both androgen receptor and PSA positive), and “benign” cell lines such as BPH1 or RWPE-1 to compare expression could have been used (312).

The androgen-stimulated experiment required several 96-plate qPCR. Regardless of care taken pipetting variations are expected in each plate and between plates. The experiments were planned to ensure each gene was represented per plate, however digital PCR would have been preferable to perform such large scale analysis, reducing the inter-plate variability (313).

Regarding the selection of the genes for the signature, the 28 genes resulted from a combination of tissue based biomarkers and in-house developed genes. For some genes in the signature there is little or no literature available (Table 4.1). For instance, the function of NT5DC1 remains unknown. Other genes in the signature (SEC61A1 and GFM1) have no data in prostate cancer. Although this is not entirely a negative factor, further understanding of the role of these genes, and understanding their function prior to inclusion in the signature may have been beneficial to the project.

Based on the RNA degradation caused by urine as discussed in Section 3.3.2, the development of a urine-based biomarker should include steps to improve the initial quality of the RNA, rather than rely on steps that can bias the results, such as pre-amplification. For instance, using RNA stabilisers like RNAlater, have shown success (314), such steps should be taken prior to RNA extraction to ensure the highest yield and quality of the RNA.

Data mining using previously published cohorts allows the study of large amounts of data. In this case, all three cohorts had different amounts of available data. All cohorts are at least five years old, and since there have been advances in the pathological classification of prostate cancer, the findings described here may not reflect current clinical practice. Additionally, there is no data on mpMRI in these cohorts, since one of the hypothesis is that this panel could aid imaging biomarkers, this data would have been beneficial to improve the panel. Another limitation with the obtained data is that long-term outcomes were not taken into account when modelling the biomarker panel. As mentioned in the introduction of this chapter, novel biomarkers only have significance if linked to clinical outcome. This panel is meant to help decide which men benefit from a biopsy and which do not, by ignoring the fact that some will have no biopsy and may still harbour significant disease (false negatives) and correcting appropriately, the panel may not be as useful clinically as expected.

In addition, each study used different technologies for sequencing. This explains why some genes were not available in all three cohorts. This also means that older sequencing technologies may have an inherent bias on the sequences used or may not be as accurate as newer kits. Both Taylor and Ross included benign samples; however, these are matched samples from a cancerous prostate; comparison between “benign” and “cancer” tissue may not reflect the state of a genuinely benign prostate in the clinical practice. This benign tissue may contain field-effect changes affecting the results (232).

3.4.3. Clinical utility

Expression of the gene panel was demonstrated in three independent cohorts (Ross-Adams, Varambally and Taylor). However, all of these rely on prostate samples. The detection of these genes in blood and urine and validation of the relative expression between cancer and proven benign samples will allow for the development of a novel biomarker that can be incorporated into the clinical practice. We hypothesised that tissue detected biomarkers could be measured in biological fluids. The next chapter will explore the pitfalls of RNA expression measurement in clinical samples, and the results on the discovery cohorts.

3.5. Lay summary

This chapter covers the study of a new panel of genes to choose patients who will benefit from a biopsy. The panel consists of 16 genes previously published by Sharma *et al.* and belong to a signature identified in tissue and linked to aggressive prostate cancer. Additionally, 12 genes currently being studied at the Whitaker lab were added, for a total of 28 genes.

Prior to analysis in human samples, the gene expression of the 28 genes was measured in four cultured cell lines. These cell lines were selected as they represent different stages of prostate cancer: from cancer that still responds to therapy to unresponsive. Additionally, these cells were stimulated to determine if the gene expression changes with androgen (a testosterone-like hormone).

To simulate conditions in the bladder, cells were incubated with urine and gene expression was compared to cells not exposed to urine. Since urine damages cells and affects the normal reading of gene expression, mechanisms to deal with the damage caused by the urine were explored. The best method was used in human samples.

Using publicly available data from three large prostate cancer studies, a measure of performance called Area Under the Curve (AUC) was used. This measurement reflects how good a test is at finding a disease, with a score from 0 to 1. A value of 1 is perfect, the test is positive in all the men with cancer and none of the men without cancer. A value of 0.5 is equal to chance, you can toss a coin and the test will have the same chance of being correct. The area under the curve was measured for each gene to answer two questions:

1. Can any of the genes predict which men have cancer?
2. Can any of the genes help classify men depending on how aggressive their cancer is?

Taking the results for each gene two panels were created, grouping the best performing genes for each question.

3.5.1. Objectives

1. Study the gene expression of the 28 genes in four prostate cancer cell lines
2. Study the gene expression of the genes after the cells are given an androgen-like hormone.
3. Study how the urine degrades the molecules used to measure gene expression (RNA).
4. Study mechanism to overcome the degradation of RNA and select the best method.
5. Using publicly available datasets select a group of genes that selects men with prostate cancer.
6. Using publicly available datasets select a panel that can distinguish between men with low grade and high grade prostate cancer.

3.5.2. Findings

1. When measuring the gene expression in the four cell lines, the genes were divided in three groups:

- a. Group 1: Gene expression higher in androgen sensitive cells.
 - b. Group 2: Gene expression is the same in all cell lines.
 - c. Group 3: Gene expression is higher in androgen insensitive cells.
2. When measuring the gene expression after cells were stimulated with a testosterone-like hormone, the genes were subdivided in three groups:
 - a. Higher expression following stimulation.
 - b. No change
 - c. Lower expression following stimulation.

Take home message: The groups of genes selected are heterogeneous. Some are highly expressed in cells that behave similar to more aggressive cancer (Group 3 – higher expression in androgen insensitive cells). Some of these genes are turned on when androgen is added to the cells, and so therapies that block androgens could turn these genes off. This is important as the selected group of genes represent different biological stages of prostate cancer.

3. The RNA is highly degraded when cells are exposed to urine.
4. The best performing solution to degraded RNA was to perform a pre-amplification step. Where the limited amount of intact RNA is multiplied, this way when the final reaction is done a better idea of the gene expression is obtained.

Take home message: Urine is a hostile environment; ideally the urine should be treated with chemicals that stop this degradation. However, this is not always possible; the urine used in this study was already collected without a preservative. The best method was found to be to multiply (pre-amplify) the available RNA to use in subsequent experiments.

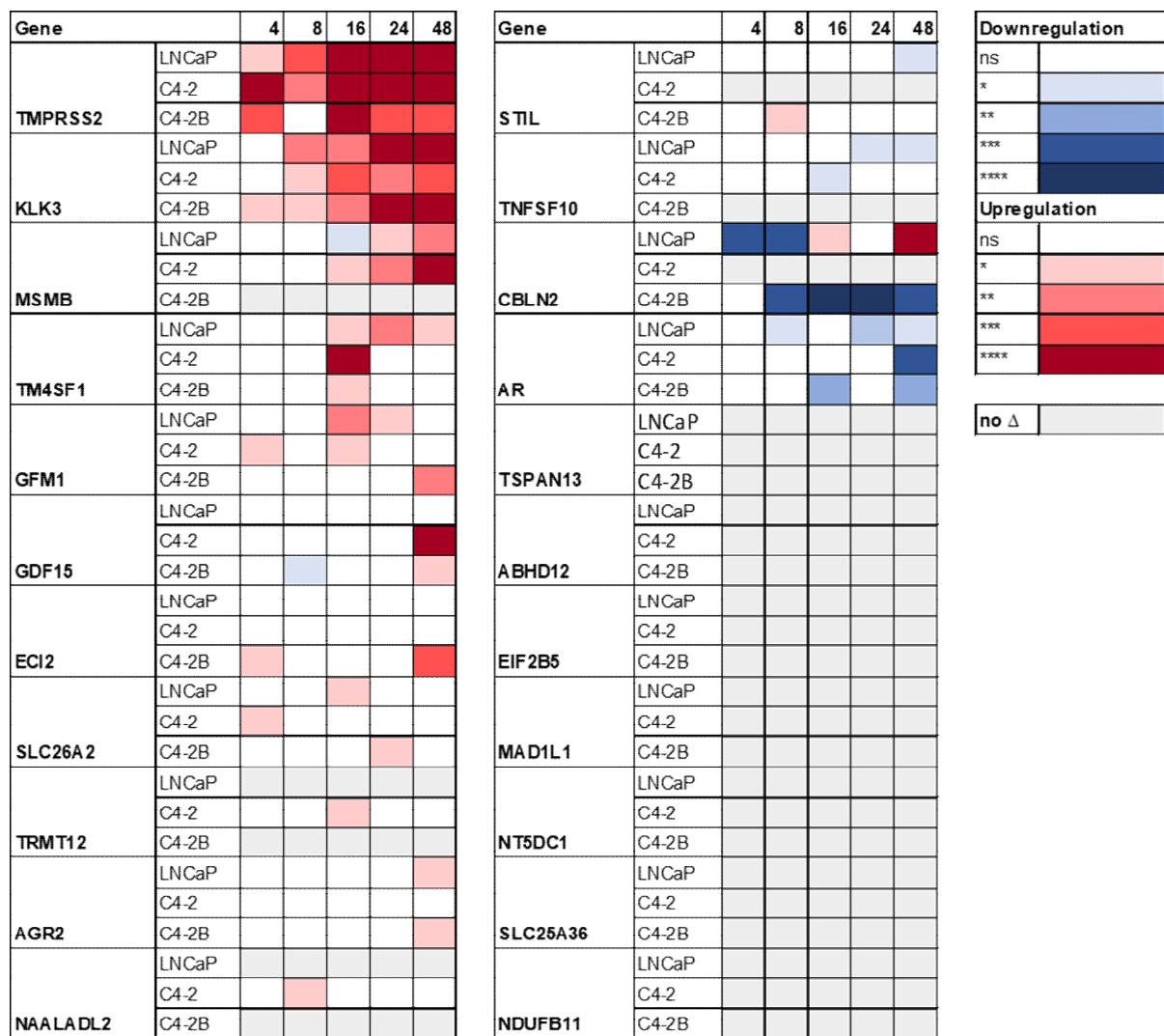
5. Two panels were created using information from publicly available datasets:
 - a. Diagnostic panel (can select men with cancer).
 - b. Risk stratification panel (can classify men depending on how aggressive their tumour is).

Take home message: Using data from other studies, we can study the performance of the genes prior to analysis in patient samples. This analysis produced two gene panels that will be studied and validated in the following chapters.

3.5.3. Limitations

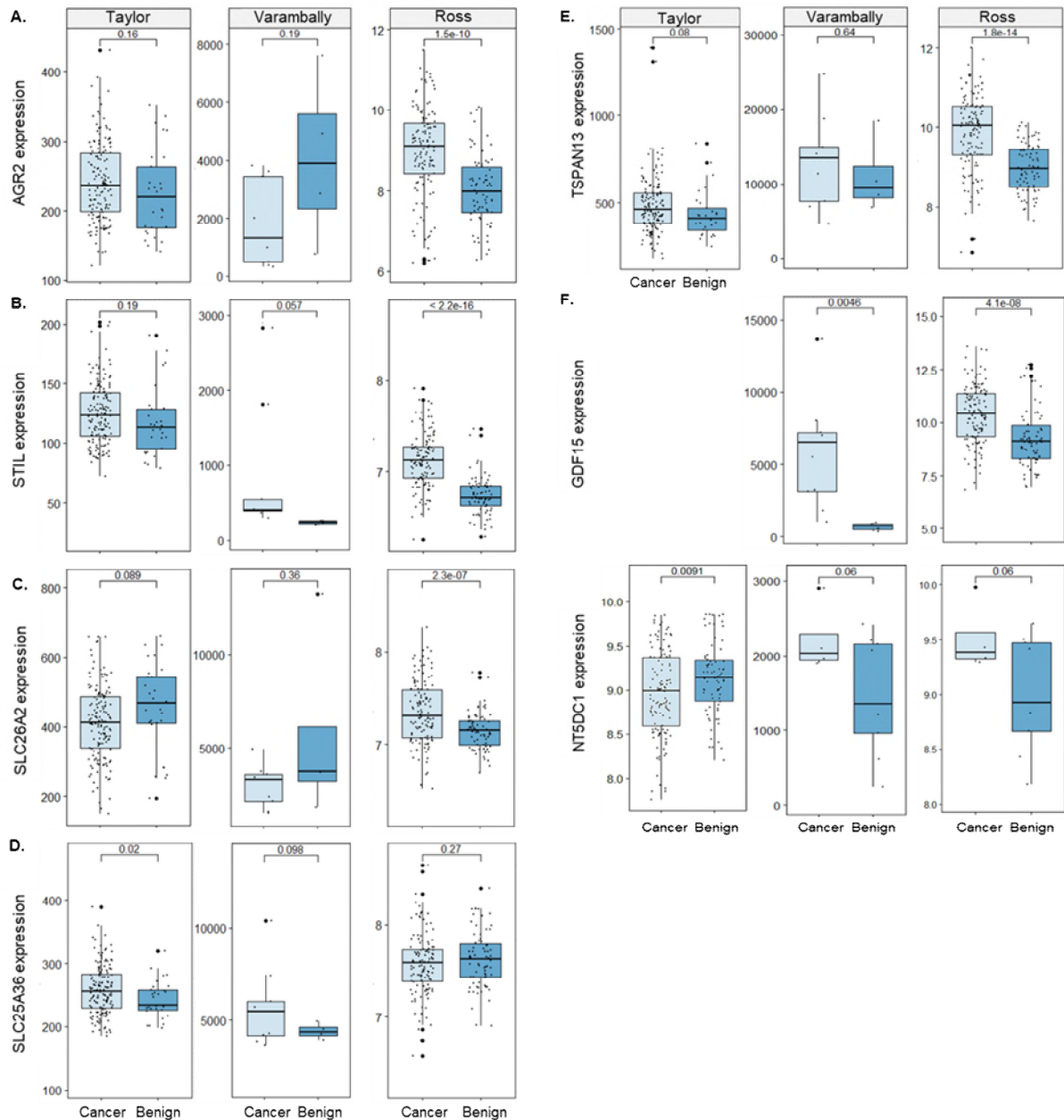
1. Cell lines are not perfect models for human disease as they grow flat in a plastic plate, whereas human cancers thrive in complex 3D models with many different cells.
2. The studies used have no MRI data and thus are not the best group of patients to try to help MRI select patients for biopsy. Additionally, these studies lacked tissue from healthy

men, and instead samples from prostate cancer patients that were free of tumour were used as control/comparison.



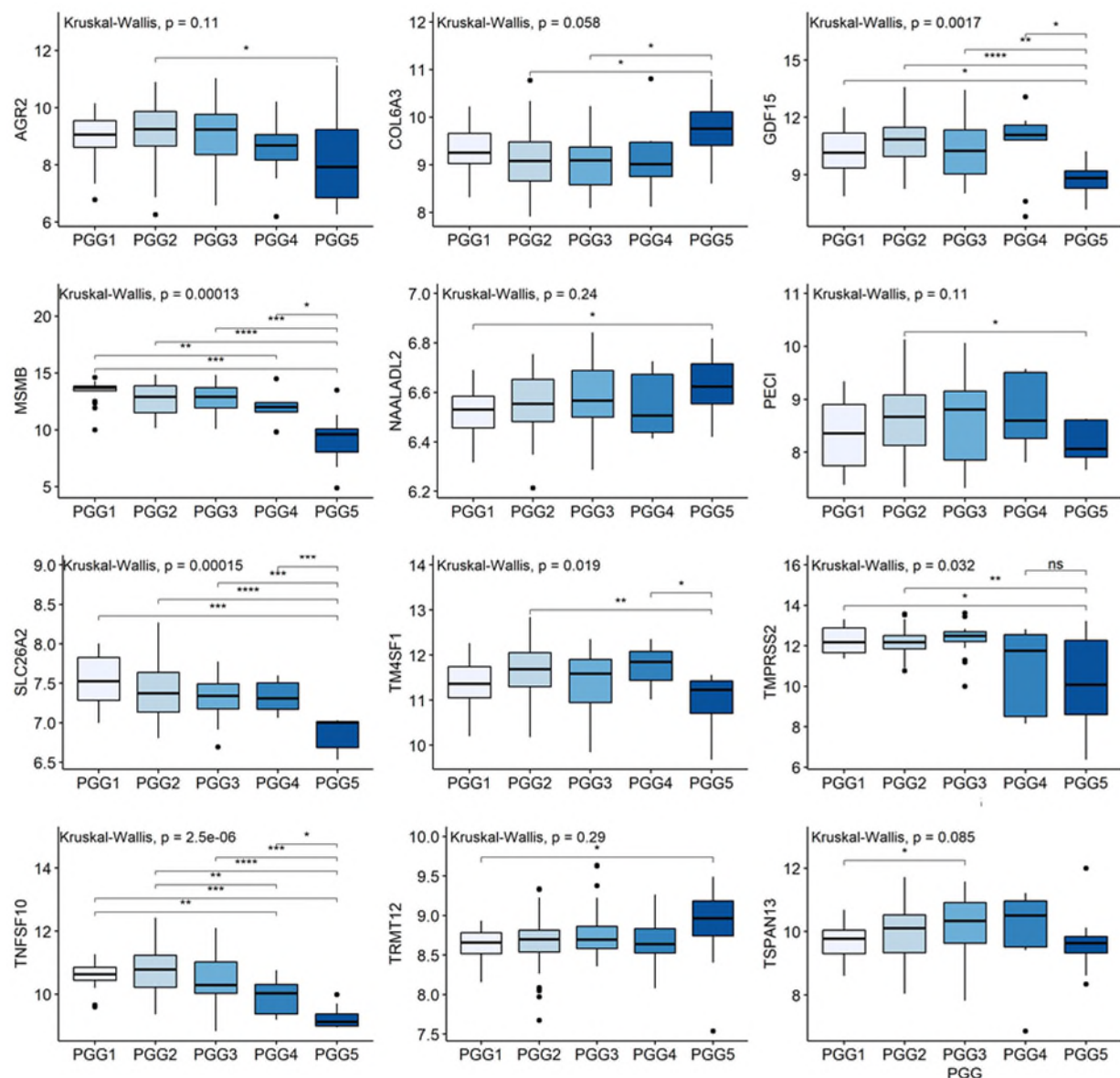
Appendix 1. Gene expression in LNCaPs, C4-2 and C4-2B cell lines following androgen stimulation.

Heatmap representing gene expression ranked by upregulation, downregulation and no change. Grey represents no androgen regulation after R1881 stimulation. Red gradient represents statistical significance if there is upregulation (KLK3, TMPRSS2, and MSMB) and blue if there is downregulation (AR, CBLN2). If there was no statistically significant difference, the areas are in white (GDF15 in LNCaP) or in grey if not change was seen (EIF2B5).



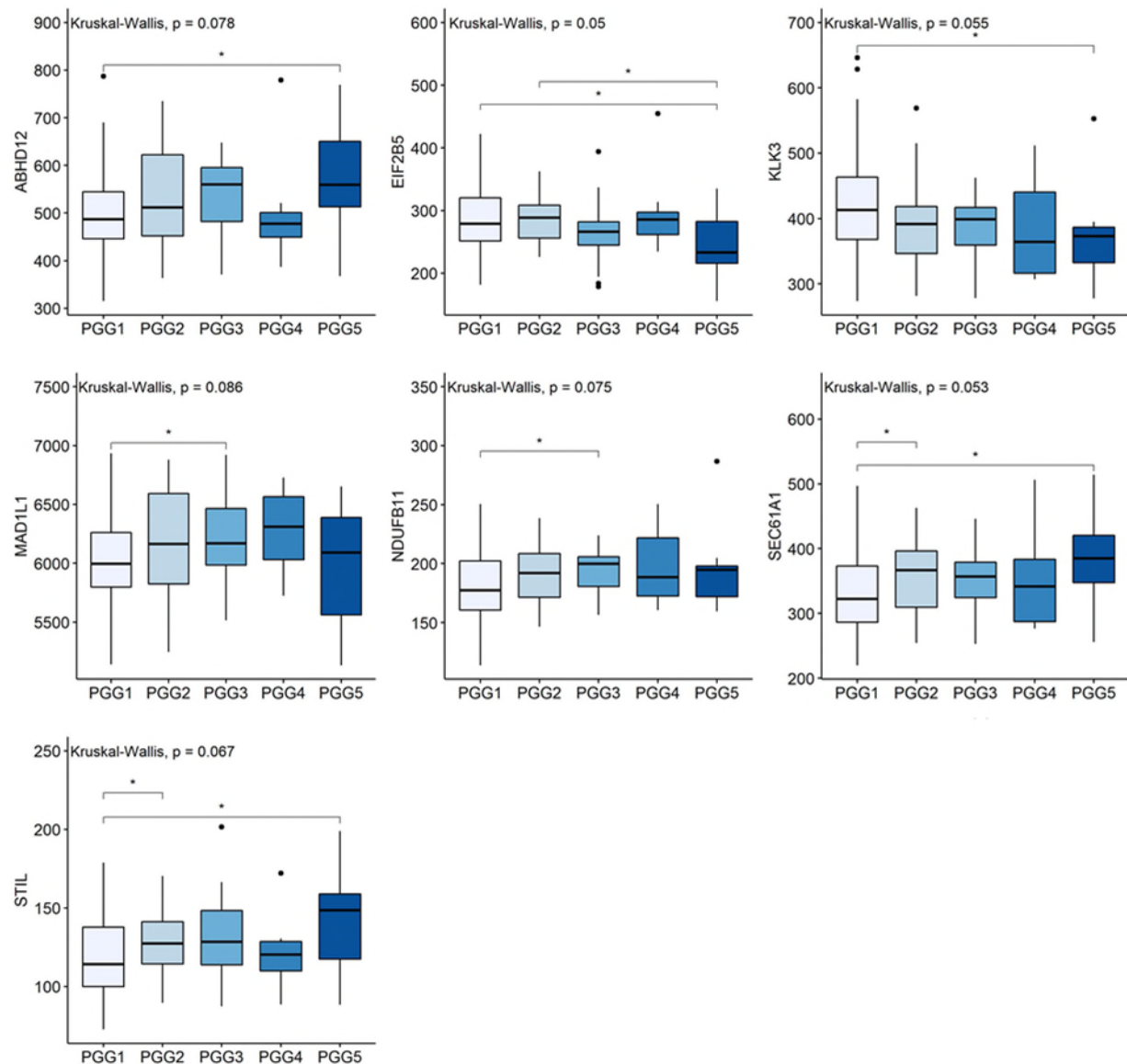
Appendix 2. Gene expression of the seven genes in the panel correlates with clinical diagnosis in three different cohorts.

The x-axis represents cancer and benign samples from each cohort, the y-axis the relative normalised expression for AGR2 (A), STIL (B), SLC26A2 (C), SLC25A36 (D), TSPAN13 (E) and GDF15 (F) and NT5DC1 (G). Note the concordance in expression in all cohorts for STIL and TSPAN13. In contrast, note concordance of SLC25A36 and SLC26A2 between Taylor and Ross-Adams samples, and concordance of AGR2 between Taylor and Varambally.



Appendix 3. Boxplots of the relative gene expression of genes in the Ross cohort.

The x-axis contains the PGG groups from one to five, the y-axis the relative gene expression. Kruskal-Wallis value displayed. If a significant difference in student t-test result is displayed. Note downward trend (significantly lower in PGG5 compared to PGG1) in AGR2, GDF15, MSMB, SLC26A2 and TNFSF10. The opposite trend is seen on NAALADL2 and TRMT12. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$



Appendix 4. Boxplots of the relative gene expression of genes in the Taylor cohort.

The x-axis contains the PGG groups from one to five, the y-axis the relative gene expression. Kruskal-Wallis value displayed. If a significant difference in student t-test result is displayed. Note downward trend (significantly lower in PGG5 compared to PGG1) in EIF2B5 and KLK3. The opposite trend is seen on SEC61A1 and STIL. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

Chapter 4. Measurement of mRNA gene panel expression in blood and urine samples in independent cohorts.

4.1. Introduction

The diagnosis of prostate cancer relies on a blood based biomarker (PSA), mpMRI and tissue examination under the microscope following a prostate biopsy (315,316). Unfortunately, these methods often over diagnose clinically insignificant disease and do not accurately identify all cases that require treatment leading to over-treatment(125,317–319). There have been many efforts to identify novel biomarker panels to improve the diagnosis and identification of “clinically significant” disease, with variable definitions and degrees of success. These biomarkers can be divided into fluidic (blood and urine), and tissue based; and further divided into diagnostic (to find the disease) and predictive (to predict risk and aid treatment decisions) (See 1.2.2).

Whilst tissue based biomarkers provide valuable information, such as the genetic makeup of the cancer (See 1.2.2.2.3), the necessity of tissue acquisition to risk stratify patients following a biopsy or extirpative surgery does not encourage for early diagnosis. Relying on tissue for risk stratification may be affected by the availability of tissue, for instance quality of nucleotides from FFPE samples and heterogeneity inherent to sampling (320–324). This dependence on tissue could be bypassed by developing minimally invasive tests on fluidic samples.

Many biomarker panels use clinical information to increase their performance, a biomarker that adds to the clinical information independently has not yet been found. For example, the STHLM3 model includes prostate exam and clinical factors (age, family history and previous biopsies) to the genetic panel and PSA results; resulting in an AUC of 0.74 for diagnosis (148). Using stepwise multivariate AUC, total PSA and risk factors combined (age, family history and previous biopsies) had an AUC of 0.58 (CI 0.56-0.60), the genetic score (risk alleles for 232 SNPs) increased the AUC by 0.02 to 0.60 (CI 0.58-0.62), addition of MSMB, MIC1, free PSA, intact PSA, and hk2 improved the AUC to 0.70 (CI 0.68-0.72). Finally, the AUC value of 0.74 (CI 0.72-0.75) was achieved by adding prostate exam information (digital rectal exam and prostate volume). Other biomarker panels are discussed on section 1.2.2.

Perhaps the search for accurate biomarkers could focus on already available biomarkers such as PSA density (PSAd = PSA value divided by prostate volume). Recent efforts to combine mpMRI information to biomarkers (219,325,326), have shown that this approach may help improve diagnosis of significant prostate cancer. For instance, PSAd has proven to be a reliable marker improving the diagnostic accuracy of mpMRI (67,136,138,139,327).

An alternative fluid for biomarker discovery is urine, as collection is painless and it is easily available. Many groups have studied optimal extraction methods, including adding stabilizer molecules such as RNAlater, or variation of extraction protocols (247,314). However, extraction of unstabilised urinary RNA can yield lower quality RNA (314,328,329). The urine used in this chapter was extracted without RNA stabilising solutions, therefore the quality of the RNA was expected to be low (See section 3.4.2). A literature review on urinary RNA preamplification prior to PCR analysis for prostate cancer biomarkers

was performed (118,294). The Whole transcriptome amplification (WTA) step was selected to produce cDNA in degraded samples (see 3.3.2.2.3). In this chapter the quality and quantity of previously extracted urinary RNA was analysed, to determine the relative gene expression of the refined diagnostic and prognostic gene panels (defined in 3.3.3.2), in a cohort of men that took part in the ProMPT trial.

Since the detection of RNA markers in blood and ensuring the disease specificity is vital, we hypothesised that tissue based biomarkers could be detected in blood and used to diagnose and risk stratify patients with prostate cancer. For this work, two panels of prostate cancer tissue based biomarkers (Figure 3-9), previously cross validated in three independent published prostate cancer tissue cohorts (220,305,306), were further tested for their diagnostic ability in two biological fluids: urine and blood. In order to do this the relative gene expression of the two gene panels were measured in urine of patients from the ProMPT study and PAXgene blood from patients from the PROMIS and INNOVATE trials (47,219). Using area under the curve values, the gene panel accuracy was determined and compared to current biomarkers such as PSA, PSA density and mpMRI.

Often, low quality RNA is obtained following extraction from blood or urine, and missing data is obtained after downstream applications, such as PCR. This may be due to RNA instability, and its susceptibility to biologically abundant RNA degrading substances (RNases, alkaline pH, for example) (330–332). Additionally, errors in processing of RNA for downstream analysis, human error and low RNA concentration can produce missing data. A potential solution is to pre-amplify samples as discussed in section 3.3.2.2.3. Regardless of the intent of obtaining a full dataset, missing data will be present, and thus, understanding missing data is part of the research process and analysis (Analysis of the patterns of missing data are discussed in section 2.2.6.3). In this chapter established statistical methods were used to understand the patterns of missing data and to populate, where appropriate, missing values for downstream analysis.

Due to the large number of samples and genes of interest, conventional qPCR was not feasible. 96 well plate qPCR is time consuming and adds to costs; in addition, inter-plate variation of experimental conditions carries additional analytical hurdles (333). Advances in micro-droplet dqPCR using the Fluidigm Biomark module allow for thousands of reactions to be performed using a single chip. This system uses nanoliter volumes to perform real time PCR by means of integrated fluidic circuits, allowing gene expression to be measured across multiple samples using a limited amount of reagents and sample.

4.1.1. Cohorts for biomarker panel discovery and validation

4.1.1.1. Cambridge ProMPT cohort

The “Prostate cancer: Mechanisms of progression and treatment” (ProMPT study) is a national collaboration effort that aims to study hormone resistant and metastatic prostate cancer (National Institute of Health and Research ID 5837). The study includes a biobank of samples of men with no cancer (presumed benign), and proven prostate cancer that attended clinic due to a raised serum PSA. Wherever possible patients had longitudinal follow up including PSA levels, mpMRI (if available), pathological information (TNM stage, Gleason score). For further demographic of patients included and clinical information see Table 4.2.

4.1.1.1.1. Urine samples cohort characteristics

The ProMPT cohorts included in this thesis included men with metastatic disease (n=7), therefore the median serum PSA and PSA density in these groups are higher when compared to the PROMIS and INNOVATE group. Samples were collected from 2003 to 2014 with variable follow up available. These samples were collected prior to the widespread use of mpMRI and thus this information is not available.

4.1.1.1.2. Blood samples cohort characteristics

In total, 89 patients with varying degrees of disease were included. This included patients with no cancer (n=10), localised disease (n=34) and metastatic disease (n=45). Clinical data was revised to ensure correct staging; a group of patients were found to be misclassified, see Table 4.1. Mainly, patients were misclassified, as the clinical data used for group subdivision was not based on the clinical status at time of samples being taken. For example, a patient was misclassified as metastatic on hormones, when at the time of collection he was hormone naïve.

| Group | Initial staging | Corrected staging |
|-------------------------------------|------------------------|--------------------------|
| Benign | 10 | 10 |
| Low risk | 11 | 11 |
| Intermediate risk | 11 | 10 |
| High risk | 11 | 15 |
| Metastatic on hormones | 16 | 11 |
| Metastatic hormone naïve | 18 | 16 |
| Metastatic hormone resistant | 12 | 16 |

Table 4.1 ProMPT patient reclassification following data review.

Groups using D'Amico classification, initial staging corresponds to classification of patients on original file. Corrected staging following review of clinical data.

4.1.1.2. Prostate MR imaging study (PROMIS) cohort at UCLH

The PROMIS study is a multicentre trial to study the ability of mpMRI to detect clinically significant prostate cancer compared to the standard of care (TRUS biopsy) and full interrogation of the prostate every 5 mm (transperineal prostate biopsy) (47). Two definitions of significance were applied: Gleason $\geq 4+3$ or more, or a MCCL ≥ 6 mm in any location (primary) or Gleason $\geq 3+4$ or a MCCL ≥ 4 mm (secondary). Patients from eleven centres took part, but only patients from UCL provided blood and urine samples for biomarker and genetic analysis (227 patients in total). Summary of clinical data is shown on Table 4.2.

This unique study includes patients at the first referral to a urologist due to suspected prostate cancer, consequently there is a large number of patients with no prostate cancer as well as low grade prostate cancer. Interestingly this offers an opportunity to differentiate men in the benign group based on common histological abnormalities such as atrophy, inflammation, high grade PIN and ASAP and the few with none of these changes (true benign). In addition, since these men have had thorough sampling of their prostates, we can assume that their negative status is as accurate as possible, short of removing the prostate for full whole-mount analysis (which is clearly impossible in a diagnostic cohort). Furthermore, the extensive sampling of these men's prostate allows for extensive pathological analysis in the cancer cohort, as well as correlation with imaging and fluid (blood and urine) mRNA expression profiles.

The follow up in this cohort is limited as the sample collection started in 2012 up to 2015, therefore it is not possible to ascertain if any of the men in the benign group went on to develop prostate cancer or had undetected prostate cancer. In addition, it has not been long enough to observe the natural history for aggressive cancers evolving to hormone resistance.

4.1.1.3. INNOVATE UCL cohort

The INNOVATE study (Combining advaNces in imaging with biOmarkers for improved diagnosis of Aggressive prostate canCEr) aimed to combine discovery and validation of known fluidic biomarkers in combination with mpMRI and a novel MRI sequence called VERDICT (219).

The cohort in this trial includes men referred to the urology clinic due to suspicion of prostate cancer, with blood and urine samples taken prior to mpMRI in addition to the VERDICT sequence. Men with suspicious lesions (MRI Likert score ≥ 3) had a targeted transperineal biopsy (if clinically indicated), those with lesions scoring Likert ≤ 2 were discharged back to their GP. This reflects the current clinical practice at UCH; nonetheless, due to the nature of the study it is not possible to ascertain that men with negative biopsies or negative mpMRI are true negatives. Clinical data can be seen on Table 4.2

| | | ProMPT (blood) | ProMPT (urine) | PROMIS (blood) | INNOVATE (blood) | p value |
|-----------------------|----------------------------|----------------------|-----------------------|---------------------|---------------------|----------|
| Disease status | Benign | 10 | 28 | 65 | 60 | |
| | Localised | 31 | 44 | 129 | 29 | |
| | Metastatic | 48 | 7 | 0 | 0 | |
| Clinical data | Age (range) | 67 (49-93) | 62.1 (39-84) | 62 (39-79) | 63 (43-81) | <0.0001 |
| | PSA (range) | 148 (2-6470) | 11.03 (1.18-95.58) | 6.77 (0.9-15) | 7.5 (0.75-37.8) | < 0.0001 |
| | Prostate volume (range) | 45 (16-160)* | N/A | 48.4 (17-126) | 54.5 (14-180) | 0.063 |
| | PSA density (range) | 1.02 (0.03-16.9)* | N/A | 0.15 (1.02-0.58) | 0.17 (0.02-2.02) | <0.0001 |
| PGG | 1 | 12 | 12 | 34 | 3 | |
| | 2 | 14 | 25 | 80 | 16 | |
| | 3 | 4 | 3 | 15 | 5 | |
| | 4 | 5 | 1 | 0 | 0 | |
| | 5 | 30 | 10 | 0 | 5 | |
| | No data | 14 | 0 | 0 | 0 | |
| Likert score (MRI) | 1 | ND | ND | 0 | 0 | |
| | 2 | ND | ND | 34 | 14 | |
| | 3 | ND | ND | 71 | 43 | |
| | 4 | ND | ND | 38 | 20 | |
| | 5 | ND | ND | 41 | 12 | |
| Total | | 89 | 79 | 194 | 89 | |

Table 4.2 Cohorts disease status, clinical, pathological and imaging data.

Three cohorts were included; the ProMPT samples originated in Cambridge University Hospital, in total 89 patients were included in the analysis. 79 urine samples correspond to a different cohort of ProMPT patients. There is no mpMRI data in this cohort (ND). The UCL cohorts are PROMIS (blood) and INNOVATE (blood), both of which have imaging data but lack metastatic samples. Disease status, clinical data (age, PSA, prostate volume and PSA density), PGG and Likert score values are plotted on this table. There was a significant difference amongst the three cohorts on the distribution of age, presenting PSA and PSA density ($p < 0.0001$, 2 way ANOVA), this can be explained by an older cohort and the presence of patients with metastasis in the ProMPT cohort. There is no statistical difference in prostate volume across the cohorts.

4.2.Objectives

4.2.1. Hypothesis 1:

Urine RNA is of sufficient quality to detect prostate RNA and can be used as a diagnostic and prognostic biomarker.

4.2.1.1. Aim 1: Evaluate whether RNA from urine cell pellets can be effectively used to measure the refined diagnostic and prognostic gene panels in a cohort of prostate cancer patients.

Objective 1: Evaluate the quality and quantity of urine RNA.

Objective 2: Measure relative gene expression of the refined diagnostic and prognostic gene panels developed in Chapter 2 in urinary mRNA.

Objective 3: Evaluate the performance of the diagnostic and prognostic panels in urinary RNA.

4.2.2. Hypothesis 2:

Blood PAXgene RNA can be used to diagnose and risk stratify prostate cancer patients.

4.2.2.1. Aim 2: Measure the RNA levels of the seven-gene panel in PAXgene blood RNA from prostate cancer patient samples.

Objective 4: Evaluate the quality and quantity of PAXgene extracted RNA.

Objective 5: Determine the relative gene expression of the two gene panels in PAXgene extracted RNA.

4.2.2.2. Aim 3: Study the impact of the gene panel as a diagnostic and prognostic test.

Objective 6: Understand the diagnostic accuracy of the gene panel in two independent clinical cohorts (PROMIS and INNOVATE), to help determine which patients will benefit from a biopsy.

Objective 7: Understand the utility of the gene panel for use as a risk stratification tool in two independent cohorts (PROMIS and INNOVATE).

4.3.Results

4.3.1. Urine as a fluid for biomarker discovery

4.3.1.1. Quality and control

Urine pellets from the PROMPT cohort of patients underwent RNA extraction as described in section 2.2.2.3. The 260/280 ratio can give an indication of organic solvent contamination introduced during sample extraction which can affect the efficiency of downstream measurements (334,335). The average 260/280 ratio of all samples was 1.69 (1.18-2.07). Carrier RNA was added at the time of extraction, which may account for a proportion of the final RNA concentration. In addition, 58 samples (73.4%) had a low 260/280 ratio as seen on Figure 4-1 A, Only 26.6% (n=21) had an acceptable 260/280 ratio. Additionally, there was a strong linear relationship between the 260/280 ratio and final concentration (R: 0.8, $p < 0.0001$) (see Figure 4-1 B). Samples with higher concentrations were found to have phenol or Qiazol contamination (peak at 230 nm) and low 260/230 ratio confirming organic contaminants.

Bioanalyzer analysis revealed that the urinary RNA was severely degraded; 62% of samples had RIN number of 1 (n =49), 18 % RIN between 1 and 2.5 (n=18), 7.5% RIN of 2.6 and no reading (n= 6 respectively), see Figure 4-1 C. This poses a challenge prior to qPCR as discussed in chapter (3.3.2).

4.3.1.2. Preamplification

As suggested by Chan *et al.* and Nguyen *et al.*, low concentration and degraded RNA can be preamplified prior to PCR (294,295). WTA was implemented on the urinary mRNA samples (For further information see section 3.3.2.2.3). This step can induce bias in the data, mainly uneven amplification of targets, producing unreliable results. Examination of the Ct values of a housekeeper gene (RPLP2) divided the samples into two groups: samples whose Ct values improved following amplification (expected outcome) and samples where there was no change or increased Ct value (undesired outcome), Figure 4-1 D and E, respectively. As a housekeeper acts as an internal control regardless of the PCR reaction, this biased amplification affects the delta Ct calculations and ultimate analysis of data. There was a reduction of the Ct value as expected in 15 samples (19%). In 57% of the cases, the Ct values increased or remained the same after amplification (n= 45), 6 samples had no Ct values before or after amplification (7.5%, n=6) and 13 samples became undetectable following amplification (16.5%) and are not plotted here. For ease of interpretation, data in Figure 4-1 D and E were plotted separately.

PCR results of sample 202-585 for three genes confirmed the biased amplification; for instance, PCA3 is undetected prior to amplification (no light blue bar), whereas both housekeepers (SDHC and RPLP2) are detected. Amplification results in detection of PCA3, whereas SHDC is undetected following amplification (no dark blue bar) (Figure 4-1 F). The WTA step was considered biased, PCR reactions were performed using unamplified RNA.

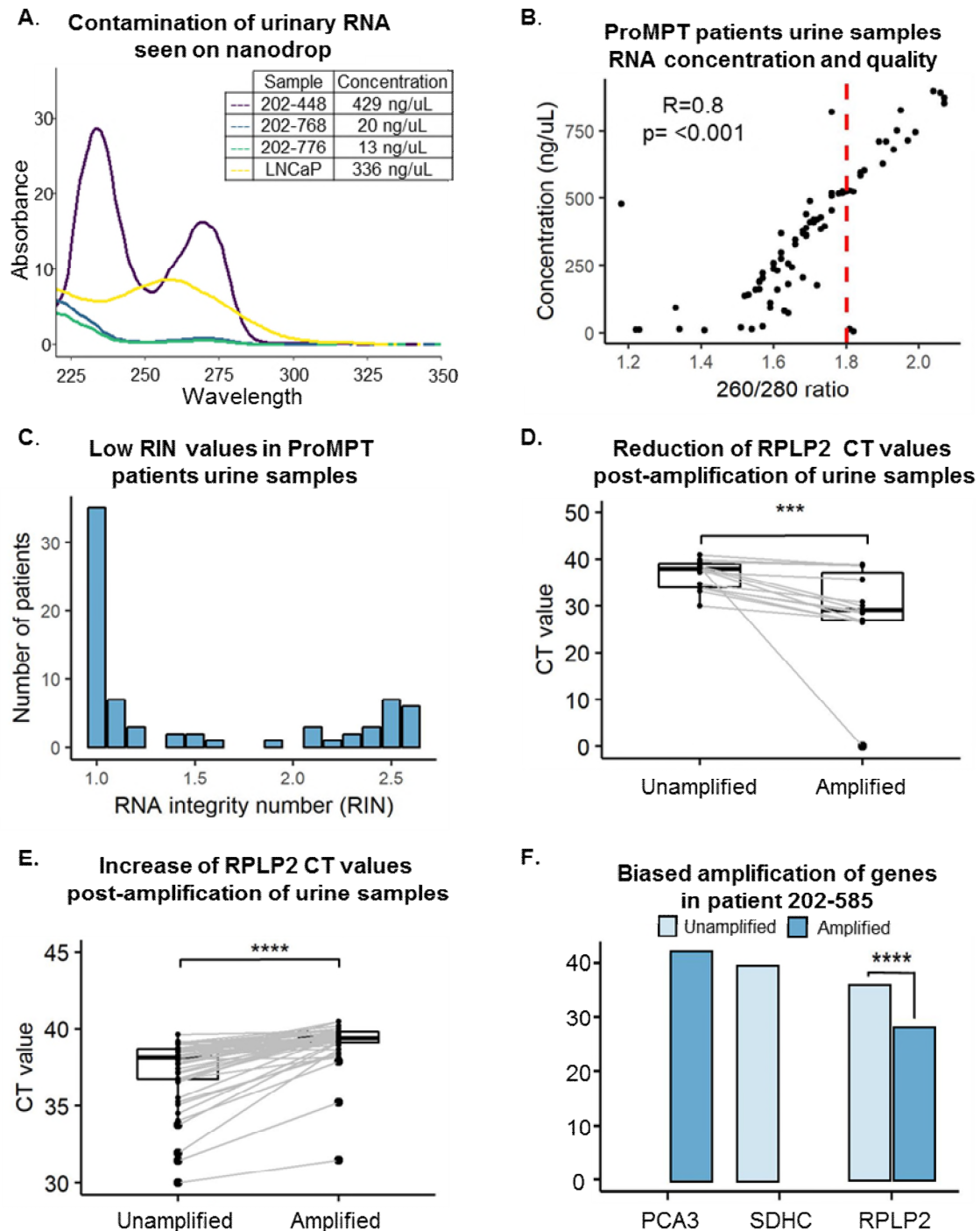


Figure 4-1 Low quality urinary mRNA in ProMPT patient samples are not suitable for unbiased preamplification prior to qPCR.

A. Optical absorbance of extracted urinary RNA for samples 202-448, 202-768, 202-766 and LNCaP. Sample 202-448 has organic compound contamination as seen by peak at 230 nm. Sample ID and nanodrop concentration on table. **B.** Dot plot comparing the RNA concentration in ng/mL (y-axis) and the 260/280 ratio (x-axis), Red dashed line represents limit of ideal 260/280 ratio as values below 1.8 are considered to be protein contaminated. **C.** Histogram representing the RIN number distribution amongst all urinary mRNA samples. **D.** Boxplot of Ct values of unamplified and amplified samples of the housekeeper gene RPLP2, where Ct values were lower following amplification and **E.** where the Ct value remained the same or increased. **F.** Barplot of Ct values for sample 202-585. PCA3 gene is not detected prior to amplification and has a high Ct value of 40 following amplification (n=3). Two housekeeper genes are plotted; SDHC became undetectable following amplification, whereas RPLP2 had a significantly lower Ct value post amplification. **** p < 0.0001.

4.3.1.3. qPCR of degraded urinary RNA

4.3.1.4. Missing data

qPCR of unamplified urine samples revealed a high proportion of missing data, fitting one of eight patterns across the technical replicates (Figure 4-2 A). The most common missing pattern was non-reads in all three replicates, followed by non-reads in two out of three replicates. The percentage of missing data per replicate can be seen on Figure 4-2 B, the gene with most missing data was STIL (96.8%-100% per replicate). When excluding the housekeeper RPLP2, a high percentage of data was missing on all genes (median 72.2%, range 42.8%-97.8%). RPLP2 had 22.0% of data missing (Figure 4-2 C). Prior to inputting missing data, Little's multivariate MCAR test was performed, the null hypothesis being that Ct values are missing independently from the data values (214,215); in this case the "missingness" cannot be explained by available or missing data and are therefore, missing completely at random (MCAR). If the p value is not below 0.05, then data is MCAR. All genes were found to be MCAR (Figure 4-2 D), with the exception of MSMB that was missing not at random (MNAR).

