UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL

Statistical Analysis of Prostate Functional Magnetic Resonance Imaging Data

Liliana Anacleto Parreira

Mestrado em Bioestatística

Trabalho de projeto orientado por:
Prof[a] Helena Mouriño

2021

Supervisora externa:
Drª Adalgisa Guerra, Hospital da Luz Lisboa

Acknowledgments

I would like to express my thanks to all those who, in some way, made this journey of mine possible and made it easier.
First of all, I would like to thank the supervisor of this dissertation, Prof. Helena, for her guidance, for her encouragement, availability and support that she has always shown. I also express my gratitude to Dr. Gisa for her availability, for her attention and always a lot of concern and for all her support in the preparation of this work.
To all my family, especially my mother who never let me give up and always supported me. To my father who allowed all access to education, without him all this would not be possible. To my boyfriend who was always by my side and helped me stay focused when I was more unmotivated.
A special thanks to my uncles who welcomed me throughout my academic career and who always treated me with the greatest love, which I will never forget.
To all my friends who helped me grow as a person and were always there for me through the good times and the bad during this time.

To all my sincere thanks.

Resumo

O cancro da próstata é o mais frequentemente diagnosticado no homem atualmente, tratando-se de uma das principais causas de morte por cancro nos homens.
As decisões a tomar sobre como os pacientes com cancro da próstata devem ser encaminhados são complicadas, e apresentam um dilema tanto para os pacientes como para os seus médicos, isto porque o cancro de próstata apresenta uma vasta gama de atividades biológicas, onde a maioria dos casos não leva a uma morte específica por cancro da próstata. Além disso, as opções atuais de tratamento para homens com cancro da próstata localizado são agressivas e têm efeitos colaterais significativos. Assim, a comunidade de pesquisa médica tem como desafios prever com precisão o comportamento de um determinado cancro da próstata, selecionar os pacientes que necessitam de terapia e tratar o cancro com o nível de intensidade apropriado, preservando a qualidade de vida do paciente.
O tratamento para este tipo de cancro que é considerado o mais eficaz é a prostatectomia radical, que consiste na remoção cirúrgica da próstata e vesículas seminais. O enorme desenvolvimento tecnológico na área da imagem de ressonância magnética (IRM) com o aparecimento de equipamentos e métodos de aquisição e processamento de imagem cada vez mais sofisticados, tem conduzido a uma utilização crescente desta técnica em áreas como apoio ao diagnóstico, estadiamento de tumores e decisão terapêutica. A excelente resolução espacial e a diversidade de contrastes usados em IRM (por exemplo, imagem anatómica e difusão) conferem a esta técnica uma elevada sensibilidade e especificidade na deteção de tumores da próstata, sendo fundamental para o planeamento de intervenções cirúrgicas robóticas minimamente invasivas.
O interesse em mapear as características histológicas/funcionais/anatómicas de forma não invasiva utilizando IRM como possível alternativa à anatomia patológica e biomarcador com valor preditivo em cirurgias da próstata tem crescido recentemente. Contudo, tais metodologias carecem de validação clínica.

Neste estudo, foi questionado se a ressonância magnética é um método eficaz para o estudo da próstata antes da prostatectomia radical (cirurgia de remoção da próstata), atribuindo estadios ao tumor da próstata com base em biomarcadores e comparando a atribuição do estadiamento tumoral antes da operação com o estadiamento tumoral real encontrado após a operação do paciente.
A ressonância magnética é um excelente exame nos casos em que já existe um diagnóstico de cancro da próstata e se pretende saber a extensão deste aos tecidos/órgãos vizinhos, ou seja, é um exame que auxilia sobretudo a observar o estadiamento do tumor. A informação obtida pelo processo de estadiamento do cancro determina o estadio da doença, o que é fundamental para o tratamento poder ser planeado. O estadiamento do tumor da próstata inclui a análise da anatomia patológica à peça operatória, exames de imagem ou medicina nuclear. Neste caso em específico, vamo-nos centrar na análise patológica e na imagem.
O sistema de estadiamento mais utilizado é o sistema de classificação TNM. Este sistema baseia-se na extensão anatómica da doença, tendo em conta as características do tumor (T), as características dos linfonodos (N), e tendo em conta a presença ou ausência de metástases (M). Cada uma destas classificações tem a sua graduação: T (T0 a T4), N (N0 a N3) e M (M0 a M1), e são analisadas separadamente. Quando as classificações T, N e M são agrupadas, ficam distribuídas em estadios, que variam de I a IV, estadios esses que expressam o nível de evolução da doença. No estadio I o tumor é pequeno, limitado à próstata e cresce lentamente, no estadio II o tumor continua limitado à próstata, mas é mais provável que se espalhe, no estadio III o cancro espalha-se para além da próstata, mas não para os linfonodos ou outros órgãos, por

fim, no estadio IV o cancro espalha-se para as áreas mais próximas (bexiga, recto, linfonodos, ou órgãos mais distante e até mesmo ossos).

Neste trabalho temos apenas em conta a classificação T, pois é a que fornece informações sobre aspetos do tumor, como o seu tamanho, o quão profundo se desenvolveu na próstata e o quanto invadiu os tecidos adjacentes. A classificação T varia de entre T0 e T4, onde cada número descreve o tamanho do tumor ou a disseminação da doença nas proximidades. Quanto maior o número de T, maior a dimensão tumor ou mais se alastrou pelos tecidos próximos.

Os tumores da próstata podem crescer bastante e afetar toda a próstata. E podem igualmente crescer através da cápsula prostática e invadir os tecidos circundantes, a este fenómeno chama-se de extensão extracapsular (ECE), e é de notar que o mesmo ocorre no estadio T3. A amostra colectada para este projeto apenas contém doentes em estadio T2 e T3, e portanto, o facto de existir ECE equivale ao estadio T3.

O estadiamento do tumor através da classificação T pode ser clínico e patológico. O estadiamento clínico é estabelecido a partir do exame físico (toque retal) e da biopsia. Atualmente, há métodos imagiológicos que ajudam a classificar o tumor localmente (é aqui que entram as ressonâncias magnéticas - estadiamento por imagem). O estadiamento patológico baseia-se nos achados cirúrgicos. Logo, é estabelecido após tratamento cirúrgico e é o que determina a extensão da doença com maior precisão, é que nos fornece o verdadeiro estadiamento final (gold standard).

Portanto, o objetivo do trabalho foi determinar se o estadiamento imagiológico consegue ser suficiente individualmente, sem que seja necessário o estadiamento patológico. A solução para este problema foi avaliar que biomarcadores da ressonância magnética são relevantes para a predição do estadio T3 patológico, estadio este que é o mais importante pois é onde ocorre a extensão extracapsular e para além disso é o que define de que modo o paciente vai ser tratado. Os pacientes cujo tumor está neste estadio são os que têm pior prognóstico entre os estadios T2 e T3, e por isso são os que têm maior probabilidade da operação não ser suficiente para tratar a doença. Basicamente, a cirurgia neste tipo de pacientes é inútil e é por isso que é tão importante avaliar a eficácia da ressonância quanto à previsão do estadio do tumor, para evitar que este tipo de pacientes seja submetido a uma cirurgia desnecessariamente.

O fator preditivo para a presença de doença extracapsular em pacientes com neoplasia prostática foi estimado através de uma regressão logística. O modelo de regressão logística foi aplicado de modo a identificar que fatores estão significativamente associados à probabilidade de ocorrência da doença extracapsular, ou seja, o modelo consegue predizer o estadiamento do tumor com base nas características observadas do paciente.
Foram incluídas variáveis de categoria clínica, que correspondem à observação clínica da extensão da doença por parte do médico com base nos resultados dos exames físicos (incluindo exame retal) e da biópsia à próstata, variáveis de imagem, variáveis que são resultantes da observação da ressonância magnética, e ainda variáveis de categoria patológica, observadas pós cirurgia. Todas estas variáveis são chamadas de biomarcadores e foram consideradas as variáveis independentes do modelo, sendo que a variável dependente foi a extensão extracapsular (variável binária onde o resultado observado é T2/T3, ou seja, não existe/existe ECE).

Chegou-se à conclusão de que o nosso modelo usando características objetivas na ressonância magnética como o comprimento do contato capsular, uma característica consensual interpretativa semântica e a classificação de Gleason na biópsia da próstata pré-tratamento é fidedigno para detetar pacientes com extensão extracapsular antes do tratamento.
A qualidade no modelo estimado foi avaliada através da curva ROC e AUC, obtendo-se o valor de 90.9%, com uma sensibilidade de 86.4% e uma especificidade de 78.4%, o que declara que se trata de um modelo com uma boa performance. Assim, o modelo estimado apresenta um bom desempenho no que diz respeito à distinção de pacientes com ECE e sem ECE, tal característica é refletida ainda no valor de AUC de uma amostra de validação (87.6%).

Palavras-Chave: Cancro da próstata, ressonância magnética, extensão extracapsular, biomarcadores, regressão logística

Abstract

Prostate cancer is the most commonly diagnosed cancer in males, being one of the main causes of cancer related death in men. Management decisions for patients with prostate cancer are complicated and present a dilemma for both patients and their clinicians as prostate cancers demonstrate a wide range of biologic activity with the majority of cases not leading to a prostate cancer-specific death. Furthermore, the current treatment options for men with localized prostate cancer are aggressive and have significant side effects, such as incontinence, rectal injury and impotence. Thus it is clear that the challenges to the medical research community are to accurately predict a given prostate cancer's behavior to, select those patients who require therapy and to treat the cancer with the appropriate level of intensity while preserving the patient's quality of life. The treatment for this type of cancer that is considered the most effective is radical prostatectomy, which consists of surgical removal of the prostate and seminal vesicles.

The enormous technological development in the area of magnetic resonance imaging (MRI) with the arising of new equipment and methods of image acquisition and processing more and more sophisticated, has led to an increasing use of this technique in areas such as diagnosis support, tumour staging and therapeutic decision. The excellent spatial resolution and the diversity of contrasts used in MRI (for example, anatomical image and diffusion) give this technique a high sensitivity and specificity in the detection of prostate tumours, being fundamental for the planning of minimally invasive robotic surgical interventions. The interest in mapping non-invasive histological/functional/anatomical features using MRI as a possible alternative to the pathological anatomy and predictive biomarker in prostate surgeries has recently increased. However, such methodologies need clinical validation. In this study, we will be questioning whether the MRI is an effective method for studying prostate cancer before operating, thus staging the prostate tumour through biomarkers, and then comparing the tumour stage attribution before the operation with the actual tumour stage found after operating the patient. The predictive factor for the presence of extracapsular disease in patients with prostatic neoplasia will also be studied. A logistic regression will be performed to identify which factors are significantly associated with the probability of occurrence of extracapsular disease, that is, a logistic regression model will be used to predict the staging of the tumour, based on the observed characteristics of the patient. Data variables included age, prostate-specific antigen (PSA), tumor length contact (TLC), tumor volume, Gleason score.

Keywords: Prostate cancer, magnetic ressonance imaging, extracapsular extension, biomarkers, logistic regression

Contents

List of Figures

List of Tables

Acronyms and Abbreviations

MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
PCa	Prostate Cancer
PI-RADS	Prostate Imaging – Reporting and Data System
CT	Computed Tomography
DRE	Digital Rectal Examination
ECE	Extracapsular Extension
EPE	Extraprostatic Extension
PSA	Prostate Specific Antigen
TRUS	Transrectal Ultrasound
mpMRI	Multiparametric Magnetic Resonance Imaging
cT	clinical T category
pT	pathologic T category
iT	imaging T category
RARP	Robotic Assisted Radical Prostatectomy
DW	Diffusion-weighted
DCE	Dynamic contrast enhanced
AIC	Akaike Information Criterion
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve

1

Introduction

First of all, it is important to mention that this work was carried out as part of a research project conducted by Hospital da Luz Lisboa, with the coordination of Adalgisa Guerra, MD.

Constant research in an area of intervention as important as Prostate Cancer (PCa) is unquestionably necessary. New ways of preventing, detecting and treating it are being studied, always taking into account the improvement in the quality of life of people with this type of cancer, during and after treatment (*Liga Portuguesa Contra o Cancro* c2015). Prostate cancer remains a compelling medical health problem. About 1 in 6 men will be diagnosed with prostate cancer during their lifetime. About 1 in 36 men will eventually die of prostate cancer. Multiparametric Magnetic Resonance Imaging (MRI) has the potential to change the usual pathway for men suspected of having PCa. Localization of prostate cancer is the most important indication for MRI and if this can be achieved, the number of men who need serial transrectal ultrasonography biopsies could be reduced and targeted biopsies to high-risk areas in the gland would be allowed. Recently, the interest to integrate MRI in radiotherapy for prostate cancer has increased considerably. MRI can contribute to all steps of the radiotherapy workflow from diagnosis, staging, and target definition to treatment follow-up. Of particular interest is the ability of MRI to provide a wide range of functional measures (Olsson et al. 2019).

Due to its superior ability to define soft tissue structures, magnetic resonance imaging is today the preferred imaging modality for several anatomical locations such as the brain, the vertebral column, the abdomen and the pelvis. This inherent ability makes MRI of interest for direct implementation in radiotherapy. Detailed anatomical description using a uniform standard nomenclature and content such as provided by the Prostate Imaging – Reporting and Data System (PI-RADS) Version 2 (Weinerb et al. 2016) in prostate cancer improves consistency in retreatment staging of patients and thereby patient selection (Dirix, Haustermans, and Vandecaveye 2014). PI-RADS is a structured reporting scheme for multiparametric prostate MRI in the evaluation of suspected prostate cancer in treatment naive prostate glands.

MRI for target definition and organ at risk identification also has clear advantages over Computed Tomography (CT) in many cases, especially for the prostate (Salembier et al. 2018; Persson et al. 2017; Tyagi et al. 2017). These advantages have led to the development of MRI-only treatment planning, where the CT-simulation is excluded, and the attenuation data are extracted from Magnetic Resonance (MR) images. Additionally, MRI offers many possibilities for treatment follow-up, including adaptive adjustments during the treatment (McPartlin et al. 2016). Besides the tissue contrast and image resolution, of particular interest is the ability of MRI to enable more or less quantitative functional information, such as diffusion, perfusion, and MR spectroscopy acting as imaging biomarkers for relevant tissue changes

related to cancer.

1.1 Prostate cancer

Cancer is a disease in which cells in the body change and grow out of control, forming a tumor. A tumor can be cancerous or benign. A cancerous tumor is malignant, meaning it can grow and spread to other parts of the body. A benign tumor means the tumor can grow but will not spread (*American Society of Clinical Oncology (ASCO)* c2005-2021).

Prostate cells are responsible for the production of prostate tissue. In their normal state, these cells grow and divide into new cells, formed as they are needed; this is called cell regeneration. When normal prostate cells age or are damaged, they die naturally. When cells lose this control mechanism and undergo changes in their genome, that is, the complete set of DNA, they become cancer cells, which do not die when they age or become damaged, and produce new cells that are not needed uncontrollably, resulting in cancer formation.

Unlike normal cells, prostate cancer cells do not respect the borders of the organ, invading the surrounding tissues, and spread to other parts of the body. This process is called metastasis, which, together with the uncontrolled growth of this type of cells, makes PCa dangerous. Prostate cancer cells can replace or deform normal tissues, causing malfunction and/or failure of not only the prostate but other vital organs.

