Predictive models for side effects following radiotherapy for prostate cancer

Juan David Ospina Arango

Juan David Ospina Arango. Predictive models for side effects following radiotherapy for prostate cancer. Signal and Image processing. Université de Rennes; Universidad nacional de Colombia, 2014. English. NNT : 2014REN1S046 . tel-01082134

HAL Id: tel-01082134
https://theses.hal.science/tel-01082134

Submitted on 12 Nov 2014

ANNÉE 2014

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

En Cotutelle Internationale avec
Universidad Nacional de Colombia, Colombie
pour le grade de
DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Mention : Traitement du Signal et Télécommunications
École doctorale Matisse

présentée par

Juan David OSPINA ARANGO

préparée à l'unité de recherche LTSI – INSERM U1099
Laboratoire Traitement du Signal et de l'Image
ISTIC UFR Informatique et Électronique

Predictive models for side effects following radiotherapy for prostate cancer

Thèse soutenue à Rennes
le 16 juin 2014

devant le jury composé de :

Gilles CRÉHANGE
PU/PH, Centre Georges François Leclerc, Dijon / *Rapporteur*

Jean-Yves TOURNERET
PU, INP-ENSEEIHT, Toulouse / *Rapporteur*

Jean-Léon LAGRANGE
PU/PH, AP-AH Hôpital Henri Mordor, Paris / *Examinateur*

Oscar ACOSTA
MCU, Université de Rennes 1, Rennes / *Examinateur*

François KAUFFMANN
MCU, Université de Caen Basse-Normandie / *Examinateur*

Raúl Alberto PÉREZ AGÁMEZ
Professeur, Universidad Nacional de Colombia, Medellin, Colombie / *Examinateur*

Renaud DE CREVOISIER
PU/PH, Centre Eugène Marquis, Université de Rennes 1, Rennes / *Directeur de thèse*

Juan Carlos CORREA MORALES
Professeur, Universidad Nacional de Colombia, Medellin, Colombie / *Co-directeur de thèse*

Acknowledgements

This thesis was made possible by a co-supervision agreement between the *Universidad Nacional de Colombia - Sede Medellín* and the *Université de Rennes 1*. I would like to acknowledge the Colombian Administrative Department for Science and Innovation - COLCIENCIAS for financially supporting this thesis under the grant scheme "*Doctorados Nacionales 2009*". This thesis was also founded by IReSP (France) under the Plan Cancer 2009-2013 Call (No. C13005N5) and the European University of Brittany (UEB, France). Without all these generous efforts it would not have been possible to develop this work.

I would like to thank Prof. Créhange, Prof. Tourneret, Prof. Lagrange, Prof. Kauffmann and Prof. Pérez for accepting to be part of my jury and for their kind and pertinent remarks that strongly contributed to improving the quality of this manuscript.

Thanks to my supervisors, Prof. Renaud de Crevoisier, Prof. Juan Carlos Correa Morales and Prof. Oscar Acosta for guiding me throughout the thesis process and for providing then necessary theoretical support to complete this thesis. What I have learned from them goes beyond the limits of academia.

The *Universidad Nacional de Colombia* helped me to build up the basis of my work upon the principles of "work and rectitude" and trained me to seek excellence in my work. In particular, the *Escuela de Estadística* provided me with a background in statistics, which enabled me to carry out this thesis. To Profs. Norman Diego Giraldo, Juan Carlos Salazar, Víctor Ignacio López, Elkin Castaño, Jesús Antonio Hernández and Aníbal Córdoba from the *Universidad Nacional de Colombia*, I express my utmost gratitude. They introduced me to the world of academia and helped me to to explore my interest in science.

I feel honored for having conducted the most part of my research work at the *Laboratoire Traitement du Signal et de l'Image* LTSI of the *Université de Rennes 1*. I met very gently people there who, with a lot of patient on their part, taught me French. In particular Muriel Diop, Soizic Charpentier and Patricia Gerouard made my days in the LTSI very pleasant and happy ones. I will always remember their kindness and welcome.

I am in debt to my colleagues Jian Zhu, Gaël Dréan, Auréline Fargeas, Geoffray Roman-Jimenez, Guillaume Cazoullat, Frédéric Commandeur, Julián Betancur and Zulma Sandoval for their unconditional support and help. They have also contributed to making this period of my life a very happy one.

My special thanks to my Irish friend Marian Lee who first showed me the French life and has have been lighting up my world for the last three years - and who proofread most of this document!

I want to thank my whole family and my Colombian friends. It is for them that I decided to undertake a PhD thesis in the first place, to build together a better country for the next generation, together.

Finally, I want to dedicate this work to my grandmother Amparo Medina.

Contents

List of Figures

List of Tables

Acronyms

AUC area under the curve 92–94, 98, 102–104, 139

CBCT cone beam computed tomography 25, 64

CT computed tomography 24, 52, 53, 64, 65, 115, 116, 129, 130

CTCAE common terminology criteria for adverse events 28, 67, 115, 129

DSH dose-surface histogram 111, 112

DVH dose-volume histogram 30, 32–37, 39, 51, 79, 89–94, 98, 102–104, 111, 112, 118, 128

EBRT external beam radiotherapy 20, 21

EUD equivalent uniform dose 34, 35

GLM general linear model 122

HOSVD high order singular value decomposition 113, 114, 116

IGRT image-guide radiotherapy 24, 25, 52, 53, 62, 64–68, 70, 78, 79, 90–92, 103

IMRT intensity-modulated radiotherapy 23, 52, 53, 56, 62, 64–68, 70, 78, 79, 91, 92, 103, 104, 111, 115, 133

LENT late effects normal tissues 26, 53, 91

LKB Lyman-Kutcher-Burman 3, 6, 90, 92–94, 98–100, 102–104, 139

LME linear mixed-effects 126

ME mixed-effects 123, 126

MRI magnetic resonance image 19

NPME nonparametric mixed-effects 39, 123, 128, 130, 133

NTCP normal tissue complication probability 3, 28, 30, 34–36, 39, 89, 90, 92–94, 98, 102–104, 111, 128, 139, 141

OAR organs at risk 20, 24

PCA principal component analysis 112–114

PSA prostate-specific antigen 17–20, 26, 39

PVD population value decomposition 39, 112, 113, 117, 118, 139

RF random forest 39, 90, 92–94, 98, 103

RF-NTCP random forest normal tissue complication probability 90, 92–94, 98, 102–104, 139, 141

RML restricted maximum likelihood 126

ROC receiver operating characteristic curve 6, 93, 98, 100

RT radiotherapy 51–53, 56, 58, 63–67, 70, 79

SOMA subjective-objective management 26, 53, 91

SVD singular value decomposition 112, 113

VBM voxel-based morphometry 112, 122, 123, 127, 129

Introduction

Prostate cancer is the most frequent type cancer among men in developed countries and the second cause of death from cancer. The estimated world population standardized incidence rate for 2011 was 125.8 in 100,000 men, which is growing with life expectation. In France, prostate cancer is the most frequent type of cancer in men. The estimated number of cases for 2011 is around 71,000 with an estimated number of deaths of 8,700 in the same year.

External beam radiotherapy (EBRT) is one of the treatments for prostate cancer. The objectives of EBRT are delivering a high dose of radiation to the tumor (prostate and seminal vesicles) while sparing the neighboring organs at risk, namely bladder and rectum. The aim of the treatment is to achieve a local control (eradication of the tumor) without side effects. The standard radiotherapy technique is the 3D conformal radiotherapy (3D-CRT). This 3D-CRT will attempt to avoid normal tissue irradiation and can be optimized using sophisticated techniques such as intensity-modulated radiotherapy (IMRT) and image-guided radiotherapy (IGRT).

The radiation delivered to normal tissues can produce complications, also called toxicities, that affect patient's quality of life. These complications may result from the interactions of the delivered dose to normal tissues, concomitant treatments (such as hormonal therapy or anticoagulant treatments) and patient characteristics (such as age, history of diabetes, previous surgeries, etc.). Different approaches can be followed to study the relationships among the factors implied in normal tissue complication following EBRT.

In radiobiological laboratories, through controlled experiments, it is possible to develop models that describe how normal cells are affected by radiation. In the clinic, however, patients are heterogeneous and carrying out controlled experiments that mimic the complexity of EBRT is complex. Indeed, clinical trials look for differences among average outcomes from different treatments. Normal tissue complication probability (NTCP) models emerge from an effort to integrate radio-biological models and observed patient outcomes. NTCP models have been designed to compare different EBRT plans but also to inform patients about the risk associated with different treatments.