Multiple random imputation was performed to input missing data, if data only was available in two readings and if the housekeeper was detected. One thousand simulations were performed and a random "complete" dataset was selected for final analysis. Data was only imputed if more than 8% of data was available; therefore, STIL was excluded as it had 97.8% missing data.

Figure 4-3 A shows the Delta Ct values of the unamplified urine samples following qPCR. As expected, a large proportion of data is missing (grey). Following multiple random imputation, the amount of missing data is reduced as seen in Figure 4-3 B and C. However, following multiple random imputation, data was mostly populated in patients with cancer; Figure 4-2 D shows the amount of missing data per patient in the benign cohort. In average 5/14 genes were detected in all patients (range 0-13). Only five patients had more than 10 data points for analysis (gene expression detected).

Relative gene expression as calculated as described in section 2.2.6.1. Prior to analysis, the data was examined to determine the best statistical test for analysis. Quantile-quantile plots (QQplot) per gene were generated; data was not equally distributed as seen in Appendix 5. Note this is especially true for the benign samples (black). Levene's test was performed to determine the variance of relative gene expression between benign and cancer samples. There was unequal variance distribution in genes AGR2, AR, ECI2, PCA3, PSA and VPS13A. The remaining genes had equal variance distribution (Appendix 5). Welch-Satterthwaite t-test was performed for the unequal variance samples and a regular student t-test for the equally distributed samples. When the relative gene expression of cancer and benign was compared, a statistically significant difference in genes AGR2, TMPRSS2, AR, ECI2, PSA, PCA3 and VPS13A was seen (Figure 4-4). Note the variation in patient numbers in the benign group.

Due to the amount of missing data in the urine samples, and imbalance between the benign and cancer group, no further analysis was performed on urine samples.

A. Possible missing data patterns

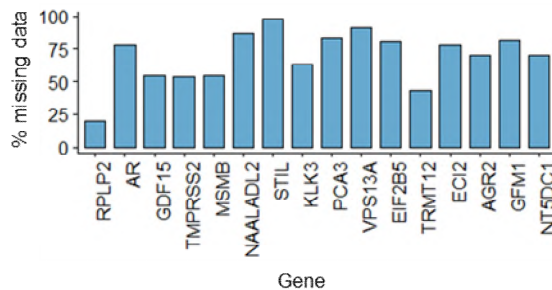
| Missing values pattern | | | |
|------------------------|----|----|----|
| | R1 | R2 | R3 |
| 1 | | | |
| 2 | | | |
| 3 | | | |
| 4 | | | |
| 5 | | | |
| 6 | | | |
| 7 | | | |
| 8 | | | |

| | |
|--|-------------|
| | Non missing |
| | Missing |

B. Percentage (%) of missing data per replicate per gene

| Gene | R1 | R2 | R3 |
|----------|------|------|-------|
| RPLP2 | 22.0 | 22.0 | 15.9 |
| AR | 77.8 | 77.8 | 77.8 |
| GDF15 | 50.7 | 58.7 | 53.9 |
| TMPRSS2 | 57.1 | 54.0 | 49.2 |
| MSMB | 58.7 | 50.8 | 54.0 |
| NAALADL2 | 85.7 | 90.5 | 84.1 |
| STIL | 96.8 | 96.8 | 100.0 |
| KLK3 | 58.7 | 61.9 | 68.3 |
| PCA3 | 85.7 | 85.7 | 79.4 |
| VPS13A | 88.9 | 92.1 | 92.1 |
| EIF2B5 | 85.7 | 77.8 | 79.4 |
| TRMT12 | 44.4 | 38.1 | 46.0 |
| ECI2 | 77.8 | 81.0 | 74.6 |
| AGR2 | 76.2 | 71.4 | 61.9 |
| GFM1 | 82.5 | 81.0 | 81.0 |
| NT5DC1 | 66.7 | 69.8 | 73.0 |

C. Total missing data per gene (%)



D. Little's MCAR test results in urine samples

| Gene | Chi square | p value | MCAR? |
|----------|------------|---------|-------|
| RPLP2 | 14.64 | 0.101 | yes |
| AR | 14.06 | 0.015 | yes |
| GDF15 | 13.41 | 0.145 | yes |
| TMPRSS2 | 21.64 | 0.01 | yes |
| MSMB | 23.06 | 0.006* | no |
| NAALADL2 | 1.104 | 0.993 | yes |
| STIL | 0 | 1 | yes |
| KLK3 | 8.5 | 0.382 | yes |
| PCA3 | 16.07 | 0.065 | yes |
| VPS13A | 6.49 | 0.261 | yes |
| EIF2B5 | 12.97 | 0.172 | yes |
| TRMT12 | 12.84 | 0.268 | yes |
| ECI2 | 11.19 | 0.189 | yes |
| AGR2 | 10.33 | 0.171 | yes |
| GFM1 | 15.68 | 0.074 | yes |
| NT5DC1 | 9.83 | 0.364 | yes |

Figure 4-2 Missing data following qPCR of unamplified urinary RNA.

A. Eight possible missing data patterns per technical repeat (R1 = repeat one, R2= repeat two and R3= repeat three). Missing data in blue and non-missing data in white. **B.** table containing percentage of missing data per technical replicate per gene. **C.** Barplot representing the average total missing data per gene, RPLP2 has only 22% missing data and STIL has 97.8% of missing data. **D.** Chi square and p values following Little's MCAR test for all genes confirming if data was MCAR or MNCAR.

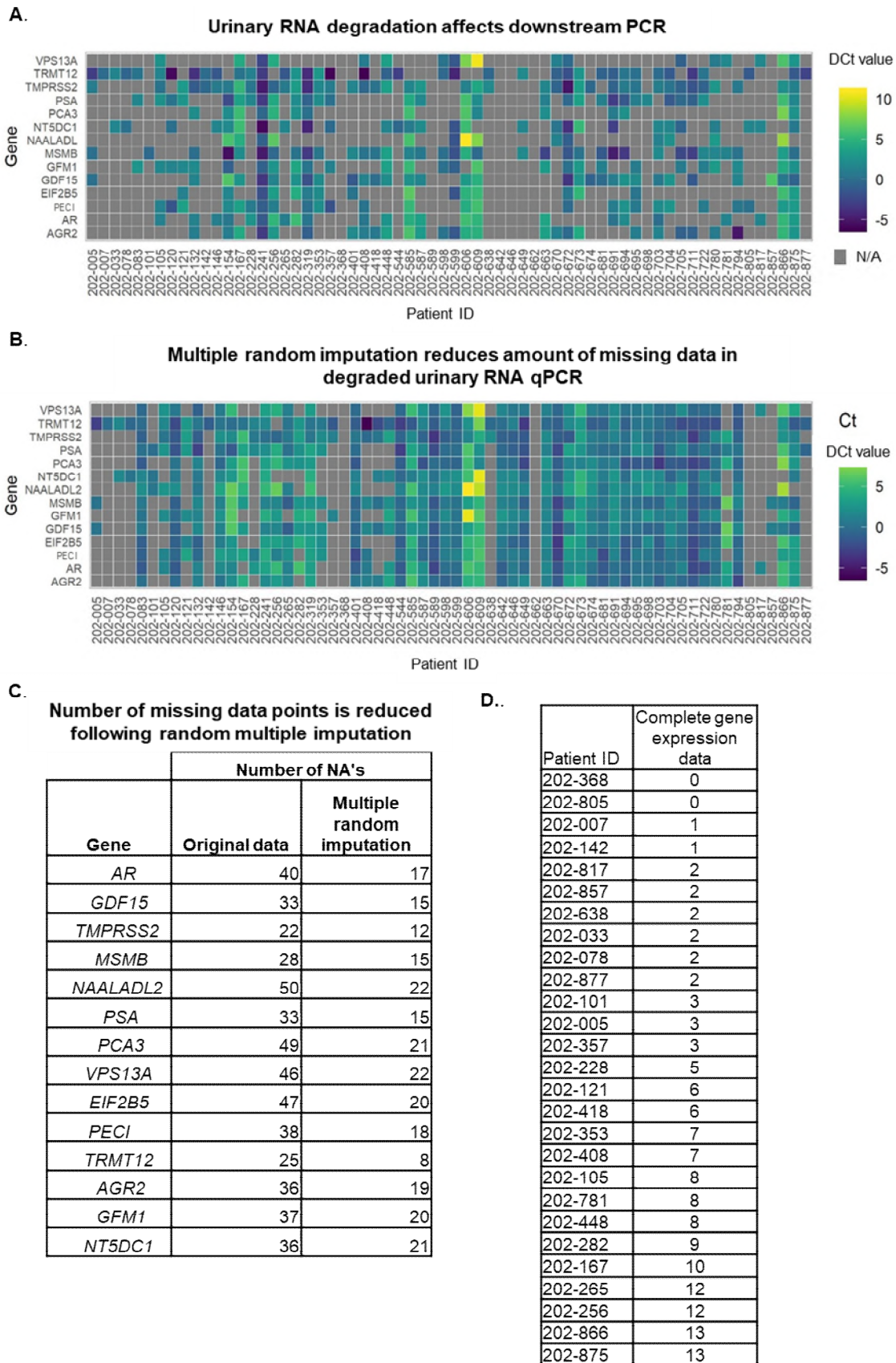
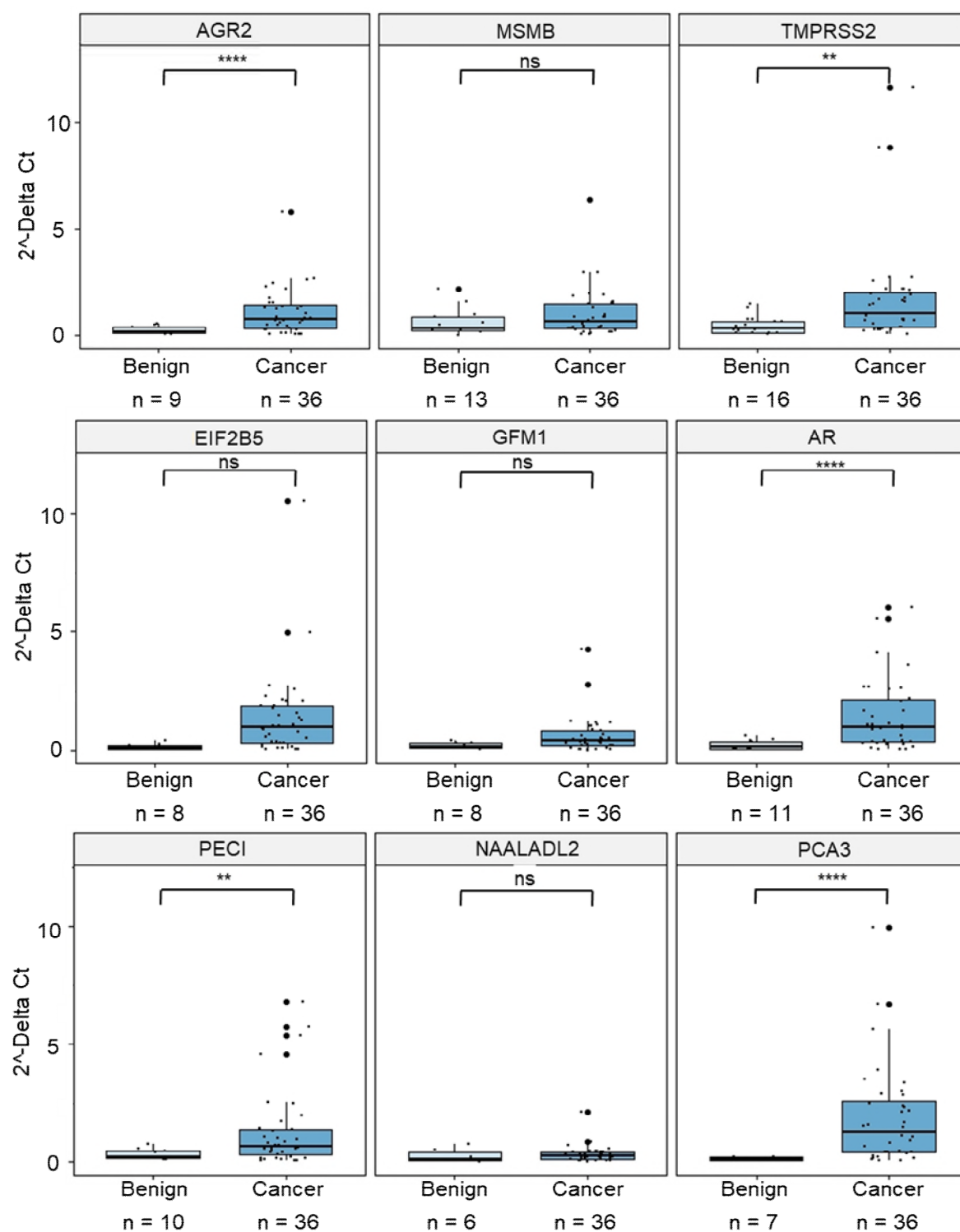


Figure 4-3 Multiple random imputation reduces the number of missing data points.

A. Heatmap representing the Delta Ct (DCt) values of 14 genes following qPCR, missing values are represented in grey. Gradient represents the DCt values from -5 to 10. **B.** Heatmap representing the

DCt values following multiple random imputation. Note reduced number of grey spaces and fewer extreme DCt values (bright yellow and dark purple) compared to A. **C.** Table containing the number of NA's prior to and after multiple random imputation. **D.** Table representing patient in the benign group and the number of genes where data was available (mean = 5, range= 0-13).



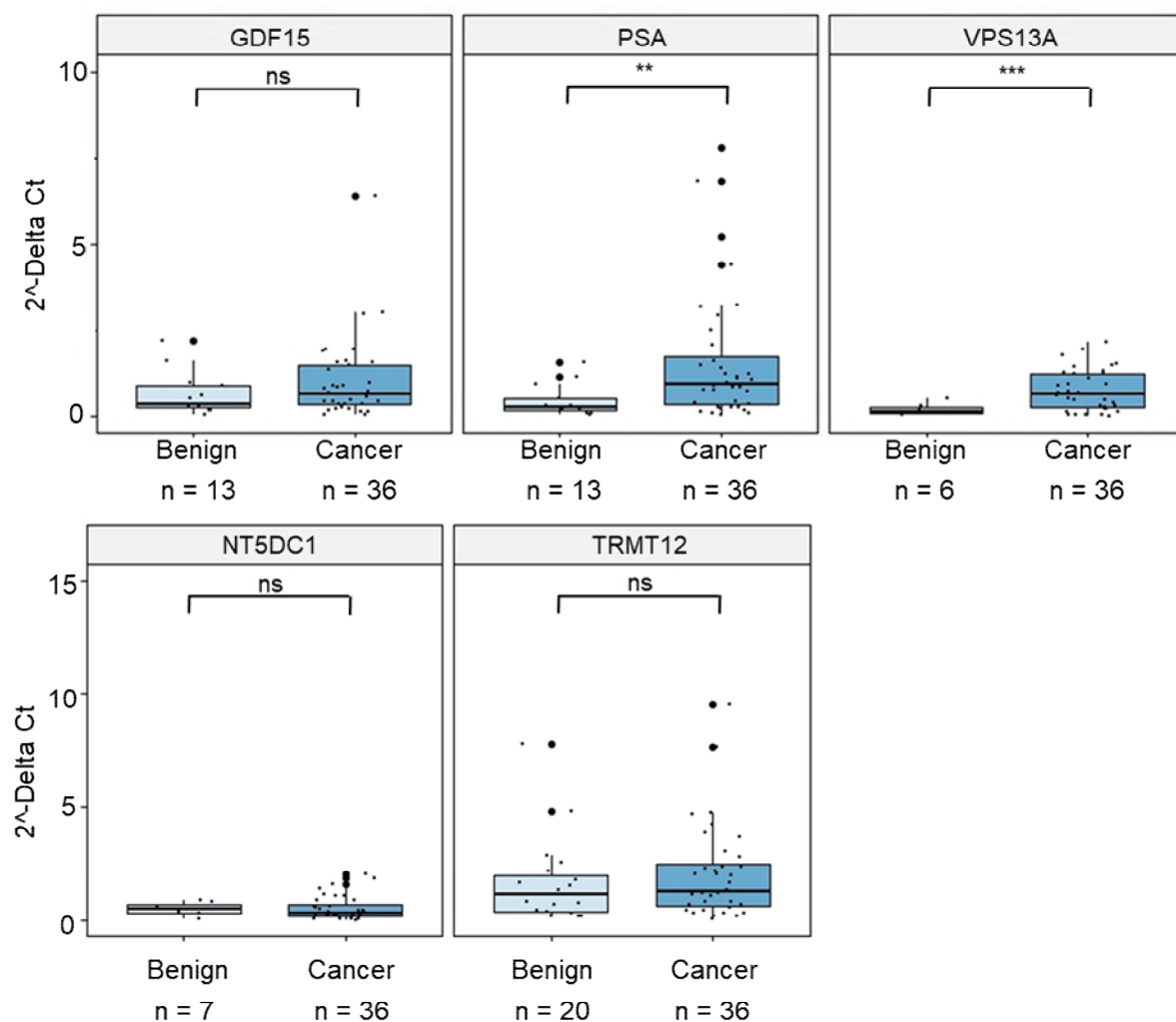


Figure 4-4 Relative gene expression of genes in urine RNA reveals 6 upregulated genes in cancer.

Boxplots comparing benign and cancer samples in urine RNA, n= number of patients for each group, $2^{-\Delta Ct}$ represented in y-axis.. Welch-Satterthwaite t-test was performed on AGR2, AR, ECI2, PCA3, PSA and VPS13A. Student's t-test was performed in EIF2B5, GDF15, MSMB, NAALADL2, NT5DC1, TMPRSS2, and TRMT12. * p < 0.05, ** < 0.01, *** p < 0.001, **** p < 0.0001, ns= not significant.

4.3.2. Studying gene expression using PAXgene whole blood

4.3.2.1. Gene expression in PAXgene RNA in patients from the ProMPT study

Gene expression of RNA extracted from PAXgene blood at Cambridge University was performed by Dr Thomas Johnston under supervision by Dr Hayley Whitaker. Data was shared for analysis. No data was available on quality and control. A concentration of 40ng was used for cDNA creation. Initial analysis of samples consisted of reviewing quality of PCR plates. This was performed by examining housekeeper variability between the samples.

Three housekeeper genes were included in three separate chips, including RPLP2, RPLP0 and SDHC. No samples showed fluorescence with RPLP0, this housekeeper was excluded. Data distribution and homogeneity of variances was examined for RPLP2 and SDHC, followed by one-way Anova or Kruskal Wallis depending on data normality (2.2.6.2). For the purpose of brevity, Ct values of RPLP2 and SHDC were found to be statistically different between different stages of disease. RPLP2 values were plotted and statistical results can be found on Figure 4-5. This data was not used for further analysis as normalisation to a housekeeper was not possible.

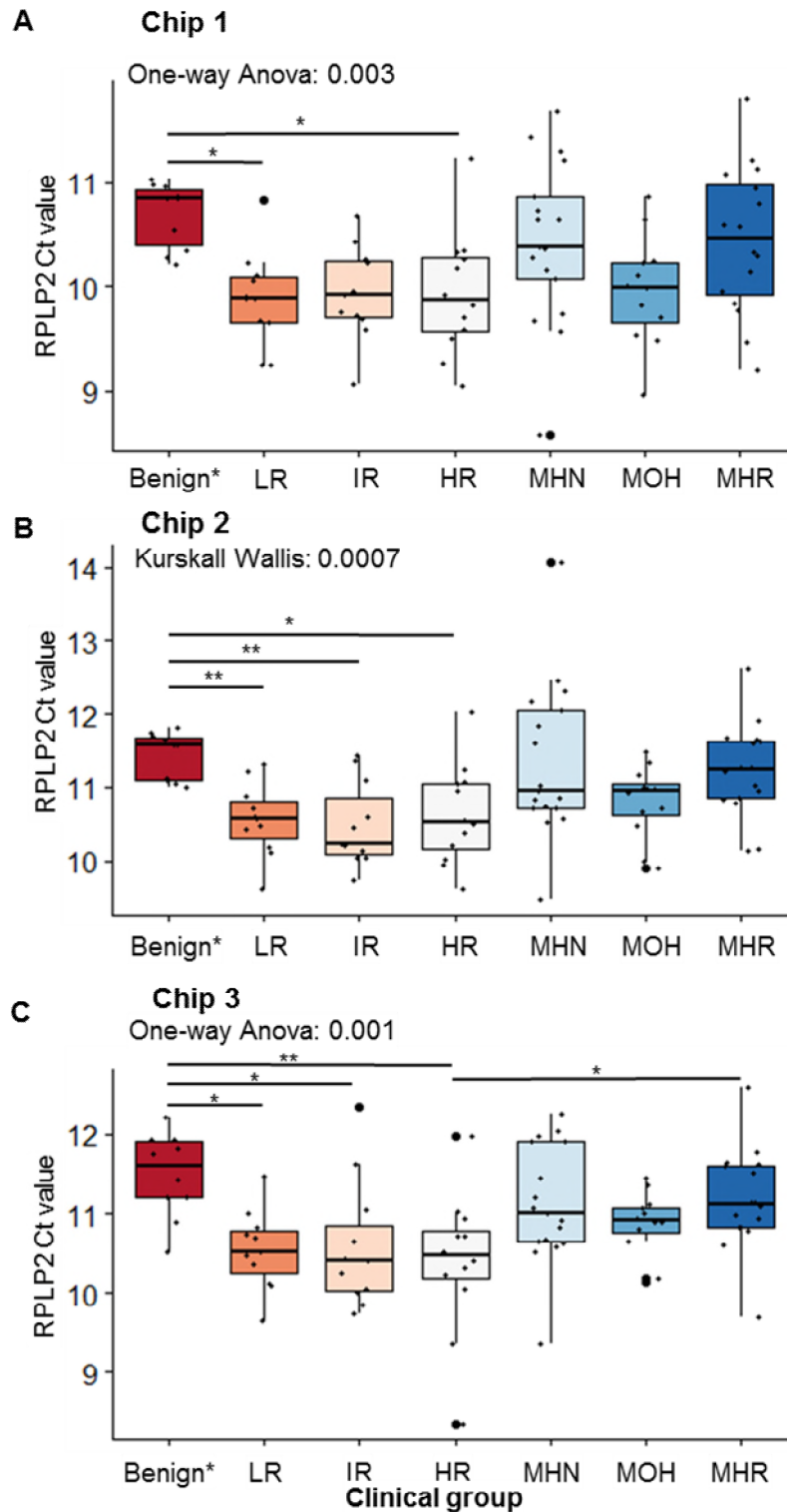


Figure 4-5 RPLP2 Ct values vary between disease stages and it is not an appropriate housekeeper gene for normalisation of PCR data.

Ct values of RPLP2 gene plotted on y-axis. D'Amico group classification on x-axis as Benign*, LR = low risk, IR= intermediate risk, HR= high risk, MHN= metastatic hormone naïve, MOH= metastatic on hormones and MHR= metastatic hormone resistant. **A** corresponds to chip1, **B** to chip 2 and **C** to chip 3. Data was normally distributed in chip1 and 3, variance distribution was homogeneous in all chips. One-way Anova was performed for chips 1 and 3, followed by Tukey post-hoc analysis. Kruskal Wallis was performed for Chip 2, followed by Nemenyi post-hoc test. $p < 0.05$, $** < 0.01$, $*** p < 0.001$, $**** p < 0.0001$.

4.3.2.2. Gene expression in PAXgene RNA in patients from the PROMIS and INNOVATE studies

4.3.2.2.1. Quality and control

Blood samples from PAXgene RNA tubes were extracted following the extraction protocols (See methods in section 2.2.2.2). In total 208 samples from PROMIS and 96 samples from INNOVATE were extracted. Final sample selection for PCR was based on 260/280 ratio, final concentration and RIN value. A low 260/280 ratio indicates contamination from reagents used in the extraction protocol, this may affect PCR efficiency. For optimal micro-droplet PCR (Fluidigm) RNA concentration in 5 μ L of reverse transcription reaction is 2.5 pg to 250 ng (336). For this study, a minimum concentration of 80 ng in 5 μ L was required. RIN values provide information on RNA integrity, as discussed in section 3.3.2.1. RIN values below 5 were excluded. Samples that did not reach the concentration and quality thresholds were excluded (n=11). For instance, sample P108042 has a very low RNA concentration as was excluded from future experiments (Figure 4-6 A).

As with urine, final concentration and 260/280 ratio were plotted (Figure 4-6 B). Density plots represent the distribution of RIN values in the INNOVATE and PROMIS cohorts (Figure 4-6 C). As two quantification methods were used (Qubit and nanodrop), comparison of final concentration values was performed (Figure 4-6 D).

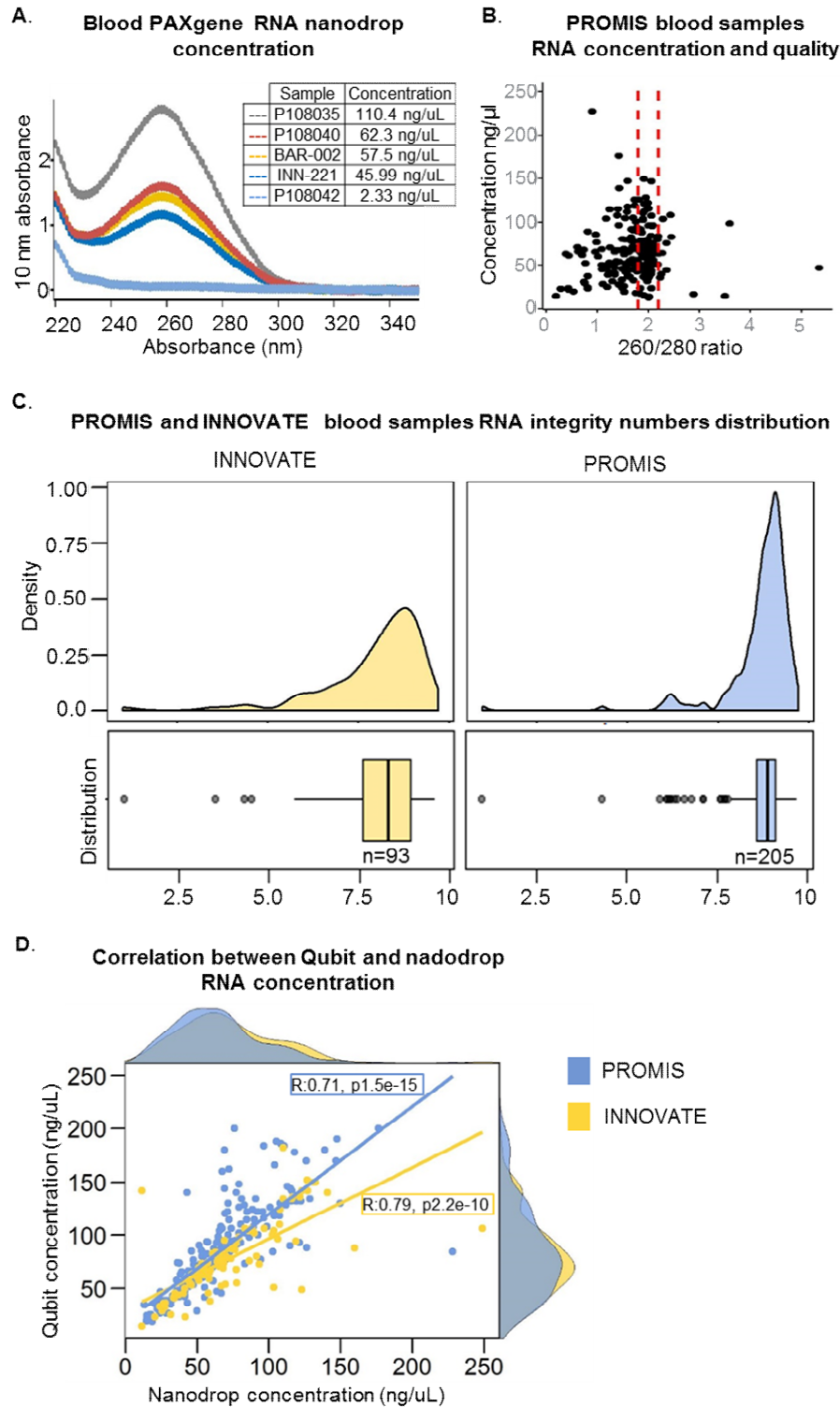


Figure 4-6 PROMIS and INNOVATE blood RNA samples quality control.

A. Optical absorbance of extracted blood PAXgene RNA for PROMIS samples P108035, P10804 and P108042 and INNOVATE samples INN-221 and BAR-002. Sample ID and nanodrop concentration on table. **B.** Dot plot comparing the RNA concentration and the 260/280 ratio from PROMIS samples; red dashed line represents limit of ideal 260/280 ratio below 1.8 and above 2.2. **C.** Density plot and boxplot representing the RIN number distribution amongst blood extracted mRNA samples. **D.** Scatterplot of nanodrop RNA concentration on the x-axis and Qubit RNA concentration on the y-axis. INNOVATE study represented in yellow, PROMIS study represented in blue. Correlation coefficient and p value in blue and yellow boxes, respectively. Density plots representing the distribution of concentration per trial.

4.3.2.2.2. Measuring the gene expression of all genes in the panel in PAXgene RNA using Fluidigm

Micro-droplet dqPCR was used to measure the panel gene expression in blood RNA from PAXgene blood (Section 2.2.4.4). Missing data analysis was performed as described in (Section 2.2.6.3). When examining all genes COL6A3, KLK3, PCA3, TMRPSS2 and AGR2 had >50% of missing data, in 47 patients all data was missing for these genes (Appendix 6A-B).

The relative gene expression data from the two panels (refined diagnostic and refined prognostic) was extracted to continue analysis. Missing data analysis was performed and plotted (See Figure 4-7). AGR2 had 78.8% of missing data; the high percentage of missing data for this gene was consistent for all technical replicates. For this reason, AGR2 was removed from future analysis, reducing the refined diagnostic panel from seven genes to six genes. All remaining genes had less than 6% of missing data and were deemed suitable for further analysis. The co-occurrence of missing data points was analysed using an UpSet plot (Figure 4-7-B), this confirmed that when AGR2 was not expressed it was in isolation (149 cases), with the exception of one patient, in whom SLC2536 was also not detected.

Data distribution was examined using Quantile-quantile plots (QQplot) per gene and Levene's test to determine the variance between benign and cancer samples (Appendix 7). Relative expression ($2^{-\Delta\Delta C_t}$) was normalised to prostate volume and results were plotted as boxplots comparing benign and cancer patients (Figure 4-8). The difference in relative expression was significant for all genes in the diagnostic signature but not on the prognostic signature (Figure 4-8 B).

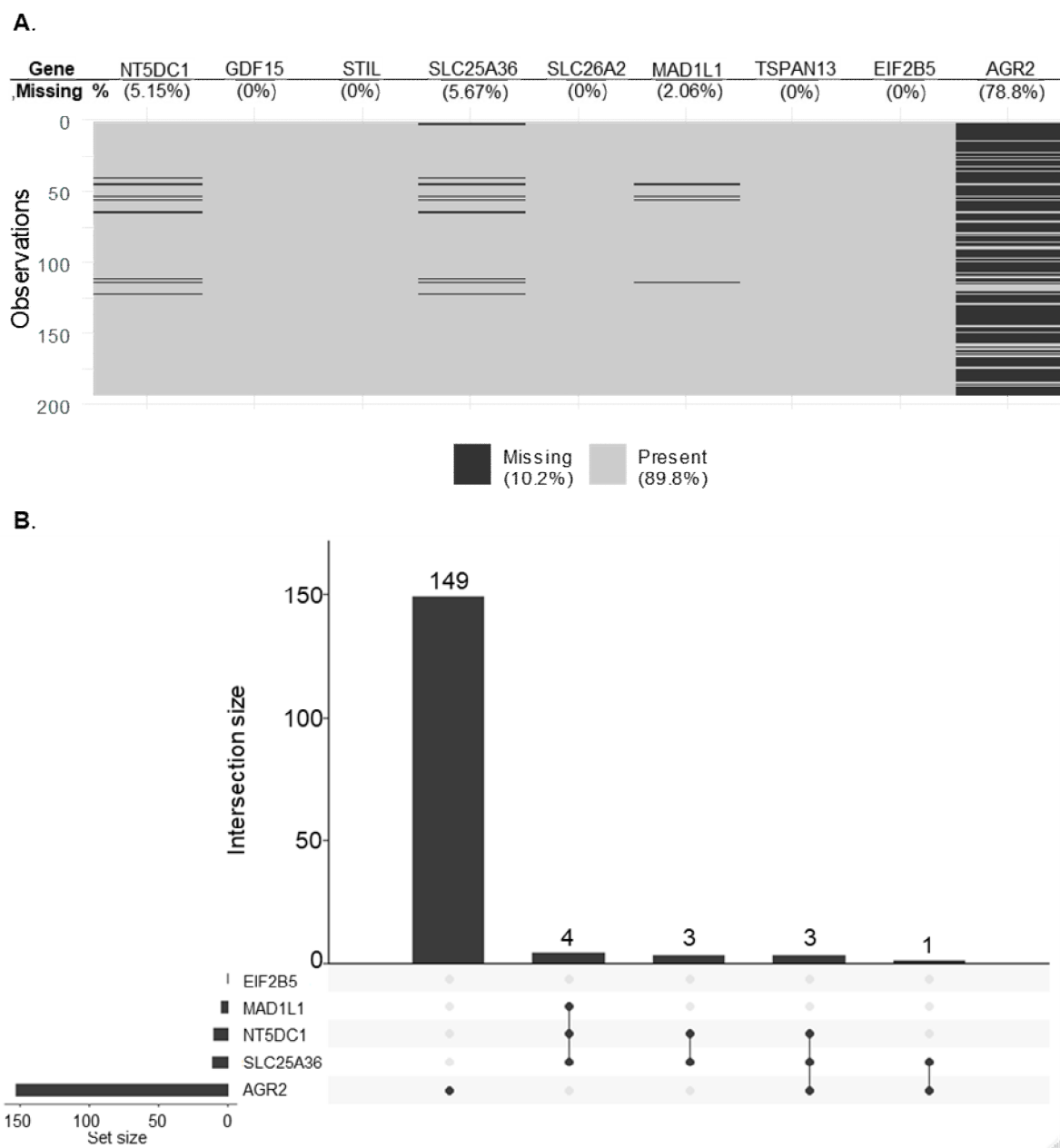


Figure 4-7 Missing data analysis of blood extracted mRNA from PAXgene blood.

A. Tile plot representing missing data analysis. Missing data is plotted in black, present data in grey. Observations are plotted in the y-axis and genes in the x-axis. Missing percentage per gene is depicted below each gene name. **B.** UpSet plot representing the patterns of missing data in EIF2B5, MAD1L1, NT5DC1, SLC25A36 and AGR2. The intersection size is plotted in the y-axis, the dots represents the parts where missing data intersects.

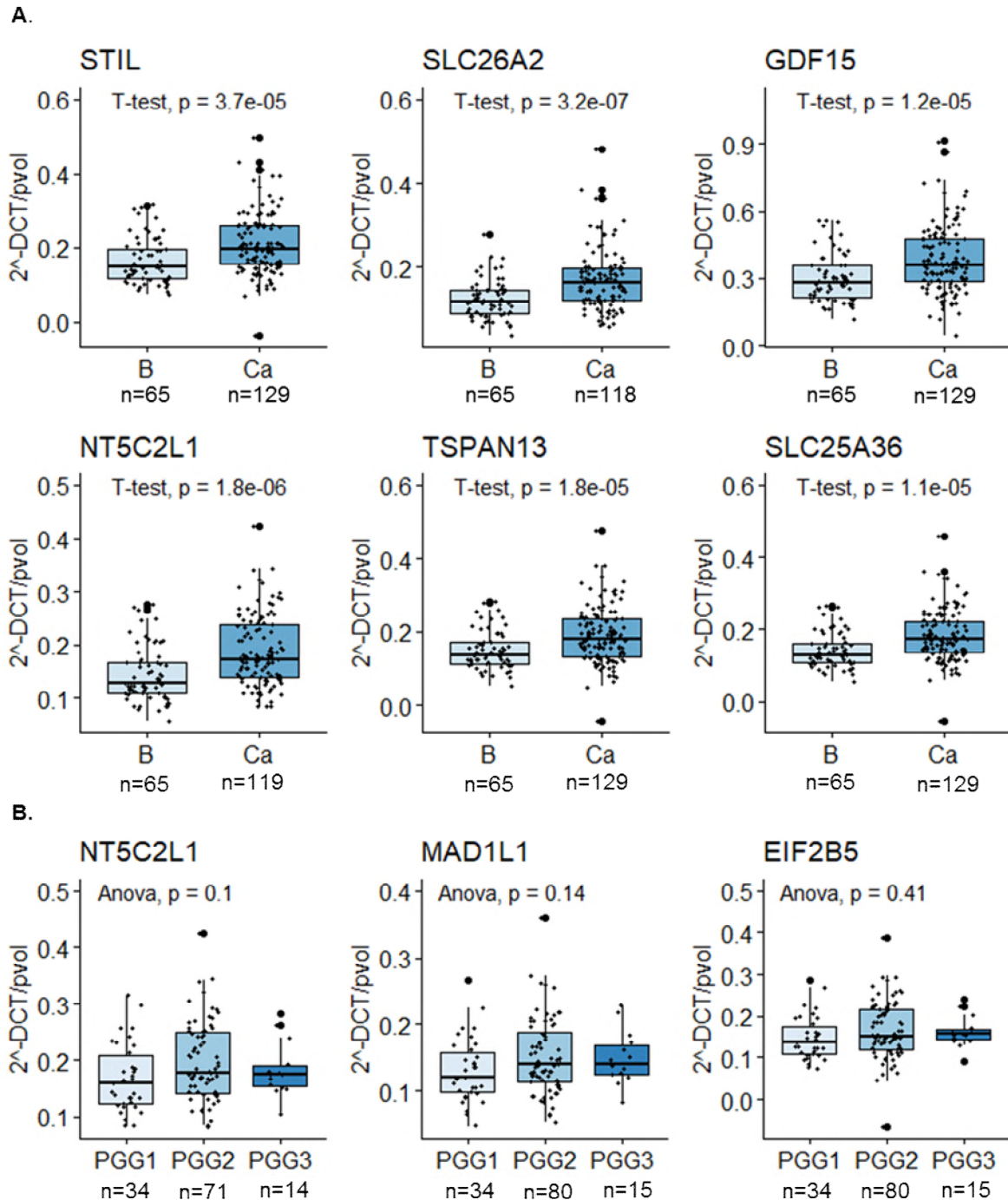


Figure 4-8 Relative gene expression of genes in PAXgene blood extracted RNA in diagnostic and prognostic signature.

A. Boxplots comparing benign and cancer samples in PAXgene blood RNA. Benign is plotted in light blue and cancer in dark blue. Normalised relative expression is plotted on y-axis (normalised by prostate volume). Benign and cancer plotted on x-axis. n = number of patients. Welch-Satterthwaite t-test was performed on GDF15 and TSPAN13. Student's t-test was performed in remaining genes. **B.** Boxplots comparing three Prostate Prognostic Groups (PGG1, PGG2 and PGG3). Normalised relative expression is plotted on y-axis (normalised by prostate volume). PGG groups are plotted on the x-axis.

4.3.3. Study of the accuracy of the diagnostic and prognostic panels.

4.3.3.1. Assessment of accuracy of biomarker panels for diagnosis and risk stratification of prostate cancer

In order to assess the applicability of the two gene panels as a combination of markers rather than individual levels of expression, the data was analysed in collaboration with the Rosenfeld lab, Jerusalem Institute of Technology, as detailed in methods section 2.2.6.5.

4.3.3.2. Assessment of diagnostic accuracy of the six-gene refined diagnostic gene panel in the PROMIS cohort for model development.

Confusion matrices were created to calculate the diagnostic accuracy estimates with 95% confidence intervals of the six-gene panel when evaluating:

1. All benign cases versus all cancer cases.
2. All benign cases versus cases meeting definition one in PROMIS (Gleason score $\geq 4+3$ or cancer core length ≥ 6 mm).
3. All benign cases versus cases meeting definition two in PROMIS (Gleason score $\geq 3+4$ or cancer core length ≥ 4 mm).

Cases that did not meet significance criteria described above were grouped with the no-cancer group. The corresponding tables are presented in Appendix 8. Area under the ROC curve plots were created for each parameter to represent diagnostic accuracy of PSA, PSA density, Likert and the six-gene panel (Figure 4-9 A). The combination of PSA_d and the six-gene panel is plotted on Figure 4-1 B. The AUC was used to determine the performance of the test to correctly classify those with and without the disease. These values were plotted in Figure 4-9 C, whilst the six-gene panel improved the diagnostic accuracy compared to PSA alone (0.68, 95% CI 0.65-0.70 compared to PSA AUC 0.64, 95% CI 0.56-0.73), PSA density performs better alone and in combination to mpMRI (AUC 0.76, 95% CI 0.69-0.83 alone and 0.85, 95% CI 0.83-0.87). A summary of the analysis of performance of the panel is depicted in Figure 4-9 D.

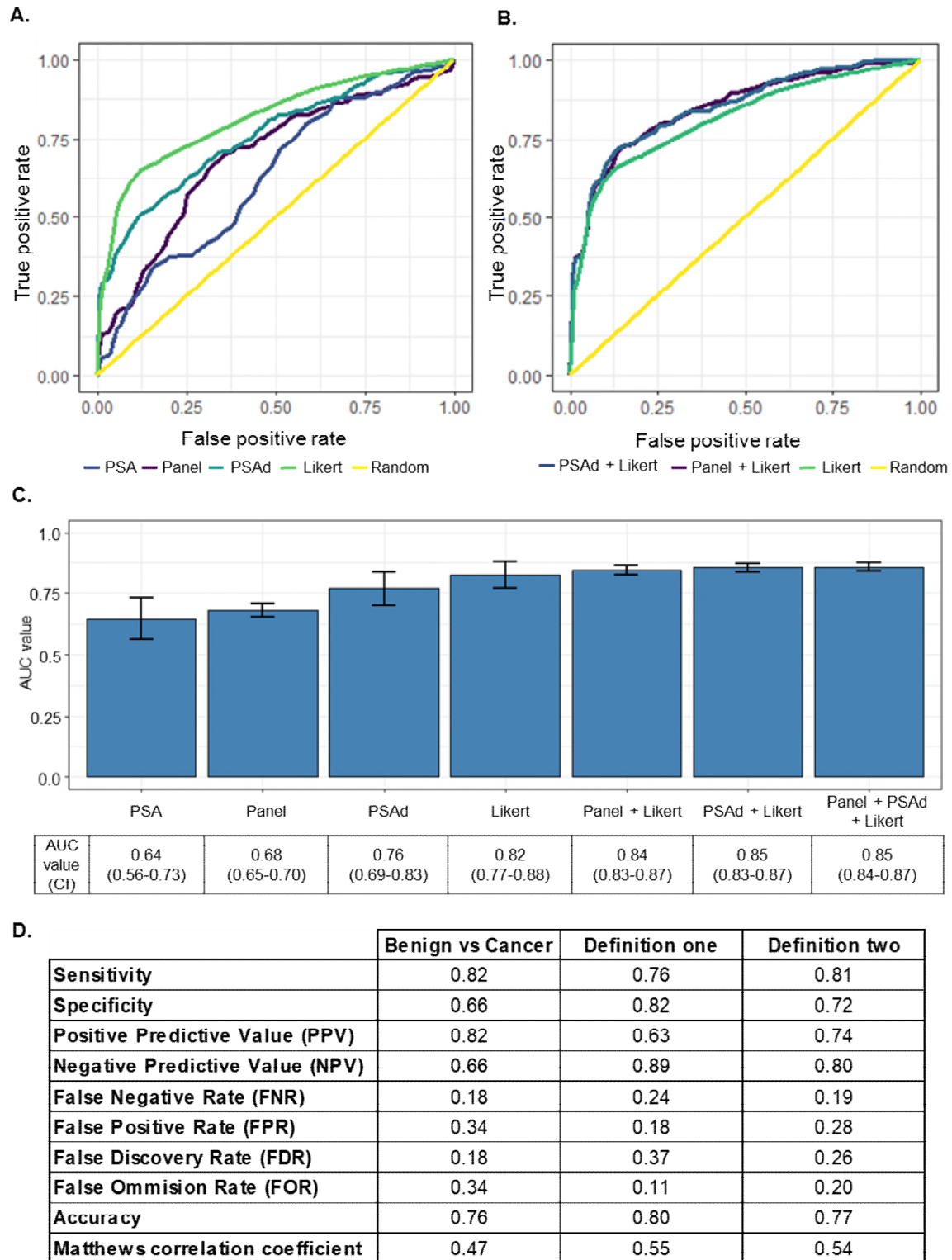


Figure 4-9 ROC curves reveal similar performance of the refined diagnostic six-gene panel compared to PSA density, but not superior.