Over time, prostate tumors can grow large and affect the entire prostate. They can also grow through the prostatic capsule and invade the surrounding tissues. This growth is called ECE or Extraprostatic Extension (EPE).

Prostate cancer is one of the most common type of cancer found in men, and often begins without symptoms. This cancer can be slow-growing, such that many men die of other diseases before the PCa causes significant problems. However, many prostate cancers are more agressive and can spread outside the confines of the prostate gland, which can be deadly. The PCa survival rate is greatly improved with early detection and personalized treatment (*UCLA Health - UCLA Prostate Cancer Program* n.d.).

1.1.1 The organ

The prostate is a gland found only in males. The walnut-sized prostate gland produces a milky-white fluid that becomes part of the semen, the liquid ejaculated during sexual activity. It is located in front of the rectum and just below the bladder (see Figure 1.1). It also wraps around the upper part of the urethra, the tube that carries urine from the bladder out of the body. On each side of the prostate are found glands that are called seminal vesicles. They make most of the fluid in semen. The seminal vesicles are sometimes removed during surgery to remove the prostate (called a radical prostatectomy).

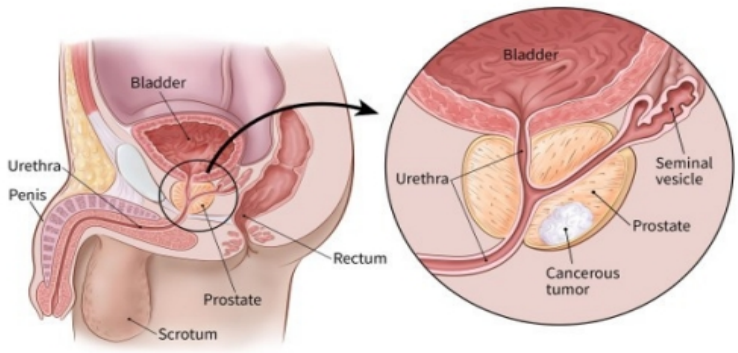


Figure 1.1: Anatomy of prostate gland (American Cancer Society) (Source: `https://www.cancer.org/cancer/prostate-cancer/about/what-is-prostate-cancer.html`)

The size of the prostate can change as a man ages. It is about the size of a walnut, in younger man, but it can be more prominent in older men.

The prostate is covered by tissue called a prostatic capsule. Inside the prostate, there are between 30 to 50 small sacs that produce and store prostatic fluid. The sacs and channels that carry the fluid constitute the glandular tissue of the prostate. Around these bags is the non-glandular tissue that contains blood vessels, lymphatic vessels, elastic fibers and muscle. The muscle allows the prostatic fluid to be moved into the urethra.

The primary function of the prostate is to make prostatic fluid. Prostatic fluid is rich in enzymes, proteins and minerals that help protect and nourish sperm. When a man is sexually aroused, the prostate pushes prostatic fluid through the ducts and into the urethra. Prostatic fluid mixes with sperm and other fluids in the urethra and is ejaculated as semen.

Hormones, including testosterone and those made by the pituitary gland and adrenal glands, help control the function of the prostate gland.

The upper part of the prostate is called the base, and the lower, narrowed part of the prostate, is called the apex. The prostate gland is divided into the right and left lobes.

The gland is surrounded by a thin layer of connective tissue called the capsule. In general, doctors divide the prostate into different areas according to its function.

The prostate consists of the capsule and three zones: the peripheral (1), central (2), and transition (3) zones (*Harvard Health Publishing - Harvard Medical School* c2021).

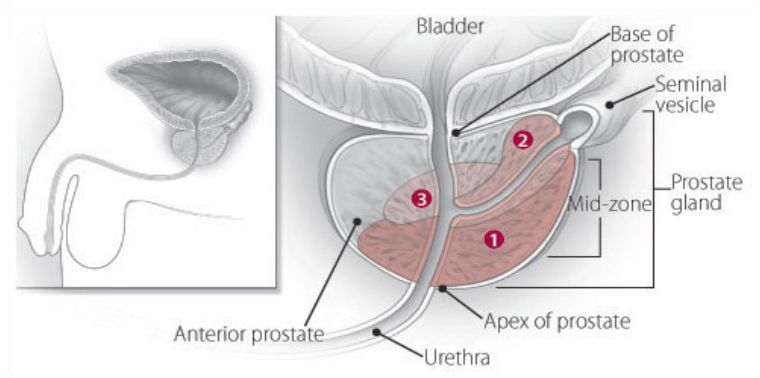


Figure 1.2: Zones of the prostate (Harvard Medical School) (Source: `https://www.health.harvard.edu/topics/prostate-healthprostate-basics`

Most prostate cancers arise in the peripheral zone, the posterior area of the prostate, which is next to the rectum. The peripheral zone is the largest area of the prostate. The doctor can easily feel it during a Digital Rectal Examination (DRE). The transition zone surrounds the part of the urethra that passes through the prostate (called the prostatic urethra). This zone gets bigger as men age, a condition called Benign Prostatic Hyperlasia (BPH). The transition zone is the innermost section of the prostate. The central zone borders the transition zone and surrounds the ejaculatory ducts, which run from the seminal vesicles to the prostatic urethra. Very few prostate cancers start in the central zone (*Canadian Cancer Society* 2021).

1.1.2 Epidemiology

Prostate cancer is now one of the major medical concerns of modern societies. PCa represents about 11,4% of all cancers. In Portugal, the number of new cases per year has been estimated around 6,609, with an approximate mortality of 1,000 patients per year (Nunes et al. 2010). It is also estimated that 1 in 6 men will be diagnosed with prostate cancer during their lifetime, but only 1 in 36 will die from this disease (*Jornal Nordeste - Seminário Regional de Informação* 2019). Several factors contribute to this, of which we highlight the fact that it is the most frequent malignant tumor among men and in this it represents the second most common cause of cancer (*Sociedade Portuguesa de Oncologia* n.d.). Prostate cancer presents

a tendency of increase in the next decades mainly due to the rise of life expectancy, taking to account that the incidence of PCa and mortality rates are strongly related to age, with the highest incidence observed in elderly men ($>$ 65 years) (Rawla 2019).

The incidence rates of this neoplasia vary considerably between countries and, according to the World Health Organization (WHO), there is a tendency for its increase, explained by the interconnection of several facts such as: the evolution of the means of diagnosis, the aging of the population or the effective increase in the number of new cases (Nunes et al. 2010). Furthermore, in recent years, systematic screening and early diagnosis campaigns have played a key role in increasing the number of cases diagnosed in the early stages (Mohler et al. 2010).

1.1.3 Risk Factors

A risk factor is anything that raises one's risk of getting a disease such as cancer. Distinct cancers have different risk factors. Some risk factors, like smoking, can be changed. Others, like a person's age or family history, cannot be changed (*American Cancer Society* c2021).

Doctors do not know exactly why prostate cancer occurs, but the following risk factors may make it more likely:

- **Age:** The risk of prostate cancer increases with age. Most cases occur in men over 65 years of age. After age 65, the risk of having prostate cancer is higher than any other type of cancer. Around 60% of prostate cancers are diagnosed in people who are 65 or older. Older adults who are diagnosed with prostate cancer can face unique challenges, specifically with regard to cancer treatment (*Cancer.Net* September);

- **Race/Ethicity:** Prostate cancer develop more often in black men and other men of African ancestry than in men of other races. And when it does develop in these men, they tend to be younger. The reasons for these racial and ethnic differences are not clear;

- **Geography:** Men living in the USA or Europe have a higher incidence of prostate cancer, with a ratio of 17 cases out of 100 men versus 2 out of 100 in Asian countries.

- **Familial history:** Prostate cancer seems to run in some families, which suggests that in some cases there may be an inherited or genetic factor. Still, most prostate cancers occur in men without a family history of it. Having a father or brother with prostate cancer more than doubles a man's risk of developing this disease. (The risk is higher for men who have a brother with the disease than for those who have a father with it). The risk is much higher for men with several affected relatives, particularly if their relatives were young when the cancer was found;

- **Diet:** A diet high in saturated fat increases the risk of prostate cancer;

- **Genome changes:** Certain genes have been known to elevate prostate cancer risks, such as BRCA1 and BRCA2 genes;

- **Obesity:** Most studies have concluded that obesity does not affect the overall risk of getting prostate cancer. However, obese men may be more likely to develop more aggressive forms of prostate cancer.

1.1.4 Diagnosis

Most prostate cancers are first found as a result of screening with a Prostate Specific Antigen (PSA) blood test, a DRE or imaging exams (transrectal prostatic ultrasound and, more recently, multiparametric prostatic magnetic resonance imaging). The actual diagnosis of prostate cancer can only be made with a prostate biopsy. However, prostate biopsy, an invasive examination with risks for the patient, should only be performed when there are suspicious changes on screening tests.

- **Prostate Specific Antigen (PSA) Test**

  PSA is a protein made by cells in the prostate gland (both normal cells and cancer cells). PSA is mostly in semen, but a small amount is also in the blood.

  The PSA blood test is used mainly to screen for PCa in men without symptoms. It is also one of the first tests done in men who have symptoms that might be caused by PCa.

  PSA in the blood is measured in units called nanograms per milliliter (ng/mL). The chance of having PCa goes up as the PSA level goes up, but there is no set cutoff point that can tell for sure if a man does or does not have prostate cancer. Many doctors use a PSA cutoff point of 4 ng/mL or higher when deciding if a man might need further testing (*American Cancer Society* c2021). The PSA level may also be elevated in other conditions that affect the prostate. As a rule of thumb, the higher the PSA level in the blood, the more likely a prostate problem is present. A low PSA is usually a sign of prostate health. However, many factors, such as age and race, can affect PSA levels. Some prostate glands make more PSA than others. Therefore, if the PSA level is high, the man might need further tests to look for prostate cancer.

  The PSA test can also be helpful if the patient has already been diagnosed with prostate cancer. More precisely:

  - In men just diagnosed with prostate cancer, the PSA level can be use together with physical exam results and tumor grade (determined on the biopsy, described further on) to help decide if other tests (such as CT scans or bone scans) are needed;
  - The PSA level is used to help determine the stage of cancer. This result can affect the treatment options since some treatments (such as surgery and radiation) are not likely to be helpful if cancer has spread to other parts of the body;
  - PSA tests are often essential for determining how well treatment evolves, and diagnosing a possible cancer recurrence after treatment.

- **Digital Rectal Examination (DRE)**

  DRE is when a health care provider inserts a gloved, lubricated finger into a man's rectum to feel the prostate for anything abnormal, such as cancer. Despite being uncomfortable, it is a quick and painless examination. PSA and DRE must be done together, as they can help to diagnose prostate cancer early before it spreads (metastasize).

- **Imaging tests**

  Better cancer-specific blood- and urine-based tests are on the horizon, as well as options for using imaging, such as MRI, to help screen and target the biopsy for prostate cancer. Regardless, they do not replace the PSA test as an important tool in the screening process.

  Prostatic ultrasound allows assessing the presence of changes in the prostatic parenchyma, its volume and the seminal vesicles' morphology. In the case of clinical suspicion of malignant neoplasm of the prostate, ultrasound may suggest areas of greater suspicion, which should be the subject of a directed biopsy. However, prostate ultrasound, supra-pubic or transrectal, should not be used as a means of screening for prostate cancer given its low diagnostic value in small volume neoplasms (*Saúdebemestar.pt* 2021).

  In recent years, the importance of **Multiparametric Magnetic Resonance Imaging (mpMRI)** for the study of the prostate in cases of suspected cancer diagnosis has been highlighted. In reality, the MRI can help the urologist to identify areas of greater or lesser probability of cancer in order to be able to direct the biopsy to those same areas. For this purpose, the images obtained on the resonance are used simultaneously with the transrectal prostatic ultrasound and, thus, the so-called fusion biopsy can be performed.

MRI is also an excellent test in cases where there is already a diagnosis of prostate cancer and if one wants to know the extent of it to the surrounding tissues/organs (it is an exam that helps above all to observe the tumor staging), and that is the main role of this report, which will be discussed later.

- **Prostate biopsy:**

  If the results of a PSA blood test, DRE, or other tests suggest that the patient might have prostate cancer, he will most likely need a **prostate biopsy**.

  A biopsy is a procedure in which small prostate samples are removed and then looked at with a microscope. A core needle biopsy is the primary method used to diagnose prostate cancer. It is usually done by a urologist. During the biopsy, the doctor usually looks at the prostate with an imaging test such as Transrectal Ultrasound (TRUS) or MRI, or a ”fusion” of the two.

  The doctor quickly inserts a thin, hollow needle into the prostate. This is done either through the rectum wall (a transrectal biopsy) or through the skin between the scrotum and anus (a transperineal biopsy). When the needle is pulled out it removes a small cylinder (core) of prostate tissue. This procedure is repeated several times. Most often the doctor will take about 12 core samples from different parts of the prostate.

  The biopsy samples are sent to a lab, where the physicians will be looked at with a microscope to see if they contain cancer cells. The results (in the form of a pathology report) might be reported as:

  - **Positive for cancer:** Cancer cells were seen in the biopsy samples.
  - **Negative for cancer** No cancer cells were seen in the biopsy samples.
  - **Suspicious:** Something abnormal was seen, but it might not be cancer.

  The biopsy is the only way to know for sure if a man has PCa. If prostate cancer is found on a biopsy, this test can also help to tell how likely cancer will grow and spread quickly, quantified by a grading system that reflects cancer characteristics. This is explained next.

1.1.5 Grading

The grade of the cancer is based on how abnormal cancer looks under the microscope. Higher grade cancers look more abnormal and are more likely to grow and spread quickly.

PCa has several types of histological differentiation and, consequently, different behaviors and aggressiveness. The commonly used histological classification is that of Gleason (Li et al. 2015).

Gleason score

The Gleason system, which has been in use for many years, assigns grades based on how much cancer looks like normal prostate tissue (*American Cancer Society* c2021). The cancerous cells fall into five distinct patterns as they change from normal cells to tumor cells. The cells are graded on a scale of 1 to 5. Grade 1 cells resemble normal prostate tissue. Cells closest to 5 are considered ”high-grade” and have mutated so much that they barely resemble normal cells (*Prostate Cancer Foundation* n.d.). Grades 2 through 4 have features in between these extremes.

Almost all cancers are grade 3 or higher; grades 1 and 2 are not often used.

The pathologist looking at the biopsy sample will assign one Gleason grade to the most predominant pattern in the patient’s biopsy and a second Gleason grade to the second most predominant pattern. These two grades are added to yield the Gleason score.

The first number assigned is the grade that is most common in the tumor. For example, if the Gleason score is written as 3+4=7, it means most of the tumor is grade 3 and less is grade 4, and they are added for a Gleason score of 7. Theoretically, Gleason scores range from 2-10. However, since the original classification, pathologists rarely assign scores 2-5, and Gleason scores assigned will range from 6 to 10, with 6 being the lowest grade cancer.

Based on the Gleason score, prostate cancers are often divided into three groups:

- Cancers with a **Gleason score of 6 or less** may be called **well-differentiated** or **low-grade**.
- Cancers with a **Gleason score of 7** may be called **moderately-differentiated** or **intermediate-grade**.
- Cancers with a **Gleason score of 8 to 10** may be called **poorly-differentiated** or **high-grade**.