This thesis aims at improving and enlarging the knowledge about NTCP models within a clinical application perspective. We propose to follow a path starting at classic statistical modeling, going through the adaption of current state-of-the-art prediction models to propose new NTCP models and finishing with the proposition of original NTCP models.

In terms of information complexity, we firstly tackle the use of patient parameters and general treatment parameters to characterize normal tissue complications following EBRT. Using Kaplan-Meier risk estimation and logistic regression, we describe bladder and rectum complications (and their associated symptoms) and we propose nomograms, a graphical tool to estimate complication probabilities and to help clinicians in treatment decision.

In a second step, we use dose-volume histograms (DHV), which are a "2D" reduction of the distribution of radiation within the patient's anatomy. Indeed, most classic NTCP models use only DVH to compute NTCP. These classic NTCP models can be seen as binary logistic regression models, which associate a high complication probability to a high dose delivered to the normal tissues. One drawback of these models is, however, the absence of inclusion of patient specific parameters. Adding new parameters to classic models implies making assumptions which are difficult to verify. An alternative is to stratify a population of patients and fitting different models to each stratum. However, in practice some strata will not meet the necessary number of patients to obtain meaningful estimators for the model parameters. In this context, we propose to use random forest, a machine-learning technique, that allows for both dosimetric and patient parameters to be simultaneously used in toxicity prediction.

In a third step, we use the "3D" dose distribution directly. As classic NTCP models use the DVH, which is a reduction of the dose distribution and the same DVH may correspond to different dose distributions, classic NTCP models lack of spatial accuracy. We have developed an image population value decomposition technique that enables to find dose patterns characterizing the differences in the dose distributions of patients presenting toxicity and those who do not. In addition, we develop a nonparametric mixed-effects model that allows for spatial (and temporal) intra-individual variability to be handled. We use this model to find anatomical regions correlated to toxicity following EBRT.

One challenge of this thesis was the use of a large cohort of patients from different French trials and centers. The process of interpreting the physician's reports and scoring the toxicity events was carried out by medical doctors who largely contributed to the development of this thesis and the interpretation of its findings.

This present thesis is a article-based thesis. After introducing the clinical context and NTCP models in Chapter I, we present in each of the following chapters one or two articles. The thesis is organized as follows:

- Chapter 2 presents two papers devoted to the classical regression modeling approach. In this chapter we describe symptom-wise rectal and bladder toxicities following EBRT. We also present nomograms for the different toxicities. These two papers are the product of a close collaboration with two medical doctors (urologist and gastroenterologist) and some of the findings were presented by them in their respective medicine thesis.

- Chapter 3 presents a new random forest NTPC (RF-NTCP) model. We develop a NTCP framework using random forest and we perform a comparison with classic

NTCP models. In this framework, DVH and patient parameters are easily integrated and the comparison of models based on the area under the curve (AUC) of the receiving operating characteristic curve (ROC) shows that RF-NTCP is a strong competitor of classic NTCP models.

- Chapter 4 presents our image population value decomposition technique and its application to find dose patters correlated to rectal toxicity following EBR.
- Chapter 5 presents our extended nonparametric mixed-effects model to a spatio-temporal framework. Using this model, we develop a framework to perform voxel-wise hypothesis testing, including they asymptotic theory that supports the use of the proposed statistic test (presented in Appendix B).

The main contributions of this thesis are:

(i) the development of new predictive models of rectal and bladder toxicities following EBRT and the demonstration that these models are strong competitors of the classic NTCP models and

(ii) the development of new image population analysis techniques and their corresponding mathematical formalisms.

Part I

Clinical context, problem definition and objectives

Clinical context, problem definition and objectives

1.1 Prostate gland and prostate cancer

The prostate is a gland in the male reproductive system. The prostate is located below the bladder and in front of the rectum (see Figure 1.1). The seminal vesicles, two irregularly-shaped glands, are connected to the base of the prostate. Between the prostate and the pubic symphysis there are fat, venous vascular structures and fibular tissues, which form the puboprostatic ligament. Various nerves located within the prostate are implicated in the erection mechanism. The urethra passes through the prostate from the bladder to the urethral opening.

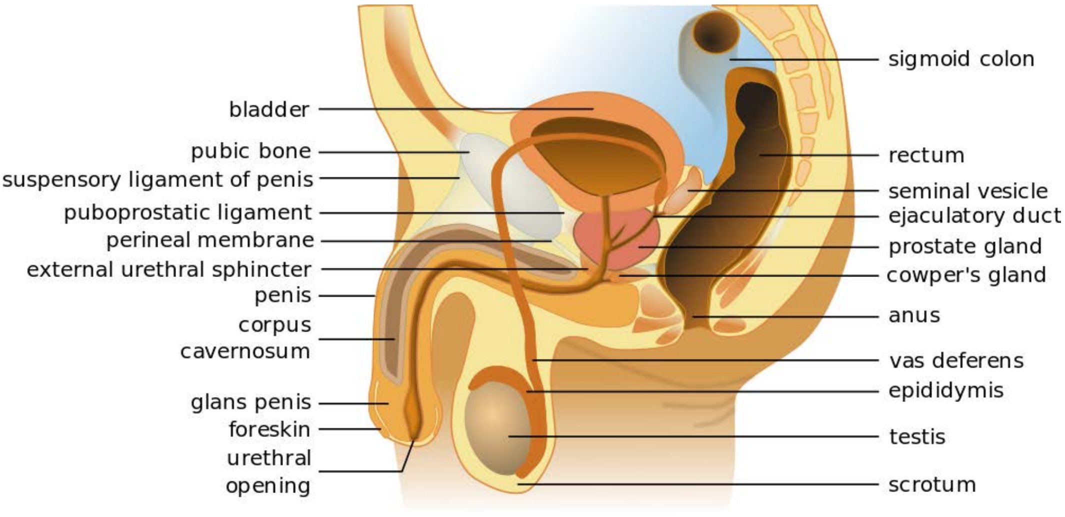


Figure 1.1: The male reproductive system (illustration by Frank H. Netter)

The prostate itself is composed of four glandular zones and a fibromuscular stroma. It is surrounded by a 3-4 mm thick capsule. It secretes a liquid involved in sperm composition and contains several enzymes, PSA. The prostate can contract to stream out sperm via the urethra, thus contributing to ejaculation.

Prostate size changes throughout life, starting its development during puberty to reach a weight of around 20g. The average prostate dimensions for an adult are 4 cm in width and 3 cm in height. Around 80% of men over 40 will undergo prostate hypertrophy, possibly compressing the urethra, leading to urinary troubles. This hypertrophy is usually benign and is part of the natural evolution of prostate, but it can also be a sign of cancer.

Prostate cancer is the most common type of male cancer in France [1]. In 2011, the estimated number of cases in France was around 71,000, with an expected number of deaths of 8,700 in the same year. The estimated world population standardized incidence rate is 125.8 in 100,000 men, which raises as life expectation increases. From 2002 to 2006 there were 39,785 new cases and 10,897 deaths from prostate cancer in Colombia, where it is the second-most frequent cancer after stomach cancer [2, 3]. For a complete review on prostate cancer screening, diagnosis and treatment (with curative intent) see for example the guidelines of the European Association of Urology (EAU) [4].

1.1.1 Diagnosis of prostate cancer

Two methods are mainly used to suspect prostate cancer: a blood test to measure the prostate-specific antigen (PSA) level and a digital rectal examination. An annual check-up is advised for men aged between 50 and 75 years [5]. This consists of a blood sample analysis to measure the PSA and of a digital rectal examination. Both sensibility and specificity of PSA test are low. Only 75% of cancers show a abnormally-high PSA value and only a quarter of these imply a positive biopsy. About 45% of men over 50 years have cancer cells within their prostates; of these, 9.5% will develop an aggressive cancer and 2.5% will die of prostate cancer [6]. In the early stage, the cancer is contained within the prostate but after it expands to include the neighboring organs and tissues as well as more distant organs such as the lymph nodes and bones. Tumor extension is usually expressed as being at certain T stage using the tumor, nodes and metastasis (TNM) classification system [7]. There are four T stages with various subcategories indicating tumor size and location:

- T1: tumor present, but not detectable clinically or with imaging:
  - T1a: Tumor found in less than 5% of prostate tissue resected (for other reasons);
  - T1b: Tumor found in more than 5% of prostate tissue resected;
  - T1c: Tumor found in a needle biopsy performed due to an elevated PSA level;
- T2: Tumor can be felt (palpated) on examination, but has not spread outside the prostate:
  - T2a: Tumor is in half or less than half of one of the prostate gland's two lobes;
  - T2b: Tumor is in more than half of one lobe, but not both;

  - T2c: Tumor is in both lobes but still inside of the prostatic capsule.
- T3: Tumor has spread throughout the prostatic capsule (if it is only half-way through, it is still classed under T2):
  - T3a: Tumor has spread through the capsule on one or both sides;
  - T3b: Tumor has invaded one or both seminal vesicles.
- T4: the tumor has invaded other nearby structures.