A. ROC graph representing the true positive rate (TPR) against the false positive rate (FPR) for the diagnosis of prostate cancer (benign vs cancer). PSA in blue, six-gene panel in purple, PSA density (PSAd) in turquoise, Likert score in green and performance of a random test in yellow. The TPR is plotted on the y-axis, the FPR on the x-axis. **B.** ROC graph for Likert alone (turquoise) and in combination to PSAd (blue) and the six-gene panel (purple). X-axis and y-axis as before. **C.** Barplot with values and confidence intervals (CI) for each parameter as error bars. Area under the curve (AUC) is plotted on the y-axis, parameters on the x-axis. Table below the plot contains AUC value and CI for

each test and corresponding combinations. **D.** Table depicting performance analysis of the test in all three conditions (Benign vs cancer, definition one and two). For information on formulas used see Materials and Methods 2.2.5.2.6

4.3.3.3. Assessment of accuracy of the refined prognostic three-gene panel for risk stratification of prostate cancer

To determine the diagnostic accuracy of the refined prognostic three-gene panel for PGG classification, ROC plots and AUC were calculated. MCCL was not taken into account for this analysis, and so none of the previously described definitions were used in this analysis. Instead, the PGG was used as follows:

1. PGG1: Gleason 3+3
2. PGG2: Gleason 3+4
3. PGG3: Gleason 4+3
4. PGG4: Gleason sum 8
5. PGG5: Gleason 9 and 10

A combined plot for the refined prognostic panel was created and its ability to discriminate between PGG in the cohort. The AUC under the ROC was calculated. All results are plotted on Figure 4-10. The three-gene panel was unable to accurately identify the different PGG classifiers.

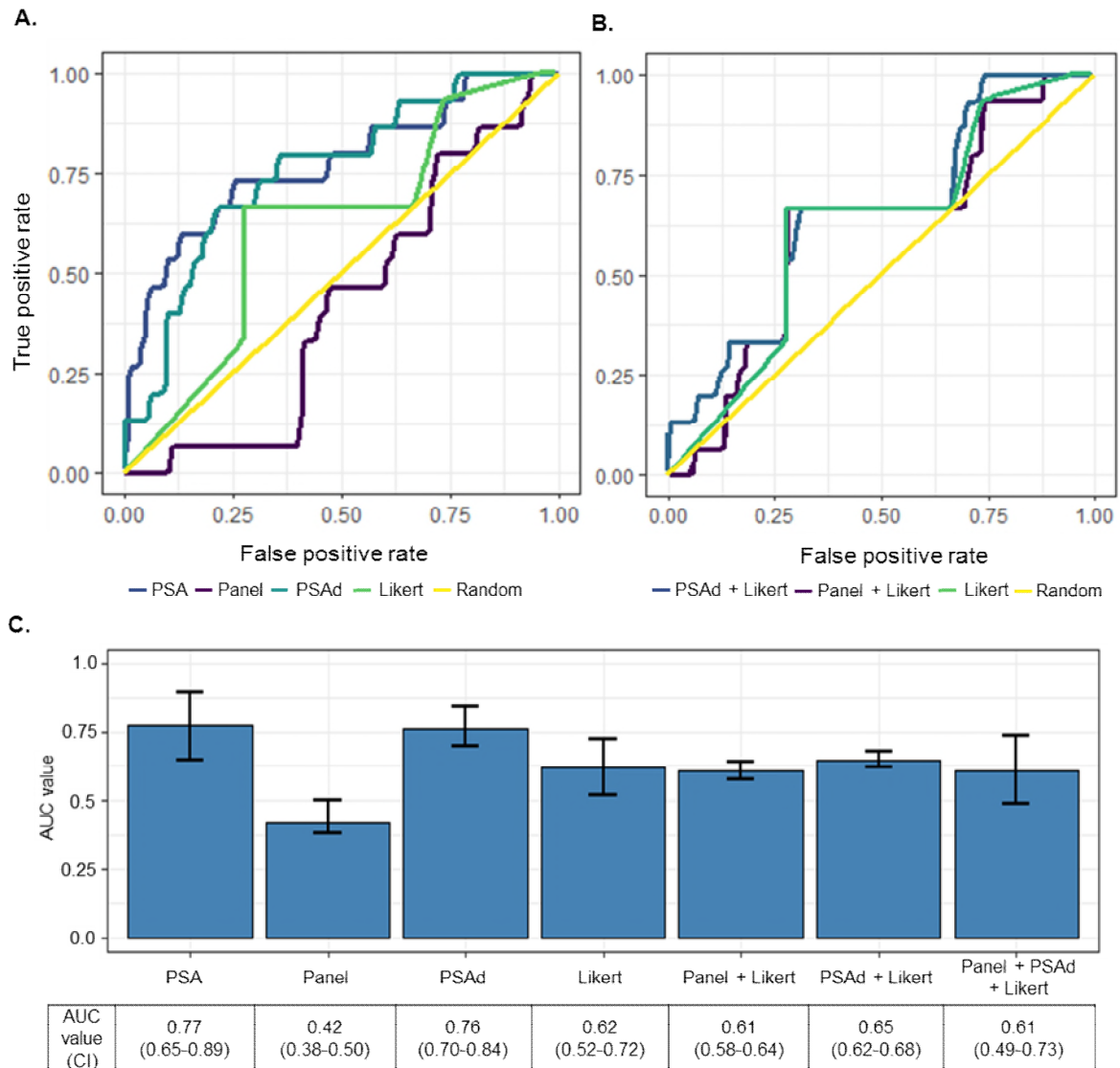


Figure 4-10 Refined prognostic three-gene panel fails to classify patients by PGG.

A. ROC graph representing the true positive rate (TPR) against the false positive rate (FPR) for the detection of PGG's. PSA in blue, three-gene panel in purple, PSA density (PSAd) in turquoise, Likert score in green and performance of a random test in yellow. The TPR is plotted on the y-axis, the FPR on the x-axis. **B.** ROC graph for Likert alone (green) and in combination to the three gene panel in purple, and PSAd (blue). X-axis and y-axis as before. **C.** Barplot with values and confidence intervals (CI) for each parameter as error bars. Area under the curve (AUC) is plotted on the y-axis, parameters on the x-axis. Table below the plot contains AUC value and CI for each test and corresponding combinations.

4.3.4. Panel validation in the INNOVATE cohort

To determine the diagnostic accuracy of the refined diagnostic six-gene panel for the diagnosis of prostate cancer in an independent cohort, the INNOVATE cohort was utilised. The logistic regression model developed with the PROMIS data (Section 4.4.1.1) was tested in INNOVATE, as a validation cohort. Data was analysed (Rosenfeld lab) and confusion matrices were created to calculate the diagnostic accuracy estimates of the six-gene panel. When evaluating the same parameters as in the PROMIS cohort:

1. All benign cases versus all cancer cases.
2. All benign cases versus cases meeting definition one in PROMIS (Gleason score $\geq 4+3$ or cancer core length ≥ 6 mm).
3. All benign cases versus cases meeting definition two in PROMIS (Gleason score $\geq 3+4$ or cancer core length ≥ 4 mm).

The corresponding tables are presented in Appendix 8. As previously, a combined ROC curve was created for each marker alone and in combination with imaging as seen on Figure 4-11 A.

The sensitivity of the panel was lower in the INNOVATE cohort compared to PROMIS, but the specificity was overall higher (Figure 4-11 D). The refined diagnostic six-gene panel was better at ruling in cancer, with a sensitivity of 45% and specificity of 88%, compared to 82% sensitivity and 66% specificity in the training set (PROMIS). The AUC was 0.65 (95% CI 0.61-0.69), only slightly lower than in PROMIS (AUC of 0.68, 95% CI 0.65-0.70), see Figure 4-10 D. The Matthews Correlation Coefficient (MCC), provides a measure of the quality of binary classifications; it yields a score of one only if the prediction obtained good results in all of the four confusion matrix (217) (see 2.2.6.5). A score of zero means the test is no better than random guess, whilst a score of one indicates perfect agreement between the test and the actual diagnosis. The panel highest MCC in the diagnostic setting was for the PROMIS (training) cohort, with a score of 0.47, lowering to 0.37 in the INNOVATE cohort.

Performance in the INNOVATE cohort of the panel for detection of definition one (Gleason $\geq 4+3$, MCCL ≥ 6 mm) had a similar result. There was a lower sensitivity of 48% (76% in PROMIS) and specificity of 88% (82% in PROMIS). The MCC was also lower in the INNOVATE cohort (0.38) compared to PROMIS (0.55).

For definition two (Gleason $\geq 3+4$, MCCL ≥ 4 mm), the panel had a 33% sensitivity and better specificity of 84% (81% and 72%, respectively in PROMIS). However, the MCC was closer to 0 with a score of 0.19 compared to 0.54 in PROMIS For all values see tables in Figure 4-9 D and Figure 4-11 D.

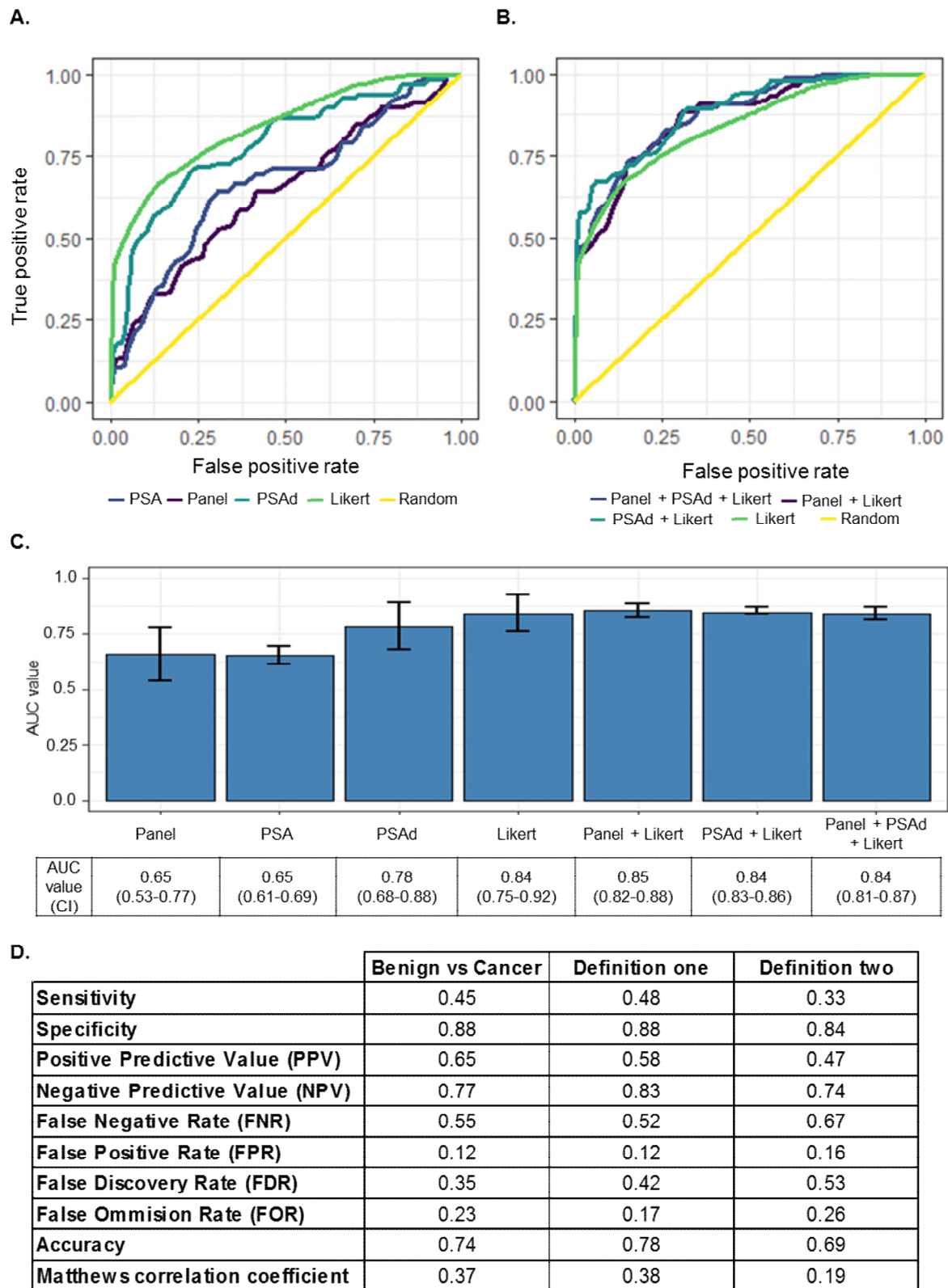


Figure 4-11 ROC curves reveal similar performance of the refined diagnostic six-gene panel in the INNOVATE cohort, but is not superior to PSAd.

A. ROC graph representing the true positive rate (TPR) against the false positive rate (FPR) for the diagnosis of prostate cancer (benign vs cancer). PSA in blue, six-gene panel in purple, PSA density (PSAd) in turquoise, Likert score in green and performance of a random test in yellow. The TPR is plotted on the y-axis, the FPR on the x-axis. **B.** ROC graph for Likert alone (turquoise) and in combination to PSAd (blue) and the six-gene panel (purple). X-axis and y-axis as before. **C.** Barplot

with values and confidence intervals (CI) for each parameter as error bars. Area under the curve (AUC) is plotted on the y-axis, parameters on the x-axis. Table below the plot contains AUC value and CI for each test and corresponding combinations. **D.** Table depicting performance analysis of the test in all three conditions (benign vs cancer, definition one and two).

4.3.5. Defining the best performing qualifiers for diagnosis and risk stratification of prostate cancer.

As seen in the previous sections, the panel performs well but not necessarily well enough to replace current clinical diagnostic and prognostic parameters. In order to understand which were the best predictors for diagnosis and diagnosis of clinically significant disease using a liquid biopsy; logistic models predicting the three separate outputs (benign vs cancer, definition one and definition two) were created, based on the PROMIS data (training set). Internal validation was performed on all combinations with predictive power set at 95% sensitivity. The best performing model for each output was selected for prediction of the three outcomes in the INNOVATE cohort. For the diagnosis of prostate cancer (cancer vs benign) the best performing model was the six-gene panel + PSA density + Likert. PSA density + Likert, performed best for definition one and two. For each patient the model predicted the presence (1) or absence (0) of cancer, Gleason 4+3 or ≥ 6 mm (definition one) or Gleason 3+4 or ≥ 4 mm (definition two). The performance of the model is represented on Figure 4-12. Both sensitivity and specificity are high when PSA density and Likert are used to find clinically significant cancer in both definitions, but specially in definition two (sensitivity 81%, specificity 76%, NPV: 90%).

| | <i>Panel + PSAd + Likert</i> | <i>PSAd + Likert</i> | <i>PSAd + Likert</i> |
|---|----------------------------------|-----------------------|-----------------------|
| | Benign vs Cancer | Definition one | Definition two |
| Sensitivity | 0.72 | 0.79 | 0.81 |
| Specificity | 0.82 | 0.78 | 0.76 |
| Positive Predictive Value (PPV) | 0.66 | 0.58 | 0.59 |
| Negative Predictive Value (NPV) | 0.86 | 0.91 | 0.90 |
| False Negative Rate (FNR) | 0.28 | 0.21 | 0.19 |
| False Positive Rate (FPR) | 0.18 | 0.22 | 0.24 |
| False Discovery Rate (FDR) | 0.34 | 0.42 | 0.41 |
| False Ommision Rate (FOR) | 0.14 | 0.09 | 0.10 |
| Accuracy | 0.79 | 0.79 | 0.78 |
| Matthews correlation coefficient | 0.53 | 0.53 | 0.53 |

Figure 4-12 PSA density (PSAd) and Likert have high sensitivity and specificity for the diagnosis of significant prostate cancer.

Table depicting performance analysis of the test in all three conditions (Benign vs cancer, definition one and two) using the best performing models following logistic models with 95% sensitivity.

4.4. Discussion

4.4.1. Summary of findings

The first aim of this chapter was to evaluate the potential utility of urine pellet mRNA to measure gene expression. The quality and concentration of the RNA extracted from urine made analysis difficult (Figure 4-1, Figure 4-2). For instance, nanodrop plots confirmed contamination probably due to the extraction method (low 260/280 ratio). Less contaminated samples (higher 260/280 ratio) had a higher concentration, but this may be secondary to carrier RNA added during the extraction procedure (see section 2.2.2.3). Additionally, all RNA was highly degraded with RIN numbers below 2.5 and more than 50% of patients with RIN values below 1.5. Based on findings described in the previous chapter (section 3.3.2.2.3), the urine RNA was pre-amplified, unfortunately the resulting relative expression profile showed the pre-amplification was biased (Figure 4-1 D-F). PCR was performed in unamplified, degraded urinary RNA.

Analysis of missing data revealed that most of the absent data belonged to men with no cancer; following multiple random imputation, the relative expression boxplots in Figure 4-4 showed a statistically significant difference in genes AGR2, TMPRSS2, AR, ECI2, PCA3 and VPS13A. However, several limitations to this work exist. First, the number of available data points in the benign group for each biomarker varied and thus it is not possible to perform meaningful comparisons or group data for accuracy analysis. Second, this cohort consists of a small group of patients with different disease stages, making study of these biomarkers in urine as a diagnostic tool unbalanced to more aggressive cancer. Last, it is not possible to confirm if men in the benign group were indeed benign, men in this benign group included those with inflammatory and benign enlargement of the prostate, and most patients were eventually lost to follow up or discharged from clinic. Since most of these had only TRUS biopsies, and only a minority had mpMRI guided transperineal biopsies, it is not possible to ascertain if any of these men had indolent disease that was not found due to sampling error, or if these men indeed represent a true benign group.

The second aim of this chapter was to determine if the refined diagnostic and prognostic panels could be detected in blood, and to study their ability to act as a diagnostic tool for the diagnosis and risk stratification of prostate cancer. Using PAXgene blood the RNA quality was markedly improved; in addition, the amount of RNA extracted was higher than that of urine without need for carrier RNA (Figure 4-6). For instance, the RIN values are above 5 for both studies, only 6 patients had RIN values below 5 and all were excluded from downstream analysis. This is in line with published studies (337,338). The missing data analysis revealed that AGR2 was missing in 78.8% of cases (Figure 4-7), for this reason, this gene was removed from further analysis. All normalised relative expression values were normalised to prostate volume as with PSA density (339).

Data from the resulting refined diagnostic six-gene, and refined prognostic three-gene panels were analysed to obtain accuracy measurements of the performance of both panels. Whilst the refined diagnostic six-gene panel was able to differentiate between benign and cancer patients, this ability was only slightly better than PSA and not better than PSA density (Figure 4-10, Figure 4-11). The prognostic

panel was unable to differentiate PGG groups (Figure 4-10). While the panel and MRI had the highest AUC values, Likert alone performed better than the panel alone and the benefit is minimal to justify using this kind of genetic testing in the clinic.

With mpMRI emerging as a tool for diagnosis and risk stratification, a genetic panel with the potential to enhance its accuracy would be helpful. Unfortunately, neither panel offers enough improvement to the AUC of mpMRI to suggest such use. In fact, the PSA density performs better than any of the panels whilst increasing the ability of mpMRI to accurately detect prostate cancer.

4.4.2. Methodological limitations

The urinary mRNA had a degree of phenol contamination (Figure 4-1 B) and low quality (seen by RIN values below 2.5 in Figure 4-1 C), this could be secondary to the extraction method used (Acid guanidinium thiocyanate-phenol-chloroform extraction). Additionally, the RNA was degraded due to exposure to urine, as explained on Chapter 3 (3.3.2.1). The large amount of missing data from the urine cohort was another limitation; this can be due to one of three reasons: no gene expression in the sample, expression is too low and is undetected or the primer used is faulty. The latter can be ruled out as these genes were successfully measured in cell lines, and in blood PAXgene mRNA. It is not possible to determine if the gene expression was too low to be detected, or if no expression exists in these samples. It is likely that with dilution in bodily fluids, even if the prostate cancer expressed this gene it would not be detectable.

Analysis of data obtained from PAXgene RNA in the ProMPT cohort, highlighted the need for quality control of each PCR run. As seen in Figure 4-5, RPLP2 variation between disease stages disqualifies the housekeeper to act as a normaliser. The experience of reviewing previously acquired data also imparted the importance of record keeping, as this analysis would not have been possible if records were lost.

Regarding the cohorts, the INNOVATE cohort characteristics reflect the current NICE guidelines and clinical practice at UCHL, where negative mpMRI patients can undergo follow up at their local GP practice. Due to the nature of the trial, it is not possible to fully ascertain that men with negative biopsies or mpMRI are true negatives. Additionally, these men only had a targeted biopsy, compared to full mapping biopsy done in PROMIS. This may account for the lower number of men with cancer in this cohort.

Regarding the PGG classification, the three-gene panel failed to accurately separate the groups. The lack of PGG4 and PGG5 patients in this particular cohort, limits the analysis of the panel's ability to stratify risk. When examining the tissue data from Chapter 3, difference in expression was seen when comparing PGG5 to lower risk PGG groups (see Appendix 3 and 4). The lack of these two groups could explain the failure of the signature. Interestingly, when inputting all 28 genes the AUC was similar to PSA density (Figure 4-10), however caution must be taken as this may reflect overfitting of the model (340,341). This improved AUC could reflect a better performance by genes that were not included in the final signature; future work could be done to study this effect.

4.4.3. Clinical application

The search for gene based biomarker panels in fluidic samples has so far had mixed results. The application of biomarker panels, like the one presented in this chapter, is expensive and time consuming. The comparison with available biomarkers such as PSA density, mpMRI and the combination of the two, showed that the performance of these for diagnosis of prostate cancer, and discrimination between the two definitions from PROMIS for clinical significance of cancer, was vastly superior to PSA alone and the genetic panel. Having said that, the panel had a good performance, and added some information to both these markers, but the use of these panels cannot be recommended beyond research practice.

There is increasing interest in finding new biomarkers that can be detected in blood, urine and other biological fluids. However, extrapolating tissue-based markers into biological fluids has so far failed to find markers that outperform cost-efficient alternatives. Based on these findings, and the increasing interest in finding minimally invasive tests that find aggressive cancer (mainly dominating Gleason 4 disease), chapter 5 focuses on attempting to understand the distribution of Gleason 4 in a small cohort of men in the PROMIS trial.

4.5. Lay summary

In chapter 3, two gene panels were created using publicly available data from the initial 28 genes. One panel for detecting prostate cancer and recommending a biopsy (diagnostic panel), and one to detect men with more aggressive disease (prognostic panel). This chapter focuses on the measurement of the two panels in blood and urine samples from patients that took part in three studies:

1. Cambridge ProMPT cohort (Prostate cancer: Mechanisms of progression and treatment), a biobank of samples of men with and without prostate cancer.
2. PROMIS (Prostate MR imaging study), a study comparing the diagnostic ability of the transrectal prostate biopsy to MRI of the prostate.
3. INNOVATE study (Combining advances in imaging with biomarkers for improved diagnosis of Aggressive prostate cancer), a study looking to combine blood and urine tests in combination to a new MRI technique.

The panels were evaluated as a group using the area under the curve measurement and, calculating parameters to measure the performance of the panels. For example, sensitivity to see how many men with the disease had a positive test and specificity to see how many men without the disease had a negative test, amongst others.

4.5.1. Objectives

1. Evaluate the quantity and quality of the RNA obtained from patient urine.
2. Measure the relative gene expression of the diagnostic and prognostic panels in urine samples from patients from the ProMPT study.
3. Evaluate quantity and quality of the RNA obtained from patient blood (ProMPT, PROMIS and INNOVATE studies)
4. Measure the relative gene expression of the diagnostic and prognostic panels in blood samples from patients from the PROMIS and INNOVATE studies.
5. Understand the diagnostic accuracy of the gene panel in two independent clinical cohorts (PROMIS and INNOVATE).
6. Understand the utility of the gene panel for use as a prognostic tool in two independent cohorts (PROMIS and INNOVATE).

4.5.2. Findings

1. The urinary RNA is severely degraded, the pre-amplification step was not successful as it “multiplied” genes unevenly, biasing the results.
2. Relative gene expression of the degraded urine RNA had lots of missing data.

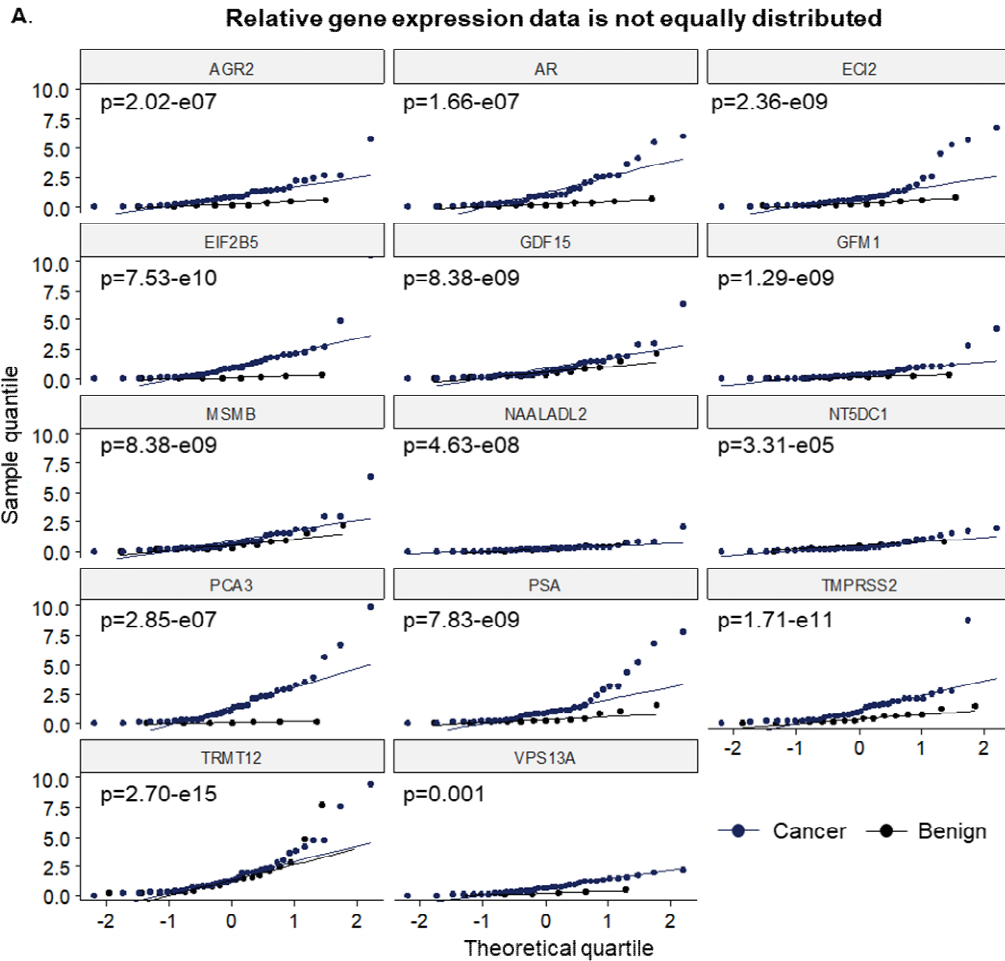
Take home message: Urinary RNA can be used to measure gene expression of these genes; however, steps must be made to protect the RNA after collection of the urine. The results obtained here showed some genes were higher in cancer compared to no cancer, however the data is limited and no meaningful conclusions can be made.

3. RNA quality from blood is adequate for gene expression studies.
4. The diagnostic panel performed very well in the PROMIS cohort. This was validated in the INNOVATE cohort with similar results. However, when compared to a commonly used biomarker (PSA density) the performance was not significantly better.
5. The risk stratification marker (to find aggressive disease) was not able to group patients according to the disease aggressiveness.

Take home message: The diagnostic panel performed well, especially in combination with MRI to select which patient would benefit from a biopsy. However, the performance was similar to that of PSA density, a biomarker already used in clinical practice.

4.5.3. Limitations

1. The urine had no preservative to avoid RNA degradation leading to a high proportion of missing data.
2. The studies used for validation lack long term follow up, this limits the prognostic value of the signature as the findings cannot be linked to outcomes such as death or recurrence after death.



B. Levene's test results per gene

| Gene | Significance | T-test |
|----------|--------------|--------|
| AGR2 | 0.023 | Welch |
| AR | 0.001 | Welch |
| ECI2 | 0.015 | Welch |
| EIF2B5 | 0.069 | T test |
| GDF15 | 0.216 | T test |
| GFM1 | 0.08 | T test |
| MSMB | 0.216 | T test |
| NAALADL2 | 0.86 | T test |
| NT5DC1 | 0.188 | T test |
| PCA3 | 0.013 | Welch |
| PSA | 0.011 | Welch |
| TMPRSS2 | 0.058 | T test |
| TRMT12 | 0.026 | T test |
| VPS13A | 0.01 | Welch |

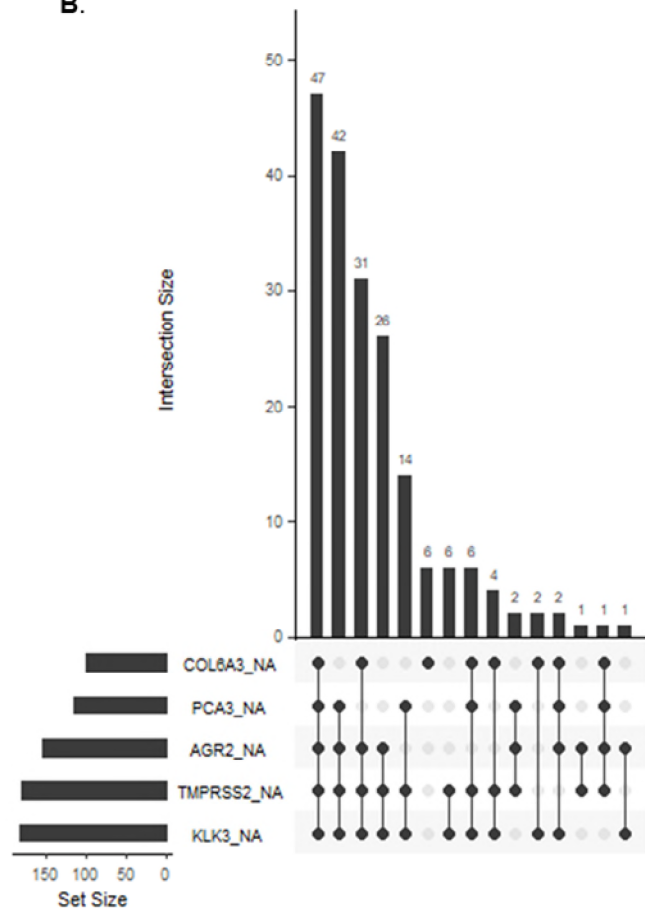
Appendix 5. Urine PCR relative gene expression data is not normally distributed and has unequal variance distribution.

A. quantile-quantile plots per gene and separated by condition, the x-axis represents the theoretical quantiles and the y-axis the sample quantiles. Shapiro-Wilk test p value is presented inside each plot ($p=$). Cancer samples are plotted in dark blue, benign samples plotted in black. **B.** Table representing the Levene's test results per gene. Eight genes have equal variance distribution, whilst six genes have unequal variance. Welch t-test was performed if variance was unequal (significance ≤ 0.05), t-test was performed if variance was equal (significance ≥ 0.05).

A.

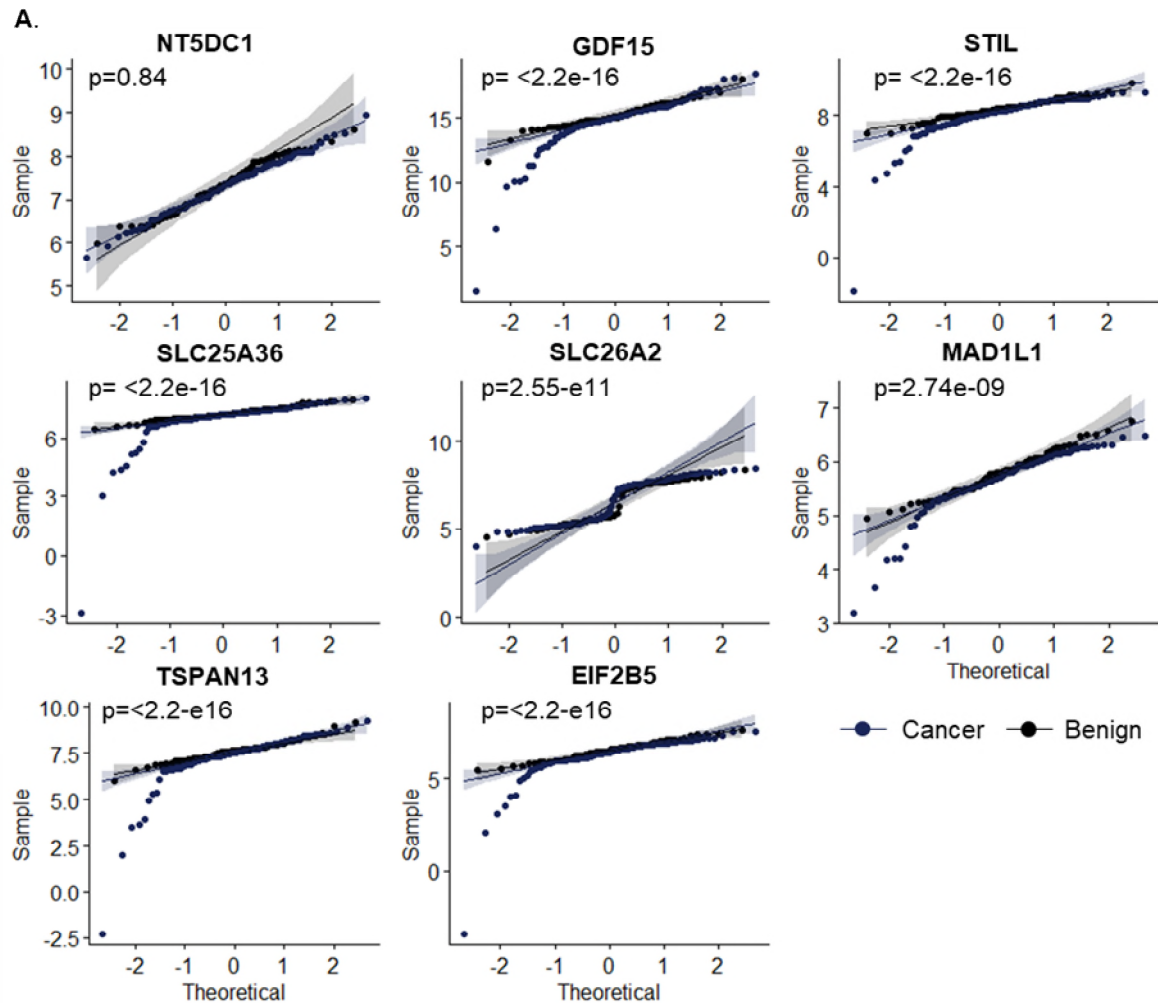
| Gene | Missing data (n) | Percentage (%) |
|----------|------------------|----------------|
| ABHD12 | 2 | 1.03 |
| AGR2 | 155 | 79.90 |
| AR | 13 | 6.70 |
| COL6A3 | 101 | 52.06 |
| EIF2B5 | 2 | 1.03 |
| GDF15 | 2 | 1.03 |
| GFM1 | 13 | 6.70 |
| GTF2H4 | 13 | 6.70 |
| KLK3 | 183 | 94.33 |
| MAD1L1 | 6 | 3.09 |
| MSMB | 33 | 17.01 |
| NAALADL2 | 14 | 7.22 |
| NDUFB11 | 12 | 6.19 |
| NT5C2L1 | 12 | 6.19 |
| NT5DC3 | 13 | 6.70 |
| PCA3 | 116 | 59.79 |
| PECI | 2 | 1.03 |
| SLC25A36 | 2 | 1.03 |
| SLC26A2 | 13 | 6.70 |
| STIL | 2 | 1.03 |
| TM4SF1 | 6 | 3.09 |
| TMPRSS2 | 182 | 93.81 |
| TNFSF10 | 2 | 1.03 |
| TRMT12 | 12 | 6.19 |
| TSPAN13 | 2 | 1.03 |

B.



Appendix 6. Missing data analysis of PAXgene blood extracted mRNA PCR in Fluidigm.

A. Table depicting missing data of Fluidigm analysed PCR mRNA, First column contains gene names, followed by number of missing values and percentage of missing data. **B.** UpSet plot representing the patterns of missing data for COL6A3, PAC3, AGR2, TMPRSS2 and KLK3 with an NA suffix. The intersection size is plotted in the y-axis, the dots represent the parts where missing data intersects.



B.

| Gene | Significance | T-test |
|----------|--------------|--------|
| NT5DC1 | 0.48 | T-test |
| GDF15 | 0.048 | Welch |
| STIL | 0.09 | T-test |
| SLC25A36 | 0.10 | T-test |
| SLC26A2 | 0.40 | T-test |
| MAD1L1 | 0.49 | T-test |
| TSPAN13 | 0.05 | Welch |
| EIF2B5 | 0.18 | T-test |

Appendix 7. Blood PCR relative gene expression data is not normally distributed and has unequal variance distribution.

A. quantile-quantile plots per gene and separated by condition, the x-axis represents the theoretical quantiles and the y-axis the sample quantiles. Shapiro-Wilk test p value is presented inside each plot ($p=$). Cancer samples are plotted in dark blue, benign samples plotted in black. **B.** Table representing the Levene's test results per gene. Two genes have equal variance distribution, whilst six genes have unequal variance. Welch t-test was performed if variance was unequal (significance ≤ 0.05), t-test was performed if variance was equal (significance ≥ 0.05).

A. Benign vs cancer

| | | Disease | | SUM |
|------|---------|---------|-------|-----|
| | | Ca (1) | B (0) | |
| Test | Pos (1) | 98 | 22 | 120 |
| | Neg (0) | 22 | 42 | 64 |
| SUM | | 120 | 64 | 184 |

B. Definition one

| | | Disease | | SUM |
|------|---------|---------|--------|-----|
| | | Yes (1) | No (0) | |
| Test | Pos (1) | 41 | 24 | 65 |
| | Neg (0) | 13 | 106 | 119 |
| SUM | | 54 | 130 | 184 |

C. Definition two

| | | Disease | | SUM |
|------|---------|---------|--------|-----|
| | | Yes (1) | No (0) | |
| Test | Pos (1) | 74 | 26 | 100 |
| | Neg (0) | 17 | 67 | 84 |
| SUM | | 91 | 93 | 184 |

E. Benign vs cancer

| | | Disease | | SUM |
|------|---------|---------|-------|-----|
| | | Ca (1) | B (0) | |
| Test | Pos (1) | 13 | 7 | 20 |
| | Neg (0) | 16 | 53 | 69 |
| SUM | | 29 | 60 | 89 |

F. Definition one

| | | Disease | | SUM |
|------|---------|---------|--------|-----|
| | | Yes (1) | No (0) | |
| Test | Pos (1) | 11 | 8 | 19 |
| | Neg (0) | 12 | 58 | 70 |
| SUM | | 23 | 66 | 89 |

G. Definition two

| | | Disease | | SUM |
|------|---------|---------|--------|-----|
| | | Yes (1) | No (0) | |
| Test | Pos (1) | 9 | 10 | 19 |
| | Neg (0) | 18 | 52 | 70 |
| SUM | | 27 | 62 | 89 |

Appendix 8. Confusion matrices for the six-gene panel for the diagnosis of prostate cancer and PROMIS definitions of significant disease in the PROMIS and INNOVATE cohorts.

A. Comparison between benign versus cancer includes all patients in the PROMIS cohort. **B.** Comparison between benign cases and men who meet criteria for definition one of PROMIS and **C.** Definition two. **D, E.** and **F** correspond to the same parameters for the INNOVATE cohort. For all of the above tables, test results are plotted as positive (1) or negative (0) on the horizontal axis. Histological diagnosis is plotted below the top tab labelled "Disease" as Cancer (1) and Benign (0). All positive and negative cases are added on the horizontal and vertical axes to obtain values of total diseased, healthy, positive and negative.

Chapter 5. Critical Evaluation Of Visual Proportion of Gleason 4 and Maximum Cancer Core Length

5.1.Introduction

In 2013 Pierorazio *et al.*, retrospectively reviewed 7850 radical prostatectomy specimens to investigate the short-term biochemical outcome using a prognostic-based scoring system called the Prostate Grading Group (PGG). By separating the Gleason sum 7 group into 3+4 and 4+3, the authors found that men with 4+3 had worse outcome defined as biochemical recurrence-free survival (8). These findings were further validated and were subsequently endorsed by the 2014 International Society of Urological Pathology Consensus Conference and the World Health Organization (WHO) (91,342,343). Prior to this, the blanket term of 'Gleason 7' cast a "grey area" connotation to Gleason 3+4 and 4+3; patients in this risk group were considered higher risk, and so not necessarily suitable for active surveillance (344,345).

The 2014 and 2019 ISUP recommendation calls for improved categorisation of the percentage of G4 (%G4) in prostate cancer to allow for better risk stratification and inform treatment decisions (9,13,346–348). The distinction between Gleason 3+4 (PGG2) and 4+3 (PGG3) is made when %G4 falls below or above 50%, respectively, as visually estimated by uropathologists (343). Additionally, the maximum amount of cancer in any core (maximum cancer core length, MCCL) has been used as a proxy for tumour volume estimation and can be used to define clinical significance (349,350).

Most histological prostate cancer burden studies have been performed in radical prostatectomy specimens or on men who have undergone transrectal systematic biopsies. The Prostate MR Imaging Study (PROMIS) includes men who are biopsy naïve whose prostates were systematically sampled every 5mm providing a unique opportunity to perform an in-depth pathologist-driven annotation and digital analysis of the pathological slides and compare this to the visually-reported %G4 and MCCL (47).

In this chapter, quantitative measurements more commonly used in the scientific field of geography were applied to prostate biopsy data, in an attempt to understand the distribution of Gleason 4 across the prostate, and how it may relate to mpMRI visibility and the index lesion. In collaboration with the Geography department at UCL, measurements of clustering and evenness were applied to the contoured digital images (Section 2.2.5.3). In geography, the study of spatial segregation, as applied to minorities and ghettos, has been used to understand racial differences to access to healthcare. Mehra *et al.* performed a review and meta-analysis studying racial segregation and adverse birth outcomes. They applied four formal measurements of segregation described in Massey and Denton's framework. These include: exposure, evenness, clustering, concentration, centralization and hyper-segregation (Table 5.1) (200,201).

For the purposes of this chapter, we will concentrate on evenness and clustering. These were considered the best two measurements to apply to the Gleason 4 data, as they measure the distribution

of Gleason 4 taking into account the region as a whole. Drawing a comparison of Gleason 4 forming “neighbourhoods” surrounded by Gleason 3, the evenness measures the proportion of Gleason 4 versus 3; an even population in this case would be formed of 50% Gleason 3 and 50% Gleason 4.

Clustering takes into account the Gleason 4 clusters (or neighbourhoods), and studies how they are grouped in relation to the Gleason 3; in other words, it determines if Gleason 4 neighbourhoods are more or less likely to be surrounded by Gleason 3 or Gleason 4, forming ghettos. The remaining measurements described in Table 5.1, these parameters relate to social exposure, proportion occupied by the minority in relation to the region and a combination of the four measurements if segregation is identified.

| | |
|--------------------------|--|
| Exposure | Degree to which minority members are exposed to minority or majority members in the same neighbourhood. |
| Evenness | Degree to which each neighbourhood has the same proportion of minority and majority members as the region and a whole. |
| Clustering | Degree to which minority neighbourhoods are contiguous and tightly clustered. |
| Concentration | Degree to which minority members occupy a small proportion on the total area as a region |
| Centralization | Degree to which a minority group is centrally located |
| Hyper-segregation | High levels of segregation on at least four aforementioned local measures |

Table 5.1 Formal measures of segregation as described by Massey and Denton.

Adapted from Mehra *et al.* Figure 2.

Graphically represented, if Gleason 4 was the “minority” (plotted in blue in Figure 5-1), and Gleason 3 was the majority (plotted in white), an even population would have a similar number of blue and white dots, spread at even intervals. The population plotted on A is less even than the population plotted on B, as the blue dots are more evenly distributed amongst the entire population. The dots on C are more clustered than those on D as they are more tightly packed, forming “ghettos”, plot D also has these clusters, however they are more spread and have fewer blue dots forming them.