Nevertheless, because many prostate cancer cases are extremely slow-growing, sometimes the Gleason system does not accurately quantify cancer risks. Patients with scores of 6 and 7 do not have a clear picture of the nature of their particular cancer (*Prostate Cancer Foundation* n.d.). Therefore, in 2014, the International Society of Urological Pathologists released supplementary guidance and a revised prostate cancer grading system, called the Grade Groups.

Grade Groups

The Grade Groups system is simpler than Gleason system, with just five grades, 1 (most likely to grow and spread slowly) through 5 (most likely to grow and spread quickly):

Table 1.1: Grade groups

Grade Groups	Gleason Score
1	≤ 6
2	$7(3+4)$
3	$7(4+3)$
4	8
5	$9-10$

The Grade Groups will likely replace the Gleason score over time, but currently one might see either the Gleason score or the Grade groups on a biopsy pathology report.

Grading cancer is one crucial component of staging cancer. While the grade describes what the actual cancer cells look like under a microscope (how they are behaving on a micro level), the stage of the cancer looks at where the cancer is present in the body (how it is behaving at the macro level) (*Prostate Cancer Foundation* n.d.).

1.1.6 Staging

Following a PCa diagnosis, staging is used to describe the extent of the disease. Prostate cancer staging is vital because it guides the treatment plan and predicts the patient's prognosis (*John Hopkins Medicine* c2021).

The stage of a PCa describes how much cancer is in the body. Cancer's stage is an assessment that takes into account a variety of factors: cancer's location, whether it has spread or metastasized, and how much it is interfering with normal body processes. Staging of PCa includes analysis of pathological anatomy of the operative part, imaging tests or nuclear medicine

A staging system is a standard way for the cancer care team to describe how far cancer has spread. The most widely used staging system for prostate cancer is the **TNM** system, developed by the American Joint Committee on Cancer (AJCC). This system is based on the anatomical extent of the disease, taking into account the size of the **tumor** (**T**), whether the cancer has spread to the lymph **nodes** (**N**), and taking into account the presence or absence of **metastases** (**M**) (*American Cancer Society* 2021).

Each of these classifications has its graduation: T (T0 to T4), N (N0 to N1) and M (M0 to M1), and are analyzed separately. The results are combined to determine the stage of cancer for each person. Stages vary from I to IV, stages that express the level of evolution of the disease:

- In **Stage I**, cancer is small, limited to the prostate and grows slowly;
- In **Stage II**, cancer remains limited to the prostate but is more likely to spread;

- In **Stage III**, cancer grows through the prostatic capsule, affects the tissues surrounding the prostate, seminal vesicles or both, but does not affect other organs;
- In **Stage IV**, cancer spreads to the nearest areas (bladder, rectum, lymph nodes, or more distant organs and even bones).

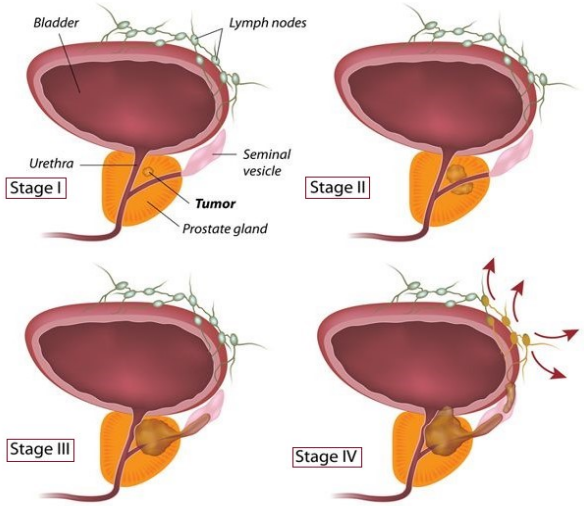


Figure 1.3: Stages of prostate cancer (Source: `https://prostatecancernewstoday.com/2015/03/04/new-target-fight-prostate-cancer-progression-found/`)

The lower the number, the less cancer has spread. Conversely, a higher number, such as stage IV, means cancer has spread more. Moreover, within a stage, an earlier letter means a lower stage. Although each person's cancer experience is unique, cancers with similar stages tend to have a similar outlook and are often treated in much the same way (*American Cancer Society* c2021).

The **T classification** provides information about the tumor's characteristics, such as its size, how deep it developed in the prostate and how much it invaded the adjacent tissues. It varies from 0 to 4 (T0, T1, T2, T3, T4), where each number describes the size of the tumor or the spread of the disease in the vicinity. The higher the number after the T, the larger the tumor size or the more it has grown and spread into nearby tissues. More precisely:

- **T0** means that there is no evidence of a primary tumor.
- T1 and T2 tumors are known as early (localized) prostate cancer.
  - **T1**: the tumor is contained in the prostate and is too small to be detected during a DRE or seen on imaging studies.
  - **T2**: the tumor is still contained in the prostate but can be detected during a DRE.
    - **T2a** - the tumor is only in half of one of the lobes of prostate.
    - **T2b** - the tumor is in more than one half of one lobe
    - **T2c** - the tumor is in both lobes.
- **T3** tumors extend outside the prostate and may be growing into tissues or organs closed by.
  - **T3a** - the tumor extends through the capsule surrounding the prostate.
  - **T3b** - the tumor has spread to the seminal vesicles.
- **T4** tumors have spread into areas close by, such as the bladder or rectum.

There are two types of T categories for prostate cancer: the **clinical T category (cT)** is the doctor's clinical observation of the extent of the disease, based on the results of the physical exam (including DRE) and prostate biopsy; nowadays, there are imaging methods that can help to classify the tumor locally (prostate MRI) - the **imaging T category (iT)**; the **pathologic T category (pT)** is based on surgical finding, and it is done after all prostate examinations carried out in the lab, so, is likely to be more accurate than the cT. The final staging (gold standard) is **pT** (*Prostate Conditions Education Council* n.d.).

The **N category** indicates if the tumor spread to nearby lymph nodes:

- **N0** if there is no cancer in nearby lymph nodes.
- **N1** if cancer has spread to nearby lymph nodes.

The absense or presence of cancer outside the prostate, or Metastasis is measured through **M category**:

- **M0** if cancer has not spread to distant sites.
- **M1** if cancer has spread to distant sites.
  - **M1a** - the tumor has spread to distant lymph nodes.
  - **M1b** - the tumor as spread to bones.
  - **M1c** - the tumor has spread to distant organs.

1.1.7 Treatment

Once diagnosed, the disease must be characterised by the risk of progression. This feature should be linked to the age of the patient and the available therapeutic options considered (*Hospital da Luz* c2021). For many men with PCa, treatment is not immediately necessary, as shown next.

Treatment for localized prostate cancer with low risk of progression:

- **Active surveillance:** the patient is kept under regular clinical surveillance with PSA testing and prostate biopsy, and the onset of curative therapy is delayed while oncological disease control is maintained;
- **Surgery:** radical prostatectomy (prostate removal), which can be performed conventionally (with a vertical incision below the navel), laparoscopically (through 5 small holes) or robotically (also through small holes but with the help of a robotic structure controlled by the surgeon);
- **Radiotherapy:** by placing radioactive seed inside the prostate (brachytherapy) or by extracorporeal administration.

Radiation and the surgical removal of the prostate are also referred to as "curative" treatments because the aim is to remove all of the tumor cells. But a few cancer cells may stay in the body, or new cancer cells might develop. For this reason, men who have had radiotherapy or surgery are still advised to have regular PSA tests (*National Center for Biotechnology Information* 2020).

Advanced prostate cancer occurs when a tumor that develops in the prostate gland spreads outside the prostate. The most common sites of prostate cancer spread are to the lymph nodes and bones. This is also called metastatic prostate cancer. Currently, no treatments can cure advanced/metastatic prostate cancer. However, there are effective ways to help slow its spread, prolong life, and control its symptoms, including immunotherapy, hormone therapy, chemotherapy, precision medicine and clinical trials (*UCLA Health* n.d.).

All treatment options carry the risk of significant side effects, including erectile dysfunction and urinary symptoms, such as needing to use the toilet more urgently or more often. For this reason, some men choose to delay treatment until there is a risk that cancer might spread (*NHS* 2021).

1.1.8 Prognosis

In general, the earlier prostate cancer is caught, the more likely it is for a man to get successful treatment and remain disease-free. The overall prognosis for prostate cancer is among the best of all cancers (*Johns Hopkins Medicine* c2021). However, it is important to keep in mind that survival rates and the likelihood of recurrence are based on averages and will not necessarily reflect any individual patient outcome. The prognosis for prostate cancer depends on many factors as already stressed before.

Approximately 80% to 85% of all prostate cancers are detected in stages I, II and III. Therefore, many men diagnosed and treated in these stages will be disease-free after five years (*John Hopkins Medicine* c2021).

However, prostate cancers detected at stage IV have an average five-year survival rate of 28% much lower than the other stages of prostate cancers. This average survival rate represents stage IV prostate cancers that have metastasized beyond nearby areas to lymph nodes, organs or bones in other parts of the body.

Because most prostate cancers are diagnosed with early screening measures and are curable, the average long-term prognosis for PCa is quite encouraging:

- **5-year relative survival rate of nearly 100% -** five years after diagnosis, the average PCa patient is about as likely as a man without PCa to still be live.
- **10-year relative survival rate of 98% -** ten years after diagnosis, the average PCa patient is just 2% less likely to survive than a man without PCa.
- **15-year relative survival rate of 95% -** fifteen years after diagnosis, the average PCa patient is 5% less likely to survive than a man without PCa.

These numbers, provided by the American Cancer Society, represent a patient's chances of survival after a specified number of years compared with healthy men's chances of survival during that same timeframe (*American Cancer Society* c2021).

1.2 Objectives

Focusing on the stage of the disease, MRI is the most accepted diagnostic method for the staging of the PCa, but is it not universally accepted as the first line method before PCa treatment. The question is whether MRI is an effective method of staging the prostate tumor before undergoing prostatectomy. For that, we will assess which are the MRI biomarkers relevant for predicting the pathological T3 stage.

T3 stage matters most, as it is where the extracapsular extension occurs, and in addition, it defines how the patient will be treated. Patients whose tumor is in this stage have the worst prognosis between stages T2 and T3, and are also more likely to have the operation not being sufficient to treat the disease. Surgery in this type of patient is useless, so it is crucial to evaluate the effectiveness of MRI to predict the stage of the tumor to prevent this type of patient (T3) from undergoing surgery unnecessarily.

Therefore, the main objective of this study is to evaluate the role of MRI in the assessment of ECE in PCa, compared to the histopathological results of the radical prostatectomy, that is, to determine if the imaging stage can be as accurate as of the pathological stage.

We will also find, in terms of probabilities, which are the determinants that contribute to the development of ECE.

1.3 Report Organization

Ending the introduction of this project, in the next chapter it is possible to find a description of the data used. Chapter 3 provides an idea of the statistical methods used for data analysis and then in Chapter 4 are the results of the analysis. Finally, the results are discussed in Chapter 5 and the conclusions in the following chapter.

2

Data

The database was constructed with the support of the Coordinator of the Urogenital Image Unit from Hospital da Luz Lisboa, Adalgisa Guerra, MD, who also helped develop the statistical model and its interpretation.

This study aimed to determine the impact of MRI semantic features classification and clinical findings to predict ECE in patients with prostate cancer.

2.1 Materials and Methods

2.1.1 Study design

A retrospective analysis of MRI was performed, comparing the clinical, prostate biopsy and MRI aspects with pathology findings in a group of patients with PCa treated with Robotic Assisted Radical Prostatectomy (RARP).

The inclusion criteria were patients with PCa (with a Gleason score ≥ 6 on presurgical prostate biopsy) operated by RARP at Hospital da Luz Lisboa between 2015 and 2018. All patients performed a mpMRI in the same MRI device (3T) and with the same protocol before RARP.

The exclusion criteria were patients with contraindication for MRI evaluation and those who did not cooperate to perform the examination at the institution (Hospital da Luz Lisboa), patients in whom the mpMRI date and the surgery were over four months and patients with MRI examination of others institutions. Examinations with imaging artefacts with no quality to perform a correct interpretation of the prostate capsule were excluded from the analysis. Patients who underwent another treatment before surgery as radiotherapy, hormonal therapy or prostate embolization were not participants in this study.

All patients in this study had in common the fact that they had PCa, the same exam (mpMRI) performed in the same device (3T) and with the same protocol, histological diagnosis reviewed by two Hospital da Luz's team members anatomopathologists and operated by RARP by a team of experienced surgeons in this surgery at that institution.

2.1.2 Magnetic Resonance Imaging Technique

All the patients underwent MR imaging using a 3T system (MAGNETOM Verio; Siemens Healthcare, Erlangen, Germany) and a pelvic phased-array coil. All examinations included multiplanar turbo spin-echo T2-weighted imaging, axial turbo spin-echo T1-weighted imaging, and axial Diffusion-weighted (DW) imaging of the prostate, with the latter performed using values of 50 and 1000 sec/mm^2 and 1500

to 2000 sec/mm^2 with inline reconstruction of the apparent diffusion coefficient (ADC) map. Then, Dynamic contrast enhanced (DCE) imaging of the prostate was performed following administration of 0.1 $mmol/kg$ of gadopentetate dimeglumine (Magnevist; Bayer HealthCare Pharmaceuticals, Montville, NJ). The contrast agent was administered as an intravenous bolus via power injector (Spectris; Medrad, Warrendale, Pa), followed by a 20 mL saline flush, both administered at a 3 mL/sec injection rate. From the DCE acquisition, a biexponential semiquantitative model was used to generate two parametric maps representing the mathematically derived maximum slope of enhancement during the contrast enhanced acquisition and the washout of contrast agent following the enhancement peak (hereafter referred to as the maximum slope and washout maps, respectively).

The images were interpreted by an investigator with eight years of experience in urological radiology, using the commercial syngo.via software by Siemens Healthineers with the standardized hanging protocol applied in two monitors. The peer review MRI analysis was done by one radiologist with three years of MRI practice. Researchers were blinded to any information regarding pathology findings to initial mpMRI results.

2.2 Data description

Patients' data were anonymized and collected in an excel database and then ordered according to the surgery date.

All variables were collected by Radiologist in different periods of the examinations and after concluded it, all data were stored in the same database.

The data under consideration were:

- **Patient's age at the date of MRI**;
- **PSA level:** value measured on the date of the MRI (ng/ml);
- **Prostate volume:** (gr);
- **Density:** is the PSA value divided by the volume of the prostate gland;
- **Gleason score:** obtained on prostate biopsy. Categorical variable divided in two groups (Less agressive: $6(3+3)$ and $7(3+4)$, More agressive: $7(4+3)$, $8(4+4)$, $9(4+5)$ and $9(5+4)$;
- **PI-RADS V2:** scores between 1 and 5. ;
- **ECE in prostatectomy specimen:** gold standard (0=False, 1=True);

In addition, it also includes variables that are the result of MRI interpretation and are called biomarkers. They are indicators of the existence or not of ECE.

- **Index lesion size:** major length in axial plane (mm);
- **Capsular contact length:** defined as the length of prostate tumor in contact with the capsule, measured in mm on axial T2 images;
- **Smooth capsular bulging:** rounded or oval projection of the external prostatic margin;
- **Capsular disruption:** disappearance of the normal capsular signal intensity (low signal intensity);
- **Unsharp margin**: disappearance of the normal differentiation between the outer prostate margin (low signal intensity) and the periprostatic tissues (normal/high signal intensity);
- **Irregular contour:** undulation of the prostatic contour;
- **Black striation periprostatic fat:** partial or complete disappearance of the periprostatic fat (high signal intensity);

- **Measurable ECE:** presence of clear periprostatic mass;
- **Rectoprostatic angle obliteration:** partial or complete destruction of the angle between the prostate and rectum.