If either the rectal examination or the blood test reveal an anomaly, an echo-guided biopsy is prescribed [8], which makes the diagnostic. Around 12 biopsy samples are taken, which allow for tumor grade assessment and therefore prognosis of the disease. The structure of cancer cells within each sample is analyzed and scored from 1 (least aggressive) to 5 (most aggressive) grade. The Gleason score is then calculated by computing the frequencies of both the most common structure and the most aggressive one. Using the results of the Gleason score, T stage, PSA and Partin's tables [9] propagation risk is computed.

An MRI is always prescribed to detect a possible expansion of the cancer outside the prostate [10]. MRI is used to assess whether the cancer has reached the outside of prostate capsule or the lymph nodes. This image is also used to assess the volume and position of the tumor. Medium- and high-risk patients usually undergo a bone scan to detect metastasis. The aforementioned tests allow for the identification of three prognostic groups, following D'Amico's classification [11]:

- Low risk: T1 to T2a stages, Gleason score ≤ 6, PSA ≤ 10 ng/mm.
- Intermediate risk: T2b stage or Gleason score $= 7$ or 10 ng/mm$\leq$PSA$\leq$10 ng/mm.
- High risk: T3 to T4 stages or Gleason score ≥ 8 or PSA ≥ 20 ng/mm.

Treatment depends on the risk group of the disease. The earlier prostate cancer is detected, the lower the risk of death from the disease or from complications relating to it. However, if detected when it was already spread to neighboring organs, the risk of death is considerably higher.

1.1.2 Common treatments for prostate cancer

Clinicians use D'Amico's classification [12] to decide upon a particular treatment, such as:

- Surgery: when prostate cancer is localized, radical prostatectomy can be an option. However, an important percentage of patients undergoing surgery later present erectile dysfunction (ranging from 10% to 100% [13]).
- Cryotherapy: as an alternative to surgery, when cancer is localized and unifocal [14], its cells can be frozen in order to kill them. This is performed by injecting a cold gas (argon) through the perineum and up to the prostate.

- Hormonal treatment: its purpose is not to cure the cancer but to slow down its growth. Prostate cancer grows when exposed to testosterone. Thus by stopping the production of testosterone and its associated hormones (androgens) the growth of the cancer can be controlled. Different hormones can be used to inhibit the production of testosterone. Hormonal treatment is usually combined with other treatments such as radiotherapy.

- Chemotherapy: drugs which affects the tumor's cellular division are advised when the cancer is widespread. This drugs are generally toxic and the objective by using them is to increase the patient's life expectation and reduce the pain due to metastasis.

- High-intensity focused ultrasound: this treatment consists of heating the prostate using ultrasound [15] in order to kill tumor cells. The main treatment device is the Ablatherm®, which combines an ultrasound transducer and an echography probe introduced to the rectum during the operation. This technique is advised for patients aged over 70 years with low-risk cancers, but is mostly used in the case of local recurrence following radiotherapy [16].

- Radiotherapy: consists of delivering high doses of radiation to the tumor zone in order to affect the reproductive capabilities of cancer cells. Radiation can be administrated via the implantation of physical devices within the prostate (brachytherapy) or via linear accelerators (external beam radiation), the latest being the least invasive procedure.

- Watchful waiting: this option consists of monitoring the disease's progression by regularly undergoing PSA tests (every three to six months), digital rectal examinations (every six to twelve months) and biopsies (every six to eighteen months). No intervention is performed that attack the tumor. Watchful waiting is usually prescribed for patients whose life expectation beyond ten years and present low-risk cancers. However, the tumor evolves, the patient will undergo treatment. About one third of patients under active surveillance will undergo curative treatment for prostate cancer.

Table 1.1 shows the distribution of prostate cancer treatments among 2,182 patients in France in 2001 [17].

1.2 External beam radiation therapy

Among the patients diagnosed with cancer, more than two thirds will receive external radiotherapy, possibly combined with a concomitant treatment (e.g., surgery or chemotherapy). External beam radiotherapy (EBRT) consists of delivering ionized rays charged with high energy photons to the tumor lesions. The objective is to deliver high doses of radiation to tumor cells whilst sparing the organs at risk (OAR). The patient is placed on a table which is above a linear particle accelerator (see Figure 1.2). To

Table 1.1: Prostate cancer treatment distribution in France in 2001

Treatment	Effective	Proportion (%)
Prostatectomy	595	27.7
Radiotherapy	513	23.8
External	429	83.6
Curietherapy	78	15.2
Unknown	6	1.2
Hormonal therapy	646	30.0
Other, watchful waiting	399	18.5
High intensity focused ultrasound	14	3.5
Unknown	28	-

allow normal tissues enough time to recover, the radiation process is carried out over several sessions in prostate cancer. During each session, just a fraction of the total prescribed radiation is delivered. The radiation units are expressed in Grays, abbreviated as Gy, which represents 1 Jule/Kg. In prostate radiotherapy for example, if 80 Gy are prescribed to the prostate, treatment can be completed in 40 sessions in eight weeks, with a fractionation scheme of two Grays per fraction (2 Gy/fraction).

The treatment strategy depends on the stage of the cancer. For an early stage tumor, the objective is to spare organs at risk, limiting side-effects by limiting the target volume and total dose. For a tumor at a more advanced stage, radiotherapy will aim to increase local control.

1.2.1 Irradiation techniques

1.2.1.1 Standard 3D conformal radiotherapy

Once it has been decided to treat the patient with EBRT, the first is the planning. In the past, the target volume was defined using X-ray images, which led to the definition of large irradiation volumes. These large volumes compensated uncertainties concerning the position and variations of different organs. This technique, known as 2D radiotherapy, was made obsolete with the development of computer tomography (CT) imaging, the latter becoming a primary tool in irradiation volume definition. Scanner images enable 3D images to be segmented, allowing for a greater clarity in viewing and identifying tumor volume. In case of prostate cancer, irradiation ballistics is usually defined by five to nine convergent beams, which conform the target volume via the modulation of a multileaf collimator located at the linear accelerator output (see Figure 1.2). This form of irradiation is known as 3D conformal radiotherapy (3D-CRT) and remains the gold standard in EBRT.