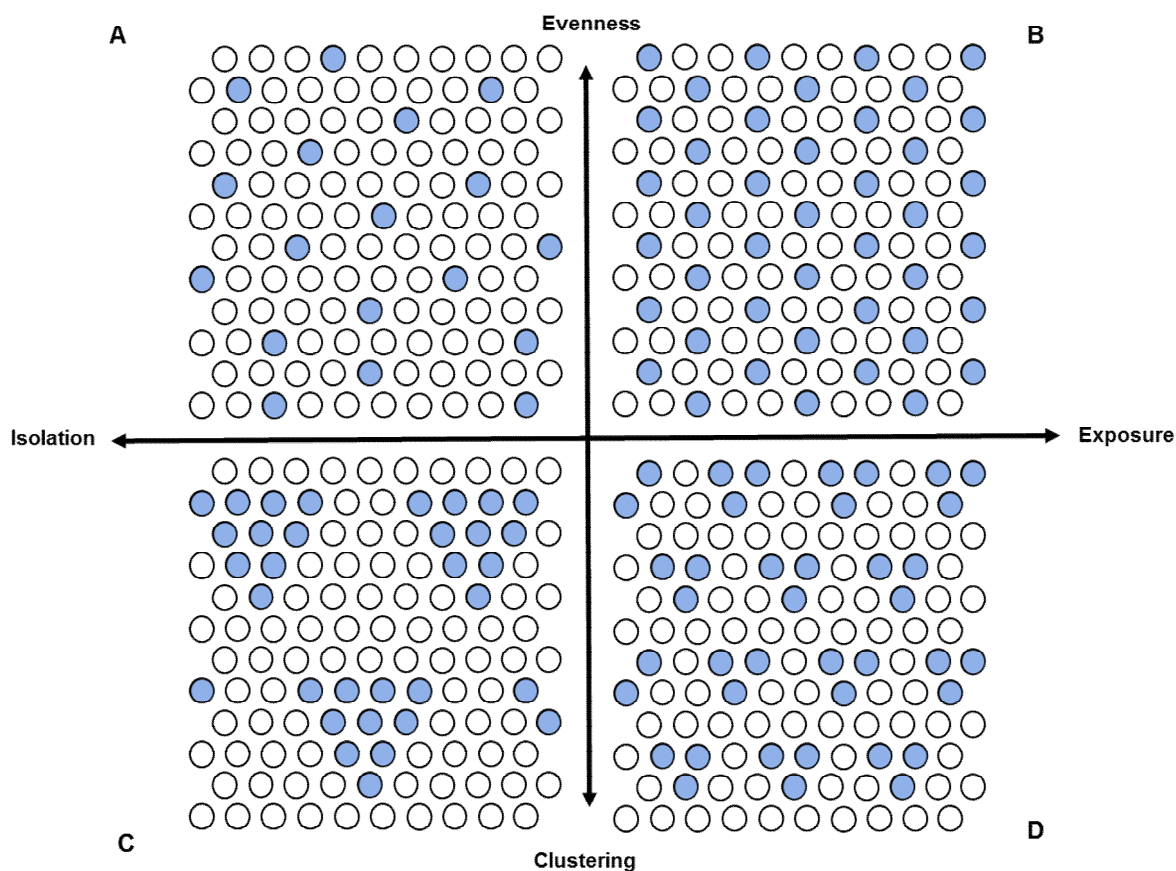


Figure 5-1 Visual representation of the measurements of spatial segregation.

Evenness plotted at the top, clustering at the bottom. Isolation and exposure plotted on the left and right, respectively. Blue dots correspond to the “minority” Gleason 4 and white dots to the “majority” Gleason 3. Adapted from Reardon & O’Sullivan, “Measurements of spatial segregation, Figure 1.

5.2. Aims and objectives

5.2.1. Hypothesis 1:

The digital measurement of MCCL differs from the visually obtained measurement.

5.2.1.1. Aim 1: Measure the MCCL using digital images of H&E prostate core biopsies and compare these to the measurement obtained using the microscope ruler.

5.2.2. Hypothesis 2:

Quantification of Gleason 4 burden differs from histopathologist visual estimation.

5.2.2.1. Aim 2: Objectively quantify the burden of Gleason 4 disease in a cohort of 30 men from the PROMIS study.

Objective 1: Determine the “true” burden of Gleason 4 by manually contouring clusters of Gleason 4 in digital images

Objective 2: Compare the percentage assessment given by pathologists to a calculated percentage based on objective 1 findings.

Objective 3: Compare impact of the visual and digital Gleason 4 percentage estimation in the index lesion.

5.2.3. Hypothesis 2:

Geographical parameters can be used to characterise the spatial distribution of Gleason 4 in prostate biopsies.

5.2.3.1. Aim 3: Understand the clustering of Gleason 4 disease by using geographical parameters.

Objective 4: Apply geographical descriptors to describe the distribution of Gleason 4.

Objective 5: Understand the randomness of distribution of Gleason 4 in prostate biopsies.

Objective 6: Compare the geographical parameters with the index lesion derived from mpMRI Likert score.

Objective 7: Compare the geographical descriptors with the Gleason 4 percentage visual estimation and digital measurement.

5.3. Results

5.3.1. Comparison of visual and digital MCCL measurement reveals discrepancy in Gleason 4 measurement.

5.3.1.1. Selection of cohort

Two-hundred and twenty-six patients from University College London Hospital took part in the PROMIS trial. Of 113 men with Gleason 7 PCa, 85 had significant disease (PROMIS definition 1: Gleason 4+3 or MCCL ≥ 6 mm). 15 patients with Gleason 3+4 and 15 patients with 4+3 disease were selected from the 85, using a random number generator on excel (Figure 5-2 A). A mean of 14.2 ± 8.05 cores per patient (IQR: 9, range: 2-34) were taken, containing with any Gleason 4. 192 H&E slides from these 30 patients were scanned as described on section 2.2.5.1. Gleason 4 and MCCL was measured as described in sections 2.2.5.2.

Comparison of age, prostate volume, presenting PSA and PSA density between the 85 patient initial cohort and the final selected 30 patients showed no statistical difference between the two cohorts in age, prostate volume and presenting PSA, and a statistical difference in PSA density (Table 5.2). Boxplots comparing clinical characteristics can be seen on Figure 5-2 B and C.

| Characteristic | Gleason 3+4 | | | Gleason 4+3 | | |
|---------------------------|--|--|---------------------------------------|--|--|---------------------------------------|
| | 85 patients <i>n</i> =67 ¹ | 30 patients <i>n</i> =15 ¹ | <i>p</i> <i>value</i> ² | 85 patients <i>n</i> =18 ¹ | 30 patients <i>n</i> =15 ¹ | <i>p</i> <i>value</i> ² |
| Age | 64 (58, 70) | 63 (55, 70) | 0.7 | 66 (62, 69) | 66 (62, 70) | > 0.9 |
| Prostate volume (cc) | 37 (27, 46) | 29 (25, 40) | 0.3 | 39 (32, 44) | 40 (30, 44) | > 0.9 |
| Presenting PSA | 7.10 (5.85, 8.90) | 7.90 (6.15, 8.80) | 0.7 | 11.05 (8.57, 12.85) | 10.50 (8.65, 12.50) | > 0.9 |
| PSA density (PSA/pvol) | 0.21 (0.15, 0.26) | 3.93 (3.29, 5.43) | <0.001 | 0.30 (0.24, 0.33) | 3.33 (3.03, 4.18) | <0.001 |

Table 5.2 Comparison of age and clinical characteristics between the 85 patient cohort and the final 30 patient cohort.

Gleason 3+4 and 4+3 separated to compare age, prostate volume, presenting PSA and PSA density. ¹ Data presented as median and IQR in parenthesis. ² Wilcoxon rank sum test performed.

5.3.1.2. Statistical analysis

Patients were divided according to the original Gleason score from the PROMIS trial into 3+4 and 4+3. The routinely performed 'visual' estimation for both measurements was used as the reference standard for all comparisons. When comparing two groups, meeting normal distribution (Shapiro-Wilk test) and same variances (F-test), a student t-test was applied. Whenever data was not normally distributed, a Mann-Whitney test was performed (for more information see section 2.2.6.2). To quantify the agreement between the two methods, the Bland Altman method was performed (351). The visual method was used as a standard for comparison; bias was defined as the average of the difference between the two methods. Limits of agreement were calculated at 95% CI.

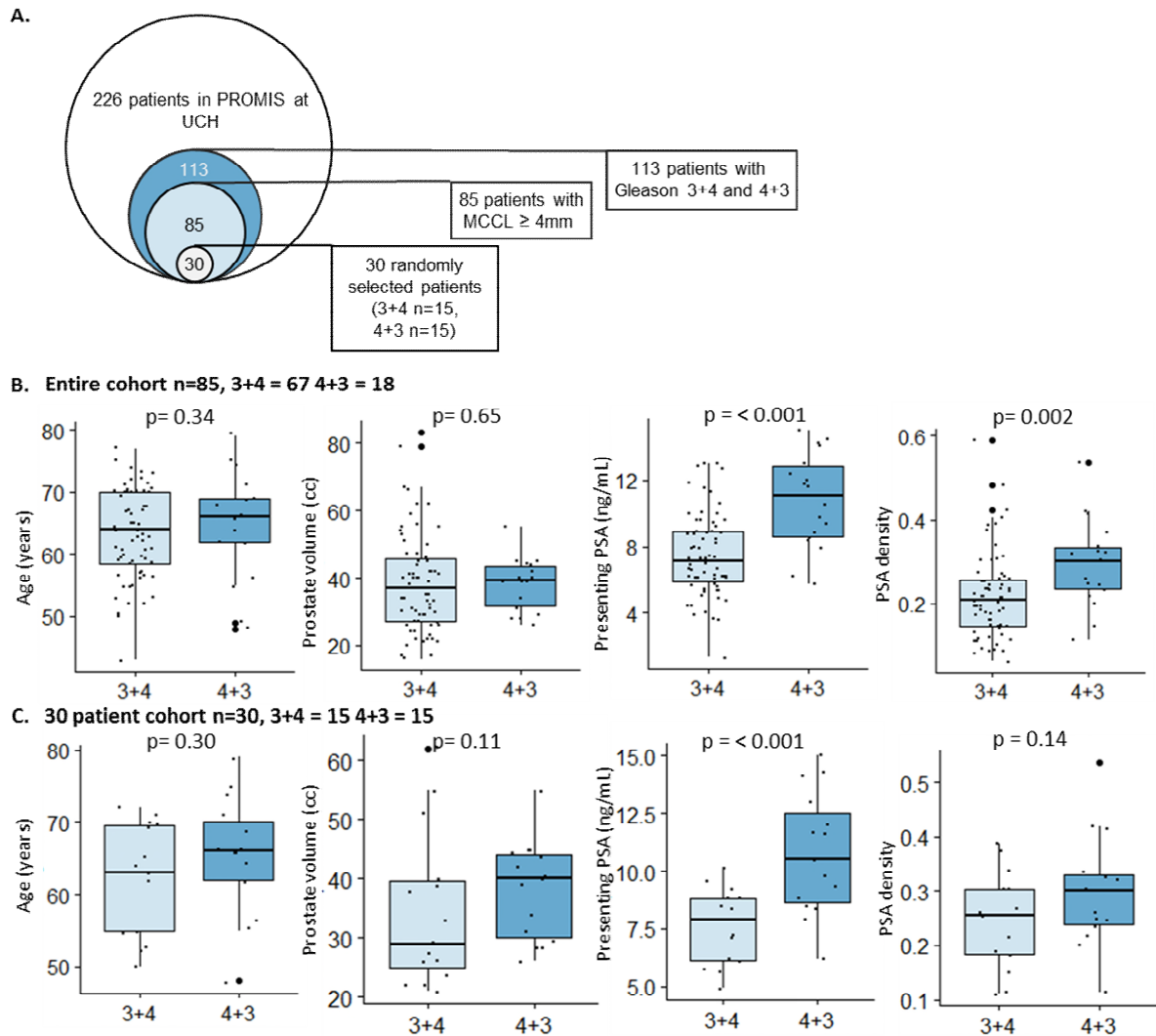


Figure 5-2 Gleason 4 PROMIS cohort selection process and demographics.

A. Euler diagram representing the selection process for 30 patient undergoing in-depth Gleason 4 analysis. **B.** Patient demographics including age, prostate volume, presenting PSA and PSA density for the entire Gleason 7 cohort and for the 30 selected patients **C.** Student t test performed on all except PSA density in the entire cohort, and prostate volume in the 30 men cohort as data was not normally distributed (Wilcoxon rank sum test performed instead).

5.3.1.3. Digital and visual MCCL measurement are comparable.

When comparing visual vs digital MCCL, in 23 of the 30 patients the difference was up to ± 2 mm; taking into account the positive and negative values, the median difference between the two measurements was 0.58 mm (range -4.12 mm to +5.52 mm, t-test, $p=0.64$) (Figure 5-3 A, B). Seven patients had measurements that differed by ≥ 2 mm between digital and visual estimation. When viewed as a density plot, there was a tendency to overestimate MCCL in the 3+4 group and under-estimate in the 4+3 group when using the visual method (Figure 5-3 C).

The Bland Altman test aids comparison of two methods measuring the same outcome, on this case comparing the gold standard (visual estimation) to the digital measurement. The test was applied to understand the degree of agreement between the two measurements. There was no systematic difference (bias) between the visual and digital assessment of MCCL, and there was no correlation between increasing MCCL and the level of disagreement between the two measurements (Figure 5-4 A). When dividing the cohort into Gleason 3+4 and 4+3 there was no significant difference in MCCL visual estimation and digital measurement ($p=0.8$ for 3+ and 0.48 for 4+3) (Figure 5-6B).

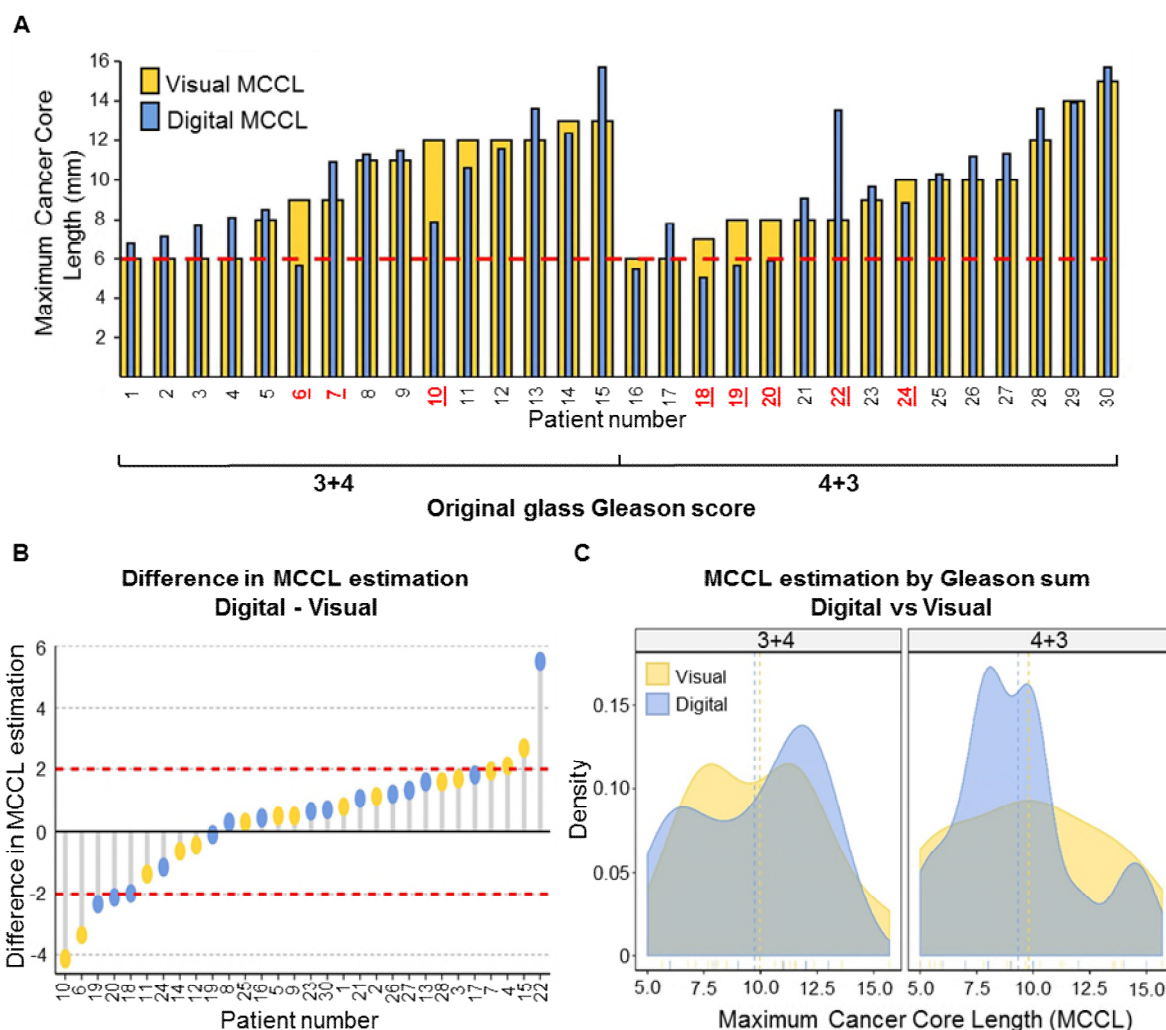


Figure 5-3 Objective measurement of MCCL and shows a discrepancy with visual measurement and pathologist estimation.

A. MCCL difference between visual and digital MCCL shows under-estimation in visual compared to digital MCCL. Bar plot of visual MCCL in yellow and digital MCCL in blue, organised by Gleason score. MCCL is plotted on the y-axis; each patient is plotted on the x-axis. Red dashed lines represent a threshold of 6 mm as the MCCL criterion for significance (PROMIS definition 1). Patients highlighted in red were over or underestimated in the original visual measurement. **B.** Waterfall plot representing the difference between visual and digital measurements as digital MCCL- visual MCCL by Gleason score (y-axis), patients plotted on the x-axis. Visual Gleason score is represented in yellow for 3+4 and blue for 4+3. Bars with a negative value represent measurements where the visual MCCL was shorter than the digital MCCL (underestimation). Bars with a positive value represent cases where the visual MCCL was higher than the digital MCCL. The difference in 80% of cases is ± 2 mm ($n=24$), red dashed line at -2 and 2 mm difference. **C.** Density plots representing the MCCL distribution between visual and digital images by Gleason scores. Y-axis represents the Kernel density estimation. The X-axis contains MCCL values. Visual MCCL score is represented in yellow and blue for the digital measurement. 4+3. The mean visual MCCL was 9.53 mm (5-15 mm) and the mean digital MCCL was 9.88 mm (5.01-15.74).

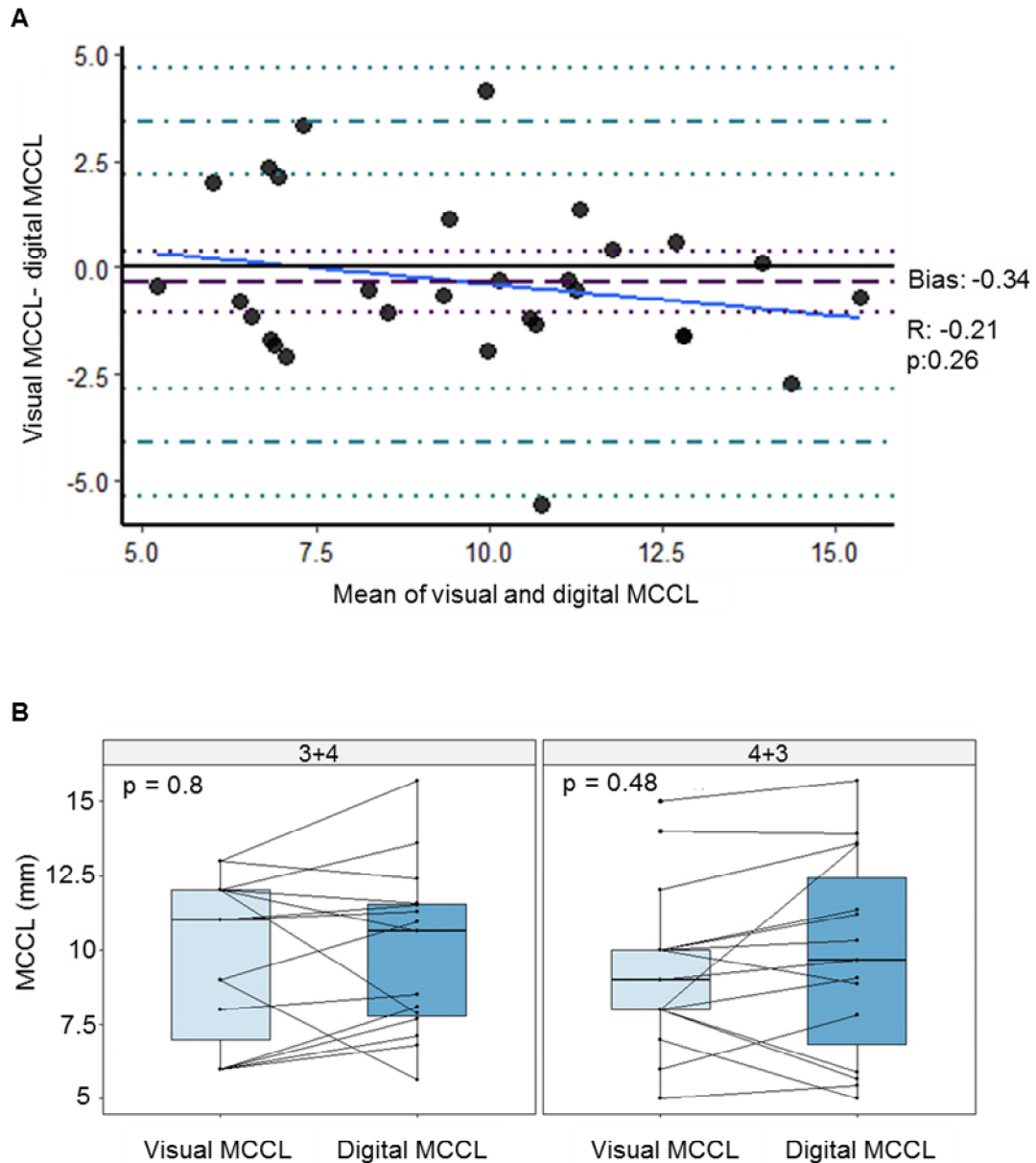


Figure 5-4 No significant bias is seen between visual and digital MCCL measurement.

A. Bland-Altman plot representing the difference in measurement in the y-axis as visual MCCL – digital MCCL. The x-axis represents the mean MCCL measurement of both techniques as (visual MCCL + digital MCCL)/2. The bold black line represents complete agreement at 0. The purple dashed line corresponds to the bias at -0.34. The dotted purple line corresponds to the bias confidence interval (-1.06 to 0.37). Dash and dotted blue lines correspond to the upper and lower limit of agreement and confidence intervals are plotted with dotted blue lines. Upper limit of agreement: 3.43 (2.18 to 4.67), lower limit of agreement: -4.12 (-5.36 to -2.87). **B.** Paired boxplot comparing MCCL values obtained from the visual and glass measurement separated by Gleason score. Lines connecting values to denote change of measurement from visual to digital. Wilcoxon signed-rank test performed to obtain p values.

5.3.1.4. Visual Gleason 4 overestimates Gleason 4 burden.

The visual %G4 overestimated Gleason 4 burden when compared to the digital assessment in all cases (Figure 5-5 A). The 4+3 group had a mean difference of +26.6% (range: 9.6%-41.9%) compared to +10.8% (range: 1.3%-24.9%) for the 3+4 group (t-test, $p=1.9 \times 10^{-5}$). The average %G4 in the patients graded 3+4 was 11.2% (range: 4.7-17.9%) compared to 30.4% (range: 12.9-50.6%) in the 4+3 group (t-test, $p<0.0001$). When pathologists were asked to assess the overall Gleason score based on the digital images (visual %G4), two patients were downgraded from their original clinical grading of 4+3 to 3+4 by both pathologists (See yellow bars of patients 23 and 18 in Figure 5-5 A).

The difference between the two measurements ranged between 1.31% and 42% (median 18% difference) Figure 5-5B. As seen on the density plot in Figure 5-5 C, the digital measurement was skewed to the left (lower percentage of Gleason 4) compared to the visual estimation.

Using the established 50% G4 threshold to designate a 4+3 cancer, and based on the digital %G4 (blue bars), only one patient (number 19 in Figure 5-5 A) would be classified as 4+3. When dividing the digital %G4 into quartiles, two patients in the original 4+3 group had less %G4 than the upper quartile of the 3+4 group (18 and 30). In other words, these two patients had less Gleason 4 than the men with the highest %G4 compromise in the original 3+4 group. Figure 5-6 A shows the Bland-Altman analysis; showing that there was a bias towards overestimation in the visual estimations as all values are located above the line of complete agreement (Complete agreement would result in a zero value). The disagreement was larger when more than 20% of G4 was present ($R: 0.79$, $p: <0.0001$).

In order to try to understand which factors may have influenced the pathologist's decision-making, several factors were correlated to the difference between visual and digital %G4. For instance, did the pathologists give a higher %G4 estimation if a patient had more blocks with G4? For this purpose, the number of blocks, cores, contours, cancer area and Gleason 4 area were correlated to the difference between the pathologist and the digital measurement. No significant correlations were found (Figure 5-6 B).

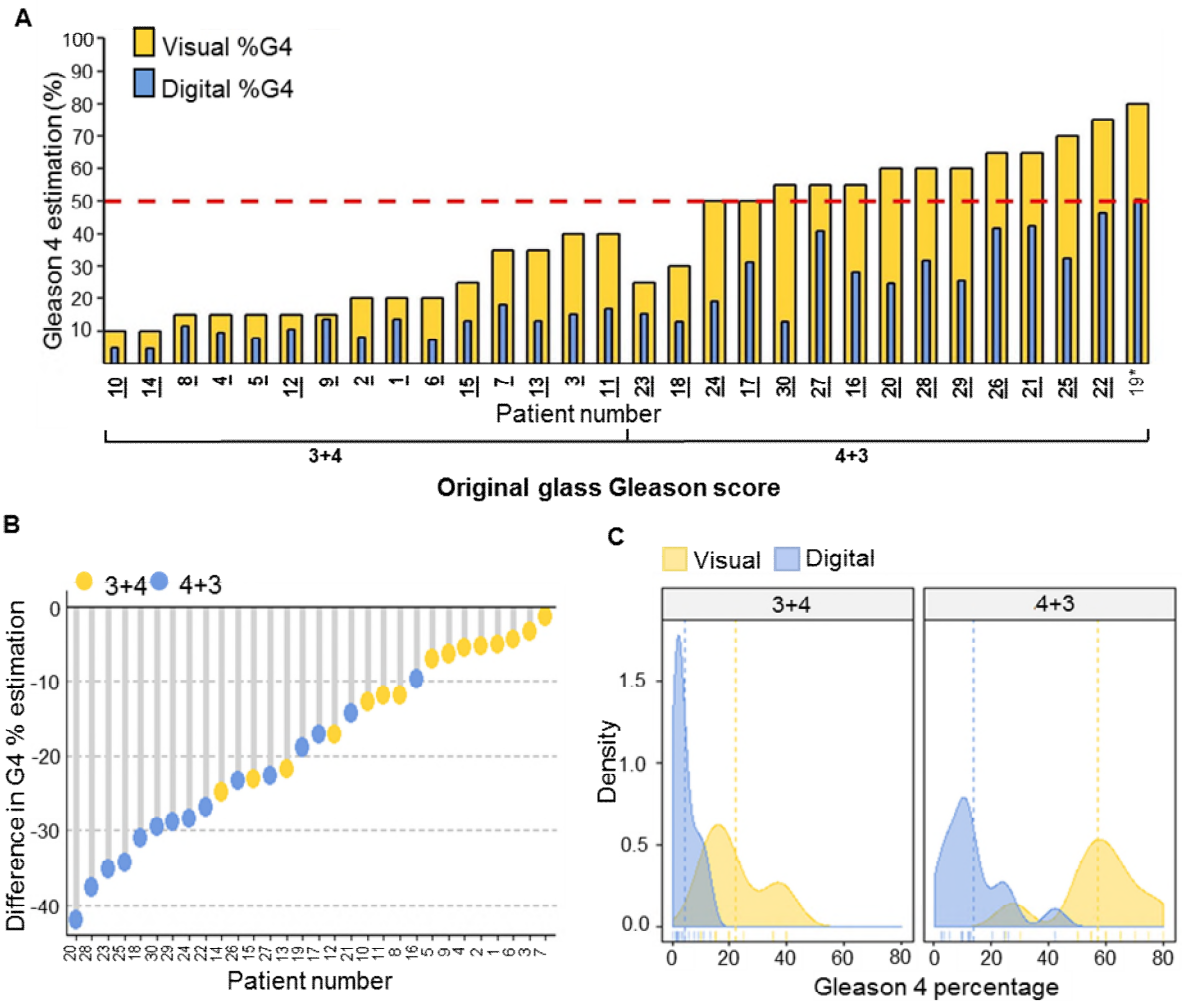


Figure 5-5 Visual Gleason 4 appraisal overestimates burden of disease.

A. Bar plot of the proportion of Gleason 4 estimation average between two uropathologists (yellow) and digital estimation (blue). %G4 is plotted on the y-axis; each patient is plotted on the x-axis. A threshold of 50% g4 for clinical significance is shown as a red dashed line. Patient number on the x-axis is highlighted in bold and underlined if the digital measurement of their %G4 would lead to reclassification based on the digital value. Patient marked with * has $\geq 50\%$ G4 in the digital measurement. **B.** Waterfall plot representing the difference between visual and digital measurements as digital G4%- visual G4% by Gleason score (y-axis), patients plotted on the x-axis. Visual Gleason score is represented in yellow for 3+4 and blue for 4+3. **C.** Density plots representing the %G4 distribution between visual and digital images by Gleason scores. Y-axis represents the Kernel density estimation. The X-axis contains %G4 values. Visual MCCL score is represented in yellow and blue for the digital measurement. 4+3. The mean visual MCCL was 9.53 mm (5-15 mm) and the mean digital MCCL was 9.88 mm (5.01-15.74).

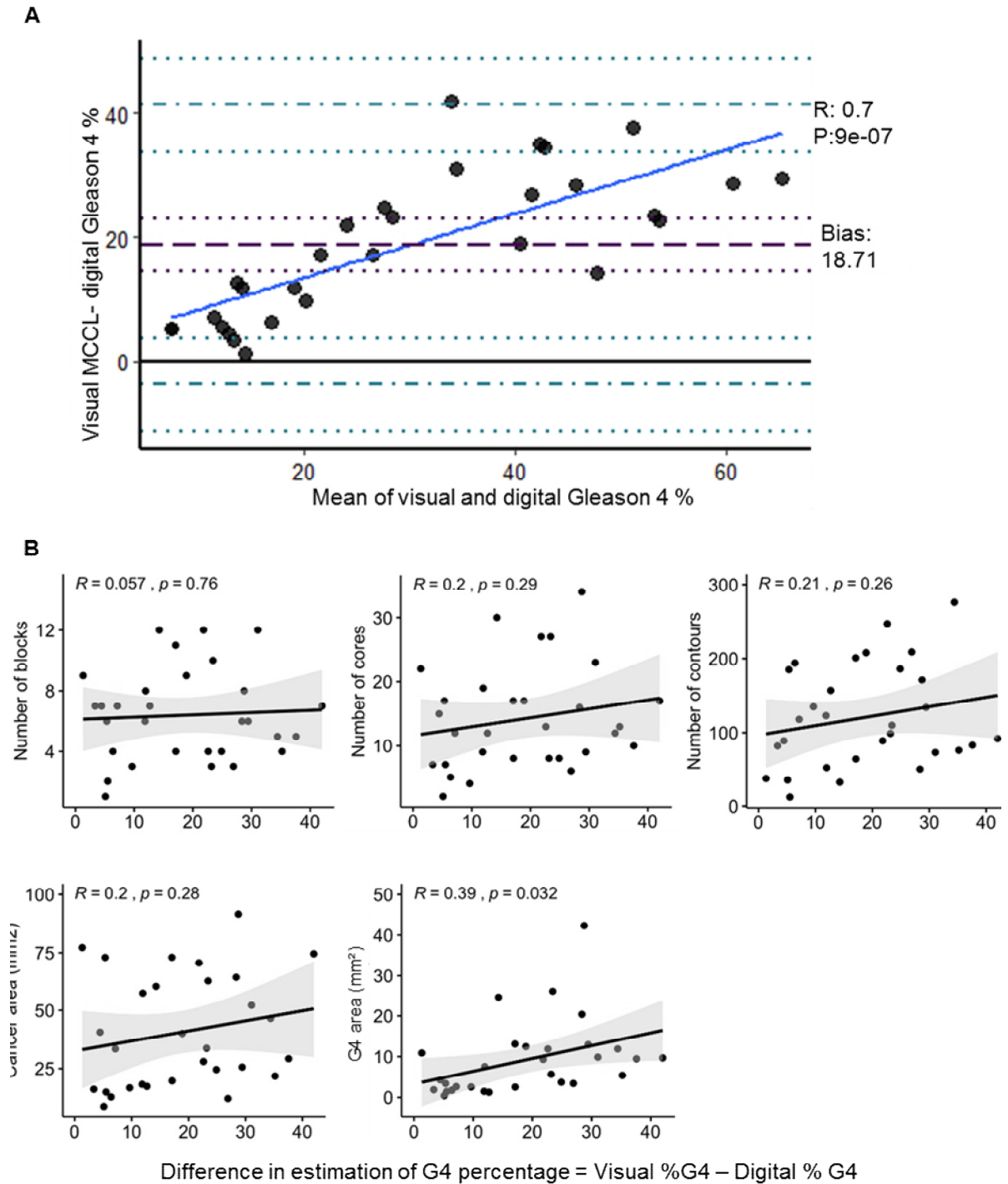


Figure 5-6 Visual estimation of Gleason 4 % overestimates Gleason 4 burden.

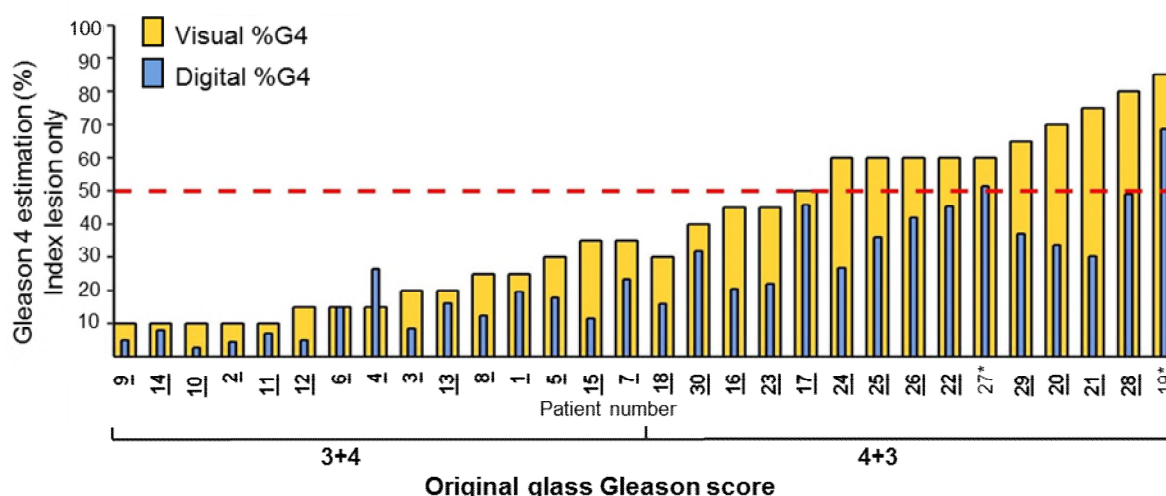
A. Bland-Altman plot representing the difference in measurement in the y-axis as visual %G4 – digital %G4. The x-axis represents the mean %G4 measurement of both techniques as (visual %G4 + digital %G4)/2. The bold black line represents complete agreement at 0. The purple dashed line corresponds to the bias at 18.71; the dotted purple line corresponds to the bias confidence interval (33.87 to 48.75). Dash and dotted blue lines correspond to the upper and lower limit of agreement and confidence intervals are plotted with dotted blue lines. Upper limit of agreement: 41.31 (33.87 to 48.75), lower limit of agreement: -3.87 (-11.31 to 3.56). Regression line is plotted as a continuous blue line. **B.** Scatter plots of the difference between visual and digital percentage G4 versus number of blocks, number of cores, number of contours, G4 area in mm² and any cancer area in mm². Regression line plotted in black, confidence interval in light grey. Pearson method was used to calculate the correlation coefficient.

5.3.1.5. Index block %G4 overestimation leads to patient reclassification

Examination of the index block (block with the highest Gleason score and MCCL), revealed equivalent findings as those seen with all tumour containing cores (Figure 5-7 A). The visual assessment of digitised images downgraded four patients' index block from 4+3 to 3+4 (patients 18, 30, 16, and 23). When examining the digital %G4, only two patients reached the 50% G4 threshold (27 and 19), and so would be the only two patients with 4+3 disease based on digital measurement. The Bland-Altman analysis revealed a similar trend to that of the overall %G4 analysis. One measurement had complete agreement between the digital and visual estimate (patient 6 in Figure 5-7 A). One patient had a higher digital estimation compared to the visual estimation (patient 4 in Figure 5-7 A). This is represented by the only dot in the negative area of Figure 5-7 B. The disagreement between measurements increased as the amount of %G4 increased (R: 0.6, p: <0.0001).

When patients were classified using the clinical significance criteria used in PROMIS in which MCCL and Gleason score were combined to derive definitions 1 ($\geq 4+3$ or $\geq 6\text{mm}$) and 2 ($\geq 3+4$ or 4mm) the digital analysis reclassified four patients' index block as lower risk. When all blocks were compared using this system, 20 patients had discrepancy between the visual and digital classification, leading to reclassification to higher or lower risk in six and fourteen patients, respectively (Figure 5-8)

A.



B.

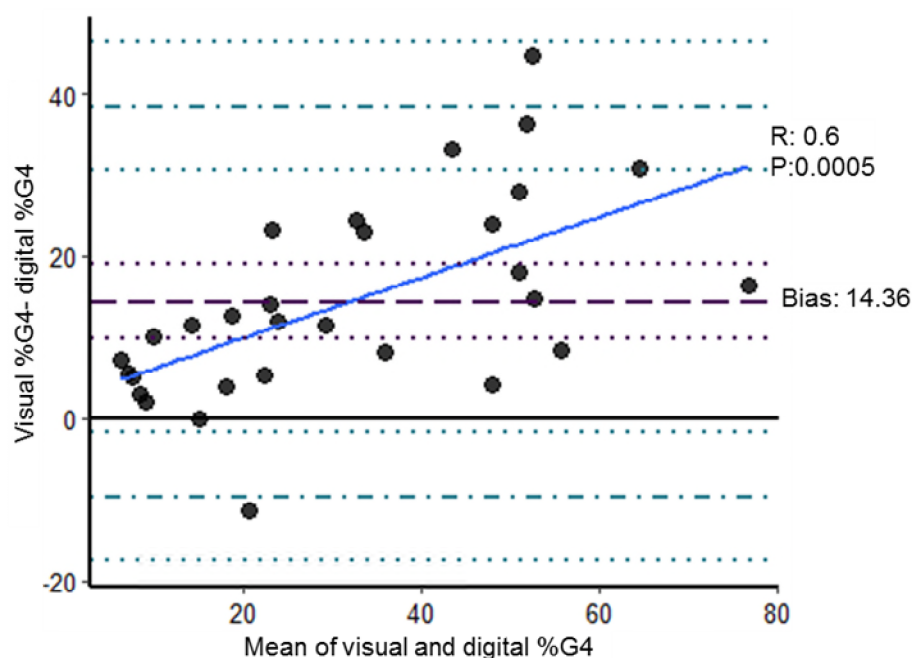


Figure 5-7 Objective measurement of Gleason 4 burden shows a discrepancy between visual measurement and the digital measurement for the index block.

A. Visual %G4 for the index block 30 patients shown in yellow overlaid with digital %G4 in blue. Patients separated by original Gleason grade grouping; 3+4 or 4+3, and organized by visual %G4. A threshold of 50% G4 for clinical significance is shown as a red dashed line. Patient number on the x-axis highlighted in bold and underlined if the objective measurement of their %G4 would cause reclassification. **B.** Bland-Altman plot representing the difference in measurement in the y-axis as visual %G4 – digital %G4. The x-axis represents the mean %G4 measurement of both techniques as (visual %G4 + digital %G4)/2. The bold black line represents complete agreement at 0. The purple dashed line corresponds to the bias at 14.36; the dotted purple line corresponds to the bias confidence interval (9.78 to 18.94). Dash and dotted blue lines correspond to the upper and lower limit of agreement and confidence intervals are plotted with dotted blue lines. Upper limit of agreement: 38.40 (30.49 to 46.32), lower limit of agreement: -9.67 (-17.59 to -1.76). The regression line is plotted as a continuous blue line.

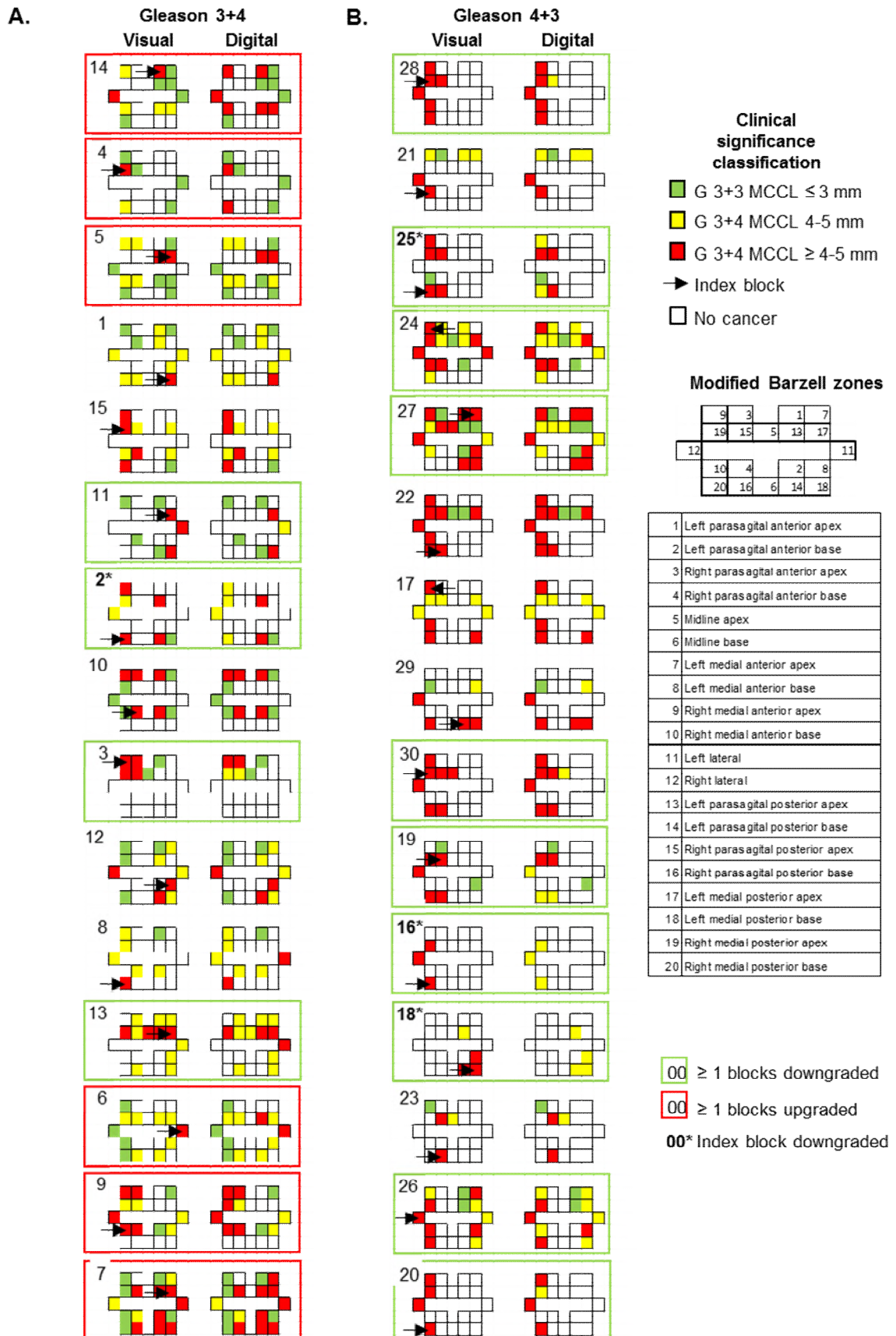


Figure 5-8 Digital assessment reclassifies patients when using the UCLH traffic light system.

Graphical representation of the 20 modified Barzell zones used for transperineal sampling in PROMIS comparing visual and digital assessment, each positive block is denoted by the traffic light system representing the UCH clinical significance classification. Gleason 3+3 and MCCL \leq 3 mm (green), Gleason 3+4 and/or MCCL 4-5 mm (yellow) and Gleason 4+3 and/or MCCL \geq 6 mm (red). The index block is highlighted with a black arrow. The table contains the anatomical description of each modified Barzell zone. Patients are separated by original Gleason grade grouping 3+4 (A) or 4+3 (B); the number corresponds to a patient number. * if index block was downgraded. Green boundary box if one or more blocks were downgraded. Red boundary box if one or more blocks were upgraded.