The objective of the study was to establish the predictive value of the semantic MRI features associated with clinical and prostate biopsy findings for predicting pathological ECE.

Between 2019 and 2020, 61 more patients were added to database, corresponding to the validation group, with similar conditions to the original group except for the MRI examination, performed outside the institution.

The study's binary outcome variable was the extracapsular extension (ECE) of prostate cancer on the pathological prostate specimen. The covariates are the remaining variables presented above, some of which have undergone changes, as is the case of Gleason score, that was obtained by dividing the index lesion into two groups: less aggressive that includes Gleason score $6(3+3)$ and Gleason score $7(3+4)$, and more aggressive that includes Gleason score $7(4+3)$, $8(4+4)$, $9(4+5)$ and $9(5+4)$.

Smooth capsular bulging, Capsular disruption, Unsharp margin, Irregular contour, Black striation periprostatic fat, Measurable ECE and Rectoprostatic angle obliteration are all binary variables (0=False; 1=True).

3

Statistical Methodology

3.1 Logistic regression

In several problems in the medical, biological, industrial, chemical and other fields, it is of great interest to check if two or more variables are related in any way. To express this relationship, it is very important to establish a mathematical model. This type of modeling is called regression. Regression analysis is a set of statistical methods used for the estimation of relationships between one or more independent variables (predictors) and a dependent variable (outcome). It can be used to build models for inference or prediction .

Regression analysis includes several variations, such as linear, multiple linear, and nonlinear. The most common are simple linear and multiple linear regressions. Nonlinear regression analysis is commonly used for more complicated data sets in which the dependent and independent variables show a nonlinear relationship (*Corporate Finance Institute* c2015-2021).

Logistic regression is the appropriate regression analysis to conduct when the response is a categorical variable (*Statistics Solutions* c2021), that is, the variable presents as possible achievements a quality (or attribute) and no longer a measurement.

The binary logistic regression model is one of the most popular of regression models (Glen 2016), it is used to describe data and to explain the relationship between one dependent binary variable and one, or more, numerical independent variables (continuous, discrete) and/or categorical.

Multinomial logistic regression (that is, when the response variable has more than two categories) is not under the scope of this work. Therefore, we will use the terminology ”logistic regression” to represent the binary logistic regression.

Instead of fitting a straight line or hyperplane, the logistic regression model uses the logistic function to fit the output of a linear equation between 0 and 1. A logistic function or logistic curve is a common S-shapped curve (sigmoid curve) defined as:

$$f(x) = \frac{1}{1 - exp(-x)} = \frac{exp(x)}{exp(x) + 1} \tag{3.1}$$

And it looks like Figure 3.1:

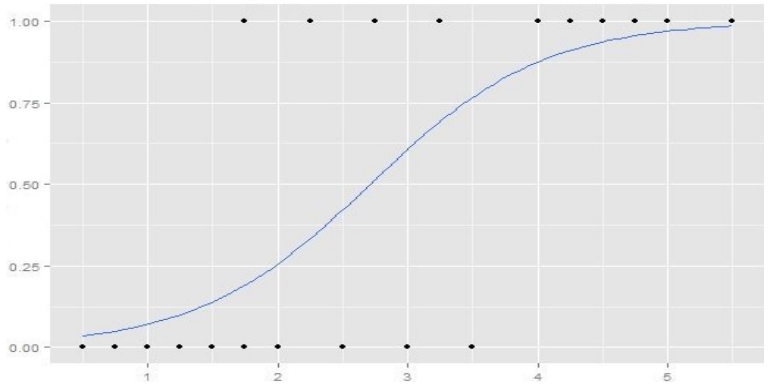


Figure 3.1: Logistic curve

The model for logistic regression analysis assumes that the outcome variable, Y, is categorical (in this work, dichotomous), but this analysis does not model this outcome variable directly. Rather, logistic regression analysis is based on probabilities associated with the values of Y. As stated above, for simplicity, and because it is the case most commonly encourtered in practice, it is assumed that Y is dichotomous, taking on values of 1 (i.e., the positive outcome, or succes) and 0 (i.e., the negative outcome, or failure).

$$Y = \begin{cases} 1 & \text{if the result is a success,} \\ 0 & \text{otherwise.} \end{cases}$$

In theory, the hypothetical, population proportion of cases for which $Y = 1$ is defined as $P(Y = 1) = \pi$. Then, the theoretical proportion of cases for which $Y = 0$ is $P(Y = 0) = 1 - \pi$.

Therefore, for each individual, we have associated a Bernoulli model.

$$Y \frown Bernoulli(1, \pi)$$

3.1.1 Simple logistic regression

This section presents the context in which the response variable has only two categories, that is, binary or dichotomous natures, and only one independent variable involved.

Let us consider a deterministic variable x and the binary random variable Y such that

- $P(Y = 1) = \pi(x)$, probability of succeeding
- $P(Y = 0) = 1 - \pi(x)$, probability of failure

It turns out that $Y \frown Bernoulli\,(\pi(x))$.

In any regression model, the key quantity is the expected value of the dependent variable given the value of the independent variable, $E[Y]$.

When working with data of a binary nature, the conditional average value must be less than or equal to one and greater than or equal to zero, $0 \leq E[Y] \leq 1$.

The mean of $Y \frown Bernoulli\,(\pi(x))$ is

$$E[Y] = \pi(x).$$

To analyze $\pi(x)$ the independent observations $x_1, ..., x_n$ are taken. Here, it is reasonable to assume, as an initial assumption, that $\pi(x)$ is a monotonous function with values between zero and one, when x varies on the real line, that is, $\pi(x)$ is a probability distribution function.

Since $\pi(.)$ varies between zero and one, a simple linear representation for π over all possible values of x is not adequate (since, otherwise, we have the linear regression model), so the logistic transformation is considered $\pi(.)$ in a linear form

$$\log\left(\frac{P(Y=1|X)}{1-P(Y=1|X)}\right)=\log\left(\frac{\pi(x)}{1-\pi(x)}\right)=g(x) \tag{3.2}$$

where,

$$g(x)=\beta_0+\beta_1 x \tag{3.3}$$

or equivalent,

$$\pi(x)=\frac{exp(g(x))}{1+exp(g(x))} \quad \text{or} \quad \pi(x)=\frac{exp(\beta_0+\beta_1 x)}{1+exp(\beta_0+\beta_1 x)} \tag{3.4}$$

where $\beta_1>0$ for π to be increasing, and $\beta_1<0$ for π to be decreasing. When x tends to infinity, $\pi(x)$ tends to zero; when $\beta_1<0$ and tends to one when $\beta_1>0$. Thus, in this way, the necessary link function for the model is defined. If $\beta_1=0$, the response variable Y is independent of the variable X.

If in the equation (3.4), $\beta_0=0$ and $\beta_1=-1$, then $\pi(x)$ has the form of the logistic distribution function of parameters $\mu=0$ and $\sigma=1$, where μ is the expected value and σ the standard deviation, hence it is taken that $\pi(x)=\frac{\exp(-x)}{1+exp(-x)}$. So, the logistic distribution function is given by:

$$F(x)=\frac{exp\left(-\frac{(x-\mu)}{\sigma}\right)}{1+exp\left(-\frac{(x-\mu)}{\sigma}\right)} \tag{3.5}$$

Then, the model parameters, β_0 and β_q, will be estimated using the maximum likelihood method. Let us consider the data of the form (x_i,y_i), $i=1,2,...,n$, where the covariates associated with each individual i are, respectively, $x_1,x_2,...,x_n$. The random variables $Y_1,Y_2,...,Y_n$ are the response variables, with $Y_i \frown Bernoulli(\pi_i)$, where

$$\pi_i\equiv\pi(x_i)=\frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)}, i=1,2,...n.$$

Where Y is a Bernoulli variable, we have that its probability function is given by

$$f(y_i)=\pi_i^{y_i}(1-\pi_i)^{1-y_i}, y_i=0,1, i=1,...,n.$$

The likelihood function is given by

$$L_y(\beta)=\prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i}, \tag{3.6}$$

where $y=(y_1,y_2,...,y_n)\in\{0,1\}^n$ is a particular value of a sample of Y and $\beta=[\beta_0,\beta_1]^T$. The log-likelihood function can be written

$$\begin{aligned} l_y(\beta_0,\beta_1) &= \log\left(\prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i}\right) \\ &= \sum_{i=1}^{n} y_i\log\left(\frac{\pi_i}{1-\pi_i}\right)+\log(1-\pi_i) \\ &= \sum_{i=1}^{n} y_i\exp(\beta_0+\beta_1 x_i)-[\log(1+\exp(\beta_0+\beta_1 x_i))] \end{aligned}$$

The value that maximizes the function $l_y(\beta_0,\beta_1)$ can be obtained by solving the following system of equations: deriving $l_y(\beta_0,\beta_1)$ in order to the model's parameters and equaling to zero we obtain the likelihood equations:

$$\begin{cases} \sum_{i=1}^{n} \left(y_i - \frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)} \right) = 0 \\ \sum_{i=1}^{n} x_i \left(y_i - \frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)} \right) = 0 \end{cases}$$

Since the equations are nonlinear, it is necessary to resort to numerical methods to find the solution, thus obtaining a maximum likelihood estimate for $\beta = [\beta_0, \beta_1]^T$, denoted by $\hat{\beta}$.

After finding the estimators of model's parameters, we should test if each covariate has a significant relationship with the response variable, that is, we want to know if the covariate is relevant for the model. The test used is:

$$H_0 : \beta_j = 0 \quad \text{vs.} \quad H_1 : \beta_j \neq 0$$

This test is called Wald test and is a way to find out if covariates in a model are significant, it means that they add something to the model. Variables that add nothing can be deleted without affecting the model in any meaningful way (*Statistics How To* c2021). So looking at the above hypotheses, if the null hypothesis is not rejected, it suggests that the variable in question can be removed without much harm to the model fit.

The Wald test is based on the asymptotic normal distribution of the maximum likelihood estimators of the model parameters, $\hat{\beta}$.

So, we reject H_0 at the significance level α if the observed value of the test statistic

$$W = \frac{\hat{\beta}_j}{\sqrt{var(\hat{\beta}_j)}},$$

w is such that $|w| > z_{1-\frac{\alpha}{2}}$, where $z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ quantile of distribution $N(0,1)$.

3.1.2 Multiple logistic regression

Now considering the case where we have a set of independent variables expressed by the vector $\mathbf{x}^T = (x_1, x_2, ..., x_p)$.

Analogously to what was presented before, the multiple logistic regression model is given by the expression that defines the probability that the event of interest occurs:

$$\pi(\mathbf{x}) = P(Y=1) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)} = \frac{\exp\left(\beta_0 + \sum_{i=1}^{p} \beta_i x_i\right)}{1 + \exp\left(\beta_0 + \sum_{i=1}^{p} \beta_i x_i\right)},$$

where β_i is the coefficient associated with the covariate x_i.

So, the *logit* of multiple logistic regression is given by

$$logit\ \pi(\mathbf{x}) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p. \tag{3.7}$$

3.1.3 Odds ratio

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occuring in the absence of that exposure.

When a logistic regression is calculated, the regression coefficient, β_1, is the estimated increase in the log odds of the outcome per unit increase in the value of the exposure. In other words, the exponential

function of the regression coefficient, $\exp(\beta_1)$, is the odds ratio associated with a one-unit increase in the exposure.

Odds ratio are used to compare the relative odds of the occurrence of the outcome of interest, given exposure to the variable of interest. The odds can also be used to determine whether a particular exposure is a risk factor for a particular outcome, and to compare the magnitude of various risk factors for that outcome (Szumilas 2010).

- $OR = 1$ is interpreted as indicating that there is no such risk factor, since the odds for the exposed are the same as those for the non-exposed;
- $OR > 1$, is interpreted as indicating that there is a risk factor, since the odds of the event occurring in response to exposure to the factor are greater than in the case of non-exposure;
- $OR < 1$, is interpreted as indicating that the odds of the event occurring in those exposed to treatment are lower than in the case of those not exposed to treatment, and thus we are in the presence of a protective factor.

The confidence interval (CI) is used to estimate the precision of the OR. A large CI indicates a low level of precision of the OR, whereas a small CI indicates a higher precision of the OR. It is important to note however, that unlike the p value, the CI does not report a measure's statistical significance. In practice, the confidence interval is often used as a proxy for the presence of statistical significance if it does not overlap the null value (e.g. $OR = 1$). Nevertheless, it would be inappropriate to interpret an OR with CI that spans the null value as indicating evidence for lack of association between the exposure and outcome.

Assuming that the covariate is binary, the odds ratio is given by the quotient between the odds of the event of interest occurring $(Y = 1)$ in individuals with $x = 1$ and the odds of this event occurring in individuals with $x = 0$. The odds of the event of interest occurring in individuals with $x = 1$ is defined by $\frac{\pi(1)}{1-\pi(1)}$. Analogously, the odds of the event of interest occurring in individuals with $x = 0$ is defined by $\frac{\pi(0)}{1-\pi(0)}$. Thus, the odds ratio is a way of comparing whether the probability of the event of interest occurring is the same for individuals with $x = 1$ or $x = 0$.

The probabilities of the event of interest occurring for the two categories of x, are given respectively by

$$\pi(1) = \frac{exp(\beta_0 + \beta_1)}{1 + exp(\beta_0 + \beta_1))} \quad \text{and} \quad \pi(0) = \frac{exp(\beta_0)}{1 + exp(\beta_0)}$$

Consequently, the odds ratio value is given by the expression

$$OR = \frac{\pi(1)\,[1 - \pi(0)]}{\pi(0)\,[1 - \pi(1)]} = \exp(\beta_1)$$

making evident the relationship between the odds ratio and the coefficient of the model. The OR value depends on the coding adopted for binary covariate x, which can be defined by any two values. Considering the coding using the generic values a and b for representing the odds of the event of interest occurring in individuals with $x = a$, the odds ratio value is given by

$$OR = \frac{\pi(a)\,[1 - \pi(b)]}{\pi(b)\,[1 - \pi(a)]} = \exp\left(\beta_1 (a - b)\right)$$

The interpretation of the odds ratio cannot be made without first knowing the codification of x. Usually, the codification adopted is defined in terms of 0 and 1, as it allows for a trivial interpretation of the parameters. In practice the odds ratio calculation is done from real data usually organized in contingency tables. From these tables we can obtain an estimate of the *odds ratio*, $\hat{OR}$, to which we can apply the logarithm obtaining an estimate of β_1 ($\log(\hat{OR}) = \hat{\beta}_1$). The 95% confidence interval for $\hat{OR}$

is obtained by exponentiating the extremes of the confidence interval of $\hat{\beta}_1$. In case the covariate has more than two categories ($k > 2$), to determine the *odds ratio* value it is necessary to use $k-1$ coding variables, called *dummy* variables. Usually the first category ($i = 1$) is considered as the reference class and takes the value zero for the $k-1$ dummy variables. In the dummy variables associated to the remaining categories ($i = 2, ..., k$) the entry i takes the value 1 and the remaining ones take the value 0. After defining these variables, the calculation of the OR is carried out analogously to the case where two categories are considered.