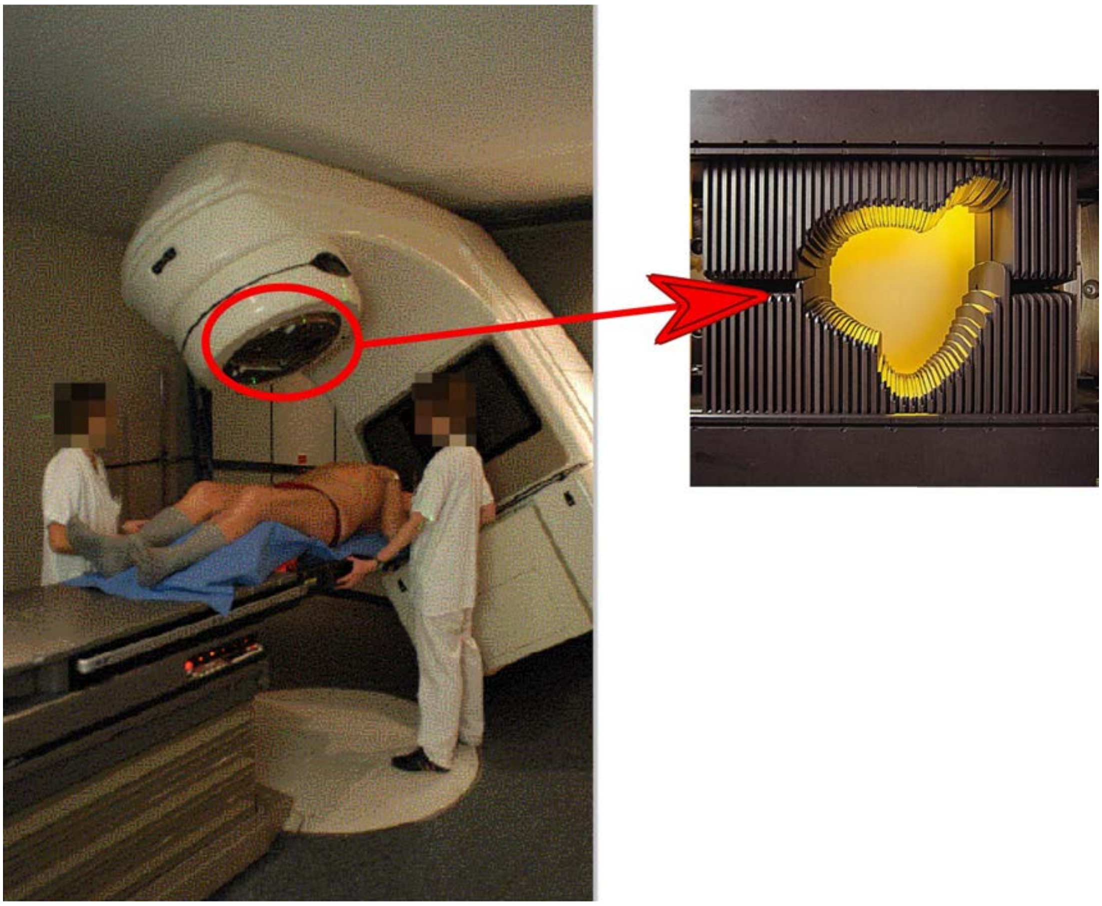


Figure 1.2: Linear accelerator (left) and multileaf collimator (right, manufactured by Varian, 120 leafs)

1.2.1.2 Intensity-modulated radiotherapy

Intensity modulated conformal radiotherapy (IMRT) appeared with the arrival of inverse-planning software. The fluence (photon amount per surface unit) is no longer homogeneous and is modulated, using the multileaf collimator, by changing the position of the leafs during the irradiation process. One advantage of this technique is obtaining concave iso-doses, an interesting property desired mainly at the prostate-rectum junction. The IMRT has become the standard irradiation technique in 30% of French treatment centers [18], including at the Centre Eugène Marquis (from 2003 onwards) which provided the medical support and background to this thesis. Treatment ballistics can be optimized using different algorithms. Figure 1.3 shows the work-flow of the Direct Machine Parameter Optimization with RayMachine using the Pinnacle® treatment planning system (TPS) manufactured by Phillips.

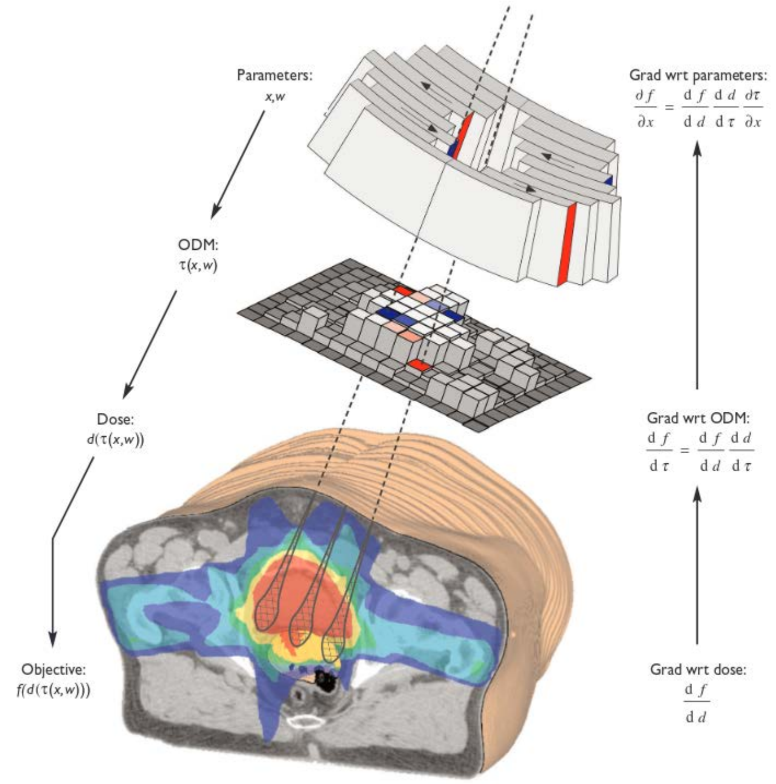


Figure 1.3: Direct Machine Parameter Optimization with RayMachine using Pinnacle (image taken from [19])

The preferences of physician are represented as a mathematical cost function, f, to be optimized. The input parameters of this function include leaf positions, some weights (w), and the fluence matrix (also called dose matrix and represented by d). Figure 1.4 shows a planned dose distribution to treat prostate cancer with IMRT. Arc-therapy is an IMRT evolution. It uses non-stationary beams that revolve around the patient during the irradiation process.

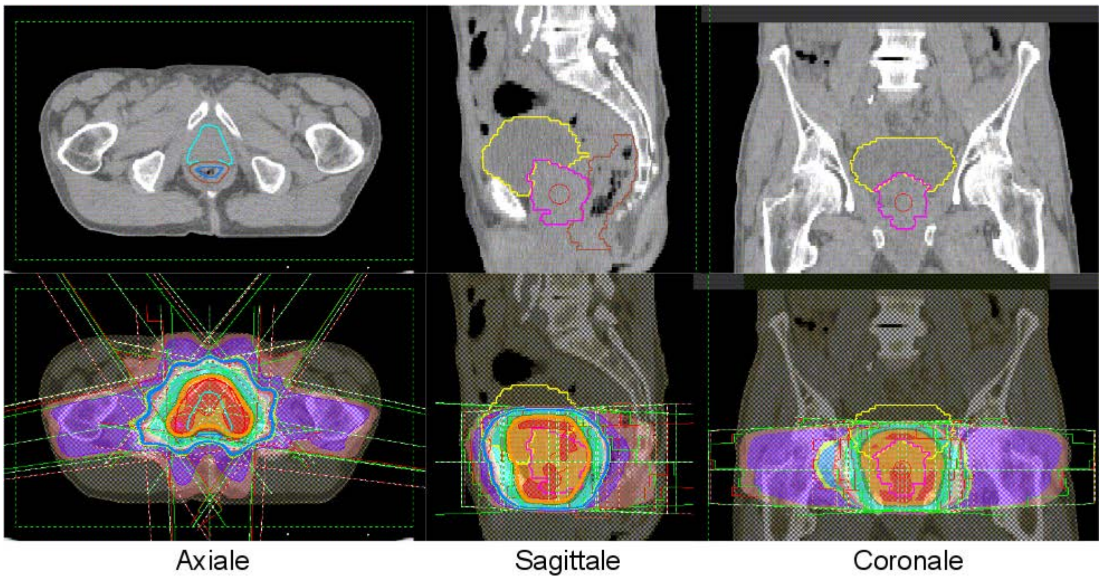


Figure 1.4: Above: Main structures (prostate, bladder and rectum) in three different views of a CT scan. Below: IMRT optimal dose distribution. Note the concave shape of the posterior part of the prostate

1.2.1.3 From bone visualization to image-guided radiotherapy

Different anatomical references can be used to set the patient in the right position. For example, markers can be tattooed on a patient's skin and then aligned at each irradiation session over the duration of the treatment using laser markers. However, as the patient might experience anatomical changes such as weight loss, the effectiveness of these skin markers is limited and may lead to setup errors, resulting in the accidental exposure of unaffected areas or, indeed, in the non-exposure of affected areas.

The portal image was developed using linear accelerators. Images are produced using high-energy X-rays delivered by the accelerator. These images allow for bone structures to be visualized, as depicted in Figure 1.5. A digitally-reconstructed radiograph (DRR) (a CT-like image) is generated by simulating irradiation in a process that mimics the geometry of the treatment. This image allows for bone structure alignment and for patient repositioning, which is intended to decrease the amount and frequency of setup errors. As this image modality implies delivering some ionized radiation to the patient, this control is usually performed at the beginning of treatment and weekly thereafter.