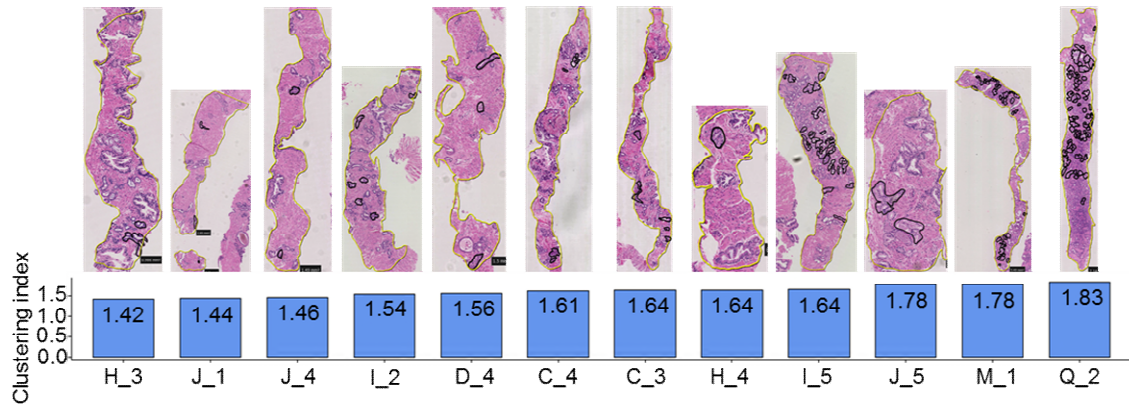
5.3.2. Geographical parameters can be used to study the spatial distribution of Gleason 4 in prostate biopsies.

Using the described methods in Section 2.2.5.3, Gleason 4 contours were converted to Geojson images to obtain measurements of clustering and evenness per core. The clustering index indicates how close together and enclaved Gleason 4 is in relation to 3, in other words, a low score indicates less segregated Gleason 4. A higher score indicates more segregated Gleason 4. In Figure 5-9 A, an example of clustering index across all Gleason 4 containing cores is depicted. Core H_3 has a lower clustering value of 1.42, as there are few, small areas of Gleason 4. Compare this to core Q_2, with a value of 1.83, with many small Gleason 4 contours.

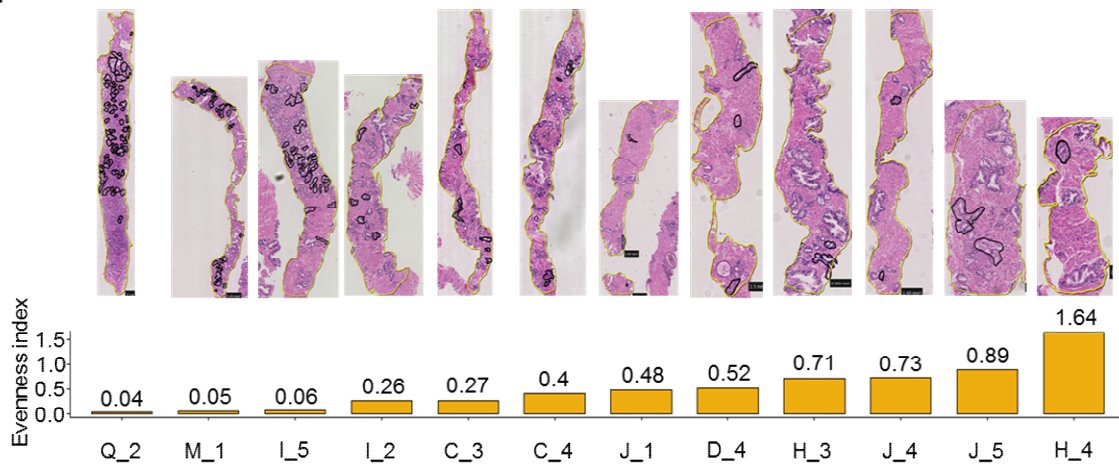
Evenness provides a measure of proportion of Gleason 4 to Gleason 3. A low evenness value indicates a near 50/50 distribution. In other words, the “population” is evenly distributed. A high value indicates a higher proportion of one or the other Gleason, so that the “population” is uneven. For example, core Q_2 in Figure 5-9 B, has a very low evenness score of 0.04, corresponding to the core in this patient with the highest amount of Gleason 4 (Figure 5-9 C). Conversely, cores with lower amounts of Gleason 4 have a higher evenness scores. For example, Core H_4 has an evenness index of 1.64 and 0.01 mm² of Gleason 4. Evenness scores above 1 are not expected and are seen where only one Gleason 4 contour was drawn (Figure 5-10).

Density plots were created to visualise the distribution of clustering and evenness indexes across twelve randomly selected patients (2.2.5.3) Overall, the distribution of the clustering index value was to the right, whereas the evenness index was to the left, as seen on Figure 5-11 A. Each patient was plotted separately to visualise the distribution of the clustering and evenness index, with the assumption that values would remain opposite (as in patients 4, 6, 8, 9 and 11). However, this was not the case; some patients had a more normally distributed curve (patient 1, for example).

A.



B.



C.

| Core | Cancer Area (mm ²) | Gleason 4 area (mm ²) | Percentage Gleason 4 (%) | Number of contours | Clustering index | Evenness index |
|------|--------------------------------|-----------------------------------|--------------------------|--------------------|------------------|----------------|
| C_3 | 3.89 | 0.14 | 3.55 | 6 | 1.64 | 0.27 |
| C_4 | 2.92 | 0.07 | 2.36 | 4 | 1.61 | 0.40 |
| D_4 | 1.50 | 0.05 | 3.17 | 3 | 1.56 | 0.52 |
| H_3 | 0.99 | 0.02 | 2.49 | 2 | 1.42 | 0.71 |
| H_4 | 0.50 | 0.01 | 2.62 | 1 | 1.64 | 1.64 |
| I_2 | 2.60 | 0.08 | 3.02 | 6 | 1.54 | 0.26 |
| I_5 | 2.60 | 0.30 | 11.55 | 28 | 1.65 | 0.06 |
| J_1 | 3.25 | 0.04 | 1.16 | 3 | 1.44 | 0.48 |
| J_4 | 1.49 | 0.02 | 1.03 | 2 | 1.46 | 0.73 |
| J_5 | 1.08 | 0.09 | 7.9 | 2 | 1.78 | 0.89 |
| M_1 | 5.49 | 0.63 | 11.42 | 34 | 1.78 | 0.05 |
| Q_2 | 7.14 | 1.28 | 17.91 | 51 | 1.83 | 0.04 |

Figure 5-9 Clustering and evenness indexes example in one patient.

Barplots representing the clustering index value (A) and Evenness index (B) on all Gleason 4 containing cores in one patient. Index value on y-axis, each core H&E and contoured core on the x axis, labelled as mentioned on Figure 2.1, (2.2.5.2.3) C. Table depicting core name, cancer area in mm², Gleason 4 area in mm², Percentage Gleason 4 as calculated depicted in Figure 2.1D, number of contours, clustering and evenness indexes.

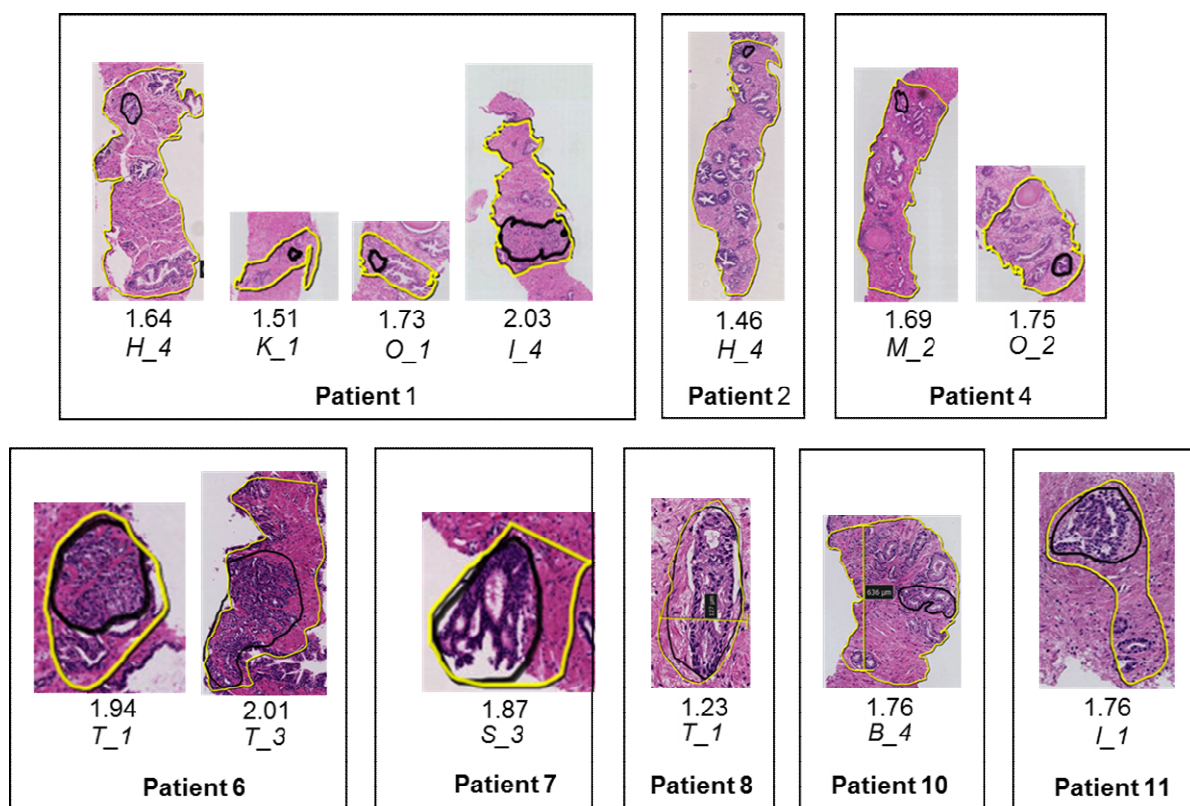


Figure 5-10 Evenness index above 1 were found on cores where a single Gleason 4 contour was drawn.

H&E images of cores with a single Gleason 4 contour in patients 1, 2, 4, 6, 7, 8, 10 and 11. Evenness index immediately below H&E image, core name in italic. Patient number in bold. Cancer area in yellow, Gleason 4 contour in black.

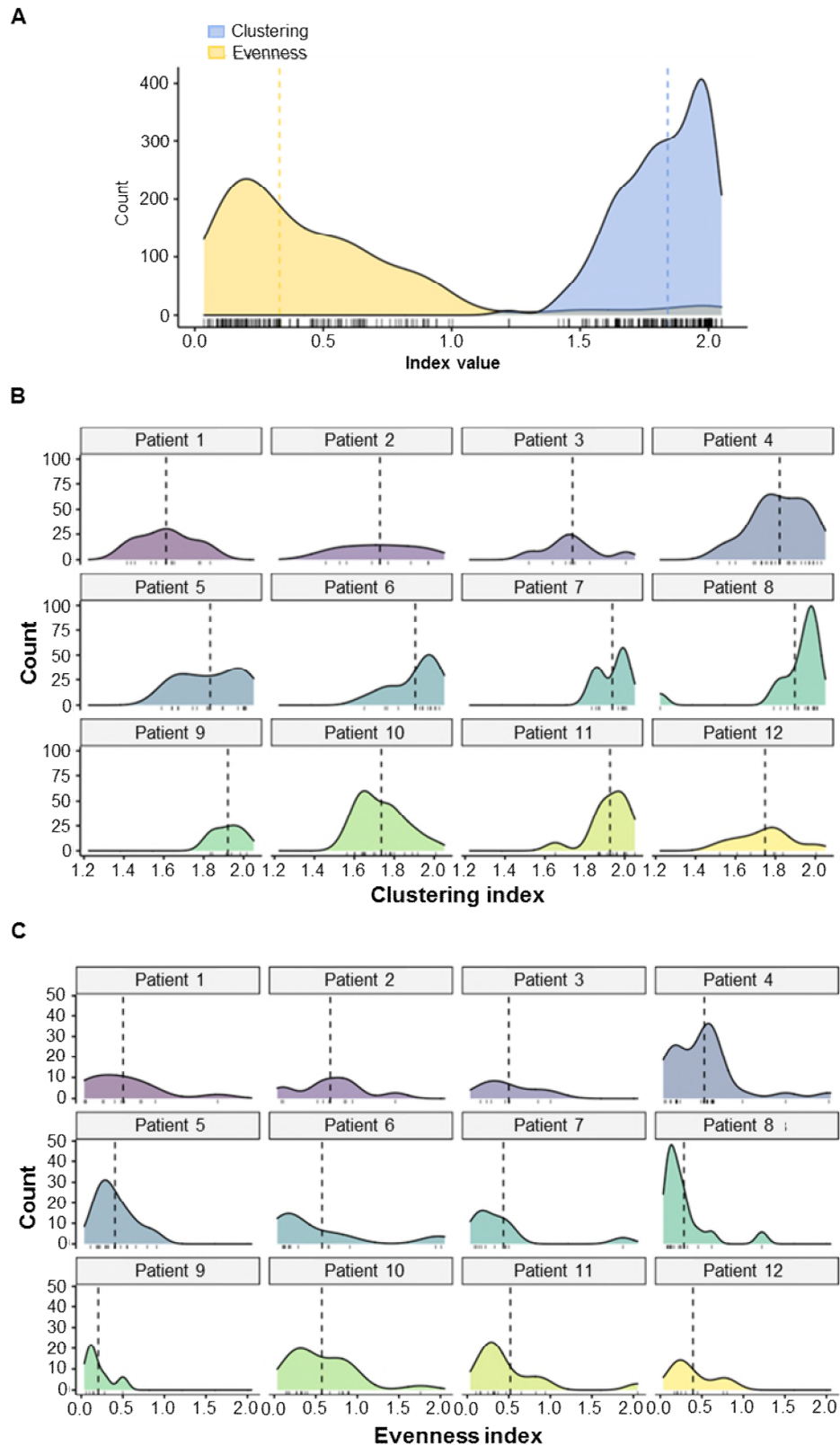


Figure 5-11 Clustering and evenness distribution in twelve patients

A. Density plot representing the overall distribution of clustering and evenness values in all patients. Count on y-axis, index values on the x-axis. Marginal rug on x-axis representing the values per core.
B. Density plot representing the clustering index values on the x-axis and a density count on the y-axis, each patient plotted separately. **C.** Density plot representing the evenness index values on the x-axis and a density count on the y-axis, each patient plotted separately.

5.3.2.1. There is no correlation between cancer area, clustering or evenness indexes and location in the prostate.

In order to understand if the total cancer, clustering or evenness values are different across the prostate zones, the data was sent for further analysis as described in section 2.2.5.3.1. Briefly, one-hot encoding was performed for Barzell zones and ordinal encoding was performed for a virtual grid overlaid on top of a schematic representation of the Barzell zones. The aim was to determine if there was any correlation between the Barzell zone or ordinal coordinate system and the total cancer area, clustering or evenness score. There was no correlation between location, ordinal coordinate system or the values studied (Table 5.3 and Table 5.4). Significant p values for cancer area and Barzell zone H/8, T/20 and clustering and zone S/19 could be secondary to random significance (352).

| Barzell zone | Cancer area | | Clustering | | Evenness | |
|---------------------|--------------------------------|----------------|--------------------------------|----------------|--------------------------------|----------------|
| | <i>Correlation coefficient</i> | <i>p value</i> | <i>Correlation coefficient</i> | <i>p value</i> | <i>Correlation coefficient</i> | <i>p value</i> |
| A = 1 | -0.326 | 0.060 | 0.133 | 0.452 | -0.244 | 0.165 |
| C = 3 | 0.014 | 0.936 | -0.031 | 0.861 | 0.031 | 0.862 |
| D = 4 | -0.189 | 0.285 | 0.086 | 0.630 | 0.171 | 0.335 |
| H = 8 | -0.378 | 0.027* | -0.276 | 0.114 | -0.053 | 0.764 |
| I = 9 | 0.118 | 0.506 | 0.272 | 0.119 | -0.298 | 0.086 |
| J = 10 | 0.023 | 0.896 | -0.292 | 0.094 | 0.037 | 0.837 |
| L = 12 | -0.300 | 0.085 | -0.035 | 0.844 | 0.008 | 0.965 |
| M = 13 | 0.013 | 0.942 | -0.242 | 0.168 | -0.067 | 0.708 |
| N = 14 | -0.021 | 0.908 | -0.112 | 0.527 | 0.147 | 0.408 |
| Q = 17 | 0.167 | 0.345 | 0.108 | 0.544 | -0.146 | 0.411 |
| S = 19 | 0.014 | 0.937 | 0.350 | 0.043* | 0.195 | 0.269 |
| T = 20 | 0.399 | 0.020* | -0.052 | 0.769 | 0.139 | 0.432 |

Table 5.3 Correlation coefficients between Barzell zone and cancer area, clustering and evenness indexes.

Barzell zones on first column, correlation coefficients and p values on subsequent columns. Values highlighted with *, represent p values <0.05.

| Ordinal coordinate system | Cancer area | | Clustering | | Evenness | |
|----------------------------------|--------------------------------|----------------|--------------------------------|----------------|--------------------------------|----------------|
| | <i>Correlation coefficient</i> | <i>p value</i> | <i>Correlation coefficient</i> | <i>p value</i> | <i>Correlation coefficient</i> | <i>p value</i> |
| Ordinal X | -0.294 | 0.091 | -0.229 | 0.192 | -0.043 | 0.810 |
| Ordinal Y | -0.165 | 0.351 | 0.303 | 0.082 | -0.320 | 0.065 |

Table 5.4 Correlation coefficients between the ordinal coordinate system and cancer area, clustering and evenness indexes.

Ordinal coordinate values on first column, correlation coefficients and p values on subsequent columns.

5.3.2.2. Clustering index is raised in high Likert mpMRI index lesion.

The index lesion on mpMRI is defined as the lesion with the highest Likert score or the largest in the event that two lesions are scored the same. The Likert score is a scale from one to five where 1 indicates a very low level of suspicion for prostate cancer; 2 a low level of suspicion; 3 is equivocal; 4, cancer is probable and 5, where the lesion is considered to definitely be cancer (353,354). The mpMRI index lesion has been shown to have good correlation to the pathological index lesion (75,355).

For this study, the index lesion on mpMRI was cognitively matched to the Barzell zone histopathological scheme, if present, data was also recorded for the satellite lesion (additional contoured abnormal are on mpMRI). The evenness and clustering indexes corresponding block or blocks were extracted. To understand if there was a significant difference in evenness and/or clustering between the different Likert scores, the index values were plotted against the three available Likert scores.

For each patient, the index lesion and satellite lesion were identified. A summary of all data is presented on Table 5.5. Patient 3 and 7 had only one lesion on mpMRI, the remaining patients had two visible lesions on mpMRI (Likert 3-5). Of these, four had values for both lesions (Patients 1, 2, 4 and 5). Whereas patients 6,8,9,10,11 and 12 had two lesions but there was no Gleason 4 identified on the matching blocks to the lesion on mpMRI. Absolute cancer area corresponds to the total cancer contoured as described on Section 2.2.5.2.2. The proportion of Gleason 3 to 4 was calculated for the index lesion alone and all the cores added.

There was a significant difference in Clustering index for Likert 3 and Likert 5 (median 1.91 (IQR: 1.73-1.97) in Likert 3 versus 1.82 (IQR: 1.86-2.00) for Likert 5, $p=0.043$) and between Likert 4 and Likert 5 (median 1.78 (IQR: 1.65-1.89) for Likert 4 and 1.82 (IQR: 1.73-2.00) for Likert 5, $p=0.0001$), see Figure 5-12 A. There was no significant difference between evenness index in the different Likert scores (Likert 3 median evenness (0.41 (IQR: 0.18-0.50), Likert 4: median 0.45 (IQR: 0.20-0.69) and Likert 5: median 0.45 (IQR: 0.18-0.50), $p=0.92$), see Figure 5-12 B.

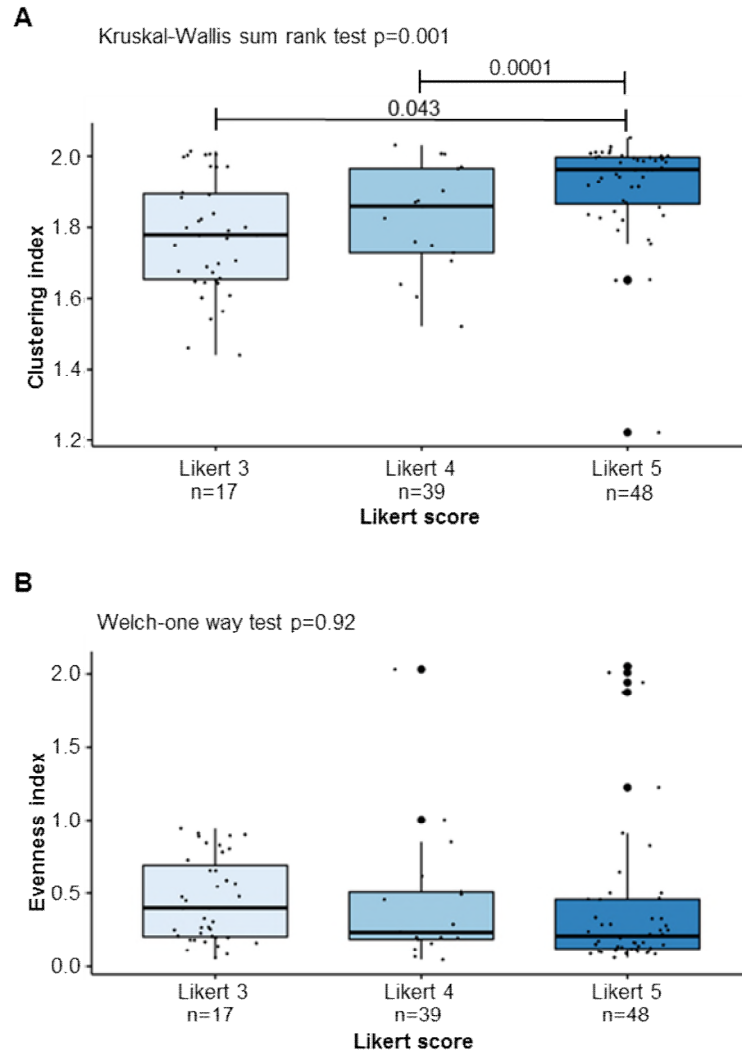


Figure 5-12 The clustering index increases with the Likert score, but not the evenness.

A. Boxplot containing data on the index lesion values of clustering. X-axis contains Likert score, y-axis the clustering index values. Kruskal-Wallis rank sum test was performed, Dunn post-hoc test performed due to inequality of group sizes. Comparison between Likert 3 and 4 was not significant ($p=0.35$). n corresponds to the number of values per block. **B.** Boxplot containing data on the index lesion values of clustering. X-axis contains Likert score, y-axis the evenness index values. Welch-one way test performed, no post-hoc test performed due to lack of significance.

| Patient number | Original Gleason | MRI | | | | | Evenness Median (Range) | Clustering Median (Range) | Absolute (mm ²) | | Proportion (%) | | Absolute total (mm ²) | | Proportion (%) total | |
|----------------|------------------|--------|----|---|---|-----|-------------------------|---------------------------|-----------------------------|-----------|----------------|-----------|-----------------------------------|-----------|----------------------|-----------|
| | | Lesion | T2 | D | C | All | | | Cancer area | Gleason 4 | Gleason 3 | Gleason 4 | Cancer area | Gleason 4 | Gleason 3 | Gleason 4 |
| Patient 1 | 3+4 | 1 | 4 | 4 | 4 | 4 | 0.48 (0.05-0.89) | 1.57 (1.44-1.78) | 11.02 | 0.51 | 95.38 | 4.62 | 33.44 | 2.63 | 92.14 | 7.86 |
| | | 2 | 3 | 4 | 4 | 4 | 0.04 (0.03-0.05) | 1.80 (1.78-1.83) | 12.63 | 1.86 | 85.28 | 14.72 | | | | |
| Patient 2 | 3+4 | 1 | 4 | 3 | 4 | 4 | 0.62 (0.09-0.94) | 1.86 (1.69-1.97) | 6.83 | 1.13 | 83.46 | 16.54 | 18.21 | 1.49 | 91.85 | 8.15 |
| | | 2 | 3 | 3 | 4 | 4 | 0.74 (0.11-1.49) | 1.57 (1.46-1.73) | 11.8 | 0.28 | 97.63 | 2.37 | | | | |
| Patient 3 | 3+4 | 1 | 2 | 1 | 3 | 3 | 0.5 (0.16-1.0) | 1.74 (1.52-2.01) | 14.96 | 1.42 | 90.51 | 9.49 | 14.96 | 1.42 | 90.51 | 9.49 |
| Patient 4 | 4+3 | 1 | 3 | 4 | 3 | 3 | 0.41 (0.05-2.03) | 1.88 (1.60-2.03) | 27.45 | 6.69 | 75.63 | 24.37 | 52.15 | 9.87 | 81.06 | 18.94 |
| | | 2 | 4 | 3 | 3 | 4 | 0.95 (0.46-1.51) | 1.66 (1.51-1.82) | 1.91 | 0.25 | 86.92 | 13.08 | | | | |
| Patient 5 | 4+3 | 1 | 4 | 4 | 4 | 4 | 0.39 (0.11-0.91) | 1.90(1.64-2.01) | 28.68 | 11.25 | 60.78 | 39.22 | 40.05 | 12.45 | 68.9 | 31.1 |
| | | 2 | 2 | 2 | 3 | 3 | 0.37 (0.29-0.56) | 1.74 (1.64-1.90) | 9.56 | 0.65 | 93.21 | 6.79 | | | | |
| Patient 6 | 4+3 | 1 | 4 | 5 | 4 | 5 | 0.60 (0.10-2.00) | 1.89 (1.65-20.3) | 18.25 | 4.11 | 77.48 | 22.52 | 21.79 | 5.41 | 75.18 | 24.82 |
| | | 2 | 3 | 3 | 4 | 4 | NA | NA | NA | NA | NA | NA | | | | |
| Patient 7 | 4+3 | 1 | 4 | 5 | 5 | 5 | 0.47 (0.09-1.87) | 1.93 (1.84-2.00) | 27.74 | 8.79 | 68.32 | 31.68 | 29.24 | 9.48 | 67.59 | 32.41 |
| Patient 8 | 4+3 | 1 | 5 | 5 | 5 | 5 | 0.26 (0.08-1.22) | 1.9 (1.22-2.01) | 63.88 | 20.31 | 68.21 | 31.79 | 64.39 | 20.39 | 68.34 | 31.66 |
| | | 2 | 2 | 3 | 3 | 3 | NA | NA | NA | NA | NA | NA | | | | |
| Patient 9 | 4+3 | 1 | 4 | 5 | 5 | 5 | 0.20 (0.06-0.50) | 1.92 (1.83-2.01) | 12.17 | 3.41 | 71.92 | 28.08 | 12.17 | 3.41 | 71.92 | 28.08 |
| | | 2 | 3 | 5 | 4 | 4 | NA | NA | NA | NA | NA | NA | | | | |
| Patient 10 | 3+4 | 1 | 3 | 5 | 4 | 4 | 0.48 (0.14-0.9) | 1.70 (1.56-1.89) | 41.96 | 1.65 | 96.07 | 3.93 | 72.99 | 3.41 | 95.33 | 4.67 |
| | | 2 | 3 | 5 | 5 | 4 | NA | NA | NA | NA | NA | NA | | | | |
| Patient 11 | 4+3 | 1 | 4 | 5 | 5 | 5 | 0.55 (0.10-2.05) | 1.95 (1.65-2.05) | 22.01 | 10.66 | 71.92 | 48.43 | 28.03 | 11.88 | 57.62 | 42.38 |
| | | 2 | 4 | 5 | 5 | 5 | NA | NA | NA | NA | NA | NA | | | | |
| Patient 12 | 4+3 | 1 | 4 | 3 | 5 | 4 | 0.34 (0.16-0.80) | 1.87 (1.60-2.00) | 17.85 | 2.52 | 85.89 | 14.11 | 19.73 | 2.54 | 87.13 | 12.87 |
| | | 2 | 2 | 1 | 3 | 3 | NA | NA | NA | NA | NA | NA | | | | |

Table 5.5 Summary of mpMRI lesions in relation to the evenness and clustering indexes and digitally obtained cancer areas and Gleason 4.

Table containing original Gleason score, index mpMRI lesion and secondary lesion (if reported). mpMRI parameters include Likert score on T2, diffusion (D), contrast (C), and overall Likert (All) per lesion. The clustering and evenness indexes median per lesion and range in parenthesis. Absolute cancer area in mm² as obtained from digital images, calculated proportion of Gleason 3 and 4. Absolute total corresponds to the total cancer area and derived Gleason 3 and 4 from all the cores (as opposed to only the index blocks). NA when no matching cancer block with cancer aligned to the mpMRI.

5.3.2.3. The clustering index correlates to the visual and digital estimation of %G4

With the aim to see if there was any relationship between the clustering/evenness indexes and the visual, digital estimation or difference of the two measurements, scatter plots were created and correlation coefficients calculated (Spearman). In all cases the clustering index had a good correlation between clustering and the visual estimation ($R: 0.61$, $p < 0.0001$), the difference between visual and digital %G4 ($R: 0.33$, $p < 0.0001$) and digital %G4 ($R: 0.59$ ($p < 0.0001$), see Figure 5-11.

There was no correlation between the evenness index and visual %G4% ($R: 0.026$ $p=0.74$), the difference between visual and digital %G4 ($R: 0.011$, $p=0.89$) or the digital %G4 ($R: 0.048$ $P=0.55$, as seen on the right column of Figure 5-13. The values for each patient were plotted separately and can be seen on Appendix 9 (clustering vs evenness), Appendix 10 (visual %G4 vs indexes), Appendix 11 (difference between visual and digital and the indexes) and Appendix 12 (digital %G4 and the indexes).

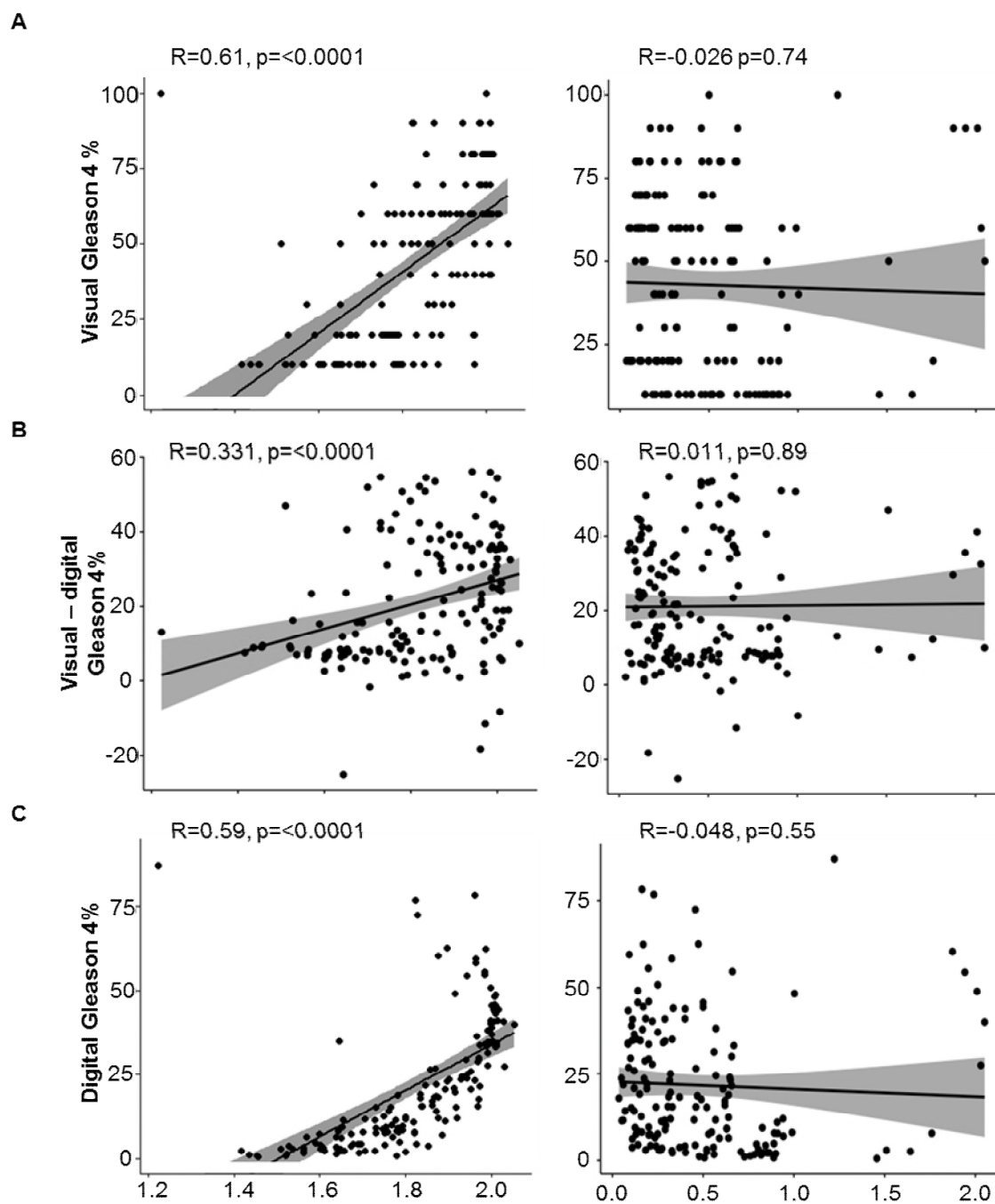


Figure 5-13 Clustering index correlates to visual, digital estimation of Gleason 4 % and the difference between the two measurements.

Scatter plots showing the correlation (or lack of) of the Clustering index on the x-axis (left column) and the evenness index (right column). Confidence interval in grey, regression line in black. A. Correlation to the visual %G4 estimation (y-axis) B. Correlation to the difference between the visual and the digital Gleason 4 % (y-axis). C. Correlation to the digital %G4 estimation (y-axis).

5.4. Discussion

5.4.1. Summary of findings

The first aim of this chapter was to compare the visual measurement of MCCL to a digitally derived measurement. The analysis shows that overall, the mean digital measurement is not significantly different to the integrated ruler measurement, and the measurements show no significant systematic difference between them; however, the variance within the data would have clinical meaning for some patients. The MCCL has importance in prostate cancer, as it can act as a proxy for tumour size. For instance, a threshold of 4mm and 6mm has been shown to correlate with 95% of lesions that have a volume higher than 0.2mL or 0.5mL, respectively (47). Demetrios *et al.* found that MCCL greater than 10mm can predict T3 disease and large tumour volumes with a hazard ratio (HR) of 5.7314 (350). Using these thresholds and taking into account the difference in the MCCL measurements, there is a potential impact on the treatment options offered. For instance, patients reclassified as having <6mm MCCL could be candidates for active surveillance instead of radical therapy (Patients 6, 18, 19 and 20 in Figure 5-3 A). Interestingly, the visual measurement of men with 3+4 disease was more likely to be greater compared to men with 4+3 disease (Figure 5-3 B). Despite these differences, the Bland-Altman analysis showed good concordance between the two measurements; thus, the accuracy of the MCCL is not compromised when a digital tool is used (See Figure 5-4 A).

The second aim was to quantify the Gleason 4 burden in a cohort of men that took part in the PROMIS study, and to compare this to the visual estimation provided by uropathologists. As previously mentioned on section 1.2.1.5, the current approach consists of visual estimation of Gleason 4 percentage and a final Gleason sum based on this percentage; if a threshold of 50% of Gleason 4 is reached, the patient is classified as 4+3, anything below 50% is classified as 3+4. This partition could arbitrarily be classifying patients into two very broad classifications. A more detailed percentage reported could better sub-classify patients and provide better tools for the management decision making. In other words, a patient with very low Gleason 4 could be actively monitored instead of being offered and undergoing surgery, which has risk of side-effects.

The total percentage of Gleason 4 when calculated digitally was much lower than expected (Figure 5-5 A), with only one patient reaching the 50% threshold to be classified as 4+3. The visual estimation of Gleason 4 percentage differed from the digital one; accurate measurement of the G4 burden has been shown to help risk-stratify patients (9,10). In a study by de Souza *et al.*, 20% of Gleason 3+4 tumours had more extensive G4 disease than the first quartile of 4+3 tumours in radical prostatectomy specimens (356). In 2014, Huang *et al.* found that 45% of men with $\leq 5\%$ of G4 in prostate biopsy had insignificant cancer in radical prostatectomy (346). Additionally, several papers have shown that tumours with lower %G4 behave closer to PGG1 tumours (8,9,12,357,358).

The visual estimation always overestimated the amount of G4 compared to a digitally calculated Gleason 4 percentage. For all of these patients, reclassification of the Gleason 4 percentage would potentially lead to a change in treatment options, and imaging follow up. For example, patient 18 was reclassified after digital assessment and would be downgraded from 4+3 of >6mm to 3+4 of <6mm

(Figure 5-5 A). The degree of disagreement between the two measurements, increased with the amount of Gleason 4 (Figure 5-6 A).

When examining the index block (Highest Gleason and/or MCCL), the discrepancy in Gleason 4 percentage remained (Figure 5-7). Additionally, when reclassifying patients using the UCL/PROMIS risk system, four patients were reclassified as lower risk (Figure 5-8).

The third aim was to use geographical descriptors to understand the distribution of Gleason 4 contours in prostate biopsies. Quantitative qualifiers have been used in the past to study tumour heterogeneity and immune cell spatial distribution. Other groups have focused on applying ecological classifiers to analyse immune infiltrate and tumour heterogeneity, in breast cancer in relation to outcome; For instance, analysis of “hotspots” using the Getis-Ord hotspot analysis of immune infiltrate, was found to correlate to better outcome in oestrogen receptor-negative breast cancer (359). Maley *et al.* used the Morisita-Horn index to quantify co-localization of cancer cells and immune cells. This index is used in ecology to measure the similarity among community structure (360,361), when applied to breast cancer, high scores indicated a better prognosis in Her-2 positive cancer (362).

In this work, a different clustering method was applied to prostate cancer biopsies to study the spatial distribution of Gleason 4. Prostate cancer is generally accepted to generate on the peripheral zone in about 80-85% of the cases (22,363,364) Transitional zone tumours are thought to be of lower grade and carry a better outcome (364–366). Here, there was no indication that either cancer area or the clustering/evenness index had any correlation with different zones within the prostate. No pattern of distribution can be found in the prostate in relation to the clustering and evenness of Gleason 4.

Regarding the index block, the evenness index seems to distinguish between cores with low density of Gleason 4 contours (Figure 5-9). However, no correlation was found between the evenness index value and the visual or digital %G4 (Figure 5-13). This may be due to outlier values >2.0 in a few cores, these can be seen on Figure 5-12 B. On the other hand, the clustering index showed a very strong correlation with the pathologist visual estimation of %G4 and the digitally derived measurement (Figure 5-13). Finally, the clustering index was shown to be significantly raised in Likert 5 lesions, compared to Likert 3 and 4 (Figure 5-12).

5.4.2. Methodological limitations

The presence of cribriform pattern was not recorded separately in this study or included in the final analysis. In addition, the pathologists retrospectively assessed %G4 on annotated images, introducing potential bias in their assessment. Additionally, no long-term follow up currently exists for the PROMIS study, this limits the application of the findings in relation to prognostic significance.

Regarding the contouring of Gleason 4, several limitations were identified: lumens and areas of stroma between glands could have had an impact on the difference in the visual and digital estimation, where pathologists can attempt to cognitively ignore this in the estimation of their final percentage; it was included in the digital calculations of area. An extra step to adjust for stromal areas was planned;

however a single value to “correct” for stroma measurement could not be obtained as the amount of stroma varied significantly between all the samples. Additionally, the manual contouring of Gleason 4 is time consuming and varies between pathologists. Future work on this would include several pathologists contouring Gleason 4 areas and comparing the outlines; additionally, previously contoured and non-contoured images could be appraised by several pathologists to try and understand the decision process.

The geographical measurement included a step where a grid was over-imposed to simulate the “neighbourhoods”, ideally each nuclei would have been annotated to avoid this step. The small sample size makes analysis and statistical inference challenging, and although some correlations to zone were significant, caution is advised as to how to interpret this (Table 5.4).

The evenness index when applied to geographical studies, produce a score between 0 and 1. When applied to this dataset the values were as high as 2.05 (Figure 5-11). The reason for this result is unknown, when inspecting cores with high evenness scores (above 1), these were found to be cores with a single Gleason 4 contour (Figure 5-10); a single Gleason 4 core (or minority) may skew the analysis and give a high evenness value. The data analysis to understand the area distribution in the prostate required averaging the evenness and clustering indexes values across cores to obtain a single measurement for a prostate zone. Due to the nested nature of the data, where one patient can have one or 20 blocks with Gleason 4, and each block can have up to 12 positive cores, with up to 50 contours each, a bootstrapping approach or hierarchical clustering approach may have yielded different results.

5.4.3. Clinical application

Integration of %G4 reporting in biopsies and radical prostatectomy specimens is already recommended (343). These findings suggest that a re-assessment of %G4 estimation may be required. Reclassification of G4 could lead to a re-evaluation of previously published biomarker and clinical studies and redefine the reference standard for research. The heterogeneity of studies of the prognostic importance of Gleason 3+4 disease as compared with Gleason 4+3 disease may be a reflection of uncertainty about how much G4 pattern disease is actually shown in specimens and is particularly relevant to treatments such as radiotherapy or ablation where there is no whole mount radical prostatectomy specimen to analyse.

As we move toward the inclusion of digital pathology in standard clinical practice, it will be essential to investigate the differences between human and digital estimation of key pathological parameters and the potential impact this could have on patient care. Whilst the type of analysis used in this work would be currently challenging to embed directly into clinical practice due to the time taken to contour each region; work is already ongoing to automate this process (12,18,367,368). Further research is also needed to develop and validate new thresholds of the burden of G4 against medium and long-term cancer control outcomes.

Application of quantitative methods in clinical practice would require validation of prediction of clinical outcome, and so a cohort with long term follow up. Study of these parameters, perhaps including

ecological measurements in radical prostatectomy specimens or in patients with long-term follow up and mpMRI.

5.5.Lay summary

This chapter focuses on digital analysis of cancer length in prostate biopsy cores, and Gleason 4 disease, and applying simple area calculations using pathologist drawings around the Gleason 4 cancer. The obtained data was then analysed using geographical parameters to try to understand the clustering of Gleason 4. These parameters are called clustering and evenness. Clustering refers to how areas of Gleason 4 are surrounded by areas of Gleason 3, and how tight and big these areas are in comparison to the whole. Evenness refers to how even the Gleason 4 is in proportion of Gleason 3, a more even population has the same amount of Gleason 3 and 4 (50/50).

5.5.1. Objectives

1. Compare the pathologist visual estimation to a digital measure derived from pathologist drawing around Gleason 4.
2. Compare the pathologist cancer length measurement to one obtained using digital measurements.
3. Apply geographical parameters to study the distribution of Gleason 4 in the prostate.

5.5.2. Findings

1. The pathologist measurement overestimates the amount of Gleason 4 in prostate biopsies.
2. The digital and visual cancer length measurement are comparable and there is no systematic difference between the two.
3. The clustering score correlates to the pathologist percentage estimation, that is they give higher Gleason 4 % measurement the more clustered the cancer is.

Take home message: The pathologist percentage Gleason 4 is classifying men as higher risk, possibly leading to unnecessary treatment. Using automated technologies, we could learn about this discrepancy and modify classification accordingly.

4. The clustering correlates with what we see on the MRI, cancers with higher clustering scores are classified by radiologists as high likelihood of cancer.
5. The evenness does not correlate to MRI results.

Take home message: MRI of the prostate is a good tool to detect clustered Gleason 4.

6. Gleason 4 appears and clusters anywhere in the prostate, and has no predilection to any zone. It was previously thought that prostate cancer was limited to the peripheral zone of the prostate. These findings dispute this.

Take home message: Prostate cancer is thought to have area predilection, the random distribution of aggressive cancer in this small cohort shows that aggressive cancer can appear anywhere in the prostate.

5.5.3. Limitations

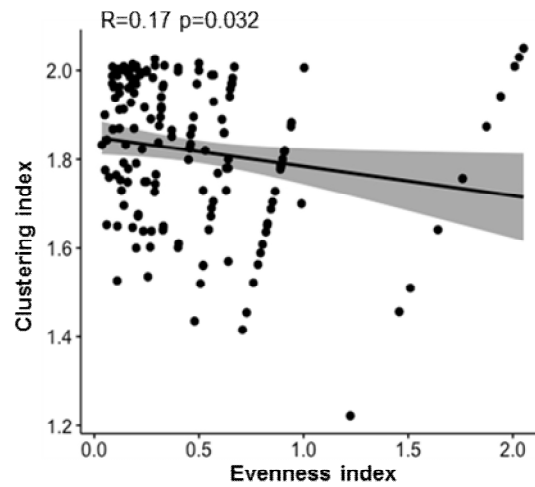
1. The sample size is very small, only 30 men. This is due to the time-consuming element of the analysis. However, automated measurement of Gleason 4 could avoid this hurdle.
2. The evenness index did not correlate to MRI or the pathologist measurement. This may be due to some cores with unusually high evenness scores. The algorithm seems unable to compute cancers with a single Gleason 4 value, and in the future, these should be excluded from the analysis.
3. Exclusion of non-cancerous cells (stroma) should be excluded from future percentage calculations. This may lower the discrepancy between the pathologist and the digital percentage.
4. Analysis of this kind should be correlated to long-term outcomes, and should be repeated in cohorts with death and recurrence data, to determine if the findings are clinically significant.