3.1.3.1 Model Selection and Validation

Studies are primarily designed to answer specific questions, using methodologies that are most likely to answer those questions. In using logistic regression model, one of the most challenging processes is selecting the explanatory variables which should be included in the model or not, even though sometimes the number of variables is very large and make the study tougher to choose among the given covariates. However, the model should be complex enough to fit the data and easier to interpret. So that, a search among models may provide clues about which predictors are most associated with the response variable.

Construction of the logistic model is carried out using maximum likelihood methods. The models based on maximum likelihood methods determine the unknown parameters by maximizing the probability of obtaining the observed data. This is equivalent to calculating the coefficients of the model in a way that best explains the data obtained. The Akaike Information Criterion (AIC) evaluates a model by how close fitted values are to the true expected values. The optimal model is the one its fitted values closest to the true outcome probabilities (Domínguez-Almendros, Benítez-Parejo, and Gonzalez-Ramirez 2011). The log likelihood of the model is the value that is maximized by the process that computes the maximum likelihood value for the coefficients which are β_i parameters.

$$AIC = -2(\log \text{likelihood} - \text{number of parameters in the model})$$

In constructing the model, we must consider the parsimonious model. To this goal we consider all the independent variables that can form part of the model (also considering the possible interactions). For selecting the variables to introduce in the model, it is advisable to previously and separately study the relationship of each factor with the dependent variable. The model should include the variables found to be statistically significant in the multiple analysis, the confounders, and the clinically relevant parameters.

There are several algorithms for the selection of predictors. The most widely used method is the *stepwise regression* (or *stepwise selection*), which consists of iteratively adding and removing predictors, in the predictive model, in order to find the subset of variables in the data set resulting in the best performing model, that is the model that lowers prediction errors (*Statistical tools for high-throughput data analysis (STHDA)* 2018).

There are three strategies of stepwise regression:

1. Forward selection, which starts with no predictors in the model, only the constant (null model), iteratively adds the most contributive predictors, and stops when the improvement is no longer statistically significant.

2. Backward selection (or backward elimination), which starts with all predictors in the model (full model), iteratively removes the least contributive predictors, and stops when you have a model where all predictors are statistically significant.

3. Stepwise selection (or sequential replacement), which is a combination of forward and backward selections. It starts with no predictors, then sequentially add the most contributive predictors (like forward selection). After adding each new variable, remove any variables that no longer provide an improvement in the model fit (like backward selection).

The logarithm of the likelihood ratio of the models is the criterion selected in each step of the construction of the stepwise model in order to determine whether a new model is to be chosen versus the current model. The smaller likelihood value, the better the model, although there is no adequate minimum value. The likelihood function is a measure of the compatibility of the data with the model, thus, if on adding a variable to the model the likelihood does not improve to a statistically significant degree, then this variable should not be included in the equation (Domínguez-Almendros, Benítez-Parejo, and Gonzalez-Ramirez 2011).

3.1.4 Classification Tables

A regression model can be statistically significant and not represent the reality under study. One of the ways to assess the model's classificatory efficiency is through classification tables (also called confusion matrix).

To construct these tables we need to calculate the estimated probabilities of the endpoint occurring and then determine the cut-off, c, for these probabilities. From the cut-off we will assume that individuals with estimated probabilities greater than c experience the endpoint and individuals with probabilities below the cut-off do not. In order to find the most appropriate value, we use graphs, such as the ROC curve, which allow us to identify the value for which the model's sensitivity and specificity are balanced.

The sensitivity of the model is defined as the probability that we predict the occurrence of the endpoint among the individuals in which it was observed. Sensitivity gives the proportion of people with a disease or a condition that have a positive test for this disease or condition. A test that is 100% sensitive means all diseased individuals are correctly identified as diseased i.e. there are no false negatives. Specificity give the proportion of people without disease that have a negative test. A test that is 100% specific means all healthy individuals are correctly identified as healthy, i.e. there are no false positives. So, specificity gives a value which is determined by the probability of predicting that the endpoint will not occur among those individuals in whom it was not observed.

Once the cut-off is determined, a table can be constructed in which the rows present the observed values and the columns the estimated values for the response variable. In a perfect model, all cases would be on the main diagonal. However, in practice it is very difficult to obtain a perfect model and therefore we will have to classify its predictive ability. The estimated model is considered good if sensitivity and specificity are higher than 80%, reasonable if these two values are between 50% and 80% and mediocre if both are lower than 50%.

The classification table or confusion matrix has the following structure (Table 3.1):

Table 3.1: Classification table or confusion matrix

		True value	
		Positive	**Negative**
Predicted value	**Positive**	TP	FP
	Negative	FN	TN

From the Table 3.1), the sensitivity and specificity values can then be calculated, as:

Sensitivity $= \frac{TP}{TP+FN}$

Specificity $= \frac{TN}{FP+TN}$

3.1.5 ROC Curve and AUC

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are powerful tools to measure and compare the performance of binary classification models. The ROC curve is a simple but robust graph that allows one to study the variation in sensitivity and specificity for different

cut-off points in the estimated probability (thresholds). The AUC is an area measure that facilitates the comparison between ROC Curves.

The analysis aims to identify or confirm the quality of the model adjustment. When we look at the graph, we can see that the ideal would be to find an area under the ROC curve close to 1, since the closer the curve is to the upper left-hand corner, the more true positives and the fewer false negatives we will have. For example, if we have an area of 0.5, we can say that the discriminatory power of the model is identical to flipping a coin to determine whether the individual has the endpoint or not. The following criteria are usually used to classify the discriminatory power of a logistic regression model (Hosmer, Lemeshow, and Sturdivant 2013):

- If $ROC = 0.5$ the model does not discriminate between individuals with and without endpoint.
- If $0.6 \leq ROC < 0.7$ the model exhibits limited discrimination.
- If $0.7 \leq ROC < 0.8$ the model exhibits acceptable discrimination.
- If $0.8 \leq ROC < 0.9$ the model exhibits excellent discrimination.
- If $ROC \geq 0.9$ the model exhibits an outstanding discrimination.

A ROC curve is produced by calculating and plotting the true positive rate (TPR) against the false positive rate (FPR) for a single classifier at a variety of threshold (*Towards Data Science* 2020).

The true positive rate, or sensitivity, is the proportion of observations that were correctly predicted to be positive out of all positive observations, can be represented as:

$$TPR = sensitivity = \frac{TP}{TP + FN}$$

where TP is the number of true positives and FN is the number of false negatives. The true positive rate is a measure of the probability that an actual positive instance will be classified as positive.

Similarly, the false positive rate, or $1-$specifity, is the proportion of observations that are incorrectly predicted to be positive out of all negative observations, and can written as:

$$FPR = 1 - specificity = \frac{FP}{FP + TN}$$

where FP is the number of false positives and TN is the number of true negatives. The false positive rate is essencially a measure of how often a "false alarm" will occur, or, how often an actual negative instance will be classified as positive.

For example, in medical testing, the true positive rate is when people are correctly identified to test positive for the disease in question.

A discrete classifier that returns only the predicted class gives a single point on the ROC space. But for probabilistic classifiers, which give a probability or score that reflects the degree to which an instance belongs to one class rather than another, we can create a curve by varying the threshold for the score. Many discrete classifiers can be converted to a scoring classifier by "looking inside" their instance statistics. For example, a decision tree determines the class of a leaf node from the proportion of instances at the node.

Therefore, ROC curve shows the trade-off between sensitivity (or TPR) and specificity (1–FPR). Classifiers that give curves closer to the top-left corner indicate a better performance. As a baseline, a random classifier is expected to give points lying along the diagonal ($FPR = TPR$). The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test is.

Note that the ROC curve does not depend on the class distribution, which makes this tool helpful for evaluating classifiers predicting rare events such as diseases or disasters. In contrast, evaluating performance using accuracy ($\frac{TP+TN}{TP+TN+FN+FP}$) would favor classifiers that consistently predict a negative outcome for rare events (*DisplayR Blog* 2018).

To compare different classifiers, it can be helpful to summarize the performance of each classifier into a single measure. One common approach is to calculate the area under the ROC curve, which is

abbreviated to AUC (*Towards Data Science* 2020). It is equivalent to the probability that a randomly chosen positive instance is ranked higher than a randomly chosen negative instance, i.e. it is equivalent to the two samples Wilcoxon rank-sum statistic .

A classifier with high a AUC can occassionally score worse in a specific region than another classifier with a lower AUC. But in practice, the AUC performs well as a general measure of predictive accuracy (*DisplayR Blog* 2018).

4

Application

This study consists of 236 male patients diagnosed with prostate cancer. The total final patients were 185, where one-quarter had ECE. The patients that were excluded from the analysis: twenty-four with contraindications for MRI evaluation and those who did not agree to perform the examination at institution; nine patients had the mpMRI and the surgery date over four months; fifteen examinations presented artifacts that hampered the prostate capsule evaluation; and four patients underwent another treatment before sugery as radiotherapy, hormonal therapy or prostate embolization.

This analysis was to test the MRI semantic features predictive value to determinate pathological ECE. Statistical analysis were performed using R (version 4.1.2).

4.1 Descriptive analysis

The response variable is a binary variable, it can take only two values. It takes value 0 if the tumor has not come out of the capsule, or value 1 if it has, i.e., if there is ECE.

The study sample was divided into two groups: the group with no ECE and the group with ECE.

In order to more easily observe the analysis of independent variables against the dependent variable were generated graphs, boxplots for continuous variables and bar charts for categorical variables.

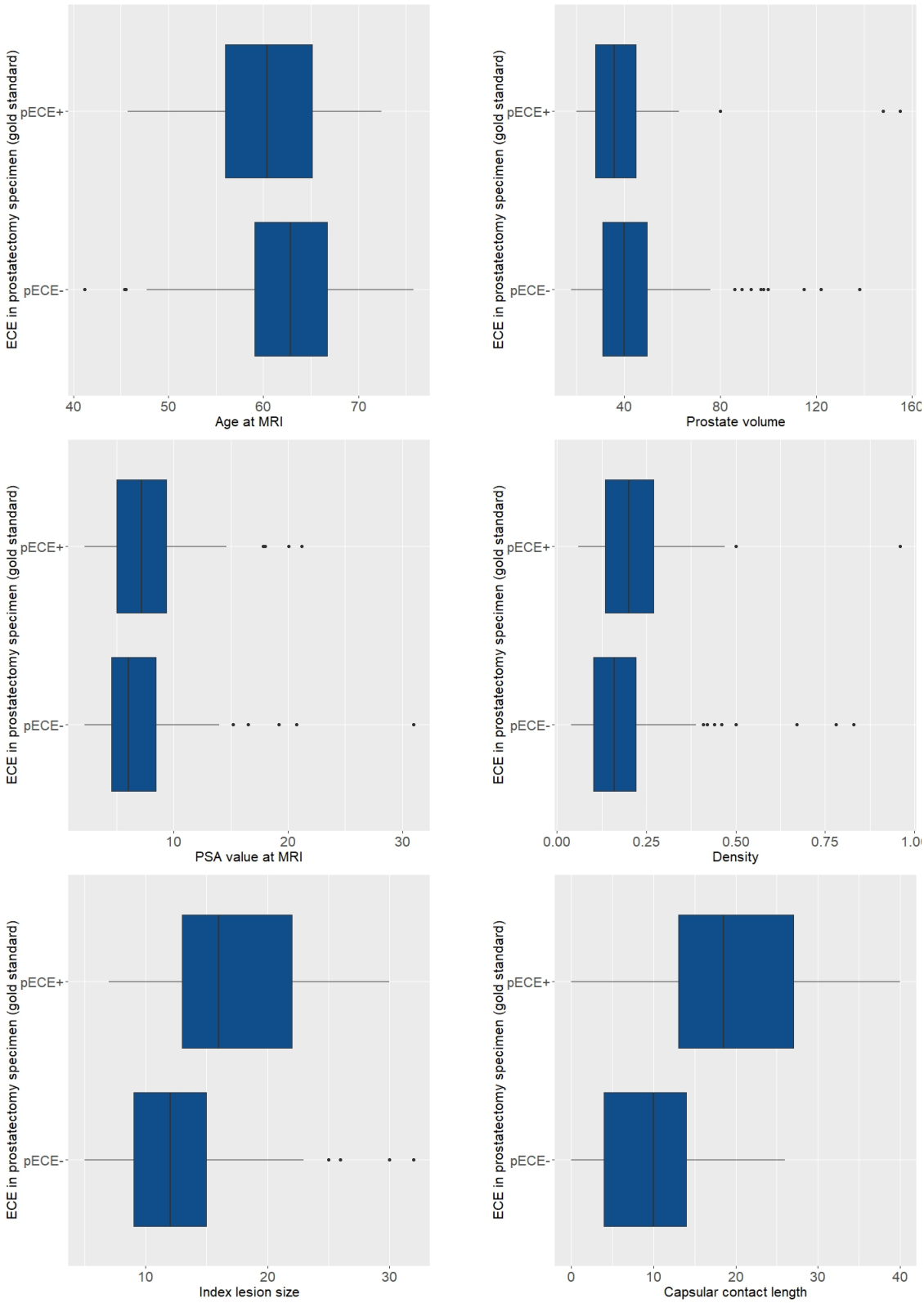
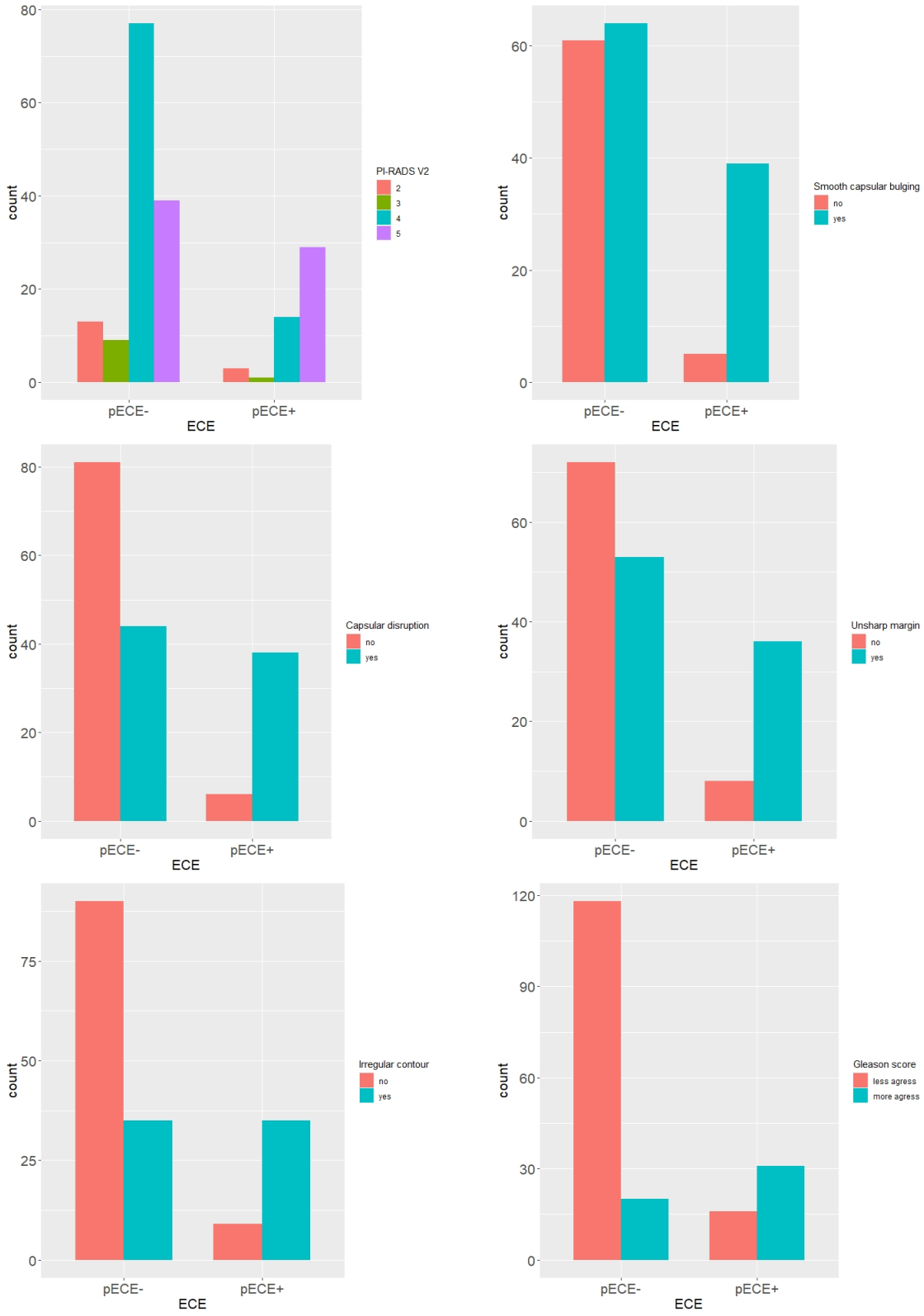