Image-guided radiotherapy (IGRT) emerges from the necessity of keeping the tumor (and not only the patient) in the same position along the different treatment sessions due to anatomical variations. The aim is to deliver the dose (in Gy) as defined in the treatment plan to the tumor, while avoiding the OAR.

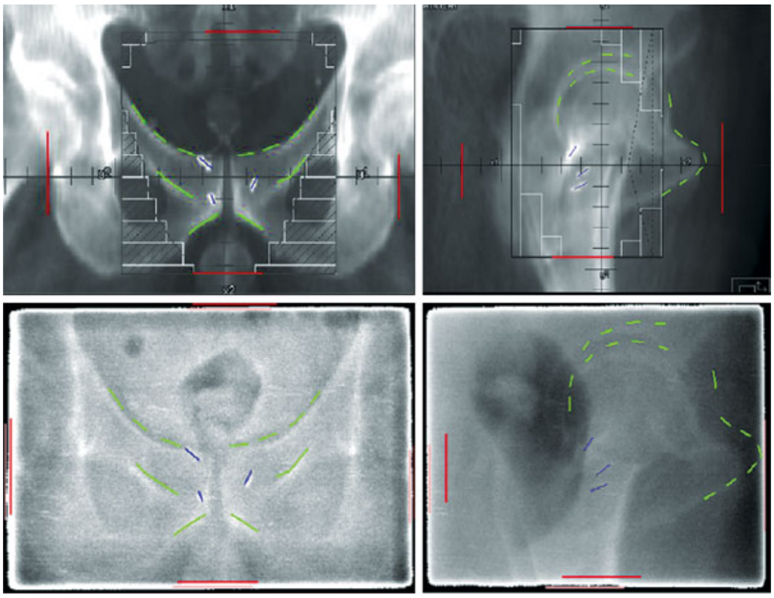


Figure 1.5: Above: DRR reference images. Below: portal image. In red: manual delineations of the irradiation field. In green: manual delineations of bone structures used to align both type of images. In blue: delineation of three gold grains implanted within the prostate

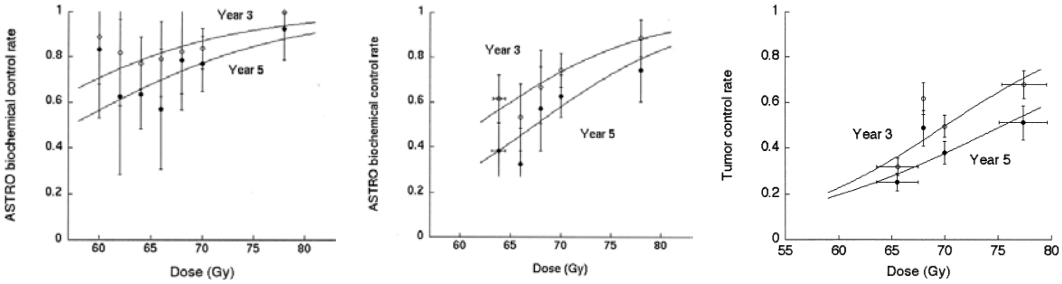
Patient repositioning according to bone structure alignment does not solve the problem of geometrical uncertainties. In most cases, the treatment plan is defined with just one planning scan. Thus, anatomical variations during treatment (or during a session) can lead to anatomical changes with respect to the bone structure referential.

In the case of prostate cancer, these deformations mostly come from prostate deformation or variations in the bladder and rectum. These anatomical variations can lead to subradiation of the target volume, namely the prostate, or to an over radiation of the organs at risk, namely the bladder and rectum, decreasing local control in the former and increasing the risk of toxicity in the latter. The need to have a precise localization of the tumor is behind the development of new image modalities (with new devices integrated to linear accelerators), such as kilovolts (kV), megavolts (MV), ultrasound or electromagnetic detection. The use of these image modalities in clinic is known as image-guided radiotherapy (IGRT). The cone beam computer tomography (CBCT), the most commonly-used image modality in IGRT, relies on a X-ray machine set to using kV energy (or MV) and a 2D detector, placed in front of the source. The acquisition system rotates around the patient and an important number of 2D projections are acquired, allowing for 3D reconstruction of the scanned area. The reconstructed 3D image is then compared to the CT taken during the planning stage.

1.2.2 Dose-effect relationships in local control

Almost all localized prostate cancer can be treated with radiotherapy. Usually, the target volume includes the prostate and seminal vesicles (sometimes the lymph nodes). In practice, once the main organs are manually delineated, security margins are added to the prostate and seminal vesicles ensemble to define a new target volume, known as the planning target volume (PTV). Other than treating the affected area, one important objective is to safeguard the preservation of the bladder, rectum, femoral head and other normal tissues.

The dose-effect relationship is based on the principle that the higher the dose, the higher the tumor-control probability (TCP) [20, 21]. Tumor control can be expressed in terms of two main types of recurrence, namely biochemical and clinical recurrence. Biochemical recurrence is defined in terms of PSA increase and clinical recurrence by a biopsy. Figure 1.6 demonstrates the dose-effect relationship as TCP (expressed in terms of biochemical control) at three and five years post-treatment. However, a limiting factor for dose escalation is normal tissue complications. In order to achieve local control and spare organs at risk, various recommendations have been proposed by different cooperative groups (see Fiorino et al [22]).



(a) Low risk (b) Medium risk (c) High risk

Figure 1.6: Dose-effect relationship of the risk of biochemical control by risk group in prostate cancer radiotherapy (taken from [21] (a,b) and from [20] (c)).

1.3 Side effects related to prostate cancer radiotherapy

Side effects following prostate cancer radiotherapy can appear as a result of normal tissue irradiation. As most radiation is delivered to the pelvic zone, complications may appear in the genitourinary (bladder) and gastrointestinal (rectum, small bowel and sigmoid) systems. In this thesis the terms “complication” and “toxicity event” are equally used to describe side effects. Complications following prostate cancer radiotherapy have an adverse impact on patients’ life quality [23]. Under prospective clinical trials, physicians record patient symptoms at each follow-up visit. Afterwards, the toxicity events are graded using standard grading scales. These scoring systems include late effects normal tissues (LENT) / subjective-objective management (SOMA) (LENT/SOMA) (see Table

Table 1.2: GTUG dose-volume constraint recommendations for target volume and organs at risk in prostate cancer radiotherapy

Volume	Notation	Definition
Target volume (PTV)	$D_{min} > 90\%$	Minimum dose to PTV must be higher than 90% of the prescribed dose.
	$V_{95} > 95\%$	The volume receiving at least 95% of the prescribed dose must be higher than 90% of the total volume.
Bladder wall (7 mm)	$D_{max} < 80\,\mathrm{Gy}$	The average dose to 1.8 cm^3 must be always lower than 80Gy.
	$V_{70} < 50\%$	The volume receiving at least 70 Gy must be lower than 50%.
Rectal wall (7 mm)	$D_{max} < 76\,\mathrm{Gy}$	The average dose to 1.8 cm^3 must always be lower than 76 Gy.
	$V_{72} < 25\%$	The volume receiving at least 72 Gy must be lower than 25%.
Femoral heads	$V_{55} < 5\%$	For each femoral head, the volume receiving at least 55 Gy must be lower than 5%.

2.2, pag. 53 and Table 2.5, pag. 65); Radiation Therapy Oncology Group and the European Organization for Research and Treatment of Cancer (RTOG/EORTC); and the common terminology criteria for adverse events (CTCAE). Most scoring systems have been shown to correlate doses received by normal tissues, although CTCAE seems to produce more toxicity events and correlates more dose-volume parameters [24, 25]. An additional effort must be made when using retrospectively collected data to match records from physicians' reports to a common terminology. For this reason, prospective studies are the basis of almost all clinical guidelines and tools. Retrospectively-collected information may be, however, used to develop and validate models.

Side effects can be divided into two categories: acute and late. Acute toxicity is associated with early events (during treatment or six months prior to the end of treatment). Late effects or late toxicity is associated with symptoms occurring after six months following the end of radiotherapy treatment.