5.5.4. Future work

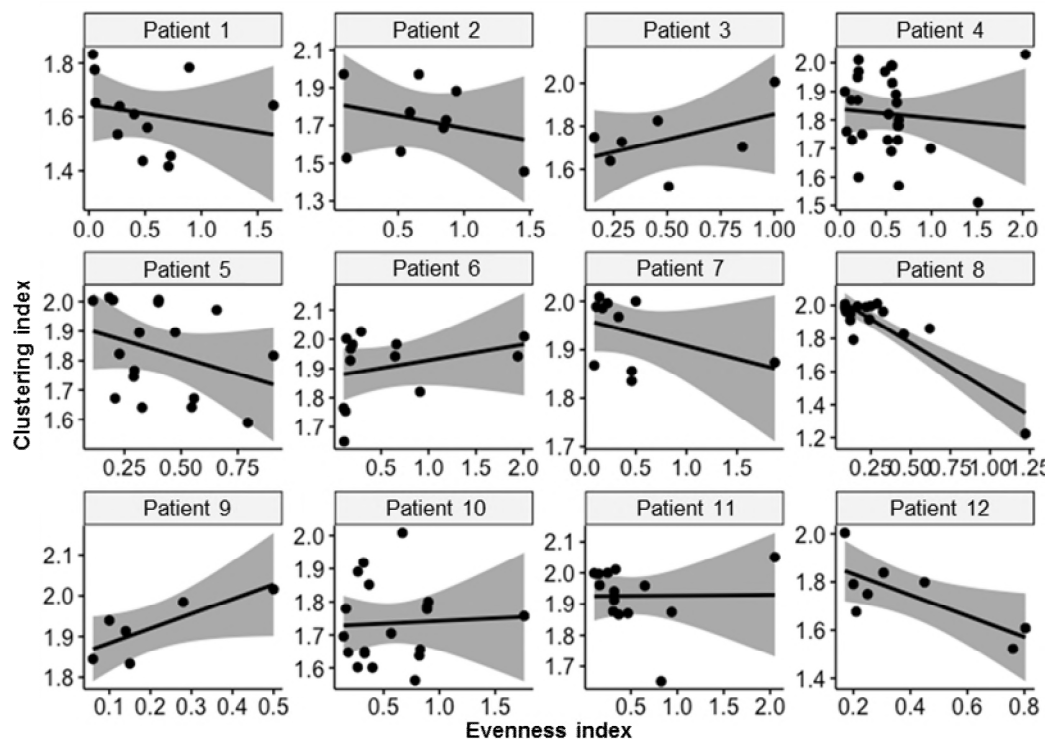
Based on the findings of this thesis, efforts should focus on understanding the role of PSA density as an adjunct biomarker to MRI to decide which patients will benefit from a biopsy. Ideally, a direct comparison between commercially available biomarkers and their performance in this cohort would better inform us on whether they perform better than PSA density.

Automation of Gleason patterns would reduce variability. Training machines through machine learning algorithms has already begun, but most of these systems are trained to give a Gleason sum score instead of percentage of Gleason. Systems would need to be taught to provide percentage estimations. The geographical parameters are already automated, and so in a larger cohort that includes death and recurrence outcomes would provide information on the biological impact of clustering and evenness.

A



B

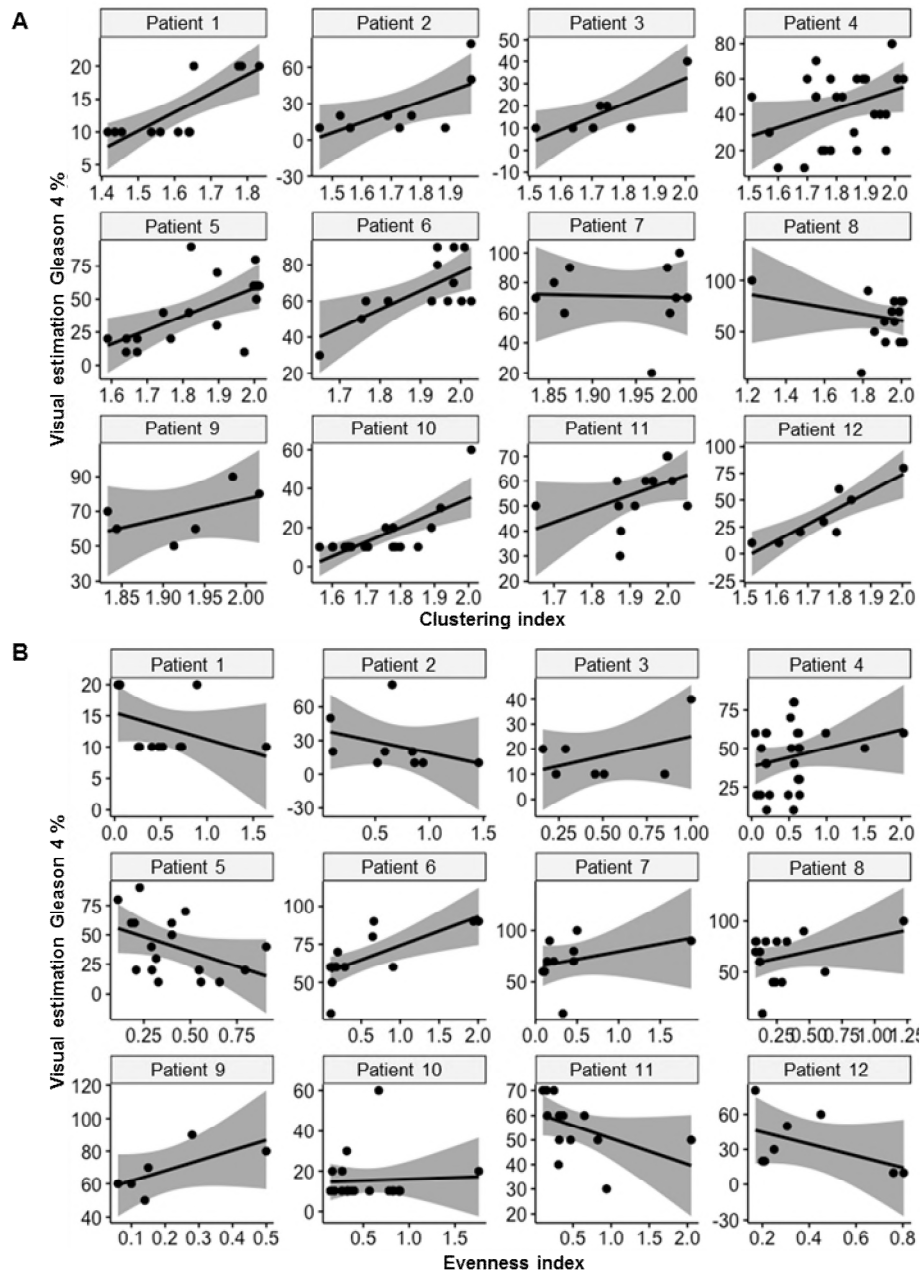


C

| Patient | Correlation coefficient (R) | p value |
|------------|-----------------------------|---------|
| Patient 1 | -0.22 | 0.49 |
| Patient 2 | -0.29 | 0.44 |
| Patient 3 | 0.49 | 0.26 |
| Patient 4 | 0.096 | 0.64 |
| Patient 5 | 0.33 | 0.19 |
| Patient 6 | 0.31 | 0.31 |
| Patient 7 | -0.4 | 0.25 |
| Patient 8 | -0.88 | <0.0001 |
| Patient 9 | 0.79 | 0.06 |
| Patient 10 | 0.053 | 0.83 |
| Patient 11 | 0.012 | 0.97 |
| Patient 12 | 0.75 | 0.03 |

Appendix 9 There is a very weak correlation between the clustering and evenness indexes.

A. Scatter plot showing the correlation (or lack of) of the Clustering index on the y-axis and the evenness index on the x-axis. Confidence interval in grey, regression line in black. **B.** Scatter plots of evenness index on the x-axis and the clustering index on the y-axis by patient. **C.** Correlation coefficients (R), and p value per patient plotted on B.



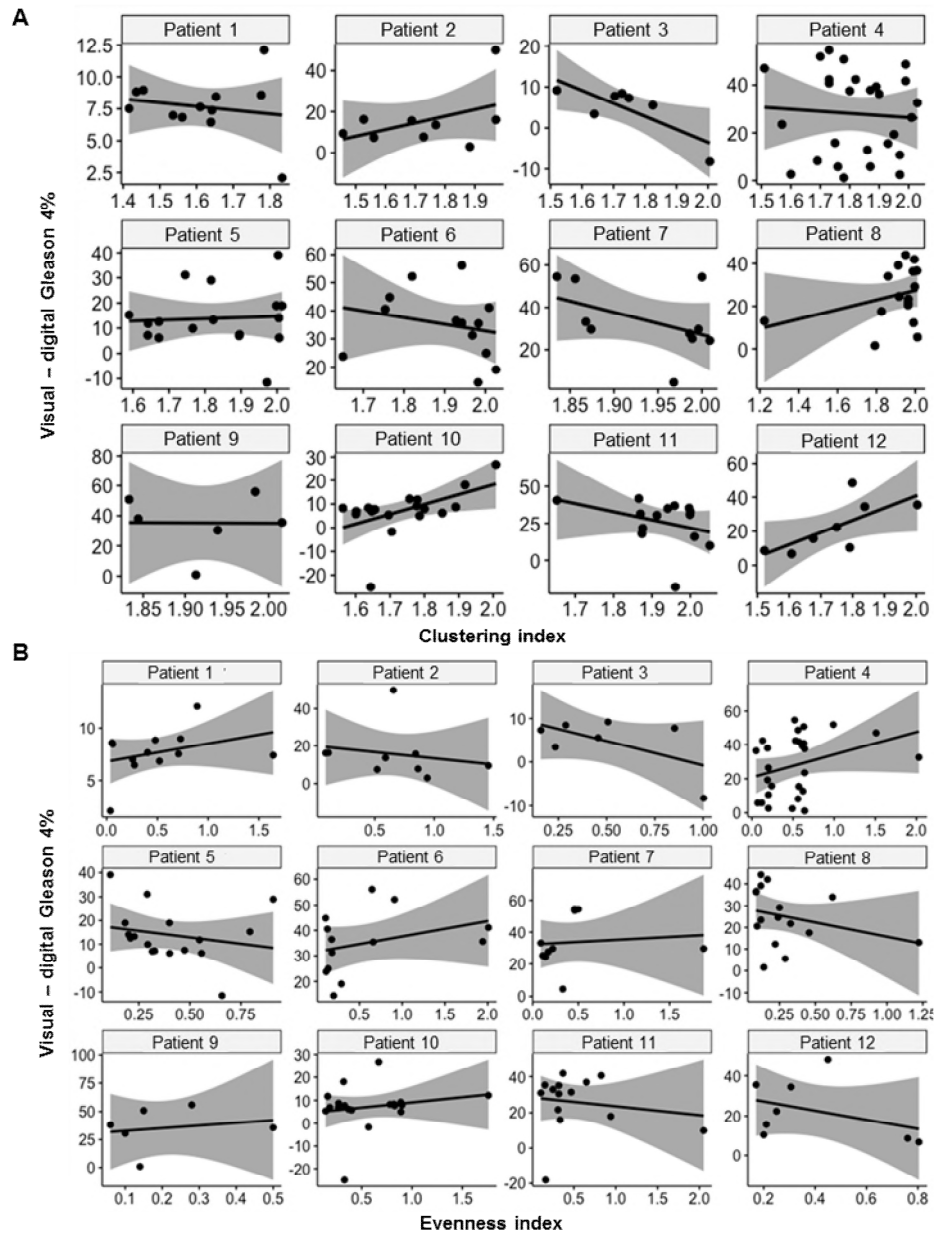
C

| Patient | Clustering index | | Evenness index | |
|------------|-----------------------------|---------|-----------------------------|---------|
| | Correlation coefficient (R) | p value | Correlation coefficient (R) | p value |
| Patient 1 | 0.8 | 0.0018 | -0.4 | 0.2 |
| Patient 2 | 0.68 | 0.045 | -0.36 | 0.34 |
| Patient 3 | 0.79 | 0.034 | 0.45 | 0.31 |
| Patient 4 | 0.36 | 0.068 | 0.25 | 0.2 |
| Patient 5 | 0.62 | 0.0075 | -0.44 | 0.078 |
| Patient 6 | 0.69 | 0.0094 | 0.72 | 0.005 |
| Patient 7 | 0.042 | 0.042 | 0.35 | 0.32 |
| Patient 8 | -0.27 | 0.32 | 0.54 | 0.19 |
| Patient 9 | 0.55 | 0.26 | 0.69 | 0.13 |
| Patient 10 | 0.73 | 0.0003 | 0.052 | 0.83 |
| Patient 11 | 0.46 | 0.094 | -0.46 | 0.1 |
| Patient 12 | 0.89 | 0.093 | -0.5 | 0.2 |

Appendix 10 There is correlation between the clustering index and the visual estimation but not with evenness.

A. Scatter plot showing the correlation (or lack of) of the Clustering index on the x-axis and the visual Gleason 4 % estimation on the x-axis, per patient. Confidence interval in grey, regression line in black.

B. Scatter plot showing the correlation (or lack of) of the Evenness index on the x-axis and the visual Gleason 4 % estimation on the x-axis, per patient. C. Correlation coefficients (R), and p value per patient plotted on A and B.

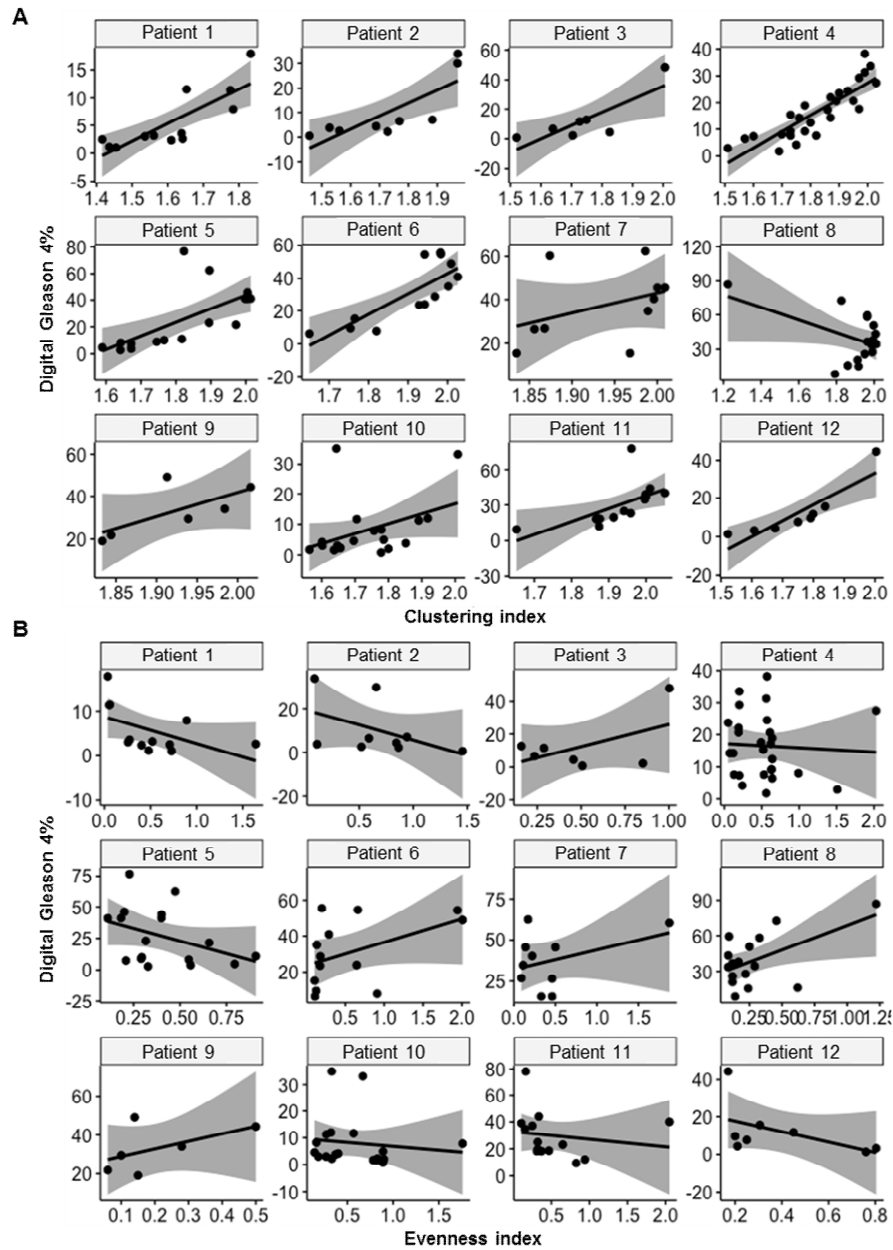


C

| Patient | Clustering index | | Evenness index | |
|------------|-----------------------------|---------|-----------------------------|---------|
| | Correlation coefficient (R) | p value | Correlation coefficient (R) | p value |
| Patient 1 | -0.18 | 0.58 | 0.34 | 0.28 |
| Patient 2 | 0.45 | 0.23 | -0.2 | 0.6 |
| Patient 3 | 0.79 | 0.03 | -0.58 | 0.17 |
| Patient 4 | 0.06 | 0.73 | 0.93 | 0.096 |
| Patient 5 | 0.06 | 0.82 | -0.21 | 0.41 |
| Patient 6 | -0.26 | 0.45 | 0.33 | 0.27 |
| Patient 7 | -0.46 | 0.33 | 0.1 | 0.78 |
| Patient 8 | 0.33 | 0.21 | -0.29 | 0.27 |
| Patient 9 | -0.0098 | 0.9 | 0.2 | 0.71 |
| Patient 10 | 0.52 | 0.02 | 0.18 | 0.45 |
| Patient 11 | -0.34 | 0.23 | -0.16 | 0.59 |
| Patient 12 | 0.7 | 0.05 | -0.38 | 0.35 |

Appendix 11 There is correlation between the clustering index and the difference between the visual and digital estimation but not with evenness.

A. Scatter plot showing the correlation (or lack of) of the Clustering index on the x-axis and the difference between the visual and digital Gleason 4 % estimation on the x-axis, per patient. Confidence interval in grey, regression line in black. **B.** Scatter plot showing the correlation (or lack of) of the Evenness index on the x-axis and the visual Gleason 4 % estimation on the x-axis, per patient. **C.** Correlation coefficients (R), and p value per patient plotted on A and B.



| Patient | Clustering index | | Evenness index | |
|------------|-----------------------------|---------|-----------------------------|---------|
| | Correlation coefficient (R) | p value | Correlation coefficient (R) | p value |
| Patient 1 | 0.81 | 0.0013 | -0.51 | 0.089 |
| Patient 2 | 0.8 | 0.0091 | -0.47 | 0.2 |
| Patient 3 | 0.83 | 0.022 | 0.52 | 0.23 |
| Patient 4 | 0.86 | <0.0001 | -0.057 | 0.78 |
| Patient 5 | 0.68 | 0.0027 | -0.39 | 0.12 |
| Patient 6 | 0.81 | 0.0088 | 0.47 | 0.11 |
| Patient 7 | 0.39 | 0.27 | 0.38 | 0.28 |
| Patient 8 | -0.48 | 0.058 | 0.55 | 0.029 |
| Patient 9 | 0.69 | 0.13 | 0.53 | 0.26 |
| Patient 10 | 0.4 | 0.089 | -0.12 | 0.64 |
| Patient 11 | 0.61 | 0.021 | -0.16 | 0.59 |
| Patient 12 | 0.88 | 0.0036 | -0.51 | 0.2 |

Appendix 12 There is correlation between the clustering index and the digital %G4 but not with evenness.

A. Scatter plot showing the correlation (or lack of) of the Clustering index on the x-axis and the digital Gleason 4 % estimation on the x-axis, per patient. Confidence interval in grey, regression line in black.
B. Scatter plot showing the correlation (or lack of) of the Evenness index on the x-axis and the digital Gleason 4 % estimation on the x-axis, per patient. C. Correlation coefficients (R), and p value per patient plotted on **A** and **B**

Chapter 6. General discussion

There is a need for a test that can accurately differentiate between men with and without prostate cancer, filling a gap in the diagnostic pathway of prostate cancer. More importantly, this test should differentiate between patients with the disease that require treatment and those who do not, to avoid unnecessary biopsies, treatment and side effects. Although mpMRI has markedly improved the diagnostic pathway, a small proportion of cancers are missed. Currently, mpMRI cannot be performed to the necessary standard in all centres, therefore there is great interest in finding a test measurable in blood or urine that can aid the biopsy decision making process.

Many companies have attempted to find such test, using PSA isoforms (PHI, 4K test) and genetic testing in blood and urine (Section 1.2.2). Unfortunately, uptake of these commercial biomarkers has been slow with very limited impact in patient care. This thesis had two main aims. First, to describe the discovery and validation of a new biomarker panel quantifiable in urine or blood, that would accurately detect men with Gleason 4+3 or ≥ 6 mm prostate cancer, while improving the accuracy of mpMRI independently of available clinical data. The second aim, was to perform a morphometric analysis of Gleason 4 distribution in prostate biopsies while comparing this to the current gold standard.

6.1. Main findings presented in this thesis

6.1.1. New biomarker panel is able to detect men with Gleason 4+3 and/or ≥ 6 mm but is not better than PSA density

The described gene panel was able to detect all cancers and clinically significant cancer using both PROMIS definitions (definition one, Gleason 4+3 or ≥ 6 mm and definition two, Gleason 3+4 or ≥ 4 mm). The panel was outperformed by PSA density in both the training and validation cohort (Figure 4-9 and 4-11).

6.1.2. Gleason 4 percentage is overestimated in the visual assessment

When comparing digitally measured Gleason 4 percentage to visual assessment by uropathologists, the visual measurement always overestimated the percentage of Gleason 4 present in a small cohort of men.

6.1.3. Gleason 4 is clustered in all regions of the prostate and correlated to mpMRI visibility

The clustering of Gleason 4 is random across the entire prostate and has no predilection for zone. Additionally, the clustering of Gleason 4 is correlated to higher Likert scores on mpMRI.

6.2. Methodological limitations

6.2.1. Biomarker panel discovery

The panel described here was developed based on tissue based biomarkers and local expertise, followed by data mining of publicly available data. The Sharma *et al.* paper developed these markers as means to find metastatic disease and not as a diagnostic tool. From the Whitaker lab signature, COL6A3, SLC25A3 and CBLN2 have been found to be higher on metastatic disease, a vastly different disease status than what the panel is aiming to find. While other panels have used tissue based biomarkers translated into urine (Select MDx), the utility of the 28 genes for diagnosis rather than risk stratification was not assessed prior to proceeding to detect in fluids.

Another limitation to the biomarker discovery described in this thesis is the use of cell lines as means to study the biological impact of these genes. Whilst the data obtained is valuable to understand the androgen dependence or independence of the genes is not necessarily reflective of the biological behaviour of these genes in patients.

The study of the utility of these genes in patients was initially performed in publicly available datasets (Taylor, Ross, and Varambally). Several limitations with this approach have been discussed in chapter 2. The main concern with this validation on retrospective datasets is the lack of “normal” samples to compare the gene expression in order to use as a diagnostic tool. Additionally, sequencing technology greatly differs from dataset to dataset, meaning significant gene expression in one cohort may be insignificant in another with better probe alignment.

Ideally, the first step in this would have included confirmation of expression in tissue samples from the PROMIS cohort. This would have allowed a comparison between true negative men (no inflammation, ASAP or HGPIN), and men with cancer. Followed by measurement in blood and urine of the genes to be able to determine if the marker detected in fluids truly corresponds to that secreted by the tumour. All cohorts used in this work are independent; therefore, it is not possible to confirm the origin of RNA measured in the fluidic samples. This has implications not only in the work presented here, but on how biomarkers are developed.

It is generally assumed that RNA measured in fluids corresponds to the disease status of the donor, in other words the RNA signatures detected in blood or urine reflect the status of the tumour regardless of their location. Whilst there is evidence of the potential source of circulating nucleic acids, including necrosis in advanced tumours, apoptotic cells and bodies (369), no such evidence exists for localised tumours. Biomarker discovery in fluids seems to be flawed on the principle of what is being detected and our lack of understanding of this mechanism. Taking advantage of the PROMIS samples resource to attempt to investigate the origin of what is measured in these fluids could have been attempted here, comparing the RNA relative expression in tissue, blood and urine of the same patients while comparing this profile to men who have no histological disease on upwards to 60 biopsy samples per patient.

Concretely, this work is limited by the assumption that tissue expression in historically obtained samples (from publicly available data), obtained with different technologies is comparable to the RNA extracted

from fluids, while being unable to prove the RNA detected truly represents a disease state. Additionally, it is not possible to determine if the urine results were true or false negatives as there is no standard of expression. That is it is not possible to establish that men with no expression of genes did in reality not express the gene of interest, as opposed to the gene was not measured due to quality of the RNA or the technology used.

6.2.2. Urine as a fluidic biomarker

Urine is an attractive fluid for biomarker discovery. Several issues were identified in the work presented in this thesis. First, the urine was highly degraded, making downstream analysis challenging. Despite attempting to simulate the degradation of cells exposed in urine, followed by pre-amplification to overcome the degradation, the resulting data was un-evenly amplified. For example, genes that are meant to act as housekeepers were undetected following pre-amplification. Following PCR, a large amount of missing data was present; making meaningful analysis of the results impossible.

When urine is used for biomarker discovery, it should be processed to allow for the highest yield of RNA to be obtained. Although an attempt was made to extract high concentrations of RNA from these samples, it is unclear why the quality of the RNA was so poor. One possibility is that no RNA stabiliser was used on the urine samples. Additionally, the urine extraction method selected has a high risk of ethanol contamination as it is left to air dry. Other methods have shown better RNA yields and quality (370). Additionally, the selection of endogenous housekeepers for this work should have included a larger number of candidates, this could have possibly allowed for use of pre-amplified urine samples instead of discarding the data altogether.

6.2.3. Obtaining meaningful outcomes from cohorts with limited follow up

In this thesis, the comparison between digitally contoured Gleason 4 areas and visual estimation of percentage showed overestimation of the latter. Only one patient met the 50% threshold to be classified 4+3. Within the methodological limitations already discussed in Chapter 5, it is difficult to define the clinical impact of this discrepancy. Not only the cohort studied is small, these patients have limited follow up. Therefore, it is not possible to determine if that patient with 50% of Gleason 4 had a worse outcome compared to one with 30% of Gleason 4. Additionally, no comparison was made with men who had a radical prostatectomy, where the 'real' total amount of Gleason 4 could have been measured.

The morphometric analysis using geographical descriptors has a similar limitation, as it is not possible to determine if patients with higher clustering have a worse outcome. The results presented here are descriptive and no clinical impact analysis can be made based on them. However, the process of analysis of clustering and evenness has been automated and can be easily applied to a larger dataset of contoured digital images.

6.2.4. Aim of the biomarker panel

The panel aimed to help determine which patients would benefit from a prostate biopsy, by identifying Gleason $\geq 4+3$ or 6 mm. the discovery cohort (PROMIS) lacked patients with PGG 4 and 5, therefore,

the panel was aiming to find Gleason 4+3 or ≥ 6 mm specifically. The UCLH PROMIS cohort reflects patients referred by a GP to urological review, but lacks men with more aggressive disease. The validation cohort (INOVATE) had patients with higher Gleason scores (5 patients with PGG5 out of 29), but was a smaller cohort compared to PROMIS (99 versus 194). Taking this into account, the cohorts available for this validation lacked sufficient patients with more aggressive disease, and may explain why the second panel of three genes (aiming to classify patients by PGG) failed to perform well.

Additionally, mpMRI can detect clinically significant disease (47,60,63). While this was not the aim of the panel, an alternative question would be to improve mpMRI by specifically focusing on men with a negative mpMRI harbouring clinically significant disease. Simply put, a panel that specifically targets invisible cancers. Although this group of patients is relatively small, a panel that finds “hidden” cancers would reduce improve cancer detection while shedding light on patients with indeterminate Likert scores.

6.2.5. Exclusion of clinical information in the biomarker panel performance review

Many commercial panels include clinical information, such as age, DRE, previous biopsies and family history to improve the performance of new panels. Here, an attempt was made to create a panel that would independently predict aggressive disease, without the need for advanced algorithms. Such a test could be performed in developing countries, where prostate cancer is under-diagnosed and where clinicians are forced to rely on surgical castration as means of treatment. A separate analysis could have been performed to understand the impact of these clinical variables on the final performance of the panel, quite possibly improving its diagnostic accuracy.

6.3. Research implications

6.3.1. Biomarker discovery

The race for finding new biomarker panels for prostate cancer has not yet found a replacement PSA. While is not the aim of this thesis to explain why this is the case, it is worth discussing the impact of some of the findings in this context. Most of the biomarker panels presented on the introduction, aim to find Gleason 7 disease. Seeing as Gleason 4 has an impact in patient outcome, accurate determination of Gleason 4 burden is of vital importance. While manual contouring of Gleason 4 in large cohorts is unrealistic, new machine learning algorithms have managed to determine different Gleason patterns in digitalised images (17,18,20,89,371). Training of new algorithms to calculate percentage of Gleason 4 and comparing this to visual estimation may lead to advancing our understanding of the impact of Gleason 4 in biomarkers and patient outcome. This may be achieved by performing a similar analysis in radical prostatectomy cohorts with long term follow up; due to the long natural history of prostate cancer alternative surrogates for overall survival could be used, such as metastasis free survival (372), followed by validation of biomarkers that aim to find Gleason 4.

6.3.2. Investigating the utility of PSA density

The performance of the panel was compared to PSA and PSA density as a point of reference. PSA density outperformed the panel, and PSA alone. The AUC was improved by the addition of PSA density to the Likert score, from 0.82 (95% CI 0.77-0.88) to 0.85 (95% CI 0.83-0.87) in the discovery cohort. In the validation cohort the AUC remained at 0.84 but the addition of PSA density reduced the 95% CI from 0.75-0.92 to 0.83-0.86.

Although the potential utility of PSA density as an adjuvant to mpMRI has been explored (327,373,374), further work is required in identifying the optimal cut-off point. While this was not a planned outcome of this work, PSA density may already offer the discriminatory and complementary characteristics that were looked for as an aim to this thesis.

6.3.3. Encouraging collaboration with physical sciences to advance our understanding of cancer

Through a collaboration with the Geography department, a quantitative description of a biological phenomenon was achieved. The spatial distribution of Gleason 4 was described using metrics developed for studying populations. These metrics correlated to mpMRI visibility, while analysis of their spatial distribution in the prostate demonstrated that Gleason 4 not only appears in every studied anatomical zone of the prostate, it is not more likely to cluster in one area compared to another. While the notion that anterior tumours are less aggressive, these findings challenge this and point to another characteristic driving the aggressiveness of anterior tumours, possibly pertaining to volume or multiplicity of lesions rather than Gleason alone.

Exploring the utility of such collaborations, where researchers from different backgrounds attempt to understand a problem may further our understanding of cancer.

6.4. Future research

6.4.1. Biomarkers development using in the PROMIS cohort

As mentioned previously, the PROMIS cohort benefits from mpMRI data, blood and urine samples and valuable tissue from systematic templated biopsies of the prostate. This valuable resource can be used to attempt to answer the question on whether what we are measuring in blood or urine truly corresponds to a higher expression in tissue. Advantage can be taken of the cohort of men in PROMIS that had no histologically observable abnormality in the examined tissue, these men are unique and can be used as a normal control instead of relying on benign tissue from radical prostatectomies. Steps to begin such analysis have already begun. A tissue microarray was constructed containing a total of 83 men with Gleason sum 7, 28 men with Gleason sum 6 and five men with no disease. The patients were randomised to one of six TMAs using a random number generator. Each TMA consists of 15 men with Gleason sum 7, five men with Gleason sum 6, five random benign cores and two liver controls. This resource can be used for analysis of protein expression using immunohistochemistry or targeted genome sequencing. The obtained data is then compared to expression data in blood and urine.

6.4.2. Study of Gleason 4 percentage estimation and impact on patient outcome

Application of an automated Gleason 4 percentage estimation can be achieved by using already developed machine learning algorithms. Digitised images of the PROMIS cohort, including that of patient who have had radical prostatectomy can then be sent to uropathologists nationwide to determine visually Gleason 4 percentage prior to and after contouring of Gleason 4. This can then be compared to prove or disprove the findings presented in this thesis. If the overestimation of visually assessed Gleason 4 is confirmed in larger datasets, reasons for this discrepancy must be discussed.

If these findings are confirmed, analysis of radical prostatectomy cohorts and long term follow up can proceed. Aiming to determine the impact of Gleason 4 percentage estimation, ideally in the context of quantitative methods, such as the geographical parameters described here, or in collaboration with other earth and biological sciences.

Bibliography

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015 Mar;65(2):87–108.
2. Fitzmaurice C, Abate D, Abbasi N, Abbastabar H, Abd-Allah F, Abdel-Rahman O, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-Adjusted life-years for 29 cancer groups, 1990 to 2017: A systematic analysis for the global burden of disease study. *JAMA Oncol*. 2019 Dec 1;5(12):1749–68.
3. NICE. Overview | Prostate cancer: diagnosis and management | Guidance | NICE [Internet]. 2019. NICE; [cited 2020 May 28]. Available from: <https://www.nice.org.uk/guidance/ng131>
4. Loeb S, Bjurlin MA, Nicholson J, Tammela TL, Penson DF, Carter HB, et al. Overdiagnosis and overtreatment of prostate cancer. Vol. 65, *European Urology*. Elsevier; 2014. p. 1046–55.
5. Cucchiara V, Cooperberg MR, Dall'era M, Lin DW, Montorsi F, Schalken JA, et al. Genomic Markers in Prostate Cancer Decision Making [Internet]. *European Urology Elsevier B.V.*; Apr 1, 2018 p. 572–82. Available from: <https://doi.org/10.1016/j.eururo.2017.10.036>
6. Nevo A, Navaratnam A, Andrews P. Prostate cancer and the role of biomarkers. *Abdom Radiol* [Internet]. 2020 Jul 1 [cited 2020 Jun 26];45(7):2120–32. Available from: <https://pubmed.ncbi.nlm.nih.gov/31720770/>
7. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The evaluation of genomic applications in practice and prevention (EGAPP) initiative: Methods of the EGAPP working group. *Genet Med*. 2009 Jan;11(1):3–14.
8. Pierorazio PM, Walsh PC, Partin AW, Epstein JI, Epstein J. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. *BJU Int*. 2013;111(5):753–60.
9. Sauter G, Steurer S, Clauditz TS, Krech T, Wittmer C, Lutz F, et al. Clinical utility of quantitative gleason grading in prostate biopsies and prostatectomy specimens. *Eur Urol* [Internet]. 2016 [cited 2019 Dec 16];69(4):592–8. Available from: <http://dx.doi.org/10.1016/j.eururo.2015.10.029>
10. Sharma M, Miyamoto H. Percent Gleason pattern 4 in stratifying the prognosis of patients with intermediate-risk prostate cancer. *Transl Androl Urol*. 2018 Sep 1;7(Suppl 4):S484–9.
11. Herman CM, Kattan MW, Otori M, Scardino PT, Wheeler TM. Primary Gleason pattern as a predictor of disease progression in Gleason score 7 prostate cancer: A multivariate analysis of 823 men treated with radical prostatectomy. *Am J Surg Pathol*. 2001;25(5):657–60.
12. Khoddami SM, Shariat SF, Lotan Y, Saboorian H, McConnell JD, Sagalowsky AI, et al. Predictive value of primary Gleason pattern 4 in patients with Gleason score 7 tumours treated with radical prostatectomy. *BJU Int* [Internet]. 2004 Jul [cited 2019 Dec 16];94(1):42–6. Available

from: <http://doi.wiley.com/10.1111/j.1464-410X.2004.04898.x>

13. Berney DM, Beltran L, Sandu H, Soosay G, Møller H, Scardino P, et al. The percentage of high grade disease in prostate biopsies significantly improves on grade groups in prediction of prostate cancer death. *Histopathology* [Internet]. 2019 Apr 29 [cited 2019 Jul 12];his.13888. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/his.13888>
14. Engers R. Reproducibility and reliability of tumor grading in urological neoplasms. Vol. 25, *World Journal of Urology*. World J Urol; 2007. p. 595–605.
15. Melia J, Moseley R, Ball RY, Griffiths DFR, Grigor K, Harnden P, et al. A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. *Histopathology*. 2006 May;48(6):644–54.
16. Yuan Y, Failmezger H, Rueda OM, Raza Ali H, Gräf S, Chin SF, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med*. 2012 Oct 24;4(157).
17. Egevad L, Swanberg D, Delahunt B, Ström P, Kartasalo K, Olsson H, et al. Identification of areas of grading difficulties in prostate cancer and comparison with artificial intelligence assisted grading. *Virchows Arch* [Internet]. 2020 Jun 15 [cited 2020 Jun 18]; Available from: <http://link.springer.com/10.1007/s00428-020-02858-w>
18. Ström P, Kartasalo K, Olsson H, Solorzano L, Delahunt B, Berney DM, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol* [Internet]. 2020 Jan 8 [cited 2020 Jan 23]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31926806>
19. Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol*. 2020 Feb 1;21(2):233–41.
20. Hossain A, Arimura H, Kinoshita F, Ninomiya K, Watanabe S, Imada K, et al. Automated approach for estimation of grade groups for prostate cancer based on histological image feature analysis. *Prostate*. 2020 Feb 1;80(3):291–302.
21. Yan C, Nakane K, Wang X, Fu Y, Lu H, Fan X, et al. Automated gleason grading on prostate biopsy slides by statistical representations of homology profile. *Comput Methods Programs Biomed* [Internet]. 2020 Oct 1 [cited 2020 Jun 26];194. Available from: <https://pubmed.ncbi.nlm.nih.gov/32470903/>
22. McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *Am J Surg Pathol*. 1988;12(12):897–906.

23. Barentsz JO, Weinreb JC, Verma S, Thoeny HC, Tempny CM, Shtern F, et al. Synopsis of the PI-RADS v2 Guidelines for Multiparametric Prostate Magnetic Resonance Imaging and Recommendations for Use HHS Public Access. *Eur Urol* [Internet]. 2016 [cited 2020 Jun 2];69(1):41–9. Available from: <http://www.acr.org/Quality-Safety/Resources/>
24. Barry MJ, Simmons LH. Prevention of Prostate Cancer Morbidity and Mortality: Primary Prevention and Early Detection. Vol. 101, *Medical Clinics of North America*. W.B. Saunders; 2017. p. 787–806.
25. Howlader, N.; Noone, AM.; Krapcho, M. et al. Cancer Statistics Review, 1975-2014 - SEER Statistics [Internet]. SEER Cancer Statistics Review - Bethesda (MD): National Cancer Institute. 2011 [cited 2020 Jun 1]. p. 1975_2008. Available from: https://seer.cancer.gov/csr/1975_2017/
26. Bratt O. Hereditary prostate cancer: Clinical aspects. Vol. 168, *Journal of Urology*. J Urol; 2002. p. 906–13.
27. Mcginley KF, Tay KJ, Moul JW. Prostate cancer in men of African origin. Vol. 13, *Nature Reviews Urology*. Nature Publishing Group; 2016. p. 99–107.
28. Rebbeck TR. Prostate Cancer Genetics: Variation by Race, Ethnicity, and Geography. Vol. 27, *Seminars in Radiation Oncology*. W.B. Saunders; 2017. p. 3–10.
29. Lebdaï S, Mathieu R, Leger J, Haillet O, Vincendeau S, Rioux-Leclercq N, et al. Metabolic syndrome and low high-density lipoprotein cholesterol are associated with adverse pathological features in patients with prostate cancer treated by radical prostatectomy. *Urol Oncol* [Internet]. 2018 Feb 1 [cited 2020 Jun 1];36(2):80.e17-80.e24. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29153942>
30. Alshaker H, Sacco K, Alfraidi A, Muhammad A, Winkler M, Pchejetski D. Leptin signalling, obesity and prostate cancer: Molecular and clinical perspective on the old dilemma. *Oncotarget* [Internet]. 2015 [cited 2020 Jun 26];6(34):35556–63. Available from: <https://pubmed.ncbi.nlm.nih.gov/26376613/>
31. Caliskan S, Kaba S, Özsoy E, Keleş MO, Koca O, Akyüz M, et al. The effect of metabolic syndrome on prostate cancer final pathology. *J Cancer Res Ther*. 2019 Mar 1;15(8):S47–50.
32. Esposito K, Chiodini P, Capuano A, Bellastella G, Maiorino MI, Parretta E, et al. Effect of metabolic syndrome and its components on prostate cancer risk: Meta-analysis. Vol. 36, *Journal of Endocrinological Investigation*. Springer International Publishing; 2013. p. 132–9.
33. Liang Z, Xie B, Li J, Wang X, Wang S, Meng S, et al. Hypertension and risk of prostate cancer: A systematic review and meta-analysis. Vol. 6, *Scientific Reports*. Nature Publishing Group; 2016.
34. Huggins C, Hodges C V. *Studies on Prostatic Cancer I. The Effect of Castration, of Estrogen*

- and of Androgen Injection on Serum Phosphatases in Metastatic Carcinoma of the Prostate*. 1941.
35. Sinowatz F, Amselgruber W, Plendl J, Kölle S, Neumüller C, Boos G. Effects of hormones on the prostate in adult and aging men and animals. *Microsc Res Tech* [Internet]. 1995 Mar 1 [cited 2020 Jun 1];30(4):282–92. Available from: <http://doi.wiley.com/10.1002/jemt.1070300404>
 36. Davey RA, Grossmann M. Androgen Receptor Structure, Function and Biology: From Bench to Bedside. Vol. 37, *Androgen Receptor Biology Clin Biochem Rev*. 2016.
 37. Massie CE, Lynch A, Ramos-Montoya A, Boren J, Stark R, Fazli L, et al. The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *EMBO J* [Internet]. 2011/05/24. 2011;30(13):2719–33. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21602788>
 38. De Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med*. 2011 May 26;364(21):1995–2005.
 39. Graf RP, Hullings M, Barnett ES, Carbone E, Dittamore R, Scher HI. Clinical Utility of the Nuclear-localized AR-V7 Biomarker in Circulating Tumor Cells in Improving Physician Treatment Choice in Castration-resistant Prostate Cancer. *Eur Urol*. 2020 Feb 1;77(2):170–7.
 40. Scher HI, Graf RP, Schreiber NA, McLaughlin B, Lu D, Louw J, et al. Nuclear-specific AR-V7 Protein Localization is Necessary to Guide Treatment Selection in Metastatic Castration-resistant Prostate Cancer. *Eur Urol*. 2017 Jun 1;71(6):874–82.
 41. Antonarakis ES, Lu C, Luber B, Wang H, Chen Y, Nakazawa M, et al. Androgen receptor splice variant 7 and efficacy of taxane chemotherapy in patients with metastatic castration-resistant prostate cancer. *JAMA Oncol* [Internet]. 2015 Aug 1 [cited 2020 Jun 26];1(5):582–91. Available from: <https://pubmed.ncbi.nlm.nih.gov/26181238/>
 42. 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* [Internet]. 1987 Oct 8 [cited 2020 Jun 1];317(15):909–16. Available from: <http://www.nejm.org/doi/abs/10.1056/NEJM198710083171501>
 43. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJJ, et al. Measurement of Prostate-Specific Antigen in Serum as a Screening Test for Prostate Cancer. *N Engl J Med* [Internet]. 1991 Apr 25 [cited 2020 Jun 1];324(17):1156–61. Available from: <http://www.nejm.org/doi/abs/10.1056/NEJM199104253241702>
 44. Etzioni R, Tsodikov A, Mariotto A, Szabo A, Falcon S, Wegelin J, et al. Quantifying the role of PSA screening in the US prostate cancer mortality decline. *Cancer Causes Control*. 2008 Mar;19(2):175–81.