Figure 4.1: Boxplots from continuous variables


count
ECE
pECE-
pECE+
PI-RADS V2
2
3
4
5
Smooth capsular bulging
no
yes
Capsular disruption
no
yes
Unsharp margin
no
yes
Irregular contour
no
yes
Gleason score
less agress
more agress

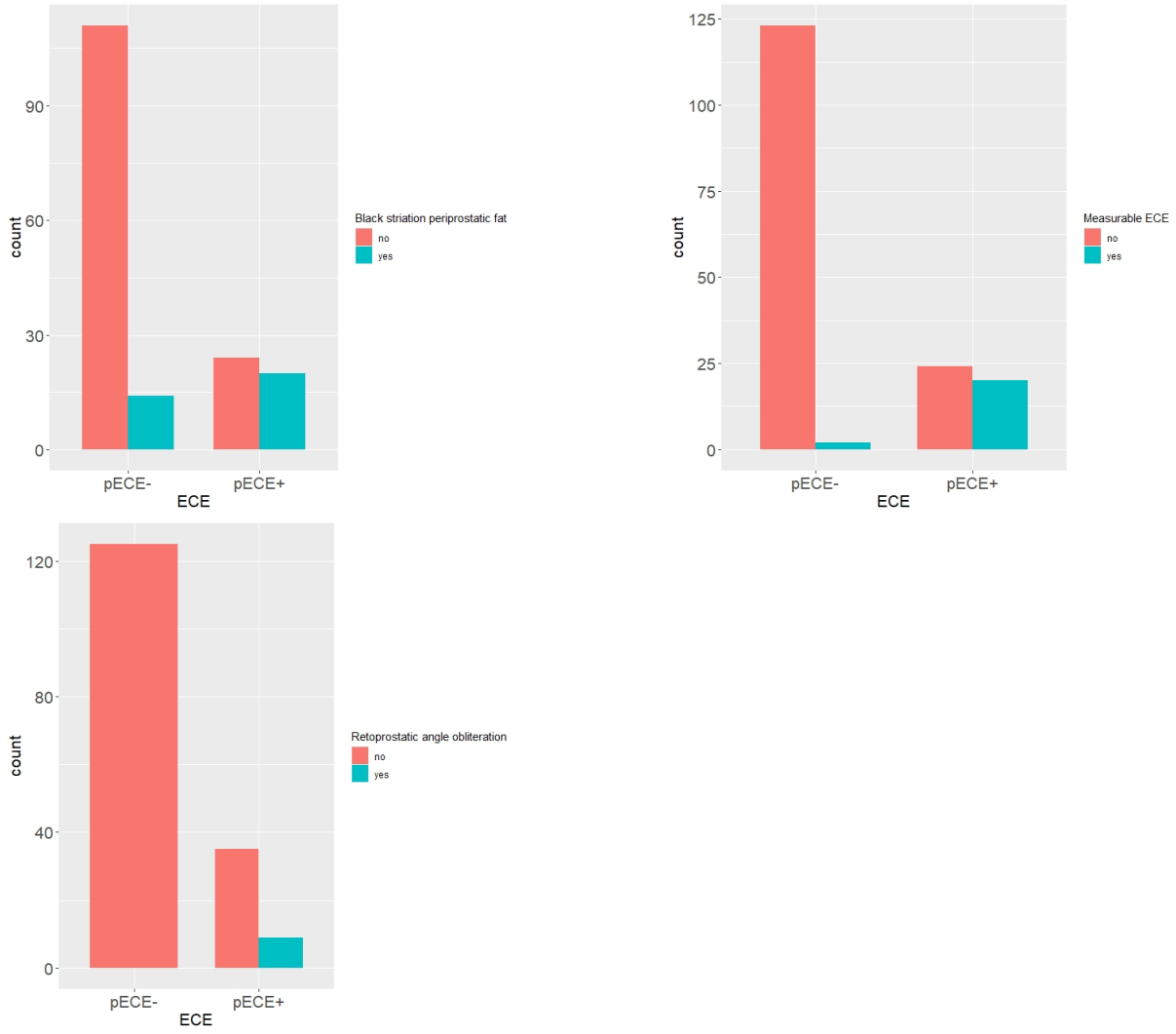


Figure 4.2: Bar charts from categorical variables

In a preliminary analysis, we were able to see that younger patients are more likely to develop ECE than older patients. Men with high values of PSA and Density, despite the small difference, are more likely to have ECE than men with low values of these two variables. Looking at the prostate volume graph, it seems that there is no great difference between the two groups, but that the group with lower values is more likely to develop ECE than the group with higher values. As for the size of the index lesion, the difference between patients whose index lesion size is larger and patients where the index lesion is not that large is evident, with the larger the lesion, the greater the risk of developing ECE. The same is true for capsular contact length, the larger the length of the tumor, the greater the chance of having extracapsular extension.

Looking at the bar charts we can conclude that a more aggressive Gleason score is more favorable to developing ECE than a less aggressive Gleason score, the same works for the variable PI-RADS V2, where scores 2 and 3 are considered less aggressive and scores 4 and 5 more aggressive. It seems that when a smooth capsular bulging, a capsular disruption, an unsharp margin, an irregular contour and a black striation periprostatic fat are present on MRI, the probability of the patient to develop ECE is greater than when these variables are not observed. Having a measurable ECE is almost certain that the patient has an ECE. With regard to the variable rectoprostatic angle obliteration, this also seems to have differences in the two groups.

To corroborate the assumptions taken from the graphical analysis, the Table 4.1 shows the descriptive analysis of each variable for each of the two study groups.

Continuous variables were described as medians, interquartile ranges (IQR) and samples ranges. The

categorical variables were expressed by counts and percentage of patients in each of their levels.

Association between categorical variables and the outcome (presence or absence of ECE) were analyzed by Fisher's exact test (two-tailed tests). Regarding the continuous variables, we used the Welch-Satterwaite unequal variances two-samples T-test to compare the means of the two groups. We performed one-tailed tests, where the alternative hypotheses were that the means of the groups with ECE were greater than the groups without ECE. In all situations, p-values less than or equal to 0.05 were considered statistically significant.

Table 4.1: Characteristics of patients by ECE in prostatectomy specimen (sample size = 185)

Variables	**pECE+** (n.º of patients = 47)	**pECE-** (n.º of patients = 138)	***p-value***
Continuous variables			
Age at MRI (years)	60.6 ± 7.0 (45.7; 72.4)	62.0 ± 6.6 (41.2; 75.8)	0.889
Prostate volume (gr)	41.5 ± 26.6 (20.0; 155.0)	44.3 ± 21.3 (18.0; 138.0)	0.741
PSA (ng/dL)	8.1 ± 4.5 (2.2; 21.2)	7.1 ± 3.9 (2.2; 31.0)	0.091
Density (ng/ml/gr)	0.23 ± 0.15 (0.06; 0.96)	0.18 ± 0.12 (0.04; 0.83)	0.043
Index lesion size (mm)	17.6 ± 6.0 (7.0; 30.0)	12.8 ± 4.9 (5.0; 32.0)	$<< 0.001$
Capsular contact length (mm)	19.9 ± 8.8 (0.0; 40.0)	9.8 ± 6.5 (0.0; 26.0)	$<< 0.001$
Categorical variables			
PI-RADS V2 *			
1	0 (0.00%)	0 (0.00%)	
2	3 (6.38%)	13 (9.42%)	
3	1 (2.13%)	9 (6.52%)	0.236
4	14 (29.79%)	77 (55.80%)	
5	29 (61.70%)	39 (28.26%)	
Smooth capsular bulging			
False	5 (11.36%)	61 (48.80%)	$<< 0.001$
True	39 (88.64%)	64 (51.20%)	
Capsular disruption			
False	6 (13.64%)	81 (64.80%)	$<< 0.001$
True	38 (86.36%)	44 (35.20%)	
Irregular contour			
False	9 (20.45%)	90 (72.00%)	$<< 0.001$
True	35 (79.55%)	35 (28.00%)	
Unsharp margin			
False	8 (18.18%)	72 (57.60%)	$<< 0.001$
True	36 (81.82%)	53 (42.40%)	
Black striation periprostatic fat			
False	24 (54.55%)	111 (88.80%)	< 0.001
True	20 (45.45%)	14 (11.20%)	
Measurable ECE			
False	24 (54.55%)	123 (98.40%)	$<< 0.001$
True	20 (45.45%)	2 (1.60%)	
Retroprostatic angle obliteration			
False	35 (79.55%)	125 (100.00%)	$<< 0.001$
True	9 (20.45%)	0 (0.00%)	
Gleason score			
Less agressive	16 (34.04%)	118 (85.51%)	$<< 0.001$
More agressive	31 (65.96%)	20 (14.49%)	

* Due to the low number of patients in categories ≤ 3 for PI-RADS V2, we combined categories 1, 2 and 3 (Less agressive), and categories 4 and 5 (More agressive), before performing Fisher's exact test.

Each continuous variable is represented as average $\pm$ standard deviation; the minimum and the maximum were also computed. Each categorical variable is described by the number of patients in each level (percentage).

In a first analysis it can be seen that the variables whose *p-value* were less than the usual significance level of 5%, are associated with the variable of interest, that is, they have an influence on the occurrence of ECE.

From the 185 patients analysed, about 25% had presence of ECE (47 patients) and 75% (138 patients) had absence of ECE. The table above (Table 4.1) summarizes the main characteristics of the sample.

The mean age at the time of the surgery was 60.6 years old ($SD = 7.0$) and 62.0 years old ($SD = 6.6$), for patients with pECE+ and pECE-, respectively, which was not statistically significant (p-value$= 0.889$). The mean prostate volume for patients which ECE is present was 41.5 gr (SD$= 26.6$), whereas this value increased to 44.3 gr (SD$= 21.3$) for patients which ECE is not present, where no statistical difference was found between the two groups (p-value$= 0.741$). The mean PSA level was 8.1 ng/dL (SD$= 4.5$) in positive ECE patients and 7.1 ng/dL (SD$= 3.9$) in no positive ECE, at the 5% level of significance, there was no difference between PSA levels (p-value$= 0.091$). Table 4.1 also demonstrates that patients with presence of ECE had higher PSA density (mean of 0.23 vs. 0.18 for patients with absence of ECE, p-value$= 0.043$), larger index lesion size (mean of 17.6 vs. 12.8 for patients with absence of ECE, p-value ≈ 0.000) and larger capsular contact length (mean of 19.9 vs. 9.8 for patients with absence of ECE, p-value ≈ 0.000).

Regarding the morphological criteria, patients with ECE showed smooth capsular bulging (88.64%), capsular disruption (86.36%), unsharp margin (81.82%), irregular contour (79.55%). In addition, almost half of the patients with ECE showed black striation periprostatic fat (45.45%), while only 20.45% of those patients showed rectroprostatic angle obliteration. About forty-five percent of the patients with ECE showed Measurable ECE, whereas almost a hundred (98.40%) of patients without ECE on pathology. Each of these predictors is significantly (p-value < 0.05) associated with extracapsular extension at pathology.

PI-RADS V2 assessment is not significantly associated with positive ECE (p-value$= 0.236$). Moreover, it is important to stress that 91.49% of patients with ECE scored 4 or 5 in PI-RADS V2, against 84.06% for the patients without ECE. Only four patients with a non-aggressive lesion (PI-RADS V2 2 and PI-RADS V2 3) demonstrate ECE on pathology.

Regarding the Gleason score, 65.96% of the patients with ECE had more aggressive Gleason scores, whereas 85.51% of the patients without ECE had less aggressive Gleason scores, representing a statistically significant association between Gleason score and ECE at MRI.

4.2 Logistic regression

The association between the covariates under study and the binary outcome variable was developed using the binary logistic regression model.

Table 4.2: Results from simple logistic regression

Variables	**Coefficient**	**SE**	**OR** (95% **CI**)	***p-value***
Age at MRI	-0.026	0.025	0.975 (0.927; 1.025)	0.317
Prostate volume	-0.011	0.010	0.989 (0.970; 1.009)	0.269
PSA	0.053	0.040	1.055 (0.975; 1.141)	0.186
Density	2.227	1.218	9.273 (0.852; 100.929)	0.068
Index lesion size	0.152	0.034	1.164 (1.089; 1.244)	$<< 0.001$
Capsular contact length	0.179	0.032	1.196 (1.124; 1.273)	$<< 0.001$
PI-RADS V2				
Less agressive	Reference			
More agressive	1.205	1.069	3.336 (0.410; 27.115)	0.260
Smooth capsular bulging				
False	Reference			
True	2.006	0.508	7.434 (2.749; 20.106)	$<< 0.001$
Capsular disruption				
False	Reference			
True	2.456	0.478	11.659 (4.573; 29.727)	$<< 0.001$
Unsharp margin				
False	Reference			
True	1.811	0.431	6.113 (2.628; 14.220)	$<< 0.001$
Irregular contour				
False	Reference			
True	2.303	0.424	10.000 (4.360; 22.935)	$<< 0.001$
Black striation periprostatic fat				
False	Reference			
True	1.888	0.415	6.607 (2.930; 14.898)	$<< 0.001$
Measurable ECE				
False	Reference			
True	3.937	0.774	51.250 (11.235; 233.774)	$<< 0.001$
Rectroprostatic angle obliteration				
False	Reference			
True	18.839	1318.727	151945758.43 (0.000; Inf)	0.989
Gleason score				
Less agressive	Reference			
More agressive	2.481	0.408	11.955 (5.369; 26.620)	$<< 0.001$

Table 4.2 displays the results from the simple logistic regression models. Through these results, it can be seen that, except for Age, Prostate volume, PSA, Density and PI-RADS V2, all the others are statistically significant in explaining the occurrence of ECE. So, the covariates Age at MRI, Prostate volume, PSA level, Density and PI-RADS V2 were not associated with ECE when analyzed individually, because their p-values are higher than the level of 5% significance.