This thesis will focus on late rectal and bladder complications. Fiorino et al [22] present a comprehensive literature review on relationships between dose and clinical variables and normal tissue effects following radiotherapy. The Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) initiative [26] provides some important literature reviews and research guidelines in normal tissue complication probability (NTCP) modeling. Of special interest are the papers focused on rectal [27] and bladder complications [28]. Rectal bleeding has been the most-studied late side effect (see Trott et al for a comprehensive review [29]) and that is why this thesis will focus on this symptom. The following complications are those to be studied:

Bladder complications

- Urinary frequency
- Dysuria (painful urination)
- Incontinence
- Hematuria (bladder bleeding)

Rectal complications

- Rectal bleeding
- Rectal incontinence
- Stool frequency

1.4 Models to predict NTCP following radiotherapy

1.4.1 Lyman's NTCP model: uniform irradiation

Clinicians used the tolerance radiation dose (the radiation dose that a normal tissue can undergo without experiencing side-effects) based on their experience, before the publication of Lyman's work [30]. The latter was an attempt to match a mathematical model with medical experience. This model was proposed in the context of uniform radiation, where the target was box-shaped and all tissues within were exposed to the same dose [31]. Lyman's NTCP model is a logistic model that enabled clinicians to interpolate data, namely by computing complication probabilities for different values of organ-volume irradiation, instead of using fixed values of tolerance doses established via consensus.

Lyman's model belongs to the class of parametric models, which means that the assumed mathematical relationship has a fixed functional form which maps independent variables (related to the target dose and the fraction of the organ volume being irradiated) to the complication probability. This model is presented in equation (1.1),

$$NTCP\left(D, V; TD_{50}\left(1\right), n, m\right) = \Phi\left(\frac{D - TD_{50}\left(V\right)}{m \times TD_{50}\left(V\right)}\right), \tag{1.1}$$

where $TD_{50}\left(V\right) = \frac{TD_{50}(1)}{V^n}$, $TD_{50}\left(1\right)$ is the dose that leads to a toxicity event probability of 0.5 for a uniform radiation of the whole organ volume, V is the irradiated partial volume, n is a parameter associated with the volume effect, m is a slope parameter and $\Phi\left(\cdot\right)$ is the cumulative distribution function of a standard normal random variable, which makes this model a nonlinear "probit" model. Note that $NTCP\left(TD_{50}\left(1\right), 1; TD_{50}\left(1\right), n, m\right) = \Phi\left(0\right) = 0.5$ and that in order to have low complication probabilities, $D < TD_{50}\left(V\right)$ is required, which implies that $DV^n < TD_{50}\left(1\right)$. Moreover, if the model parameters are known, it is possible to give explicit formulas for calculating the fraction of volume that irradiated with dose D results in a complication probability p_0, as in equation (1.2). It is also possible to calculate the dose that, with a partially irradiated fraction of volume V, results in a complication probability p_0, as in equation (1.3),

$$V = \left[\frac{TD_{50}\left(1\right)\left(1 + m\Phi^{-1}\left(p_0\right)\right)}{D}\right]^{1/n}, \tag{1.2}$$

$$D = \left(1 + m\Phi^{-1}\left(p_0\right)\right) TD_{50}\left(V\right). \tag{1.3}$$

The model parameters can be obtained using maximum likelihood. The main remark to make is that with this model it is possible to make decisions concerning the treatment planning [32], and that it is consistent with some basic assumptions:

1. Increasing the dose increases the complication probability.
2. Increasing the irradiated fraction of volume increases the complication probability.
3. For different normal tissues the relationship between the target dose, the fraction of uniformly-irradiated volume and the complication probability is different and is reflected by different organ-specific parameters.

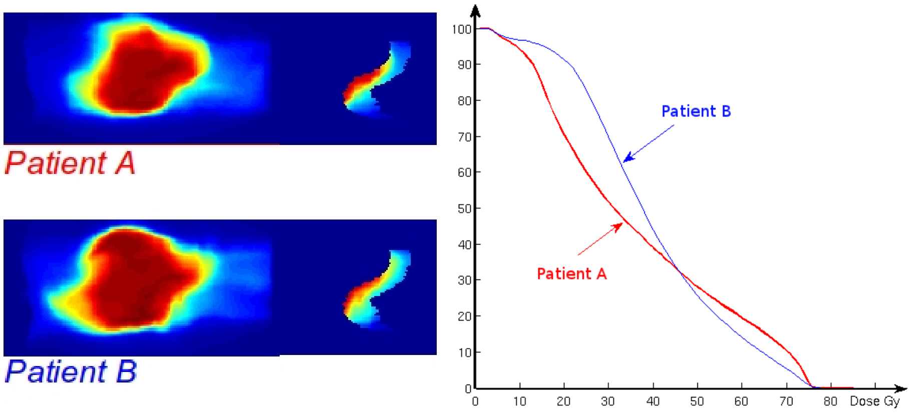
Equivalent formulas to those presented in equations (1.2) and (1.3), can also be given just by changing $\Phi^{-1}\left(p_0\right)$ with $logit\left(p_0\right) := \ln\left(\frac{p_0}{1-p_0}\right)$. Equation (1.4) shows a model reparametrization in the context of generalized non linear models,

$$NTCP_{Ly} = f\left(\beta_0 + \beta_1 D V^{\beta_2}\right), \tag{1.4}$$

with $f\left(\cdot\right)$ being either the logistic function or the cumulative standard normal distribution. It is easy to verify that $m = -1/\beta_0$, $n = \beta_2$ and $TD_{50}\left(1\right) = -\beta_0/\beta_1$. Hereafter, the model presented in equation (1.4) will be referred to as Lyman's model.

1.4.2 Lyman's NTCP model: nonuniform irradiation

Dose distribution evolved with the emergence of 3D-conformal radiotherapy (3D-CRT) and was no longer uniform. The objectives of this technique were to achieve a dose distribution adapted to the shape of the clinical target volume and to spare the organs at risk. The information contained in the dose distribution is condensed in the DVH. The DVH is represented as a non-increasing function of the dose that matches any given dose value with the fraction of the organ volume receiving at least that dose. This standard definition of the DVH is also known as cumulative DVH. Figure 1.7-(a) shows two 2D views of two 3D-CRT dose distributions to the rectum for two different patients, Patient A and Patient B, while Figure 1.7-(b) shows the corresponding rectal DVHs.



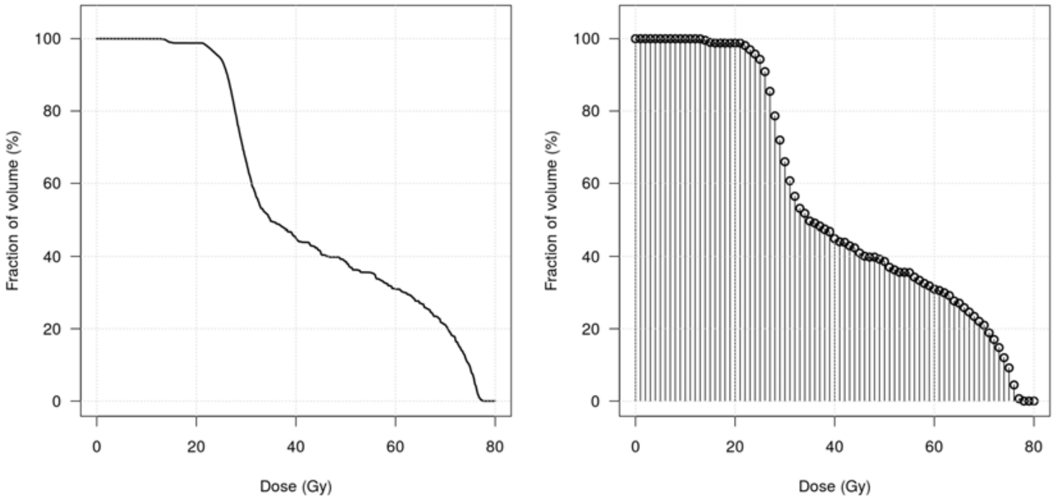
(a) A 2D view of a 3D dose distribution. (b) Corresponding DVHs.

Figure 1.7: An example of the DVH for rectum in two different patients.

One strategy of using DVH in NTCP estimations consists of reducing the DVH to a maximum dose and an effective volume [33]. This effective volume is defined as the fraction of the organ volume that should be irradiated to the maximum dose, namely D_{max}, in order to have the same total dose.