45. Bittner N, Merrick G, Taira A, Bennett A, Schattel A, Butler W, et al. Location and Grade of Prostate Cancer Diagnosed by Transperineal Template-guided Mapping Biopsy after Negative Transrectal Ultrasound-guided Biopsy. *Am J Clin Oncol Cancer Clin Trials*. 2018 Aug 1;41(8):723–9.
46. Omer A, Lamb AD. Optimizing prostate biopsy techniques. Vol. 29, *Current Opinion in Urology*. Lippincott Williams and Wilkins; 2019. p. 578–86.
47. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet [Internet]*. 2017;389(10071):815–22. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28110982>
48. Loeb S, Van Den Heuvel S, Zhu X, Bangma CH, Schröder FH, Roobol MJ. Infectious complications and hospital admissions after prostate biopsy in a European randomized trial. *Eur Urol*. 2012;61(6):1110–4.
49. Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, et al. Systematic review of complications of prostate biopsy. Vol. 64, *European Urology*. *Eur Urol*; 2013. p. 876–92.
50. Gershman B, Zietman AL, Feldman AS, McDougal WS. Transperineal template-guided prostate biopsy for patients with persistently elevated PSA and multiple prior negative biopsies. *Urol Oncol Semin Orig Investig*. 2013 Oct 1;31(7):1093–7.
51. Satoh T, Matsumoto K, Fujita T, Tabata KI, Okusa H, Tsuboi T, et al. Cancer core distribution in patients diagnosed by extended transperineal prostate biopsy. *Urology*. 2005 Jul 1;66(1):114–8.
52. Huang LH, Lin PH, Tsai KW, Wang LJ, Huang YH, Kuo HC, et al. The effects of storage temperature and duration of blood samples on DNA and RNA qualities. *PLoS One*. 2017 Sep 1;12(9).
53. Jones JS, Patel A, Schoenfield L, Rabets JC, Zippe CD, Magi-Galluzzi C. Saturation technique does not improve cancer detection as an initial prostate biopsy strategy. *J Urol*. 2006 Feb;175(2):485–8.
54. 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 [Internet]*. 2012 Feb 1 [cited 2020 Jun 2];109(3):367–71. Available from: <http://doi.wiley.com/10.1111/j.1464-410X.2011.10355.x>
55. Borghesi M, Ahmed H, Nam R, Schaeffer E, Schiavina R, Taneja S, et al. Complications After Systematic, Random, and Image-guided Prostate Biopsy [figure presented]. Vol. 71, *European Urology*. Elsevier B.V.; 2017. p. 353–65.

56. 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 May 1;11(3):192–9.
57. Meyer AR, Joice GA, Schwen ZR, Partin AW, Allaf ME, Gorin MA. Initial Experience Performing In-office Ultrasound-guided Transperineal Prostate Biopsy Under Local Anesthesia Using the PrecisionPoint Transperineal Access System. *Urology*. 2018 May 1;115:8–13.
58. Hung AY, Levy L, Kuban DA. Stage T1c prostate cancer: A heterogeneous category with widely varying prognosis. *Cancer J*. 2002 Nov 1;8(6):440–4.
59. Lomas DJ, Ahmed HU. All change in the prostate cancer diagnostic pathway [Internet]. Vol. 17, *Nature Reviews Clinical Oncology*. Nature Research; 2020 [cited 2020 Jun 26]. p. 372–81. Available from: <https://www.nature.com/articles/s41571-020-0332-z>
60. Moldovan PC, Van den Broeck T, Sylvester R, Marconi L, Bellmunt J, van den Bergh RCN, et al. What Is the Negative Predictive Value of Multiparametric Magnetic Resonance Imaging in Excluding Prostate Cancer at Biopsy? A Systematic Review and Meta-analysis from the European Association of Urology Prostate Cancer Guidelines Panel. *Eur Urol*. 2017 Aug 1;72(2):250–66.
61. Drost F-JH, Osses DF, Nieboer D, Steyerberg EW, Bangma CH, Roobol MJ, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *Cochrane Database Syst Rev* [Internet]. 2019 Apr 25 [cited 2020 Jun 2];(4). Available from: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012663.pub2/full>
62. Schouten MG, van der Leest M, Pokorný M, Hoogenboom M, Barentsz JO, Thompson LC, et al. Why and Where do We Miss Significant Prostate Cancer with Multi-parametric Magnetic Resonance Imaging followed by Magnetic Resonance-guided and Transrectal Ultrasound-guided Biopsy in Biopsy-naïve Men? *Eur Urol* [Internet]. 2017 Jun 1 [cited 2020 Jun 26];71(6):896–903. Available from: <https://pubmed.ncbi.nlm.nih.gov/28063613/>
63. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med* [Internet]. 2018 May 10 [cited 2020 Jun 2];378(19):1767–77. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa1801993>
64. Porpiglia F, Manfredi M, Mele F, Cossu M, Bollito E, Veltri A, et al. Diagnostic Pathway with Multiparametric Magnetic Resonance Imaging Versus Standard Pathway: Results from a Randomized Prospective Study in Biopsy-naïve Patients with Suspected Prostate Cancer. *Eur Urol*. 2017 Aug 1;72(2):282–8.
65. Panebianco V, Barchetti F, Sciarra A, Ciardi A, Indino EL, Papalia R, et al. Multiparametric magnetic resonance imaging vs. standard care in men being evaluated for prostate cancer: A

- randomized study. *Urol Oncol Semin Orig Investig*. 2015 Jan 1;33(1):17.e1-17.e7.
66. Siddiqui MM, Rais-Bahrami S, Turkbey B, George AK, Rothwax J, Shakir N, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. *JAMA - J Am Med Assoc*. 2015 Jan 27;313(4):390–7.
 67. Brizmohun Appayya M, Adshead J, Ahmed HU, Allen C, Bainbridge A, Barrett T, et al. National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection – recommendations from a UK consensus meeting. 2018 [cited 2020 Jun 2]; Available from: www.bjui.org
 68. Davies C, Castle JT, Stalbow K, Haslam PJ. Prostate mpMRI in the UK: the state of the nation. Vol. 74, *Clinical Radiology*. W.B. Saunders Ltd; 2019. p. 894.e11-894.e18.
 69. Smith CP, Harmon SA, Barrett T, Bittencourt LK, Law YM, Shebel H, et al. Intra- and interreader reproducibility of PI-RADSV2: A multireader study. *J Magn Reson Imaging [Internet]*. 2019 Jun 1 [cited 2020 Jun 26];49(6):1694–703. Available from: <https://pubmed.ncbi.nlm.nih.gov/30575184/>
 70. Padhani AR, Barentsz J, Villeirs G, Rosenkrantz AB, Margolis DJ, Turkbey B, et al. PI-RADS Steering Committee: The PI-RADS Multiparametric MRI and MRI-directed Biopsy Pathway. *Radiology [Internet]*. 2019 [cited 2020 Jul 2];292(2):464–74. Available from: <https://pubmed.ncbi.nlm.nih.gov/31184561/>
 71. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol [Internet]*. 2016 Jan 1 [cited 2020 Jul 2];69(1):16–40. Available from: <http://dx.doi.org/10.1016/j.eururo.2015.08.052>
 72. Park KJ, Choi SH, Lee JS, Kim JK, Kim M. Inter-reader Agreement in Prostate Imaging Reporting and Data System Version 2 for Prostate Cancer: A Systematic Review and Meta-Analysis. *J Urol [Internet]*. 2020 Jun 18 [cited 2020 Jul 2]; Available from: <http://www.jurology.com/doi/10.1097/JU.0000000000001200>
 73. Muller BG, Shih JH, Sankineni S, Marko J, Rais-Bahrami S, George AK, et al. Prostate cancer: Interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric mr imaging1. *Radiology*. 2015 Dec 1;277(3):741–50.
 74. Asvadi NH, Afshari Mirak S, Mohammadian Bajgiran A, Khoshnoodi P, Wibulpolprasert P, Margolis D, et al. 3T multiparametric MR imaging, PIRADSV2-based detection of index prostate cancer lesions in the transition zone and the peripheral zone using whole mount histopathology as reference standard. *Abdom Radiol*. 2018 Nov 1;43(11):3117–24.
 75. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, et al. Interobserver reproducibility of the PI-RADS version 2 lexicon: A multicenter study of six

- experienced prostate radiologists. *Radiology* [Internet]. 2016 Sep 1 [cited 2020 Jun 5];280(3):793–804. Available from: <https://pubs.rsna.org/doi/abs/10.1148/radiol.2016152542>
76. Latifoltojar A, Appayya MB, Barrett T, Punwani S. Similarities and differences between Likert and PIRADS v2.1 scores of prostate multiparametric MRI: a pictorial review of histology-validated cases. Vol. 74, *Clinical Radiology*. W.B. Saunders Ltd; 2019. p. 895.e1-895.e15.
 77. Zawaideh JP, Sala E, Pantelidou M, Shaida N, Koo B, Caglic I, et al. Comparison of Likert and PI-RADS version 2 MRI scoring systems for the detection of clinically significant prostate cancer. *Br J Radiol* [Internet]. 2020 Jun 11 [cited 2020 Jul 2];20200298. Available from: <https://pubmed.ncbi.nlm.nih.gov/32479105/>
 78. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate [Internet]. Vol. 17, *Modern Pathology*. 2004 [cited 2020 Jun 4]. p. 292–306. Available from: <https://www.nature.com/articles/3800054>
 79. Chen N, Zhou Q. The evolving gleason grading system. Vol. 28, *Chinese Journal of Cancer Research*. AME Publishing Company; 2016. p. 58–64.
 80. Pan CC, Potter SR, Partin AW, Epstein JI. The prognostic significance of tertiary Gleason patterns of higher grade in radical prostatectomy specimens: A proposal to modify the Gleason grading system. *Am J Surg Pathol*. 2000 Apr;24(4):563–9.
 81. Singh R V, Agashe SR, Gosavi A V, Sulhyan KR. Interobserver reproducibility of Gleason grading of prostatic adenocarcinoma among general pathologists. *Indian J Cancer* [Internet]. 2011 Oct [cited 2020 Jun 4];48(4):488–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22293266>
 82. Allsbrook WC, Mangold KA, Johnson MH, Lane RB, Lane CG, Amin MB, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: Urologic pathologists. *Hum Pathol*. 2001;32(1):74–80.
 83. Netto GJ, Eisenberger M, Epstein JI. Interobserver variability in histologic evaluation of radical prostatectomy between central and local pathologists: Findings of TAX 3501 multinational clinical trial. *Urology*. 2011 May;77(5):1155–60.
 84. Berney DM, Algaba F, Camparo P, Comp  rat E, Griffiths D, Kristiansen G, et al. The reasons behind variation in Gleason grading of prostatic biopsies: Areas of agreement and misconception among 266 european pathologists. *Histopathology*. 2014 Feb;64(3):405–11.
 85. Egev  d L, Ahmad AS, Algaba F, Berney DM, Boccon-Gibod L, Comp  rat E, et al. Standardization of Gleason grading among 337 European pathologists. *Histopathology*. 2013 Jan;62(2):247–56.
 86. Kweldam CF, Nieboer D, Algaba F, Amin MB, Berney DM, Billis A, et al. Gleason grade 4

- prostate adenocarcinoma patterns: an interobserver agreement study among genitourinary pathologists. *Histopathology* [Internet]. 2016 Sep 1 [cited 2020 Jun 4];69(3):441–9. Available from: <http://doi.wiley.com/10.1111/his.12976>
87. Ryu HS, Jin MS, Park JH, Lee S, Cho J, Oh S, et al. Automated gleason scoring and tumor quantification in prostate core needle biopsy images using deep neural networks and its comparison with pathologist-based assessment. *Cancers (Basel)*. 2019 Dec 1;11(12).
 88. Lucas M, Jansen I, Savci-Heijink CD, Meijer SL, de Boer OJ, van Leeuwen TG, et al. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Arch*. 2019 Jul 1;475(1):77–83.
 89. Kott O, Linsley D, Amin A, Karagounis A, Jeffers C, Golijanin D, et al. Development of a Deep Learning Algorithm for the Histopathologic Diagnosis and Gleason Grading of Prostate Cancer Biopsies: A Pilot Study. *Eur Urol Focus*. 2019 Nov;
 90. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *Eur Urol* [Internet]. 2016 Mar 1 [cited 2019 Feb 11];69(3):428–35. Available from: <https://www.sciencedirect.com/science/article/pii/S0302283815005576?via%3Dihub>
 91. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40(2):244–52.
 92. Georgescu I, Gooding RJ, Doiron RC, Day A, Selvarajah S, Davidson C, et al. Molecular characterization of Gleason patterns 3 and 4 prostate cancer using reverse Warburg effect-associated genes. *Cancer Metab*. 2016;4(1).
 93. Sowalsky AG, Kissick HT, Gerrin SJ, Schaefer RJ, Xia Z, Russo JW, et al. Gleason score 7 prostate cancers emerge through branched evolution of clonal Gleason pattern 3 and 4. *Clin Cancer Res*. 2017 Jul 15;23(14):3823–33.
 94. Lavery HJ, Droller MJ. Do gleason patterns 3 and 4 prostate cancer represent separate disease states? Vol. 188, *Journal of Urology*. J Urol; 2012. p. 1667–75.
 95. Sweat SD, Bergstralh EJ, Slezak J, Blute ML, Zincke H. Competing risk analysis after radical prostatectomy for clinically nonmetastatic prostate adenocarcinoma according to clinical Gleason score and patient age. *J Urol*. 2002 Aug 1;168(2):525–9.
 96. Traxer O, Gettman MT, Napper CA, Scott DJ, Jones DB, Roehrborn CG, et al. Prognostic factors for survival of patients with pathological gleason score 7 prostate cancer: Differences in outcome between primary gleason grades 3 and 4. *J Urol*. 2001 Nov;166(5):1692–7.

97. Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3+4 versus Gleason score 4+3 tumor at radical prostatectomy. *Urology*. 2000;56(5):823–7.
98. Matoso A, Epstein JI. Grading of Prostate Cancer: Past, Present, and Future. Vol. 17, Current Urology Reports. Current Medicine Group LLC 1; 2016. p. 1–6.
99. Van Leenders GJLH, Van Der Kwast TH, Grignon DJ, Evans AJ, Kristiansen G, Kweldam CF, et al. The 2019 International Society of Urological Pathology (ISUP) Consensus Conference on Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2020;
100. Tan N, Margolis DJA, McClure TD, Thomas A, Finley DS, Reiter RE, et al. Radical prostatectomy: Value of prostate MRI in surgical planning. Vol. 37, Abdominal Imaging. Abdom Imaging; 2012. p. 664–74.
101. Marengo J, Orczyk C, Collins T, Moore C, Emberton M. Role of MRI in planning radical prostatectomy: what is the added value? *World J Urol*. 2019 Jul 1;37(7):1289–92.
102. Kilcoyne A, Price MC, McDermott S, Harisinghani MG. Imaging on nodal staging of prostate cancer. Vol. 13, Future Oncology. Future Medicine Ltd.; 2017. p. 551–65.
103. Pomykala KL, Farolfi A, Hadaschik B, Fendler WP, Herrmann K. Molecular Imaging for Primary Staging of Prostate Cancer. Vol. 49, Seminars in Nuclear Medicine. W.B. Saunders; 2019. p. 271–9.
104. Hernandez DJ, Nielsen ME, Han M, Partin AW. Contemporary Evaluation of the D'Amico Risk Classification of Prostate Cancer. *Urology*. 2007 Nov;70(5):931–5.
105. Cooperberg MR, Pasta DJ, Elkin EP, Litwin MS, Latini DM, Duchane J, 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. *J Urol*. 2005;173(6):1938–42.
106. Brajtbord JS, Leapman MS, Cooperberg MR. The CAPRA Score at 10 Years: Contemporary Perspectives and Analysis of Supporting Studies. *Eur Urol* [Internet]. 2017 May 1 [cited 2020 Jun 26];71(5):705–9. Available from: <http://www.europeanurology.com/article/S0302283816305772/fulltext>
107. Mohler J, Bahnson RR, Boston B, Busby JE, D'Amico A, Eastham JA, et al. NCCN clinical practice guidelines in oncology: prostate cancer. *J Natl Compr Canc Netw* [Internet]. 2010;8(2):162–200. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20141676>
108. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework [Internet]. 2001 p. 89–95. Available from: <http://doi.wiley.com/10.1067/mcp.2001.113989>

109. FDA-NIH Biomarker Working Group. BEST (Biomarkers , EndpointS , and other Tools). Updat Sept 25 [Internet]. 2017 [cited 2020 Jun 26];(Md):55. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK326791/>
110. Freedland SJ, Moul JW. Prostate Specific Antigen Recurrence After Definitive Therapy. Vol. 177, *Journal of Urology*. J Urol; 2007. p. 1985–91.
111. Sandler HM, Eisenberger MA. Assessing and Treating Patients With Increasing Prostate Specific Antigen Following Radical Prostatectomy. *J Urol*. 2007 Sep;178(3 SUPPL.).
112. Roberts SG, Blute ML, Bergstralh EJ, Slezak JM, Zincke H. PSA doubling time as a predictor of clinical progression after biochemical failure following radical prostatectomy for prostate cancer. *Mayo Clin Proc*. 2001;76(6):576–81.
113. Epstein JI, Walsh PC, Brendler CB. Radical prostatectomy for impalpable prostate cancer: The Johns Hopkins experience with tumors found on transurethral resection (stages T1A and T1B) and on needle biopsy (stage T1C). In: *Journal of Urology*. 1994. p. 1721–9.
114. Page EC, Bancroft EK, Brook MN, Assel M, Hassan Al Battat M, Thomas S, et al. Interim Results from the IMPACT Study: Evidence for Prostate-specific Antigen Screening in BRCA2 Mutation Carriers. *Eur Urol*. 2019 Dec 1;76(6):831–42.
115. Nicholson A, Mahon J, Boland A, Beale S, Dwan K, Fleeman N, et al. The clinical effectiveness and cost-effectiveness of the PROGENSA® prostate cancer antigen 3 assay and the prostate health index in the diagnosis of prostate cancer: A systematic review and economic evaluation. *Health Technol Assess (Rockv)*. 2015 Oct 1;19(87):1–191.
116. Sathianathen NJ, Kuntz KM, Alarid-Escudero F, Lawrentschuk NL, Bolton DM, Murphy DG, et al. Incorporating Biomarkers into the Primary Prostate Biopsy Setting: A Cost-Effectiveness Analysis. *J Urol* [Internet]. 2018 Dec 1 [cited 2020 Jun 9];200(6):1215–20. Available from: <http://www.jurology.com/doi/10.1016/j.juro.2018.06.016>
117. Dijkstra S, Govers TM, Hendriks RJ, Schalken JA, Van Criekinge W, Van Neste L, et al. Cost-effectiveness of a new urinary biomarker-based risk score compared to standard of care in prostate cancer diagnostics – a decision analytical model. *BJU Int*. 2017 Nov 1;120(5):659–65.
118. Sanda MG, Feng Z, Howard DH, Tomlins SA, Sokoll LJ, Chan DW, et al. Association between combined TMPRSS2:ERG and PCA3 RNA urinary testing and detection of aggressive prostate cancer. *JAMA Oncol*. 2017 Aug 1;3(8):1085–93.
119. Aubry W, Arnold Willis :, Grant Bagley :, Willis Iii M, Andrew Layton, Ba ; Budget Impact Model: Epigenetic Assay Can Help Avoid Unnecessary Repeated Prostate Biopsies and Reduce Healthcare Spending [Internet]. Vol. 6, *American Health & Drug Benefits I*. [cited 2020 Jun 17]. Available from: www.AHDBOnline.com

120. Health Quality Ontario. Prolaris Cell Cycle Progression Test for Localized Prostate Cancer: A Health Technology Assessment. Ont Health Technol Assess Ser [Internet]. 2017 [cited 2020 Jun 26];17(6):1–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28572867>
121. Roth JA, Ramsey SD, Carlson JJ. Cost-Effectiveness of a Biopsy-Based 8-Protein Prostate Cancer Prognostic Assay to Optimize Treatment Decision Making in Gleason 3 + 3 and 3 + 4 Early Stage Prostate Cancer. *Oncologist* [Internet]. 2015 Dec 19 [cited 2020 Jun 17];20(12):1355–64. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1634/theoncologist.2015-0214>
122. Albala D, Kemeter MJ, Febbo PG, Lu R, John V, Stoy D, et al. Health Economic Impact and Prospective Clinical Utility of Oncotype DX® Genomic Prostate Score. *Rev Urol*. 2016;18(3):123–32.
123. Lobo JM, Trifiletti DM, Sturz VN, Dicker AP, Buerki C, Davicioni E, et al. Cost-effectiveness of the Decipher Genomic Classifier to Guide Individualized Decisions for Early Radiation Therapy After Prostatectomy for Prostate Cancer. *Clin Genitourin Cancer*. 2017 Jun 1;15(3):e299–309.
124. Andriole GL, Crawford ED, Grubb III RL, Buys SS, Chia D, Church TR, et al. Prostate Cancer Screening in the Randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: Mortality Results after 13 Years of Follow-up. *JNCI J Natl Cancer Inst* [Internet]. 2012 Jan 6;104(2):125–32. Available from: <https://doi.org/10.1093/jnci/djr500>
125. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012 Mar 15;366(11):981–90.
126. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for prostate cancer US Preventive services task force recommendation statement. *JAMA - J Am Med Assoc*. 2018 May 1;319(18):1901–13.
127. Greene KL, Albertsen PC, Babaian RJ, Carter HB, Gann PH, Han M, et al. Prostate specific antigen best practice statement: 2009 update. *J Urol* [Internet]. 2009 Nov [cited 2020 Jun 5];182(5):2232–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19781717>
128. AAFP. Summary of Recommendations for Clinical Preventive Services Introduction to AAFP Summary of Recommendations For Clinical Preventive Services. 2017.
129. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level ≤ 4.0 ng per milliliter. *N Engl J Med*. 2004 May 27;350(22):2239–2246+2321.
130. Thompson IM, Ankerst DP, Chi C, Lucia MS, Goodman PJ, Crowley JJ, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/mL or lower. *J Am Med Assoc*. 2005 Jul;294(1):66–70.

131. Mistry K, Cable G. Meta-analysis of prostate-specific antigen and digital rectal examination as screening tests for prostate carcinoma. *J Am Board Fam Pract.* 2003 Mar;16(2):95–101.
132. Karazanashvili G, Abrahamsson PA. Prostate specific antigen and human glandular kallikrein 2 in early detection of prostate cancer. Vol. 169, *Journal of Urology.* J Urol; 2003. p. 445–57.
133. Nishiya M, Miller GJ, Lookner DH, Crawford ED. Prostate specific antigen density in patients with histologically proven prostate carcinoma. *Cancer.* 1994 Dec 1;74(11):3002–9.
134. Ohori M, Dunn JK, Scardino PT. Is prostate-specific antigen density more useful than prostate-specific antigen levels in the diagnosis of prostate cancer? *Urology.* 1995;46(5):666–71.
135. Cuocolo R, Stanzione A, Rusconi G, Petretta M, Ponsiglione A, Fusco F, et al. PSA-density does not improve bi-parametric prostate MR detection of prostate cancer in a biopsy naïve patient population. *Eur J Radiol.* 2018 Jul 1;104:64–70.
136. Distler FA, Radtke JP, Bonekamp D, Kesch C, Schlemmer HP, Wieczorek K, et al. The Value of PSA Density in Combination with PI-RADS™ for the Accuracy of Prostate Cancer Prediction. *J Urol.* 2017 Sep 1;198(3):575–82.
137. Nordström T, Akre O, Aly M, Grönberg H, Eklund M. Prostate-specific antigen (PSA) density in the diagnostic algorithm of prostate cancer. *Prostate Cancer Prostatic Dis.* 2018 Apr 1;21(1):57–63.
138. Hansen NL, Barrett T, Koo B, Doble A, Gnanapragasam V, Warren A, et al. The influence of prostate-specific antigen density on positive and negative predictive values of multiparametric magnetic resonance imaging to detect Gleason score 7-10 prostate cancer in a repeat biopsy setting. *BJU Int [Internet].* 2017 May [cited 2020 Jan 29];119(5):724–30. Available from: <http://doi.wiley.com/10.1111/bju.13619>
139. Washino S, Okochi T, Saito K, Konishi T, Hirai M, Kobayashi Y, et al. Combination of prostate imaging reporting and data system (PI-RADS) score and prostate-specific antigen (PSA) density predicts biopsy outcome in prostate biopsy naïve patients. *BJU Int [Internet].* 2017 Feb 1 [cited 2020 Jan 10];119(2):225–33. Available from: <http://doi.wiley.com/10.1111/bju.13465>
140. Vickers AJ, Cronin AM, Roobol MJ, Savage CJ, Peltola M, Pettersson K, et al. A four-kallikrein panel predicts prostate cancer in men with recent screening: Data from the european randomized study of screening for prostate cancer, Rotterdam. *Clin Cancer Res [Internet].* 2010 Jun 15 [cited 2017 Aug 24];16(12):3232–9. Available from: <http://clincancerres.aacrjournals.org/content/clincanres/16/12/3232.full.pdf>
141. Benchikh A, Savage C, Cronin A, Salama G, Villers A, Lilja H, et al. A panel of kallikrein markers can predict outcome of prostate biopsy following clinical work-up: An independent validation study from the European Randomized Study of Prostate Cancer screening, France. *BMC*

- Cancer [Internet]. 2010 Nov 25 [cited 2020 Jun 8];10(1):635. Available from: <http://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-10-635>
142. Vickers AJ, Cronin AM, Aus G, Pihl CG, Becker C, Pettersson K, et al. A panel of kallikrein markers can reduce unnecessary biopsy for prostate cancer: Data from the European Randomized Study of Prostate Cancer Screening in Göteborg, Sweden. BMC Med [Internet]. 2008 Jul 8 [cited 2020 Jun 8];6(1):19. Available from: <http://bmcmecine.biomedcentral.com/articles/10.1186/1741-7015-6-19>
 143. McDonald ML, Parsons JK. 4-Kallikrein Test and Kallikrein Markers in Prostate Cancer Screening. Vol. 43, Urologic Clinics of North America. W.B. Saunders; 2016. p. 39–46.
 144. Russo GI, Regis F, Castelli T, Favilla V, Privitera S, Giardina R, et al. A Systematic Review and Meta-analysis of the Diagnostic Accuracy of Prostate Health Index and 4-Kallikrein Panel Score in Predicting Overall and High-grade Prostate Cancer. Vol. 15, Clinical Genitourinary Cancer. Elsevier Inc.; 2017. p. 429-439.e1.
 145. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. Eur Urol. 2017 Apr 1;71(4):618–29.
 146. Catalona WJ, Partin AW, Sanda MG, Wei JT, Klee GG, Bangma CH, et al. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. J Urol. 2011 May;185(5):1650–5.
 147. Gnanapragasam VJ, Burling K, George A, Stearn S, Warren A, Barrett T, et al. The Prostate Health Index adds predictive value to multi-parametric MRI in detecting significant prostate cancers in a repeat biopsy population. Sci Rep. 2016 Oct 17;6.
 148. Gronberg H, Adolfsson J, Aly M, Nordstrom T, Wiklund P, Brandberg Y, et al. Prostate cancer screening in men aged 50-69 years (STHLM3): a prospective population-based diagnostic study. Lancet Oncol [Internet]. 2015;16(16):1667–76. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26563502>
 149. Bussemakers MJG, Van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HFM, Schalken JA, et al. DD3: A New Prostate-specific Gene, Highly Overexpressed in Prostate Cancer. CANCER Res [Internet]. 1999 [cited 2017 Aug 24];59:5975–9. Available from: <http://cancerres.aacrjournals.org/content/canres/59/23/5975.full.pdf>
 150. Groskopf J, Aubin SMJ, Deras IL, Blase A, Bodrug S, Clark C, et al. APTIMA PCA3 molecular urine test: Development of a method to aid in the diagnosis of prostate cancer. Clin Chem. 2006 Jun;52(6):1089–95.

151. Marks LS, Fradet Y, Lim Deras I, Blase A, Mathis J, Aubin SMJ, et al. PCA3 Molecular Urine Assay for Prostate Cancer in Men Undergoing Repeat Biopsy. *Urology*. 2007 Mar;69(3):532–5.
152. Vlaeminck-Guillem V, Ruffion A, André J, Devonec M, Paparel P. Urinary Prostate Cancer 3 Test: Toward the Age of Reason? Vol. 75, *Urology*. *Urology*; 2010. p. 447–53.
153. Whitman EJ, Groskopf J, Ali A, Chen Y, Blase A, Furusato B, et al. PCA3 Score Before Radical Prostatectomy Predicts Extracapsular Extension and Tumor Volume. *J Urol* [Internet]. 2008 Nov 17 [cited 2020 Jun 9];180(5):1975–9. Available from: <http://www.jurology.com/doi/10.1016/j.juro.2008.07.060>
154. Nakanishi H, Groskopf J, Fritsche HA, Bhadkamkar V, Blase A, Kumar SV, et al. PCA3 Molecular Urine Assay Correlates With Prostate Cancer Tumor Volume: Implication in Selecting Candidates for Active Surveillance. *J Urol*. 2008 May;179(5):1804–10.
155. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PFA, et al. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. *Eur Urol* [Internet]. 2008 Nov [cited 2020 Jun 9];54(5):1081–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18602209>
156. Ploussard G, Haese A, Van Poppel H, Marberger M, Stenzl A, Mulders PFA, et al. The prostate cancer gene 3 (PCA3) urine test in men with previous negative biopsies: does free-to-total prostate-specific antigen ratio influence the performance of the PCA3 score in predicting positive biopsies? *BJU Int* [Internet]. 2010 Oct 1 [cited 2020 Jun 29];106(8):1143–7. Available from: <http://doi.wiley.com/10.1111/j.1464-410X.2010.09286.x>
157. Porpiglia F, Cantiello F, De Luca S, Manfredi M, Veltri A, Russo F, et al. In-parallel comparative evaluation between multiparametric magnetic resonance imaging, prostate cancer antigen 3 and the prostate health index in predicting pathologically confirmed significant prostate cancer in men eligible for active surveillance. *BJU Int* [Internet]. 2016 Oct 1 [cited 2020 Jun 29];118(4):527–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/26350955/>
158. Hessels D, Van Gils MPMQ, Van Hooij O, Jannink SA, Witjes JA, Verhaegh GW, et al. Predictive value of PCA3 in urinary sediments in determining clinico-pathological characteristics of prostate cancer. *Prostate* [Internet]. 2010 Jan 1 [cited 2020 Jun 29];70(1):10–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/19708043/>
159. Van Gils MPMQ, Hessels D, Hulsbergen-van De Kaa CA, Witjes JA, Jansen CFJ, Mulders PFA, et al. Detailed analysis of histopathological parameters in radical prostatectomy specimens and PCA3 urine test results. *Prostate* [Internet]. 2008 Aug [cited 2020 Jun 9];68(11):1215–22. Available from: <https://pubmed.ncbi.nlm.nih.gov/18500693/>
160. Hegde J V., Veruttipong D, Said JW, Reiter RE, Steinberg ML, King CR, et al. Prostate Cancer Antigen 3 Score Does Not Predict for Adverse Pathologic Features at Radical Prostatectomy or

- for Progression-free Survival in Clinically Localized, Intermediate- and High-risk Prostate Cancer. *Urology* [Internet]. 2017 Sep 1 [cited 2020 Jun 29];107:171–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/28552819/>
161. Busetto GM, De Berardinis E, Sciarra A, Panebianco V, Giovannone R, Rosato S, et al. Prostate cancer gene 3 and multiparametric magnetic resonance can reduce unnecessary biopsies: Decision curve analysis to evaluate predictive models. *Urology*. 2013;82(6):1355–62.
 162. Nilsson J, Skog J, Nordstrand A, Baranov V, Mincheva-Nilsson L, Breakefield XO, et al. Prostate cancer-derived urine exosomes: A novel approach to biomarkers for prostate cancer. *Br J Cancer*. 2009 May 19;100(10):1603–7.
 163. Donovan MJ, Noerholm M, Bentink S, Belzer S, Skog J, O'Neill V, et al. A molecular signature of PCA3 and ERG exosomal RNA from non-DRE urine is predictive of initial prostate biopsy result. *Prostate Cancer Prostatic Dis*. 2015 Dec 1;18(4):370–5.
 164. McKiernan J, Donovan MJ, O'Neill V, Bentink S, Noerholm M, Belzer S, et al. A novel urine exosome gene expression assay to predict high-grade prostate cancer at initial biopsy. *JAMA Oncol*. 2016 Jul 1;2(7):882–9.
 165. Leyten GHJM, Hessels D, Smit FP, Jannink SA, De Jong H, Melchers WJG, et al. Identification of a candidate gene panel for the early diagnosis of prostate cancer. *Clin Cancer Res*. 2015 Jul 1;21(13):3061–70.
 166. Van Neste L, Partin AW, Stewart GD, Epstein JI, Harrison DJ, Van Criekinge W. Risk score predicts high-grade prostate cancer in DNA-methylation positive, histopathologically negative biopsies. *Prostate*. 2016 Sep 1;76(12):1078–87.
 167. Hendriks RJ, van der Leest MMG, Dijkstra S, Barentsz JO, Van Criekinge W, Hulsbergen-van de Kaa CA, et al. A urinary biomarker-based risk score correlates with multiparametric MRI for prostate cancer detection. *Prostate* [Internet]. 2017 Oct 1 [cited 2020 Jun 10];77(14):1401–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/28853167/>
 168. Pepe P, Dibenedetto G, Pepe L, Pennisi M. Multiparametric MRI versus selectMDX accuracy in the diagnosis of clinically significant PCA in men enrolled in active surveillance. *In Vivo (Brooklyn)*. 2020;34(1):393–6.
 169. Tomlins SA, Day JR, Lonigro RJ, Hovelson DH, Siddiqui J, Kunju LP, et al. Urine TMPRSS2:ERG Plus PCA3 for Individualized Prostate Cancer Risk Assessment. *Eur Urol*. 2016 Jul 1;70(1):45–53.
 170. Leyten GHJM, Hessels D, Jannink SA, Smit FP, de Jong H, Cornel EB, et al. Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. *Eur Urol* [Internet]. 2014 Mar [cited 2020 Jun

- 10];65(3):534–42. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23201468>
171. Zhou CK, Young D, Yeboah ED, Coburn SB, Tettey Y, Biritwum RB, et al. TMPRSS2:ERG Gene Fusions in Prostate Cancer of West African Men and a Meta-Analysis of Racial Differences. *Am J Epidemiol* [Internet]. 2017 Dec 15;186(12):1352–61. Available from: <https://academic.oup.com/aje/article/186/12/1352/3867257>
 172. Wojno KJ, Costa FJ, Cornell RJ, Small JD, Pasin E, Van Criekinge W, et al. Reduced Rate of Repeated Prostate Biopsies Observed in ConfirmMDx Clinical Utility Field Study. *Am Heal drug benefits* [Internet]. 2014 May [cited 2020 Jun 29];7(3):129–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24991397>
 173. Partin AW, Van Neste L, Klein EA, Marks LS, Gee JR, Troyer DA, et al. Clinical validation of an epigenetic assay to predict negative histopathological results in repeat prostate biopsies. *J Urol*. 2014 Oct 1;192(4):1081–7.
 174. Waterhouse RL, Van Neste L, Moses KA, Barnswell C, Silberstein JL, Jalkut M, et al. Evaluation of an Epigenetic Assay for Predicting Repeat Prostate Biopsy Outcome in African American Men. *Urology*. 2019 Jun 1;128:62–5.
 175. Stewart GD, Van Neste L, Delvenne P, Delrée P, Delga A, McNeill SA, et al. Clinical utility of an epigenetic assay to detect occult prostate cancer in histopathologically negative biopsies: Results of the MATLOC study. *J Urol*. 2013 Mar;189(3):1110–6.
 176. Shore ND, Kella N, Moran B, Boczek J, Bianco FJ, Crawford ED, et al. Impact of the Cell Cycle Progression Test on Physician and Patient Treatment Selection for Localized Prostate Cancer. *J Urol*. 2016;195(3):612–8.
 177. Cuzick J, Stone S, Fisher G, Yang ZH, North B V., Berney DM, et al. Validation of an RNA cell cycle progression score for predicting death from prostate cancer in a conservatively managed needle biopsy cohort. *Br J Cancer* [Internet]. 2015 Jul 28 [cited 2017 Aug 24];113(3):382–9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4522632/pdf/bjc2015223a.pdf>
 178. Cuzick J, Berney DM, Fisher G, Mesher D, Møller H, Reid JE, et al. Prognostic value of a cell cycle progression signature for prostate cancer death in a conservatively managed needle biopsy cohort. *Br J Cancer*. 2012 Mar 13;106(6):1095–9.
 179. Sommariva S, Tarricone R, Lazzeri M, Ricciardi W, Montorsi F. Prognostic Value of the Cell Cycle Progression Score in Patients with Prostate Cancer: A Systematic Review and Meta-analysis. *Eur Urol* [Internet]. 2014/12/08. 2014; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25481455>
 180. Eggener SE, Rumble RB, Armstrong AJ, Morgan TM, Crispino T, Cornford P, et al. Molecular biomarkers in localized prostate cancer: ASCO guideline. *J Clin Oncol*. 2020 May

- 1;38(13):1474–94.
181. Wibmer AG, Robertson NL, Hricak H, Zheng J, Capanu M, Stone S, et al. Extracapsular extension on MRI indicates a more aggressive cell cycle progression genotype of prostate cancer. *Abdom Radiol*. 2019 Aug 15;44(8):2864–73.
 182. Shipitsin M, Small C, Choudhury S, Giladi E, Friedlander S, Nardone J, et al. Identification of proteomic biomarkers predicting prostate cancer aggressiveness and lethality despite biopsy-sampling error. *Br J Cancer*. 2014 Jul 17;111(6):1201–12.
 183. Blume-Jensen P, Berman DM, Rimm DL, Shipitsin M, Putzi M, Nifong TP, et al. Biology of Human Tumors Development and Clinical Validation of an In Situ Biopsy-Based Multimarker Assay for Risk Stratification in Prostate Cancer. *Clin Cancer Res* [Internet]. [cited 2017 Aug 24];21(11). Available from: <http://clincancerres.aacrjournals.org/content/clincanres/21/11/2591.full.pdf>
 184. Knezevic D, Goddard AD, Natraj N, Cherbavaz DB, Clark-Langone KM, Snable J, et al. Analytical validation of the Oncotype DX prostate cancer assay – a clinical RT-PCR assay optimized for prostate needle biopsies. *BMC Genomics* [Internet]. 2013 [cited 2017 Aug 24];14. Available from: <http://www.biomedcentral.com/1471-2164/14/690>
 185. Klein EA, Yousefi K, Haddad Z, Choeurng V, Buerki C, Stephenson AJ, et al. A genomic classifier improves prediction of metastatic disease within 5 years after surgery in node-negative high-risk prostate cancer patients managed by radical prostatectomy without adjuvant therapy. *Eur Urol* [Internet]. 2015 Apr 1 [cited 2020 Jun 16];67(4):778–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25466945>
 186. Kornberg Z, Cowan JE, Westphalen AC, Cooperberg MR, Chan JM, Zhao S, et al. Genomic Prostate Score, PI-RADS™ version 2 and Progression in Men with Prostate Cancer on Active Surveillance. *J Urol*. 2019 Feb 1;201(2):300–6.
 187. Erho N, Crisan A, Vergara IA, Mitra AP, Ghadessi M, Buerki C, et al. Discovery and Validation of a Prostate Cancer Genomic Classifier that Predicts Early Metastasis Following Radical Prostatectomy. 2013 [cited 2017 Aug 24];8(6). Available from: <http://www.nrc-cnrc.gc.ca/eng/>
 188. Spratt DE, Yousefi K, Dehesi S, Ross AE, Den RB, Schaeffer EM, et al. Individual patient-level meta-Analysis of the performance of the decipher genomic classifier in high-risk men after prostatectomy to predict development of metastatic disease. *J Clin Oncol*. 2017 Jun 20;35(18):1991–8.
 189. Nguyen PL, Haddad Z, Ross AE, Martin NE, Dehesi S, Lam LLC, et al. Ability of a Genomic Classifier to Predict Metastasis and Prostate Cancer-specific Mortality after Radiation or Surgery based on Needle Biopsy Specimens. *Eur Urol*. 2017 Nov 1;72(5):845–52.

190. Tosoian JJ, Birer SR, Jeffrey Karnes R, Zhang J, Davicioni E, Klein EE, et al. Performance of clinicopathologic models in men with high risk localized prostate cancer: impact of a 22-gene genomic classifier. *Prostate Cancer Prostatic Dis.* 2020;
191. Ross AE, Feng FY, Ghadessi M, Erho N, Crisan A, Buerki C, et al. A genomic classifier predicting metastatic disease progression in men with biochemical recurrence after prostatectomy. *Prostate Cancer Prostatic Dis* [Internet]. 2014 Mar [cited 2020 Jun 29];17(1):64–9. Available from: [/pmc/articles/PMC4332821/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/24811111/)
192. Taylor AS, Morgan TM, Wallington DG, Chinnaiyan AM, Spratt DE, Mehra R. Correlation between cribriform/intraductal prostatic adenocarcinoma and percent Gleason pattern 4 to a 22-gene genomic classifier. *Prostate.* 2020 Feb 1;80(2):146–52.
193. Greenland NY, Cowan JE, Zhang L, Carroll PR, Chan E, Stohr BA, et al. Expansile cribriform Gleason pattern 4 has histopathologic and molecular features of aggressiveness and greater risk of biochemical failure compared to glomerulation Gleason pattern 4. *Prostate.* 2020 May 1;80(8):653–9.
194. Palsdottir T, Nordström T, Aly M, Jäderling F, Clements M, Grönberg H, et al. A Unified Prostate Cancer Risk Prediction Model Combining the Stockholm3 Test and Magnetic Resonance Imaging. *Eur Urol Oncol.* 2019 Sep 1;2(5):490–6.
195. Kayano PP, Carneiro A, Castilho TML, Sivaraman A, Claros OR, Baroni RH, et al. Comparison of Gleason upgrading rates in transrectal ultrasound systematic random biopsies versus US-MRI fusion biopsies for prostate cancer. *Int Braz J Urol.* 2018;44(6):1106–13.
196. Alqahtani S, Wei C, Zhang Y, Szewczyk-Bieda M, Wilson J, Huang Z, et al. Prediction of prostate cancer Gleason score upgrading from biopsy to radical prostatectomy using pre-biopsy multiparametric MRI PIRADS scoring system. *Sci Rep.* 2020 Dec 1;10(1):1–9.
197. Thach DC, Lin B, Walter E, Kruzelock R, Rowley RK, Tibbetts C, et al. Assessment of two methods for handling blood in collection tubes with RNA stabilizing agent for surveillance of gene expression profiles with high density microarrays. *J Immunol Methods* [Internet]. 2003 Dec 1 [cited 2019 Nov 16];283(1–2):269–79. Available from: <https://www.sciencedirect.com/science/article/pii/S0022175903004034?via%3Dihub>
198. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009 Apr;42(2):377–81.
199. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. Vol. 95, *Journal of Biomedical Informatics*. Academic Press Inc.; 2019. p. 103208.