Afterwards, we adjusted a multiple logistic regression model with all covariates under analysis.

First of all, it is necessary to stress a few points. Given that Density is the quotient between PSA and the Prostate volume, it was decided to exclude the variable Density from the regression analysis due to collinearity issues.

Also, the variables Capsular contact length and Index lesion size demonstrated high collinearity. Therefore, we excluded Index lesion size from the regression analysis. It is worth noting that the Capsular

contact length associate with the total size and is more relevant for the ECE due to the proximity to the tissues outside the prostate. The lesions in the central zone could be huge, but if they do not contact capsule, the likelihood to get ECE is low.

The variable Black striation periprostatic fat is highly correlated with Irregular contour, so, it was excluded too from the analysis.

Thus, the model selection was done through the stepwise method, which led to model described in 4.3.

Table 4.3: Final model from multiple logistic regression were the variables were selected by stepwise procedure using AIC

Variables	**Coefficient**	**SE**	**OR** (95% **CI)**	***p-value***
Intercept	-2.701	0.773	0.067 (0.015; 0.305)	$<< 0.001$
Prostate volume	-0.033	0.015	0.968 (0.940; 0.996)	0.027
Capsular contact length	0.093	0.047	1.097 (1.001; 1.203)	0.048
Capsular disruption				
False	Reference			
True	1.112	0.576	3.040 (0.983; 9.402)	0.054
Measurable ECE				
False	Reference			
True	1.652	0.991	5.217 (0.749; 36.357)	0.095
Gleason score				
Less agressive	Reference			
More agressive	1.823	0.518	6.188 (2.242; 17.081)	$<< 0.001$

Based on the multiple logistic regression model (Table 4.3) we found the following significant predictors for ECE at pathology: Prostate volume, Capsular contact length, Capsular disruption, Measurable ECE and Gleason score.

Interpreting in terms of odds ratio, except the Prostate volume, which is a protector factor for ECE, all the others variables in Table 4.3 are risk factors for pathological ECE because the estimated odds ratio are larger than one.

For every increase of one unit in the Capsular contact length, the odds of pathological ECE increased 1.1 times. A patient with Capsular disruption was 3 more likely of having pathological ECE, when compared to a patient without Capsular disruption. Patients with Measurable ECE had 5.2 times higher likelihood of showing pathological ECE, than those with no Measurable ECE. The odds of pathological ECE being present in patients with Gleason scores ≥ 4 was 6.2 times higher than in patients with Gleason scores ≤ 3.

Variance Inflation Factors (VIFs) associated with these covariates are reported in Table 4.4, which do not reveal the existence of multicollinearity.

Table 4.4: Variance Inflation Factors

Variable	**VIF**
Prostate volume	1.267
Capsular contact length	1.663
Capsular disruption	1.155
Measurable ECE	1.455
Gleason score	1.116

Moving now to model validation, one method to evaluate the logistic regression model is the Hosmer-Lemeshow goodness of fit test, that is based on dividing the sample up according to their predicted probabilities, or risks. Specifically, based on the estimated parameter values $\hat{\beta}_0, \hat{\beta}_1, .., \hat{\beta}_p$, for each

observation in the sample the probability that $Y = 1$ is calculated, based on each observation's covariate values:

$$\hat{\pi} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + .. + \hat{\beta}_p X_p)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + .. + \hat{\beta}_p X_p)}$$

Essentially it is a chi-square goodness of fit for grouped data, so the observations in the sample are then split into g groups according to their predicted probabilities. Usually that $g = 10$. Then the first group consists of the observations with the lowest 10% predicted probabilities. The second group consists of the 10% of the sample whose predicted probabilities are next smallest, etc. (*The Stats Geek* 2014).

Since this is a chi-square goodness of fit test, we need to calculate the HL statistic:

$$\sum_{i=0}^{1} \sum_{j=1}^{g} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

where g are the number of groups, and where O_{ij} denotes the number of observed $Y = i$ observations in the gth group and similarly, E_{ij} denotes the expected number of i observations in group j.

The test used is chi-square with g–2 degrees of freedom. A significant test indicates that the model is not a good fit and a non-significant test indicates a good fit (*Real Statistics using Excel* c2021).

In this case, the test statistic was 4.963 with an associated p-value of 0.761. Therefore, the Hosmer and Lemeshow test showed the estimated model fitted quite well the data.

To assess the model's classificatory efficiency was built a classification table, also known as a confusion matrix, for that was calculated a threshold and the value that most appropriated was 0.2043616. After finding this limit, it was possible to build the desired table, which is displayed in Table 4.5.

Table 4.5: Confusion matrix

		True value	
		Positive	**Negative**
Predicted value	**Positive**	$TP = 38$	$FP = 27$
	Negative	$FN = 6$	$TN = 98$

From the Table 4.5 it is possible to calculate the measures to assess the quality of the model, like:

Sensitivity $= \frac{TP}{TP+FN} = 0.8636$

Specificity $= \frac{TN}{FP+TN} = 0.7840$

Positive Predicted Value (PPV) $= \frac{TP}{TP+FP} = 0.5846$

Negative Predicted Value (NPV) $= \frac{TN}{FN+TN} = 0.9423$

Accuracy $= \frac{TP+TN}{TP+FP+TN+FN} = 0.8047$

To evaluate the performance of the multiple logistic regression model in distinguishing patients who presented ECE from those who did not, it was computed the empirical and theoretical ROC curves and AUC (Figure 4.3). Theoretical is based on the assumption that outcome have an equal chance of occurring while empirical is based on the observations of an experiment (Budu 2013).

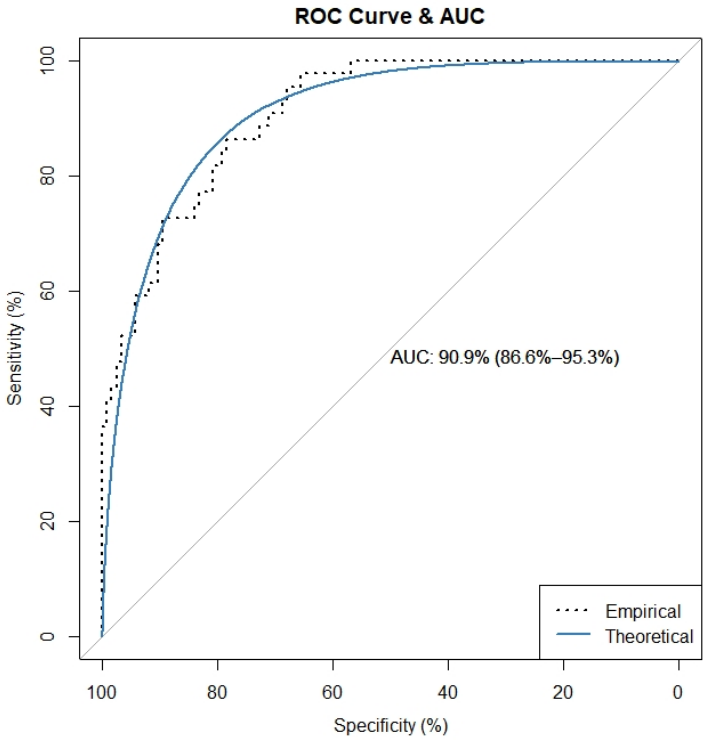


Figure 4.3: Empirical and theoretical ROC curves and AUC from test sample

The Table 4.6 displays the main results:

Table 4.6: Results of final model quality assessment measures

Measure	**Value** (95% **CI**)
Area Under the Curve (AUC)	90.9 (85.4, 94.4)
Sensitivity	86.4 (76.2, 96.5)
Specificity	78.4 (71.2, 85.6)
Positive Predicted Value (PPV)	58.5 (43.9, 73.0)
Negative Predicted Value (NPV)	94.2 (90.1, 98.3)

In conclusion, the model has a good performance in distinguishing patients with and without ECE (AUC $= 0.909$). The estimated model shows high sensitivity (86.4%) and moderate specificity (78.4%).

In order to have a better validation of the model, the final multiple logistic regression model was also evaluated using a validation set.

There is a total of 61 patients in the final model validation sample.

With the purpose of understanding if this sample would be representative of the original sample, a small comparison was made between some variables, applying the T test to compare two samples through the mean (Table 4.7).

Table 4.7: Comparison of some variables of the test sample against the validation sample

Variables	**pECE+**			**pECE-**		
	Test (n=47)	Validation (n=20)	*p-value*	Test (n=138)	Validation (n=41)	*p-value*
Age at MRI	60.6 ± 7.0	64.2 ± 5.8	0.034	62.0 ± 6.6	61.2 ± 6.7	0.495
Prostate volume	41.5 ± 26.6	46.8 ± 26.9	0.466	44.3 ± 21.3	49.4 ± 22.9	0.206
PSA	8.1 ± 4.5	8.3 ± 3.0	0.829	7.1 ± 3.9	6.3 ± 5.7	0.423
Density	0.23 ± 0.15	0.20 ± 0.08	0.335	0.18 ± 0.12	0.14 ± 0.11	0.027
Index lesion size	17.6 ± 6.0	18.6 ± 7.6	0.637	12.8 ± 4.9	11.6 ± 4.9	0.165
Capsular contact length	19.9 ± 8.8	18.6 ± 11.2	0.653	9.8 ± 6.5	8.9 ± 6.9	0.465

From Table 4.7, it is concluded that the validation sample is in fact sufficiently representative of the original sample, because in the T test for comparison of means of two independent samples we obtain practically all p-values greater than 5%, which means that the hypothesis that there are no differences between the two samples is not rejected.

So we can proceed with the computation of the empirical and theoretical ROC curves and AUC, which yielded the following results (Figure 4.32):

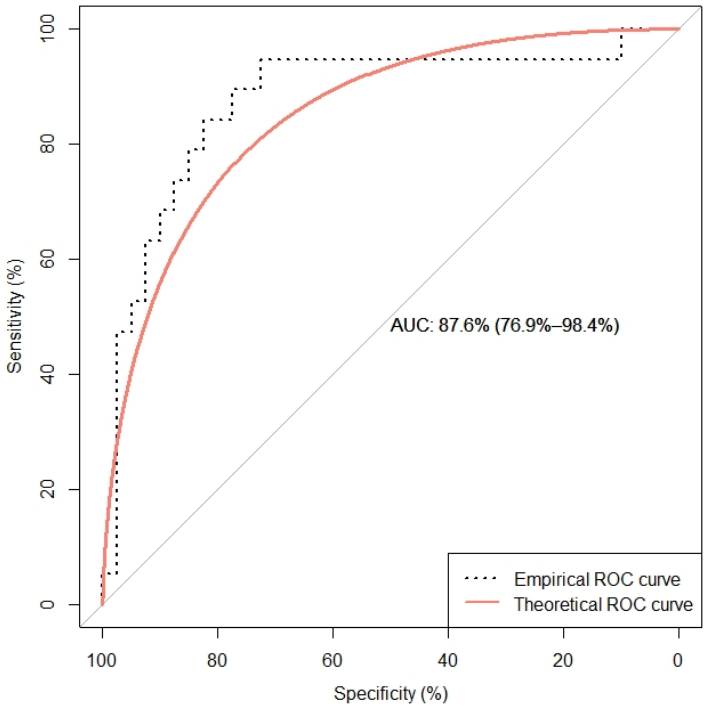


Figure 4.4: Empirical and theoretical ROC curves and AUC from validation sample

From Figure 4.4 the AUC value of this sample was 87.6%, which is very close to the value obtained by the test sample. Therefore, the estimated model has an excellent performance regarding distinguishing patients with and without ECE.

5

Discussion

Several multiparametric magnetic resonance imaging features have been described in the literature as associated with extracapsular extension on pathology. These include curvilinear contact length, capsular irregularity, capsular bulging, obliteration of rectoprostatic angle, asymmetry of the neurovascular bundles, invasion of the periprostatic fat, MRI measurable ECE and seminal vesicle invasion. The original lexicon PI-RADS score, published by the European Society of Urogenital Radiology (ESUR), also proposed a five-point grading scale based on MRI features to predict the presence of ECE in patients with prostate cancer (Barentsz et al. 2012). However, validation studies for these markers are still sparse (Pesapane et al. 2020; Schieda et al. 2015; Weinreb et al. 2016).

Schieda et al. 2015 employed a standardized reporting system, a five-point probability scale for assessment of ECE, using a slight modification of the PI-RADS criteria proposed by the ESUR on the staging of prostate cancer. They concluded that PI-RADS criteria improved the sensitivity of MRI for ECE (24% no standardize report to 50.5% PI-RADS system) without great reducing specificity (75% to 68%) leading to improvement in overall diagnostic accuracy.

In our study, we evaluated the impact of each classical semantic mpMRI feature proposed by ESUR on predicting ECE. MRI semantic features proved to be significant individual predictors of ECE in univariate analysis, and its frequency in patients with ECE are not very different from other studies (Mehralivand et al. 2019; Pesapane et al. 2020).

We confirmed that semantic features considered ”late signs of ECE” as capsular disruption, irregular contour, black striation periprostatic fat, rectoprostatic angle obliteration and measurable ECE are very uncommon in absence of ECE and are usually observed in late stages and aggressive tumors. These features are observed in late stages and represent the presence of more invasive cluster of neoplastic cells in the prostate capsule leading to irregular, disruption of capsule and later on the infiltration to periprostatic fat, which in the end can be seen macroscopically at MRI as measurable ECE (Pesapane et al. 2020).

Based on the multiple logistic regression model (Table 4.3), we found that capsular contact length, capsular disruption, measurable ECE and Gleason score on biopsy represent the most impacting features in ECE prediction. Our data partially overlaps the recent grading system to access the preoperative presence of ECE proposed by (Mehralivand et al. 2019). They elaborated an MRI-derived ECE grading system that included curvilinear contact length, capsular bulging, irregular capsular contour and ECE visible at MRI.

The capsular contact length and index lesion are objective variables associated with the interest variable, ECE. The mean index lesion length and capsular contact length were significantly higher for patients with ECE (17.6 and 19.9 versus 12.8 and 9.8, in positive and negative ECE patients, respectively). This result is in line with other studies that proved that contact capsular length is an independent and reproducible predictor of ECE condition (Kim et al. 2020). The optimal cut-off value predict ECE on pathology has not yet established among the various studies and varies between 10 and 20 mm, being more common between 11 and 13 mm (Kim et al. 2020). In our opinion, the global assessment of contact capsular length and its itegration with the other model's covariables is the key for obtaining the individual probability of ECE.

The measurable ECE on MRI was seen less than half of positive ECE cases (45.5%). However, almost all patients with measurable ECE were positive ECE, except only two patients (1.6%). Reviewing these two false positive cases we concluded that in one of them there was hemorrhage hampering the interpretation of the images and in the other there was granulomatous prostatitis coexisting with prostate cancer.

Thus, when measurable ECE is the only semantic feature that allow us to confirm the presence of ECE. However, measurable ECE is a relatively late marker of ECE and become primarily visible in advanced stages of prostate cancer. Therefore, it should be kept in mind that its absence does not rule out extraprostatic extension. In institution most advanced, prostate cancer patients were not proposed to radical prostatectomy and the majoraty of our cases operated had minimal (< 5 mm) periprostatic extension in pathology. This explains why measurable ECE is observed only less than 50% of all positive ECE cases.