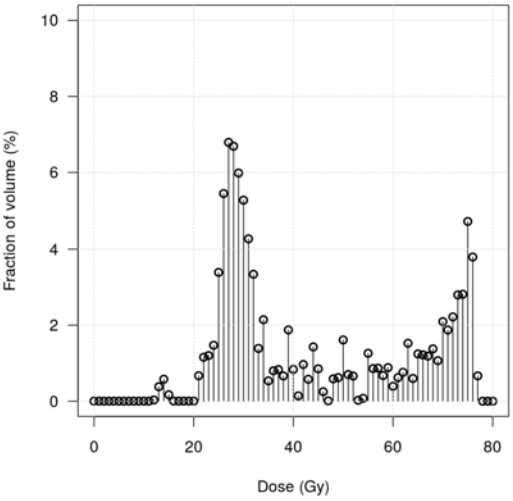
To illustrate this concept, it is useful to see the DVH as a discrete function. For a given patient it is possible to represent cumulative DVHs as a set of points (D_i, V_i), $i = 1, 2, \ldots, N$. There, V_i represents the fraction of the organ volume that receives at least a dose equal to D_i Gy (1 Gy=1 Joule/Kg), and $(D_i)_i$ is a partition of the interval $[0, D_{max}]$. The differential DVH is obtained by defining $v_i = V_i - V_{i+1}$, which represents the fraction of the volume receiving a dose between D_i and D_{i+1} Gy. For simplicity, it is assumed that v_i is the fraction of volume receiving a dose of D_i Gy and that $v_N = V_N$ (or equivalently $V_{N+1} = 0$). Figure 1.8-(a), -(b) and -(c) represent a rectal DVH, its discretized version with one Gy bin step and the differential version.

Then, one can make the following assumption: if the fraction of volume v_i receives D_{max} Gy instead of D_i Gy, a smaller volume than v_i should be irradiated in order to



(a) Rectal DVH from a conformal plan

(b) Discrete version



(c) Differential version

Figure 1.8: Rectal cumulative DVH, discrete version of the cumulative DVH and discrete version of the corresponding differential DVH

keep the same quantity of radiation. This smaller volume is called the effective volume, v_{eff_i}. A power law is assumed to find this new v_{eff_i} [33], leading to equation (1.5),

$$v_{eff_i} = v_i \left(\frac{D_i}{D_{\max}} \right)^n , \tag{1.5}$$

where n is a size parameter. According to equation (1.5), the effective volume V_{eff} representing the fraction of the organ volume that should be irradiated with a dose D_{max}, keeping the same energy, can be calculated as in equation (1.6),

$$V_{eff} = \sum_{i=1}^{N} v_{eff_i} . \tag{1.6}$$

Lyman's model (equation (1.4)) can be evaluated using (D_{max}, V_{eff}) instead of (D, V). The main assumption behind this strategy is that both the original DVH and simplest-squared DVH (D_{max}, V_{eff}) lead to the same complication probabilities. The same parameter n is used in both equations (1.1) and (1.5), meaning that it is assumed that the power relationship that holds for tolerance doses also holds for non-uniformly irradiated volumes. Thus, the model using V_{eff} looks similar to the model using the uniformly-irradiated volume V.

1.4.3 NTCP and volume effects

Other strategies have been proposed to take into account volume effects [34]. Let us assume that $P(D, V)$ is the complication probability of an organ with a fraction V of its volume being uniformly-irradiated with a dose of D Gy. This fraction is defined with respect to a reference volume V_0 (for example the total organ volume) whose associated complication probability is known when irradiated using some specific dose D_0 Gy. Thus, $P(D_0, 1) = P(D_0, V_0/V_0)$ is known. The probability that the organ avoids injury is then $1 - P(D, v)$, with $v = V/V_0$. If the organ is subdivided into N (equal) subvolumes, the probability of each subvolume avoiding injury is then $1 - P(D, 1/N)$. If we assume that a single injured subvolume represents an organ injury and that subvolumes are independent among themselves, the probability that all parts scape injury can be calculated as $[1 - P(D, 1/N)]^N$. This former probability must be equal to $1 - P(D, 1)$, leading to equation (1.7):

$$1 - P(D, 1) = [1 - P(D, 1/N)]^N . \tag{1.7}$$

If only $M < N$ subvolumes are irradiated with a dose D Gy, the probability injury to the whole organ is:

$$1 - P(D, M/N) = [1 - P(D, 1/N)]^M . \tag{1.8}$$

Defining $v = M/N (= V/V_0)$, together with equations (1.7) and (1.8), yields the expression given in equation (1.9),

$$P(D, v) = 1 - [1 - P(D, 1)]^v . \tag{1.9}$$

Notice that $P\left(D,1\right)$ in equation (1.9) is assumed to be known. Thus, it can be considered as the only model parameter.

For a non-homogeneous dose distribution, let $\left(D_i,v_i\right)_{i=1}^{N}$, $i=1,2,\ldots,N$, be the differential DVH (see Figure 1.8-(c)). Define $P\left(\left(D_i,v_i\right)_i^N\right)$ as the probability that this DVH causes injury to the organ. Then, once again assuming independence among the N subvolumes irradiated with different doses, the probability of the whole organ avoiding injury can be expressed as in equation (1.10),

$$1-P\left(\left(D_i,v_i\right)_{i=1}^{N}\right)=\prod_{i=1}^{i=N}\left[1-P\left(D_i,v_i\right)\right]. \tag{1.10}$$

Using equation (1.9) it can be said that $P\left(D_i,v_i\right)=1-\left[1-P\left(D_i,1\right)\right]^{v_i}$, which results in equation (1.11),

$$P\left(\left(D_i,v_i\right)_{i=1}^{N}\right)=1-\prod_{i=1}^{i=N}\left[1-P\left(D_i,1\right)\right]^{v_i}, \tag{1.11}$$

provided that $P\left(D_i,1\right)$ is known for $i=1,2,\ldots,N$. Note that $\frac{dP(D,v)}{dv}=-\left[1-P\left(D,1\right)\right]^{v}\times\ln\left(1-P\left(D,1\right)\right)$, then expanding $P\left(D,v\right)$ around $v=0$ and by taking into account the fact that for x enough small $\ln\left(1-x\right)\approx -x$, equation (1.9) can be approximated as in equation (1.12) for enough small $P\left(D,1\right)$,

$$P\left(D,v\right)=v\times P\left(D,1\right), \tag{1.12}$$

which, when used in equation (1.11) as well as the fact that $1-\prod_{k=1}^{N}\left|a_k\right|\approx\sum_{k=1}^{N}\left|a_k\right|$, provided that $\left|a_k\right|<<1$, $k=1,2,\ldots,N$, yields equation (1.13),

$$P\left(\left(D_i,v_i\right)_{i=1}^{N}\right)=\sum_{i=1}^{i=N}v_i\times P\left(D_i,1\right). \tag{1.13}$$

However, $P\left(D_i,1\right)<<1$(indeed very close to zero) is a necessary condition for this approximation to hold. It is also important to note that, in equation (1.9), $P\left(D,1\right)$ was assumed to be a fixed parameter. Moreover, in equation (1.13) one must know the value of $P\left(D_i,1\right)$ for all doses D_i, $i=1,2,\ldots,N$. This is solved by assuming the relationship of equation (1.14),

$$P\left(D,1\right)=\frac{1}{1+\left(TD\left(50,1\right)/D\right)^{k}}, \tag{1.14}$$

where k is a fixed parameter (known) and $TD\left(50,1\right)$ is as previously defined. Under the assumption that $P\left(D_i,1\right)<<1$, the approximation $P\left(D,1\right)=\left(D/TD\left(50,1\right)\right)^{k}$ allows the expressions given in equations (1.15) and (1.16) to be expressed for homogeneous and non-homogeneous irradiation patterns, respectively,

$$P\left(D,v\right)=v\times\left(D/TD\left(50,1\right)\right)^{k}, \tag{1.15}$$

$$P\left((D_i, v_i)_{i=1}^{N}\right) = \sum_{i=1}^{i=N} v_i \times (D_i/TD\,(50,1))^k\,. \tag{1.16}$$

Equation (1.11) remains a more accurate model and does not use the strong assumption $P\,(D_i, 1) << 1$. This is referred to in the literature as Shulteiss' model. The expression in equation (1.13) can be seen as a discretization of an integral of the form $\int P\,(D,v)\,dv$ (assuming a continuous differential DVH), and remains the same as the formalism of the complication probability factor (CPF) [35] and whose application in radiotherapy planning has been demonstrated [36]. The CPF assumes a linear dependency on volume and that $P\,(D_i, 1) << 1$. The CPF can be considered the first attempt to directly include complication probability in the treatment planning process. Its deduction follows similar steps as explained above.