200. Mehra R, Boyd LM, Ickovics JR. Racial residential segregation and adverse birth outcomes: A systematic review and meta-analysis. Vol. 191, Social Science and Medicine. Elsevier Ltd; 2017. p. 237–50.
201. Massey DS, Denton NA. The Dimensions of Residential Segregation. Soc Forces. 1988 Dec;67(2):281.
202. Reardon SF, O’Sullivan D. 3. Measures of Spatial Segregation. Sociol Methodol [Internet]. 2004 Dec 24 [cited 2020 Jun 29];34(1):121–62. Available from: <http://journals.sagepub.com/doi/10.1111/j.0081-1750.2004.00150.x>
203. Gillies S. geojson · PyPI [Internet]. [cited 2020 May 13]. Available from: <https://pypi.org/project/geojson/#data>
204. Van Rossum G. Centrum voor Wiskunde en Informatica Python tutorial. 1995 Jan.
205. Gillies S. The Fiona User Manual — Fiona 2.0dev documentation [Internet]. [cited 2020 May 13]. Available from: <https://fiona.readthedocs.io/en/latest/manual.html>
206. Gillies S. The Shapely User Manual [Internet]. 2013 [cited 2020 May 13]. Available from: <https://shapely.readthedocs.io/en/latest/manual.html>
207. Gillies S. Rtree: Spatial indexing for Python — Rtree 0.7.0 documentation [Internet]. 2018 [cited 2020 May 13]. Available from: <https://toblerity.org/rtree/>
208. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. Methods [Internet]. 2001 Dec 1 [cited 2019 Nov 15];25(4):402–8. Available from: <https://www.sciencedirect.com/science/article/pii/S1046202301912629?via%3Dihub>
209. Kassambara A. Package “ggpubr” Type Package Title “ggplot2” Based Publication Ready Plots. 2018 [cited 2020 Jun 29]; Available from: <https://rpkgs.datanovia.com/ggpubr/>
210. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. 2013. Available from: <http://www.r-project.org/>
211. Team RC. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2019. Available from: <https://www.r-project.org/>
212. Pohlert T. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). 2014; Available from: <https://cran.r-project.org/package=PMCMR>
213. Giavarina D. Understanding Bland Altman analysis. Biochem Medica. 2015;25(2):141–51.
214. Dong Y, Peng C-YJ. Principled missing data methods for researchers. Springerplus [Internet]. 2013 Dec [cited 2019 Sep 2];2(1):222. Available from:

<http://www.ncbi.nlm.nih.gov/pubmed/23853744>

215. Li C. Little's test of missing completely at random. Vol. 13, The Stata Journal. 2013.
216. Li Q, Birkbak NJ, Györfy B, Szallasi Z, Eklund AC. Jetset: Selecting the optimal microarray probe set to represent a gene. BMC Bioinformatics [Internet]. 2011 Dec 15 [cited 2020 May 11];12(1):474. Available from: <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-474>
217. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics [Internet]. 2020 Jan 2 [cited 2020 May 7];21(1):6. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-019-6413-7>
218. Donovan JL, Young GJ, Walsh EI, Metcalfe C, Lane JA, Martin RM, et al. A prospective cohort and extended comprehensive-cohort design provided insights about the generalizability of a pragmatic trial: the ProtecT prostate cancer trial. J Clin Epidemiol. 2018 Apr 1;96:35–46.
219. Johnston E, Pye H, Bonet-Carne E, Panagiotaki E, Patel D, Galazi M, et al. INNOVATE: A prospective cohort study combining serum and urinary biomarkers with novel diffusion-weighted magnetic resonance imaging for the prediction and characterization of prostate cancer. BMC Cancer. 2016;16(1).
220. Ross-Adams H, Lamb A, Dunning M, Halim S, Lindberg J, Massie C, et al. Integration of copy number and transcriptomics provides risk stratification in prostate cancer: A discovery and validation cohort study. EBioMedicine [Internet]. 2015 Sep 1 [cited 2020 Jun 29];2(9):1133–44. Available from: <https://pubmed.ncbi.nlm.nih.gov/26501111/>
221. Lapointe J, Li C, Higgins JP, Van De Rijn M, Bair E, Montgomery K, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A [Internet]. 2004 Jan 20 [cited 2020 Jun 29];101(3):811–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/14711987/>
222. Glinsky G V., Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. J Clin Invest [Internet]. 2004 [cited 2020 Jun 29];113(6):913–23. Available from: [/pmc/articles/PMC362118/?report=abstract](http://pmc/articles/PMC362118/?report=abstract)
223. Batra JS, Girdhani S, Hlatky L. A Quest to Identify Prostate Cancer Circulating Biomarkers with a Bench-to-Bedside Potential. J biomarkers [Internet]. 2014 [cited 2019 Oct 11];2014:321680. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26317031>
224. Clinton TN, Bagrodia A, Lotan Y, Margulis V, Raj G V, Woldu SL. Tissue-based biomarkers in prostate cancer. Expert Rev Precis Med drug Dev [Internet]. 2017 [cited 2019 Apr 1];2(5):249–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29226251>

225. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004 Dec 30;351(27):2817–26.
226. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* [Internet]. 2018 Jul 12 [cited 2020 May 26];379(2):111–21. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa1804710>
227. Clark-Langone KM, Sangli C, Krishnakumar J, Watson D. Translating tumor biology into personalized treatment planning: Analytical performance characteristics of the Oncotype DX® Colon Cancer Assay. *BMC Cancer*. 2010 Dec 23;10.
228. Gray RG, Quirke P, Handley K, Lopatin M, Magill L, Baehner FL, et al. Validation Study of a Quantitative Multigene Reverse Transcriptase–Polymerase Chain Reaction Assay for Assessment of Recurrence Risk in Patients With Stage II Colon Cancer. *J Clin Oncol* [Internet]. 2011 Dec 10 [cited 2020 May 26];29(35):4611–9. Available from: <http://ascopubs.org/doi/10.1200/JCO.2010.32.8732>
229. Klein EA, Cooperberg MR, Magi-Galluzzi C, Simko JP, Falzarano SM, Maddala T, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol* [Internet]. 2014/05/20. 2014;66(3):550–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24836057>
230. Cullen J, Rosner IL, Brand TC, Zhang N, Tsiatis AC, Moncur J, et al. A biopsy-based 17-gene genomic prostate score predicts recurrence after radical prostatectomy and adverse surgical pathology in a racially diverse population of men with clinically low- and intermediate-risk prostate cancer. *Eur Urol*. 2015 Jul 1;68(1):123–31.
231. Cooper CS, Eeles R, Wedge DC, Loo P, Gundem G, Alexandrov LB, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet* [Internet]. 2015;47(4):367–72. Available from: <http://dx.doi.org/10.1038/ng.3221>
232. Nonn L, Ananthanarayanan V, Gann PH. Evidence for field cancerization of the prostate. *Prostate* [Internet]. 2009 Sep 15 [cited 2019 Mar 6];69(13):1470–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19462462>
233. Sharma NL, Massie CE, Ramos-Montoya A, Zecchini V, Scott HE, Lamb AD, et al. The androgen receptor induces a distinct transcriptional program in castration-resistant prostate cancer in man. *Cancer Cell*. 2013;23(1):35–47.
234. Heinlein CA, Chang C. Androgen receptor in prostate cancer. *Endocr Rev* [Internet]. 2004;25(2):276–308. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/15082523>

235. Eiberger B, Schilling K. Cerebellins: capstones to bridge the synaptic cleft. *J Neurochem* [Internet]. 2012;121(5):697–9. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22288876>
236. Qiao J, Fang C-Y, Chen S-X, Wang X-Q, Cui S-J, Liu X-H, et al. Stroma derived COL6A3 is a potential prognosis marker of colorectal carcinoma revealed by quantitative proteomics. *Oncotarget* [Internet]. 2015 Oct 6 [cited 2019 Mar 29];6(30):29929–46. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26338966>
237. Bruzzese F, Hagglof C, Leone A, Sjoberg E, Roca MS, Kiflemariam S, et al. Local and systemic protumorigenic effects of cancer-associated fibroblast-derived GDF15. *Cancer Res* [Internet]. 2014;74(13):3408–17. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24780757>
238. Toyoshima M, Howie HL, Imakura M, Walsh RM, Annis JE, Chang AN, et al. Functional genomics identifies therapeutic targets for MYC-driven cancer. *Proc Natl Acad Sci U S A* [Internet]. 2012;109(24):9545–50. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22623531>
239. Ulmert D, O'Brien MF, Bjartell AS, Lilja H. Prostate kallikrein markers in diagnosis, risk stratification and prognosis. *Nat Rev Urol* [Internet]. 2009;6(7):384–91. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19578355>
240. Whitaker HC, Shiong LL, Kay JD, Gronberg H, Warren AY, Seipel A, et al. N-acetyl-L-aspartyl-L-glutamate peptidase-like 2 is overexpressed in cancer and promotes a pro-migratory and pro-metastatic phenotype. *Oncogene* [Internet]. 2014;33(45):5274–87. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24240687>
241. Whitaker HC, Warren AY, Eeles R, Kote-Jarai Z, Neal DE. The potential value of microseminoprotein-beta as a prostate cancer biomarker and therapeutic target. *Prostate* [Internet]. 2010;70(3):333–40. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19790236>
242. Di Noia MA, Todisco S, Cirigliano A, Rinaldi T, Agrimi G, Iacobazzi V, et al. The human SLC25A33 and SLC25A36 genes of solute carrier family 25 encode two mitochondrial pyrimidine nucleotide transporters. *J Biol Chem* [Internet]. 2014;289(48):33137–48. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/25320081>
243. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia* [Internet]. 2008;10(2):177–88. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18283340>
244. Cheng I, Levin AM, Tai YC, Plummer S, Chen GK, Neslund-Dudas C, et al. Copy number alterations in prostate tumors and disease aggressiveness. *Genes Chromosom Cancer* [Internet]. 2012;51(1):66–76. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21965145>
245. Sobel RE, Sadar MD. CELL LINES USED IN PROSTATE CANCER RESEARCH: A

- COMPENDIUM OF OLD AND NEW LINES—PART 1. [cited 2017 Sep 18]; Available from: http://ac.els-cdn.com/S0022534705604673/1-s2.0-S0022534705604673-main.pdf?_tid=24b4622a-9c3d-11e7-a64b-00000aab0f6b&acdnat=1505717388_10e2269ea8f400c90df0144bf8478dd0
246. Wu X, Gong S, Roy-Burman P, Lee P, Culig Z. Current mouse and cell models in prostate cancer research.
 247. Martínez-Fernández M, Paramio JM, Dueñas M. RNA Detection in Urine: From RNA Extraction to Good Normalizer Molecules. 2016 Jan 1 [cited 2019 Nov 16];18(1):15–22. Available from: <https://www.sciencedirect.com/science/article/pii/S1525157815002202?via%3Dihub>
 248. Gillet J-P, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. *J Natl Cancer Inst* [Internet]. 2013 Apr 3 [cited 2019 Mar 6];105(7):452–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23434901>
 249. Na D, Son H, Gsponer J. Categorizer: a tool to categorize genes into user-defined biological groups based on semantic similarity [Internet]. 2014 [cited 2019 Apr 1]. Available from: <http://www.biomedcentral.com/1471-2164/15/1091>
 250. Baijer J, Déchamps N, Perdry H, Morales P, Kerns S, Vasilescu A, et al. TNFSF10/TRAIL regulates human T4 effector memory lymphocyte radiosensitivity and predicts radiation-induced acute and subacute dermatitis. *Oncotarget* [Internet]. 2016 Apr 19 [cited 2019 Mar 28];7(16):21416–27. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26982083>
 251. Hartwig T, Montinaro A, Von S, El-Bahrawy MA, Quezada SA, Correspondence HW, et al. The TRAIL-Induced Cancer Secretome Promotes a Tumor-Supportive Immune Microenvironment via CCR2 In Brief Hartwig et al. show that endogenous TRAIL signaling in cancer cells induces a FADD-dependent secretome that promotes the accumulation of M2-like immune cells and tumor growth via host CCR2. The TRAIL-Induced Cancer Secretome Promotes a Tumor-Supportive Immune Microenvironment via CCR2. *Mol Cell* [Internet]. 2017 [cited 2019 Mar 28];65:730–42. Available from: <http://dx.doi.org/10.1016/j.molcel.2017.01.021>
 252. Wang D, Lu J, Tindall DJ. Androgens regulate TRAIL-induced cell death in prostate cancer cells via multiple mechanisms. *Cancer Lett* [Internet]. 2013 Jul 10 [cited 2019 Mar 13];335(1):136–44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23402817>
 253. Armstrong AP, Miller RE, Jones JC, Zhang J, Keller ET, Dougall WC. RANKL acts directly on RANK-expressing prostate tumor cells and mediates migration and expression of tumor metastasis genes. *Prostate* [Internet]. 2008;68(1):92–104. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18008334>
 254. Jones DH, Nakashima T, Sanchez OH, Kozieradzki I, Komarova S V, Sarosi I, et al. Regulation

- of cancer cell migration and bone metastasis by RANKL. *Nature* [Internet]. 2006;440(7084):692–6. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/16572175>
255. Tsukasaki K, Miller CW, Greenspun E, Eshaghian S, Kawabata H, Fujimoto T, et al. Mutations in the mitotic check point gene, MAD1L1, in human cancers. *Oncogene* [Internet]. 2001 May 22 [cited 2019 Mar 13];20(25):3301–5. Available from: <http://www.nature.com/articles/1204421>
 256. Zhong Q, Liu Z-H, Lin Z-R, Hu Z-D, Yuan L, Liu Y-M, et al. The RARS-MAD1L1 Fusion Gene Induces Cancer Stem Cell-like Properties and Therapeutic Resistance in Nasopharyngeal Carcinoma. 2017 [cited 2019 Mar 13]; Available from: <http://clincancerres.aacrjournals.org/content/clincanres/early/2017/11/11/1078-0432.CCR-17-0352.full.pdf>
 257. Rabinowicz N, Mangala LS, Brown KR, Checa-Rodriguez C, Castiel A, Moskovich O, et al. Targeting the centriolar replication factor STIL synergizes with DNA damaging agents for treatment of ovarian cancer. *Oncotarget* [Internet]. 2017 Apr 18 [cited 2019 Mar 13];8(16):27380–92. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28423708>
 258. Patwardhan D, Mani S, Passemard S, Gressens P, El Ghouzzi V. STIL balancing primary microcephaly and cancer. *Cell Death Dis* [Internet]. 2018 [cited 2019 Mar 13];9(2):65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29352115>
 259. Wu X, Xiao Y, Yan W, Ji Z, Zheng G. The human oncogene SCL/TAL1 interrupting locus (STIL) promotes tumor growth through MAPK/ERK, PI3K/Akt and AMPK pathways in prostate cancer. *Gene* [Internet]. 2019 Feb 20 [cited 2019 Feb 22];686:220–7. Available from: <https://www.sciencedirect.com/science/article/pii/S0378111918311910?via%3Dihub>
 260. Fitzgerald J, Holden P, Hansen U. The expanded collagen VI family: new chains and new questions. *Connect Tissue Res* [Internet]. 2013;54(6):345–50. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/23869615>
 261. Schenck J, Fitzgerald J. Analysis of COL6A3, COL6A5 and COL6A6 Gene Expression in Breast and Prostate Tumors. *Adv Clin Transl Res* [Internet]. [cited 2019 Mar 29];2(2):1–10. Available from: <https://www.onjourn.org/open-access-article/analysis-of-col6a3-col6a5-and-col6a6-gene-expression-in-breast-and-prostate-tumors-100011.html>
 262. Thorsen K, Sorensen KD, Brems-Eskildsen AS, Modin C, Gaustadnes M, Hein AM, et al. Alternative splicing in colon, bladder, and prostate cancer identified by exon array analysis. *Mol Cell Proteomics* [Internet]. 2008;7(7):1214–24. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18353764>
 263. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature* [Internet]. 2012;487(7406):239–43. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22722839>

264. Mandal RK, Kapoor R, Mittal RD. Polymorphic variants of DNA repair gene XRCC3 and XRCC7 and risk of prostate cancer: A study from North Indian population. *DNA Cell Biol* [Internet]. 2010 Nov 25 [cited 2019 Mar 29];29(11):669–74. Available from: <http://www.liebertpub.com/doi/10.1089/dna.2010.1047>
265. Barry KH, Koutros S, Andreotti G, Sandler DP, Burdette LA, Yeager M, et al. Genetic variation in nucleotide excision repair pathway genes, pesticide exposure and prostate cancer risk. *Carcinogenesis* [Internet]. 2012;33(2):331–7. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22102698>
266. Zhao L, Lee BY, Brown DA, Molloy MP, Marx GM, Pavlakis N, et al. Identification of candidate biomarkers of therapeutic response to docetaxel by proteomic profiling. *Cancer Res* [Internet]. 2009;69(19):7696–703. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19773444>
267. Brown DA, Lindmark F, Stattin P, Balter K, Adami HO, Zheng SL, et al. Macrophage inhibitory cytokine 1: a new prognostic marker in prostate cancer. *Clin Cancer Res* [Internet]. 2009;15(21):6658–64. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19843661>
268. Itkonen HM, Brown M, Urbanucci A, Tredwell G, Lau CH, Barfeld S, et al. Lipid degradation promotes prostate cancer cell survival. *Oncotarget* [Internet]. 2017 Jun 13 [cited 2019 Feb 27];8(24):38264–75. Available from: <http://www.oncotarget.com/fulltext/16123>
269. Yoshie H, Sedukhina AS, Minagawa K, Oda K, Ohnuma S, Yanagisawa N, et al. A bioinformatics-to-clinic sequential approach to analysis of prostate cancer biomarkers using TCGA datasets and clinical samples: a new method for precision oncology? *Oncotarget* [Internet]. 2017 Nov 21 [cited 2019 Mar 13];8(59):99601–11. Available from: <http://www.oncotarget.com/fulltext/20448>
270. Sebastian A, Hum NR, Hudson BD, Loots GG. Cancer-Osteoblast Interaction Reduces Sost Expression in Osteoblasts and Up-Regulates lncRNA MALAT1 in Prostate Cancer. *Microarrays (Basel)* [Internet]. 2015;4(4):503–19. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/27600237>
271. Tian S, Tao K, Hu J, Liu Z, Ding X, Chu Y, et al. The prognostic value of AGR2 expression in solid tumours: a systematic review and meta-analysis. *Sci Rep* [Internet]. 2017 Dec 14 [cited 2019 Mar 13];7(1):15500. Available from: <http://www.nature.com/articles/s41598-017-15757-z>
272. Hu Z, Gu Y, Han B, Zhang J, Li Z, Tian K, et al. Knockdown of AGR2 induces cellular senescence in prostate cancer cells. *Carcinogenesis* [Internet]. 2012 Jun 1 [cited 2019 Mar 13];33(6):1178–86. Available from: <https://academic.oup.com/carcin/article-lookup/doi/10.1093/carcin/bgs141>
273. Zhang J-S, Gong A, Cheville JC, Smith DI, Young CYF. AGR2, an androgen-inducible secretory protein overexpressed in prostate cancer. *Genes, Chromosom Cancer* [Internet]. 2005 Jul 1 [cited 2019 Mar 13];43(3):249–59. Available from: <http://doi.wiley.com/10.1002/gcc.20188>

274. Bu H, Schweiger MR, Manke T, Wunderlich A, Timmermann B, Kerick M, et al. Anterior gradient 2 and 3 - Two prototype androgen-responsive genes transcriptionally upregulated by androgens and by oestrogens in prostate cancer cells. *FEBS J* [Internet]. 2013 Mar 1 [cited 2019 Feb 27];280(5):1249–66. Available from: <http://doi.wiley.com/10.1111/febs.12118>
275. Gottlieb B, Beitel LK, Wu JH, Trifiro M. The androgen receptor gene mutations database (ARDB): 2004 update. *Hum Mutat* [Internet]. 2004;23(6):527–33. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/15146455>
276. Brooke GN, Bevan CL. The role of androgen receptor mutations in prostate cancer progression. *Curr Genomics* [Internet]. 2009;10(1):18–25. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19721807>
277. Tan MH, Li J, Xu HE, Melcher K, Yong EL. Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol Sin* [Internet]. 2015;36(1):3–23. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24909511>
278. Balk SP, Knudsen KE. AR, the cell cycle, and prostate cancer. *Nucl Recept Signal* [Internet]. 2008;6:e001. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18301781>
279. Hillebrand AC, Pizzolato LS, Branchini G, Brum IS. Androgenic modulation of AR-Vs. *Endocrine* [Internet]. 2018 Nov 19 [cited 2019 Nov 16];62(2):477–86. Available from: <http://link.springer.com/10.1007/s12020-018-1682-5>
280. Ogasawara D, Ichu T-A, Vartabedian VF, Benthuyssen J, Jing H, Reed A, et al. Selective blockade of the lyso-PS lipase ABHD12 stimulates immune responses in vivo. *Nat Chem Biol* [Internet]. 2018 Dec 12 [cited 2019 Mar 13];14(12):1099–108. Available from: <http://www.nature.com/articles/s41589-018-0155-8>
281. Rodriguez V, Chen Y, Elkahloun A, Dutra A, Pak E, Chandrasekharappa S. Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of *TRMT12* in breast cancer. *Genes, Chromosom Cancer* [Internet]. 2007 Jul 1 [cited 2019 Mar 28];46(7):694–707. Available from: <http://doi.wiley.com/10.1002/gcc.20454>
282. Jiao Y, Fu Z, Li Y, Meng L, Liu Y. High EIF2B5 mRNA expression and its prognostic significance in liver cancer: a study based on the TCGA and GEO database. *Cancer Manag Res* [Internet]. 2018 Nov 20 [cited 2019 Mar 13];Volume 10:6003–14. Available from: <https://www.dovepress.com/high-eif2b5-mrna-expression-and-its-prognostic-significance-in-liver-c-peer-reviewed-article-CMAR>
283. Wang Z-L, Li B, Luo Y-X, Zhou H, Yang J-H, Correspondence L-HQ, et al. Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers *Cell Reports Resource Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers. CellReports* [Internet]. 2018 [cited 2019 Mar 13];22:286–98. Available from:

<https://doi.org/10.1016/j.celrep.2017.12.035>

284. Lamb R, Harrison H, Hulit J, Smith DL, Lisanti MP, Sotgia F. Mitochondria as new therapeutic targets for eradicating cancer stem cells: Quantitative proteomics and functional validation via MCT1/2 inhibition. *Oncotarget* [Internet]. 2014 Nov 30 [cited 2019 Mar 13];5(22):11029–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25415228>
285. Arencibia JM, Martín S, Pérez-Rodríguez FJ, Bonnin A. Gene expression profiling reveals overexpression of TSPAN13 in prostate cancer. *Int J Oncol* [Internet]. 2009 Jan 1 [cited 2019 Mar 6];34(2):457–63. Available from: http://www.spandidos-publications.com/ijo/article.jsp?article_id=ijo_34_2_457
286. Yan C, Kim YH, Kang HW, Seo SP, Jeong P, Lee IS, et al. Urinary nucleic acid TSPAN13-to-S100A9 ratio as a diagnostic marker in prostate cancer. *J Korean Med Sci*. 2015;30(12):1784–92.
287. Jaiswal RK, Kumar P, Kumar M, Yadava PK. hTERT promotes tumor progression by enhancing TSPAN13 expression in osteosarcoma cells. *Mol Carcinog* [Internet]. 2018 Aug 1 [cited 2019 Mar 28];57(8):1038–54. Available from: <http://doi.wiley.com/10.1002/mc.22824>
288. Lai J, An J, Nelson CC, Lehman ML, Batra J, Clements JA. Analysis of androgen and anti-androgen regulation of KLK-related peptidase 2, 3, and 4 alternative transcripts in prostate cancer. *Biol Chem* [Internet]. 2014 Sep 1 [cited 2019 Nov 16];395(9):1127–32. Available from: <http://www.degruyter.com/view/j/bchm.2014.395.issue-9/hsz-2014-0149/hsz-2014-0149.xml>
289. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nat Rev Cancer* [Internet]. 2008;8(4):268–78. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18337732>
290. Dimberg LY, Towers CG, Behbakht K, Hotz TJ, Kim J, Fosmire S, et al. A Genome-Wide Loss-of-Function Screen Identifies SLC26A2 as a Novel Mediator of TRAIL Resistance. *Mol Cancer Res* [Internet]. 2017 Apr 1 [cited 2019 Mar 6];15(4):382–94. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28108622>
291. Hornung T, Logie SC, Bondre AS, Zarkovic J, Tinder TT, Maher V, et al. Identification of exosomal proteins from a prostate cancer cell line by Adaptive Dynamic Artificial Polyligand Targeting (ADAPT). *J Clin Oncol* [Internet]. 2016 May 20 [cited 2019 Mar 29];34(15_suppl):e23050–e23050. Available from: http://ascopubs.org/doi/10.1200/JCO.2016.34.15_suppl.e23050
292. Squire JA. TMPRSS2-ERG and PTEN loss in prostate cancer. *Nat Genet* [Internet]. 2009;41(5):509–10. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19399032>
293. Hagglof C, Hammarsten P, Stromvall K, Egevad L, Josefsson A, Stattin P, et al. TMPRSS2-

- ERG expression predicts prostate cancer survival and associates with stromal biomarkers. *PLoS One* [Internet]. 2014;9(2):e86824. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24505269>
294. Nguyen PN, Violette P, Chan S, Tanguay S, Kassouf W, Aprikian A, et al. A panel of TMPRSS2:ERG fusion transcript markers for urine-based prostate cancer detection with high specificity and sensitivity. *Eur Urol* [Internet]. 2011;59(3):407–14. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21144642>
 295. Chan SW, Nguyen PN, Violette P, Brimo F, Taguchi Y, Aprikian A, et al. Early detection of clinically significant prostate cancer at diagnosis: a prospective study using a novel panel of TMPRSS2:ETS fusion gene markers. *Cancer Med* [Internet]. 2013;2(1):63–75. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24133629>
 296. Schoenborn JR, Nelson P, Fang M. Genomic profiling defines subtypes of prostate cancer with the potential for therapeutic stratification. *Clin Cancer Res* [Internet]. 2013;19(15):4058–66. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/23704282>
 297. Sjoblom L, Saramaki O, Annala M, Leinonen K, Nattinen J, Tolonen T, et al. Microseminoprotein-Beta Expression in Different Stages of Prostate Cancer. *PLoS One* [Internet]. 2016;11(3):e0150241. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/26939004>
 298. Sutcliffe S, De Marzo AM, Sfanos KS, Laurence M. MSMB variation and prostate cancer risk: clues towards a possible fungal etiology. *Prostate* [Internet]. 2014 May [cited 2019 Mar 29];74(6):569–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24464504>
 299. Allioli N, Vincent S, Vlaeminck-Guillem V, Decaussin-Petrucci M, Ragage F, Ruffion A, et al. TM4SF1, a novel primary androgen receptor target gene over-expressed in human prostate cancer and involved in cell migration. *Prostate* [Internet]. 2011 Aug 1 [cited 2019 Feb 21];71(11):1239–50. Available from: <http://doi.wiley.com/10.1002/pros.21340>
 300. Lin C-I, Merley A, Sciuto TE, Li D, Dvorak AM, Melero-Martin JM, et al. TM4SF1: a new vascular therapeutic target in cancer. *Angiogenesis* [Internet]. 2014 Oct 2 [cited 2019 Mar 13];17(4):897–907. Available from: <http://link.springer.com/10.1007/s10456-014-9437-2>
 301. Ho ME, Quek S-I, True LD, Morrissey C, Corey E, Vessella RL, et al. Prostate cancer cell phenotypes based on AGR2 and CD10 expression. *Mod Pathol* [Internet]. 2013 Jun [cited 2019 Mar 13];26(6):849–59. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23348903>
 302. Brady LK, Wang H, Radens CM, Bi Y, Radovich M, Maity A, et al. Transcriptome analysis of hypoxic cancer cells uncovers intron retention in EIF2B5 as a mechanism to inhibit translation. Christofk H, editor. *PLOS Biol* [Internet]. 2017 Sep 29 [cited 2019 Mar 13];15(9):e2002623. Available from: <http://dx.plos.org/10.1371/journal.pbio.2002623>

303. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Vol. 27, *Molecular Aspects of Medicine*. Pergamon; 2006. p. 126–39.
304. Ruiz-Villalba A, Van Pelt-Verkuil E, Gunst QD, Ruijter JM, Van Den Hoff MJ. Amplification of nonspecific products in quantitative polymerase chain reactions (qPCR). *Biomol Detect Quantif* [Internet]. 2017 [cited 2020 Jun 29];14:7–18. Available from: <http://dx.doi.org/10.1016/j.bdq.2017.10.001>
305. Varambally S, Yu J, Laxman B, Rhodes DR, Mehra R, Tomlins SA, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* [Internet]. 2005 Nov [cited 2020 Jun 29];8(5):393–406. Available from: <https://pubmed.ncbi.nlm.nih.gov/16286247/>
306. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell* [Internet]. 2010;18(1):11–22. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/20579941>
307. Li J, Liu Y, Kim T, Min R, Zhang Z. Gene Expression Variability within and between Human Populations and Implications toward Disease Susceptibility. 2010 [cited 2019 Mar 6]; Available from: www.ploscompbiol.org
308. Husaini Y, Lockwood GP, Nguyen T V., Tsai VW-W, Mohammad MG, Russell PJ, et al. Macrophage Inhibitory Cytokine-1 (MIC-1/GDF15) Gene Deletion Promotes Cancer Growth in TRAMP Prostate Cancer Prone Mice. Koul HK, editor. *PLoS One* [Internet]. 2015 Feb 19 [cited 2019 Mar 6];10(2):e0115189. Available from: <http://dx.plos.org/10.1371/journal.pone.0115189>
309. Bansal N, Kumar D, Gupta A, Chandra D, Sankhwar SN, Mandhani A. Relevance of MIC-1 in the Era of PSA as a Serum Based Predictor of Prostate Cancer: A Critical Evaluation. *Sci Rep* [Internet]. 2017 Dec 4 [cited 2019 Mar 6];7(1):16824. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29203798>
310. GC B, EA E, MJ E. Cell Line Cross-Contamination: How Aware Are Mammalian Cell Culturists of the Problem and How to Monitor It? *In Vitro Cell Dev Biol Anim*. 2004;40(7).
311. Zeeberg BR, Kohn KW, Kahn A, Larionov V, Weinstein JN, Reinhold W, et al. Concordance of gene expression and functional correlation patterns across the NCI-60 cell lines and the Cancer Genome Atlas glioblastoma samples. *PLoS One*. 2012 Jul 26;7(7).
312. Cunningham D, You Z. In vitro and in vivo model systems used in prostate cancer research.
313. Taylor SC, Nadeau K, Abbasi M, Lachance C, Nguyen M, Fenrich J. The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. Vol. 37, *Trends in Biotechnology*. Elsevier Ltd; 2019. p. 761–74.
314. Medeiros M, Sharma VK, Ding R, Yamaji K, Li B, Muthukumar T, et al. Optimization of RNA

- yield, purity and mRNA copy number by treatment of urine cell pellets with RNeasy. *J Immunol Methods*. 2003 Aug 1;279(1–2):135–42.
315. Litwin MS, Tan HJ. The diagnosis and treatment of prostate cancer: A review. Vol. 317, *JAMA - Journal of the American Medical Association*. American Medical Association; 2017. p. 2532–42.
 316. Bjurlin MA, Carter HB, Schellhammer P, Cookson MS, Gomella LG, Troyer D, et al. Optimization of Initial Prostate Biopsy in Clinical Practice: Sampling, Labeling and Specimen Processing. *J Urol* [Internet]. 2013 Jun [cited 2020 May 1];189(6):2039–46. Available from: <http://www.jurology.com/doi/10.1016/j.juro.2013.02.072>
 317. Crawford ED, Grubb R, Black A, Andriole GL, Chen MH, Izmirlian G, et al. Comorbidity and mortality results from a randomized prostate cancer screening trial. *J Clin Oncol*. 2011 Feb 1;29(4):355–61.
 318. Draisma G, Boer R, Otto SJ, van der Crujisen IW, Damhuis RAM, Schröder FH, et al. Lead times and overdiagnosis due to prostate-specific antigen screening: Estimates from the European randomized study of screening for prostate cancer. *J Natl Cancer Inst*. 2003 Jun 18;95(12):868–78.
 319. Telesca D, Etzioni R, Gulati R. Estimating Lead Time and Overdiagnosis Associated with PSA Screening from Prostate Cancer Incidence Trends. *Biometrics* [Internet]. 2008 Mar 1 [cited 2020 May 1];64(1):10–9. Available from: <http://doi.wiley.com/10.1111/j.1541-0420.2007.00825.x>
 320. Groelz D, Sobin L, Branton P, Compton C, Wyrich R, Rainen L. Non-formalin fixative versus formalin-fixed tissue: A comparison of histology and RNA quality. *Exp Mol Pathol*. 2013 Feb 1;94(1):188–94.
 321. Antonov J, Goldstein DR, Oberli A, Baltzer A, Pirotta M, Fleischmann A, et al. Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. *Lab Invest*. 2005 Aug;85(8):1040–50.
 322. Chung JY, Braunschweig T, Hewitt SM. Optimization of recovery of RNA from formalin-fixed, paraffin-embedded tissue. *Diagnostic Mol Pathol*. 2006 Dec;15(4):229–36.
 323. Dietrich D, Uhl B, Sailer V, Holmes EE, Jung M, Meller S, et al. Improved PCR Performance Using Template DNA from Formalin-Fixed and Paraffin-Embedded Tissues by Overcoming PCR Inhibition. Castresana JS, editor. *PLoS One* [Internet]. 2013 Oct 14 [cited 2020 May 1];8(10):e77771. Available from: <https://dx.plos.org/10.1371/journal.pone.0077771>
 324. Okello JBA, Zurek J, Devault AM, Kuch M, Okwi AL, Sewankambo NK, et al. Comparison of methods in the recovery of nucleic acids from archival formalin-fixed paraffin-embedded autopsy tissues. *Anal Biochem*. 2010 May 1;400(1):110–7.
 325. Falagario UG, Beksac AT, Martini A, Cumarasamy S, Gupta A, Prasad S, et al. Defining Prostate

- Cancer at Favorable Intermediate Risk: The Potential Utility of Magnetic Resonance Imaging and Genomic Tests. *J Urol* [Internet]. 2019 Jul 1 [cited 2020 May 4];202(1):102–6. Available from: <http://www.jurology.com/doi/10.1097/JU.0000000000000134>
326. Grönberg H, Eklund M, Picker W, Aly M, Jäderling F, Adolfsson J, et al. Prostate Cancer Diagnostics Using a Combination of the Stockholm3 Blood Test and Multiparametric Magnetic Resonance Imaging. *Eur Urol*. 2018 Dec 1;74(6):722–8.
 327. Kotb AF, Spaner S, Crump T, Hyndman ME. The role of mpMRI and PSA density in patients with an initial negative prostatic biopsy. *World J Urol* [Internet]. 2018 Dec 1 [cited 2020 Jan 10];36(12):2021–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/29808301/>
 328. Monteiro MB, Santos-Bezerra DP, Thieme K, Passarelli M, Machado UF, Lin CJ, et al. Optimization of total RNA isolation from human urinary sediment. *Clin Chim Acta*. 2016 Nov 1;462:158–61.
 329. Wang G, Szeto CC. Quantification of gene expression in urinary sediment for the study of renal diseases. Vol. 12, *Nephrology*. 2007. p. 494–9.
 330. Hynie I, Meuffels M, Poznanski WJ. Determination of Phosphodiesterase I Activity in Human Blood Serum. *Clin Chem*. 1975 Sep 1;21(10):1383–7.
 331. Lüthje J, Ogilvie A. 5'-Nucleotide phosphodiesterase isoenzymes in human serum: Quantitative measurement and some biochemical properties. *Clin Chim Acta*. 1987 May 15;164(3):275–84.
 332. Bryzgunova OE, Laktionov PP. Extracellular Nucleic Acids in Urine: Sources, Structure, Diagnostic Potential. Vol. 7, *Acta Naturae*. Russian Federation Agency for Science and Innovation; 2015. p. 48–54.
 333. Ruijter JM, Ruiz Villalba A, Hellemans J, Untergasser A, van den Hoff MJB. Removal of between-run variation in a multi-plate qPCR experiment. *Biomol Detect Quantif*. 2015 Sep 1;5:10–4.
 334. Glasel JA. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques*. 1995;18(1):62–3.
 335. Schrader C, Schielke A, Ellerbroek L, John R. PCR inhibitors - occurrence, properties and removal. *J Appl Microbiol* [Internet]. 2012 Nov 1 [cited 2020 May 4];113(5):1014–26. Available from: <http://doi.wiley.com/10.1111/j.1365-2672.2012.05384.x>
 336. Pavšič J, Žel J, Milavec M. Assessment of the real-time PCR and different digital PCR platforms for DNA quantification Coefficient of variation dPCR Digital PCR Fast master mix TaqMan® Fast Virus 1-Step Master Mix gDNA Genomic DNA Gene master mix TaqMan® Gene Expression Master Mix HCMV Human cytomegalovirus LOD Limit of detection LOQ Limit of quantification MIQE Minimum Information for Publication of Quantitative Real-Time PCR Experiments NTC

- Negative template control. *Anal Bioanal Chem*. 2016;408:107–21.
337. Weber DG, Casjens S, Rozynek P, Lehnert M, Zilch-Schöneweis S, Bryk O, et al. Assessment of mRNA and microRNA stabilization in peripheral human blood for multicenter studies and biobanks. *Biomark Insights*. 2010;2010(5):95–102.
 338. Tang R, She Q, Lu Y, Yin R, Zhu P, Zhu L, et al. Quality control of RNA extracted from PAXgene Blood RNA tubes after different storage periods. *Biopreserv Biobank*. 2019 Oct 1;17(5):477–82.
 339. Benson MC, Ihn Seong Whang, Pantuck A, Ring K, Kaplan SA, Olsson CA, et al. Prostate specific antigen density: A means of distinguishing benign prostatic hypertrophy and prostate cancer. *J Urol*. 1992;147(3 II):815–6.
 340. Taylor JMG, Ankerst DP, Andridge RR. Validation of Biomarker-Based Risk Prediction Models. *Clin Cancer Res* [Internet]. 2008 Oct 1 [cited 2020 May 7];14(19):5977–83. Available from: <http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-07-4534>
 341. Hernández B, Parnell A, Pennington SR. Why have so few proteomic biomarkers “survived” validation? (Sample size and independent validation considerations). *Proteomics* [Internet]. 2014 Jul 1 [cited 2020 May 7];14(13–14):1587–92. Available from: <http://doi.wiley.com/10.1002/pmic.201300377>
 342. Epstein JI, Allsbrook WC, Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol* [Internet]. 2005 Sep [cited 2020 Jun 4];29(9):1228–42. Available from: <http://journals.lww.com/00000478-200509000-00015>
 343. Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part B: Prostate and Bladder Tumours. *Eur Urol*. 2016;70(1):106–19.
 344. Klotz L, Zhang L, Lam A, Nam R, Mamedov A, Loblaw A. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. *J Clin Oncol* [Internet]. 2010;28(1):126–31. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19917860>
 345. Ploussard G, Isbarn H, Briganti A, Sooriakumaran P, Surcel CI, Salomon L, et al. Can we expand active surveillance criteria to include biopsy Gleason 3+4 prostate cancer? A multi-institutional study of 2,323 patients. *Urol Oncol Semin Orig Investig*. 2015 Feb 1;33(2):71.e1-71.e9.
 346. Huang CC, Kong MX, Zhou M, Rosenkrantz AB, Taneja SS, Melamed J, et al. Gleason score 3+4=7 prostate cancer with minimal quantity of gleason pattern 4 on needle biopsy is associated with low-risk tumor in radical prostatectomy specimen. *Am J Surg Pathol*. 2014;38(8):1096–101.
 347. Cole AI, Morgan TM, Spratt DE, Palapattu GS, He C, Tomlins SA, et al. Prognostic Value of Percent Gleason Grade 4 at Prostate Biopsy in Predicting Prostatectomy Pathology and

- Recurrence. *J Urol* [Internet]. 2016 Aug [cited 2019 Jul 12];196(2):405–11. Available from: <http://www.jurology.com/doi/10.1016/j.juro.2016.01.120>
348. Stark JR, Perner S, Stampfer MJ, Sinnott JA, Finn S, Eisenstein AS, et al. Gleason Score and Lethal Prostate Cancer: Does $3 + 4 = 4 + 3$? *J Clin Oncol* [Internet]. 2009 Jul 20 [cited 2019 Jul 12];27(21):3459–64. Available from: <http://ascopubs.org/doi/10.1200/JCO.2008.20.4669>
 349. Ahmed HU, Hu Y, Carter T, Arumainayagam N, Lecornet E, Freeman A, et al. Characterizing Clinically Significant Prostate Cancer Using Template Prostate Mapping Biopsy. *J Urol* [Internet]. 2011 Aug [cited 2019 Jul 12];186(2):458–64. Available from: <http://www.jurology.com/doi/10.1016/j.juro.2011.03.147>
 350. Simopoulos DN, Sisk AE, Priester A, Felker ER, Kwan L, Delfin MK, et al. Cancer core length from targeted biopsy: an index of prostate cancer volume and pathological stage. *BJU Int* [Internet]. 2019 Feb 24 [cited 2019 Jul 12]; Available from: <http://doi.wiley.com/10.1111/bju.14691>
 351. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol* [Internet]. 2003 Jul [cited 2019 Dec 16];22(1):85–93. Available from: <http://doi.wiley.com/10.1002/uog.122>
 352. Kim J, Bang H. Three common misuses of P values. Vol. 7, *Dental Hypotheses*. Medknow Publications; 2016. p. 73–80.
 353. Shin T, Smyth TB, Ukimura O, Ahmadi N, de Castro Abreu AL, Ohe C, et al. Diagnostic accuracy of a five-point Likert scoring system for magnetic resonance imaging (MRI) evaluated according to results of MRI/ultrasonography image-fusion targeted biopsy of the prostate. *BJU Int* [Internet]. 2018 Jan 1 [cited 2020 May 20];121(1):77–83. Available from: <http://doi.wiley.com/10.1111/bju.13972>
 354. Harada T, Abe T, Kato F, Matsumoto R, Fujita H, Murai S, et al. Five-point Likert scaling on MRI predicts clinically significant prostate carcinoma Urology imaging. *BMC Urol*. 2015 Sep 4;15(1).
 355. Russo F, Regge D, Armando E, Giannini V, Vignati A, Mazzetti S, et al. Detection of prostate cancer index lesions with multiparametric magnetic resonance imaging (mp-MRI) using whole-mount histological sections as the reference standard. *BJU Int* [Internet]. 2016 Jul 1 [cited 2020 May 20];118(1):84–94. Available from: <http://doi.wiley.com/10.1111/bju.13234>
 356. de Souza MF, de Azevedo Araujo ALC, da Silva MT, Athanazio DA. The Gleason pattern 4 in radical prostatectomy specimens in current practice - Quantification, morphology and concordance with biopsy. *Ann Diagn Pathol* [Internet]. 2018 Jun 1 [cited 2019 Aug 27];34:13–7. Available from: <https://www.sciencedirect.com/science/article/pii/S1092913417303325>
 357. Berg KD, Roder MA, Brasso K, Vainer B, Iversen P. Primary Gleason pattern in biopsy Gleason

- score 7 is predictive of adverse histopathological features and biochemical failure following radical prostatectomy. *Scand J Urol* [Internet]. 2014 Apr 29 [cited 2019 Dec 16];48(2):168–76. Available from: <http://www.tandfonline.com/doi/full/10.3109/21681805.2013.821628>
358. Sato S, Kimura T, Yorozu T, Onuma H, Iwatani K, Egawa S, et al. Cases Having a Gleason Score 3+4=7 with <5% of Gleason Pattern 4 in Prostate Needle Biopsy Show Similar Failure-free Survival and Adverse Pathology Prevalence to Gleason Score 6 Cases in a Radical Prostatectomy Cohort. *Am J Surg Pathol*. 2019;43(11):1560–5.
 359. Nawaz S, Heindl A, Koelble K, Yuan Y. Beyond immune density: Critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Mod Pathol*. 2015 Jun 1;28(6):766–77.
 360. Morisita M. Measuring of dispersion of individuals and analysis of the distributional patterns. *Japanese J Ecol*. 1961 Dec 1;11(6):252.
 361. Hayes JJ, Castillo O. Geo-Information A New Approach for Interpreting the Morisita Index of Aggregation through Quadrat Size. *mdpi.com* [Internet]. [cited 2020 May 21]; Available from: www.mdpi.com/journal/ijgi
 362. Maley CC, Koelble K, Natrajan R, Aktipis A, Yuan Y. An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. *Breast Cancer Res* [Internet]. 2015 Sep 22 [cited 2020 May 21];17(1):131. Available from: <https://breast-cancer-research.biomedcentral.com/articles/10.1186/s13058-015-0638-4>
 363. Augustin H, Erbersdobler A, Graefen M, Jaekel T, Haese A, Huland H, et al. Differences in biopsy features between prostate cancers located in the transition and peripheral zone. *BJU Int* [Internet]. 2003 Apr 1 [cited 2020 Jun 5];91(6):477–81. Available from: <http://doi.wiley.com/10.1046/j.1464-410X.2003.04140.x>
 364. King CR, Ferrari M, Brooks JD. Prognostic significance of prostate cancer originating from the transition zone. *Urol Oncol Semin Orig Investig*. 2009 Nov;27(6):592–7.
 365. Lee JJ, Thomas I-C, Nolley R, Ferrari M, Brooks JD, Leppert JT. Biologic differences between peripheral and transition zone prostate cancer. *Prostate* [Internet]. 2015 Feb 1 [cited 2020 Jun 4];75(2):183–90. Available from: <http://doi.wiley.com/10.1002/pros.22903>
 366. Sakai I, Harada KI, Kurahashi T, Yamanaka K, Hara I, Miyake H. Analysis of differences in clinicopathological features between prostate cancers located in the transition and peripheral zones. *Int J Urol*. 2006 Apr;13(4):368–72.
 367. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare [Internet]. *Nature Medicine* Nature Publishing Group; Jan, 2019 p. 24–9. Available from: <http://www.nature.com/articles/s41591-018-0316-z>

368. Nir G, Hor S, Karimi D, Fazli L, Skinnider BF, Tavassoli P, et al. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. *Med Image Anal.* 2018 Dec 1;50:167–80.
369. Tzimagiorgis G, Michailidou EZ, Kritis A, Markopoulos AK, Koudou S. Recovering circulating extracellular or cell-free RNA from bodily fluids. Vol. 35, *Cancer Epidemiology*. Elsevier; 2011. p. 580–9.
370. Bradley MS, Boudreau MH, Grenier C, Huang Z, Murphy SK, Siddiqui NY. Urine RNA Processing in a Clinical Setting: Comparison of 3 Protocols. *Female Pelvic Med Reconstr Surg* [Internet]. 2019 May 1 [cited 2020 Jun 25];25(3):247–51. Available from: <https://europepmc.org/articles/PMC5975100>
371. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol.* 2019 Nov 1;16(11):703–15.
372. Xie W, Regan MM, Buyse M, Halabi S, Kantoff P, Sartor O, et al. Metastasis-free survival is a strong Surrogate of overall survival in localized prostate cancer. *J Clin Oncol* [Internet]. 2017 Sep 20 [cited 2020 Oct 11];35(27):3097–104. Available from: [/pmc/articles/PMC5652387/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/32063483/)
373. Roscigno M, Stabile A, Lughezzani G, Pepe P, Galosi AB, Naselli A, et al. The Use of Multiparametric Magnetic Resonance Imaging for Follow-up of Patients Included in Active Surveillance Protocol. Can PSA Density Discriminate Patients at Different Risk of Reclassification? *Clin Genitourin Cancer* [Internet]. 2020 [cited 2020 Jun 25]; Available from: <https://pubmed.ncbi.nlm.nih.gov/32493676/>
374. Han C, Liu S, Qin XB, Ma S, Zhu LN, Wang XY. MRI combined with PSA density in detecting clinically significant prostate cancer in patients with PSA serum levels of 4~10 ng/mL: Biparametric versus multiparametric MRI. *Diagn Interv Imaging* [Internet]. 2020 Apr 1 [cited 2020 Jun 25];101(4):235–44. Available from: <https://pubmed.ncbi.nlm.nih.gov/32063483/>