As stated in literature, our model confirms that the combination of MRI features with Gleason score on pre-treatment biopsy is superior than imaging features alone to predict ECE (Patel and Gnanapragasam 2016).

The Gleason score is the variable that defines the aggressivity of the prostate cancer. Therefore, it improves the diagnostic accuracy of the disease's aggressivity and it has been incorporated in many predictive nomograms to detect ECE (Shieh et al. 2020).

Our estimated model shows high sensitivity (86.4%) and moderade specificity (78.4%), which is in accordance with the recent metanalysis of Kim et al. 2020, but slightly differs from the metanalysis of Rooij et al. 2016. The explanation is related to the use of objective parameters, like contact capsular length, in addition to subjective semantic features to predict ECE. This feature may have important clinical implications as a high sensitivity or high specificity interpretation may be preferred, depending on the clinical scenario. For example, high sensitivity is required when selecting patients to enroll in active surveillance programs or choosing candidates for radical prostatectomy with neurovascular bundle sparing. On the other hand, high specificity could be favored when the objective is to avoid potential curative treatment delay.

The clinical covariates age, prostate volume and index lesion PI-RADS score were not associated with ECE (p-values > 0.05) as also referred to in previous studies (Mehralivand et al. 2019).

About PI-RADS score, most of positive ECE lesions were PI-RADS 5 (62%) and negative ECE were PI-RADS 4 (56%). Only three peripheral lesions PI-RADS 2, not seen in diffusion-weighted images, were ECE is present. These three lesions were reviewed by another senior Radiologist, who agreed with the score PI-RADS 2, and by the Pathologist. Both were very small subcapsular lesions (6 mm) with an aggressive Gleason score $9(5+4)$. Although PI-RADS is not statistically significant in differentiating patients with or without ECE, we concluded that in the absence of lesions PI-RADS 4 or 5, the probability of having ECE is very low (2%).

The covariate PSA alone was not significant. However, patients with ECE showed increased PSA density (mean of 0.23 vs. 0.18, in patients withou ECE, p-value$= 0.043$), meaning that PSA is only

significant when associated with prostate volume. Patients with smaller prostates and higher PSA are more likely to have a more aggressive disease than patients with the same PSA but larger prostate volume. This is an expected result as PSA density is the ratio between PSA and prostate volume. Therefore, PSA alone and PSA density could never be together in the same regression model due to collinearity. Preliminary analyses revealed that the inclusion of PSA density, along with the other covariates under study, led to weaker multiple estimation results than those obtained by adding PSA. Therefore, the variable PSA density was discarded from the modelling process.

Our model presented a good agreement between readers for the presence of contact capsular length and measurable ECE (more than 92%), but it was not good enough for capsular disruption. The other imaging covariables are very heterogeneous between readers and were not significant in the multiple model.

This model accurately distinguishes patients with ECE from the others. The AUC reflects this feature: 87.6% for the validation group, very similar to the 90.9% for the original group.

Our study has some limitations. First, the patient sample is relatively small, and it was performed in a single center, using the same MRI protocol and a 3T MRI scanner. To overcome this bias, the authors introduced 61 patients whose MRI examination was performed outside, in a different MRI equipment, in order to increase the model's robustness. Moreover, a sampling bias results from selecting the prostatectomy specimen as the histopathological reference standard, since the more advanced cases were not proposed to prostatectomy.

Second, we focused only on the index lesion identified on MRI, and it was only prostate lesion correlated with pathology. Therefore, we did not take into account the tumor multifocality on MRI and pathology.

Third, our model uses measurable ECE as a determinant MRI semantic feature to detect ECE, consensual between readers, however with a lower prevalence in patients with presence of ECE, because it is seen mainly in more advanced cases. The other more prevalent semantic feautures were not significant in detecting early cases of ECE and were not so consensual between readers. Thus, more work should be done to detect microscopic ECE pre-operatively, maybe using objective markers, as contact capsular length, or new artificially intelligence analysis.

In conclusion, our model using objective features on MRI as contact capsular length, a semantic interpretative consensual feature, and a Gleason score of pre-treatment prostate biopsy is accurate to detect patients with extracapsular extension before the treatment.

6

Conclusion

This master's project aimed to understand the role of magnetic resonance in the staging of prostate tumors before prostatectomy, aiming above all to determine which are the predictors of extracapsular involvement in MRI (stage T3), because this is the stage that defines the outcome of the patient.

To this end, this analysis was based on a set of variables, including: clinical variables, variables that derived from the doctor's clinical observation of the extent of the disease, based on the results of the physical examination (including digital rectal examination) and biopsy of the prostate; image variables, resulting from the MRI interpretation; and pathological variables, variables observed after the prostate removal surgery.

Firstly, data was collected and then the database was organized. Then, a preliminary analysis was carried out, which involved descriptive and graphical analysis.

Finally, most importantly, a logistic regression model was applied to predict the development of tumor ECE based on observed characteristics of the patients. It was concluded that a model with prostate volume, capsular contact length, capsular disruption, measurable ECE and Gleason score is accurate to detect patients with ECE before the treatment. This model has a good performance in distinguishing patients with ECE from the others and the AUC reflects this feature (90.9%). In order to test the validity of the model, an evaluation was carried out using an external sample, which showed us an area under the curve of 87.6%, which was very close to the value obtained by the principal sample. Therefore, it is proven that the use of objective MRI features, such as capsular contact length, a consensus interpretive semantic feature, and a pre-treatment prostate biopsy Gleason score, is accurate to detect patients with extracapsular extension before treatment.

This is a study that can be used for future investigations. There is already another study derived from this one which is already underway, the main goal of which being to analyze the influence of all the variables already mentioned on the evaluation of the biochemical relapse in patients with a prostate tumor operated by radical prostatectomy.

Bibliography

American Cancer Society (c2021). URL: https://www.cancer.org/cancer/prostate-cancer.html.

American Cancer Society (c2021). URL: https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/survival-rates.html.

American Cancer Society (Oct. 2021). URL: https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/staging.html.

American Society of Clinical Oncology (ASCO) (c2005-2021). URL: https://www.cancer.net/cancer-types/prostate-cancer/introduction.

Barentsz, Jelle O et al. (Apr. 2012). "ESUR prostate MR guidelines 2012". In: *European Radiology* 22.4, pp. 746–757. DOI: https://doi.org/10.1007/s00330-011-2377-y.

Budu, Emmanuella (June 2013). *Medium*. URL: https://medium.com/@emmabudu/the-difference-between-empirical-and-theoretical-probability-d42938aa8b7.

Canadian Cancer Society (Feb. 2021). URL: https://www.cancer.ca/en/cancer-information/cancer-type/prostate/prostate-cancer/the-prostate/?region=on.

Cancer.Net (Apr. September). URL: https://www.cancer.net/cancer-types/prostate-cancer/risk-factors-and-prevention/.

Corporate Finance Institute (c2015-2021). URL: https://corporatefinanceinstitute.com/resources/knowledge/finance/regression-analysis/.

Dirix, Piet, Karin Haustermans, and Vincent Vandecaveye (July 2014). "The Value of Magnetic Resonance Imaging for Radiotherapy Planning". In: *Seminars in Radiation Oncology* 24, pp. 151–159. DOI: https://doi.org/10.1016/j.semradonc.2014.02.003.

DisplayR Blog (July 2018). URL: https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/.

Domínguez-Almendros, S., N. Benítez-Parejo, and A.R. Gonzalez-Ramirez (May 2011). "Allergologia et immunopathologia: Logistic Regression Models". In: *Elsevier Doyma* 39(5), pp. 295–305.

Glen, Stephanie (Sept. 2016). "Wald Test: Definition, Examples, Running the Test". In: URL: https://towardsdatascience.com/what-is-logistic-regression-d3f59ef1a9b.

Harvard Health Publishing - Harvard Medical School (c2021). URL: https://www.health.harvard.edu/blog/prostate-gland.

Hosmer, David W., Stanley Lemeshow, and Rodney X. Sturdivant (2013). *Applied Logistic Regression*. 3rd ed. Wiley, pp. 173–177.

Hospital da Luz (c2021). URL: https://www.hospitaldaluz.pt/en/health-guide/health-dictionary/C/70/prostate-cancer-causes-treatment.

John Hopkins Medicine (c2021). URL: https://www.hopkinsmedicine.org/health/conditions-and-diseases/prostate-cancer/prostate-cancer-stages.

Johns Hopkins Medicine (c2021). URL: https://www.hopkinsmedicine.org/health/conditions-and-diseases/prostate-cancer/prostate-cancer-prognosis.

Jornal Nordeste - Seminário Regional de Informação (2019). URL: https://www.jornalnordeste.com/olho-clinico/cancro-da-prostata-sinais-de-alerta-prevencao-e-diagnostico.

Kim, Tae Hyung et al. (June 2020). "The Diagnostic Performance of the Length of Tumor Capsular Contact on MRI for Detecting Prostate Cancer Extraprostatic Extension: A Systematic Review and

Meta-Analysis". In: *Korean Journal of Radiology* 21.6, pp. 684–694. DOI: https://doi.org/10.3348/kjr.2019.0842.

Li, Long-Cheng et al. (2015). "Molecular Basis of Prostate Cancer". In: *Elsevier*, pp. 549–560. DOI: https://doi.org/10.1016/B978-1-4557-4066-6.00038-X.

Liga Portuguesa Contra o Cancro (c2015). URL: https://www.ligacontracancro.pt/cancro-da-prostata/.

McPartlin, A.J. et al. (June 2016). "MR-guided prostate adaptive radiotherapy - A systematic review". In: *Radiotherapy Oncology* 119, pp. 371–380. DOI: https://doi.org/10.1016/j.radonc.2016.04.014.

Mehralivand, Sherif et al. (Mar. 2019). "A Grading System for the Assessment of Risk of Extraprostatic Extension of Prostate Cancer at Multiparametric MRI". In: *Radiology* 290.3, pp. 709–719. DOI: https://doi.org/10.1148/radiol.2018181278.

Mohler, James et al. (Feb. 2010). "NCCN clinical practice guidelines in oncology: prostate cancer". In: *J Natl Compr Canc Netw.* 8, pp. 162–200. DOI: 10.6004/jnccn.2010.0012.

National Center for Biotechnology Information (Mar. 2020). URL: https://www.ncbi.nlm.nih.gov/books/NBK487255/.

NHS (Oct. 2021). URL: https://www.nhs.uk/conditions/prostate-cancer/.

Nunes, Pedro et al. (2010). "Registo Nacional de Cancro da Próstata, em Portugal - ReNaCaP". In: *Acta Urologica* 3, pp. 39–45.

Olsson, Lars E. et al. (Jan. 2019). "Basic concepts and applications of functional magnetic resonance imaging for radiotherapy of prostate cancer". In: *Physics and Imaging in Radiation Oncology* 9, pp. 50–57. DOI: https://doi.org/10.1016/j.phro.2019.02.001.

Patel, Keval M and Vincent J Gnanapragasam (Dec. 2016). "Novel concepts for risk stratification in prostate cancer". In: *Journal of Clinical Urology* 9.2 Suppl, pp. 18–23. DOI: https://doi.org/10.1177/2051415816673502.

Persson, Emilia et al. (Nov. 2017). "MR-OPERA: A Multicenter/Multivendor Validation of Magnetic Resonance Imaging-Only Prostate Treatment Planning Using Synthetic Computed Tomography Images". In: *International Journal of Radiation Oncology*Biology*Physics* 99, pp. 692–700. DOI: https://doi.org/10.1016/j.ijrobp.2017.06.006.

Pesapane, Filippo et al. (Jan. 2020). "T-staging of prostate cancer: Identification of useful signs to standardize detection of posterolateral extraprostatic extension on prostate MRI". In: 59.1, pp. 1–7. DOI: https://doi.org/10.1016/j.clinimag.2019.08.007.

Prostate Cancer Foundation (n.d.). URL: https://www.pcf.org/about-prostate-cancer/diagnosis-staging-prostate-cancer/gleason-score-isup-grade/.

Prostate Conditions Education Council (n.d.). URL: https://www.prostateconditions.org/about-prostate-conditions/prostate-cancer/newly-diagnosed/staging.

Rawla, Prashanth (Apr. 2019). "Epidemiology of Prostate Cancer". In: *World Journal of Oncology* 10, pp. 63–89. DOI: 10.14740/wjon1191.

Real Statistics using Excel (c2021). URL: https://www.real-statistics.com/logistic-regression/hosmer-lemeshow-test/.

Rooij, Maarten de et al. (Aug. 2016). "Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis". In: *European Urology* 70.2, pp. 233–245. DOI: https://doi.org/10.1016/j.eururo.2015.07.029.

Salembier, Carl et al. (Apr. 2018). "ESTRO ACROP concensus guideline on CT- and MRI-based target volume delineation for primary radiation therapy of localized prostate cancer". In: *Radiotherapy and Oncology* 127, pp. 49–61. DOI: https://doi.org/10.1016/j.radonc.2018.01.014.

Saúdebemestar.pt (Apr. 2021). URL: https://www.saudebemestar.pt/pt/clinica/urologia/cancro-da-prostata/.

Schieda, Nicola et al. (Oct. 2015). "Evaluation of the European Society of Urogenital Radiology (ESUR) PI-RADS scoring system for assessment of extra-prostatic extension in prostatic carcinoma". In: 84.10, pp. 1843–1848. DOI: https://doi.org/10.1016/j.ejrad.2015.06.016.

Shieh, Alice C et al. (Dec. 2020). “Extraprostatic extension in prostate cancer: primer for radiologists”. In: *Abdominal Radiology (New York)* 45.12, pp. 4040–4051. DOI: https://doi.org/10.1007/s00261-020-02555-x.

Sociedade Portuguesa de Oncologia (n.d.).

Statistical tools for high-throughput data analysis (STHDA) (Mar. 2018). URL: http://www.sthda.com/english/articles/37-model-selection-essentials-in-r/154-stepwise-regression-essentials-in-r/.

Statistics How To (c2021). URL: https://www.statisticshowto.com/probability-and-statistics/regression-analysis/.

Statistics Solutions (c2021). URL: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/what-is-logistic-regression/.

Szumilas, Magdalena (Aug. 2010). “Explaning Odds Ratios”. In: *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 19(3), pp. 227–229.

The Stats Geek (Feb. 2014). URL: https://thestatsgeek.com/2014/02/16/the-hosmer-lemeshow-goodness-of-fit-test-for-logistic-regression/.

Towards Data Science (Sept. 2020). URL: https://towardsdatascience.com/understanding-the-roc-curve-and-auc-dd4f9a192ecb.

Tyagi, Neelam et al. (July 2017). “Clinical workflow for MR-only simulation and planning in prostate”. In: *Radiat Oncol* 12, pp. 1–12. DOI: https://doi.org/10.1186/s13014-017-0854-4.

UCLA Health (n.d.). URL: https://www.uclahealth.org/urology/prostate-cancer/advanced-prostate-cancer.

UCLA Health - UCLA Prostate Cancer Program (n.d.). URL: https://www.uclahealth.org/urology/prostate-cancer/what-is-prostate-cancer.

Weinerb, Jeffrey C. et al. (Jan. 2016). “PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2”. In: *European Urology* 69, pp. 16–40. DOI: https://doi.org/10.1016/j.eururo.2015.08.052.

Weinreb, Jeffrey C et al. (Jan. 2016). “PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2”. In: *European Radiology* 69.1, pp. 16–40. DOI: https://doi.org/10.1016/j.eururo.2015.08.052.