1.4.4 Classic NTCP models

The key idea is that by linking clinical experience, collected data or that based on some biological assumptions, it is possible to derive models relating treatment parameters (e.g., dose distribution) to the outcome (toxicity probability). The NTCP models most frequently found in the literature are, as follows:

1. The Lyman-Burman-Kutcher (LKB) model [30, 37]: Presented in equation (1.17), model parameters are TD_{50}, n and m,

$$NTCP = \Phi^{-1}\left(\frac{D_{max} - TD\,(50,v)}{mTD\,(50,v)}\right), TD\,(50,v) = TD_{50} \times \left(\sum_i v_i D_i^{1/n}\right)^{-n}. \tag{1.17}$$

2. The Logit EUD model [34]: this is a logistic regression model with an equivalent uniform dose (EUD) reduction method. Presented in equation (1.18) with parameters TD_{50}, k and n,

$$NTCP = \left[1 + \left(\frac{TD_{50}}{EUD}\right)^k\right]^{-1}, EUD = \left(\sum_i v_i D_i^{1/n}\right)^{-n}. \tag{1.18}$$

3. Schultheiss [34]: Presented in equation (1.19) with parameters TD_{50} and k, the main assumption here is the "serial" behavior of the organ. In practice, this model expresses complication probability as the probability of failure in a serial system:

$$NTCP = 1 - \prod_i \left[1 - \left[1 + \left(\frac{TD_{50}}{D_i}\right)^k\right]^{-1}\right]^{v_i}. \tag{1.19}$$

4. The Poisson model [38]: Presented in equation (1.20) with parameters TD_{50}, γ and n, its EUD is like that of equation (1.18),

$$NTCP = 2^{-\exp\left(e\gamma\left(1-\frac{EUD}{TD_{50}}\right)\right)}. \tag{1.20}$$

5. The Kälman model [38]: Presented in equation (1.21) with parameters TD_{50}, γ and s,

$$NTCP = \left[1 - \prod_i \left[1 - PD_i^s\right]^{v_i}\right]^{1/s}, PD_i = 2^{-\exp\left(e\gamma\left(1-\frac{D_i}{TD_{50}}\right)\right)}. \tag{1.21}$$

6. The Parallel model: Presented in equation (1.22) with parameters TD_{50}, u_{50}, σ and k, this model, stated by Niemierko and Goitain [39] and Jackson et al [40], relies on the assumption that some normal tissues behave like parallel units [41]

$$NTCP = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^f e^{\frac{-(u-u_{50})^2}{2\sigma^2}}, f = \sum_i v_i p\left(D_i\right), p\left(D_i\right) = \left[1 + \left(\frac{TD_{50}}{D_i}\right)^k\right]^{-1}. \tag{1.22}$$

1.4.5 Parameter estimation of classic NTCP models

Identifying parameters for NTCP models can be performed via maximum likelihood. If M patients are available and we let R_i be the indicator function, thus for the i-th patient is 1 if he presents toxicity and 0 otherwise. The likelihood for any of the aforementioned models can be expressed as in equation (1.23),

$$L\left(\theta|data\right) = \prod_{i=1}^{M} NTCP\left(DVH_i|\theta\right)^{R_i}\left(1 - NTCP\left(DVH_i|\theta\right)\right)^{1-R_i}, \tag{1.23}$$

where $NTCP\left(DVH_i|\theta\right)$ represents the complication probability for the i-th patient whose treatment information is given as DVH_i (possibly represented as $(D_i, V_i)_{i=1}^{N}$) and θ stands for the model parameters (for example, for the LKB model we have $\theta = (TD_{50}, n, m)$). By *data* we mean all patient DVHs. The log-likelihood is given in equation (1.24),

$$l\left(\theta|data\right) = \sum_{i=1}^{M} R_i \times \log NTCP\left(DVH_i|\theta\right) + \left(1 - R_i\right) \times \log\left(1 - NTCP\left(DVH_i|\theta\right)\right). \tag{1.24}$$

The negative of this function must be minimized for the NTCP models at each clinical endpoint in order to obtain the parameter estimates and the 95% confidence intervals can be derived profiling the log-likelihood function [42].

Although many studies have shown a correlation between dose, volume and rectal toxicity [43, 44, 45, 46, 47, 48], they lack spatial accuracy and are not able to correlate treatment outcome with the specific dose pattern. While reducing the rich three-dimensional information of the dose to a two-dimensional pattern such as DVH, the potential correlation that may exist between toxicity and irradiation at small portions of the rectum may not be detected. Some attempts have been made to introduce the notion of spatiality [49] but only within the dose-volume space. Another major issue is that most NTCP models use information from planned treatments, which have been shown to differ from actual delivered treatments [50]. In this thesis we will not tackle this issue.

1.4.6 Alternative NTCP models

The objective of NTCP models is to help clinicians best assess, in order to select the best combination of treatment parameters. As can be noted from the given definition of NTCP models (equations (1.17) through (1.22)), patient characteristics (such as age, medical history, etc.) are not explicitly considered. One strategy for including patient characteristics in classic NTCP models is stratifying the population and estimating model parameters at each stratum [51, 52]. One disadvantage of this approach however is that each subgroup does not always have enough patients to estimate the parameters. Other approaches include the use of principal component analysis (PCA) [53] to extract the most relevant information from DVHs to predict toxicity. The use of artificial neural networks (ANN) in predicting toxicity has also been explored [54, 55]. Nomograms, charts to calculate probabilities from logistic regression models, have also been proposed [56].

Most NTCP models consider toxicity events occurring before a fixed time following treatment (e.g. 3-year rectal bleeding or 5-year overall bladder toxicity). This approach has been selected in this thesis in order to compare our results with most of publications on toxicity prediction. However, other approaches taking into account individual patient follow-up and censoring in the LKB model, have been reported in the literature [57].

1.4.7 From DVH to dose-distribution studies to predict toxicity

Few attempts have been made to study the full 3D distribution to predict toxicity following radiotherapy. The advantage of DVH-based approaches is that, by reducing the spatial 3D dose distribution to the DVH, patient morphology is dropped from the analysis. Patients undergoing the same treatment can be then used to identify parameters that subsequently allow for toxicity prediction. It is possible, however, to construct two different dose distributions that lead to the same DVH. Thus, if a specific spatial pattern in 3D dose distribution is responsible of some type of toxicity, DVH-based models fail to detect it. The main disadvantage is that DVH-based NTCP models lack spatial accuracy. To overcome this disadvantage, a vectorized DVH (called VDVH) was recently proposed [58]. For a set of dose ranges, the mean value of the x, y, and z coordinates of the voxels in each range as well as the boxplot of the distance from each voxel in each range to some structure of interest (the planning target volume for example) are represented in the same plot that the DVH. The information provided for this VDVH has not been used yet to predict toxicity following prostate cancer radiotherapy.

A strategy to compare dose distribution is the use of dose-surface maps [59]. Using some geometrical and anatomical criteria, points of the dose distribution are extracted to build up dose maps. By simulating clusters of damaged units [59] or by thresholding this dose maps[60] sets of vector features are generated and then correlated to some toxicity endpoint.

Using the dose-surface maps, Buettner et al have shown that late complications in the rectum are not only related to volumetric aspects of the dose, but particularly to the shape of the dose distribution [60]. More recently, spatial considerations were incorporated by parameterizing the 3D dose patterns [61]. In this way, by selecting a limited set of predictive features, their method outperforms classical models based on DVH. This approach still lies in the reduction of feature dimensionality by fitting analytic functions to each dose shape. Nevertheless, subtle underlying correlations may exist between toxicity and dose distribution at voxel level. Analogous works for the bladder have been reported by Heemsbergen et al [62] and Witte et al [63]. Producing voxel-wise statistical models of toxicity might therefore help to discover more about these relationships and would highlight heterogeneous intra-organ sensitivity.

On the other hand, finding a spatial pattern that takes into account all the patients' dose distribution is challenging. The challenge resides in the fact that each patient has a different morphology, making comparisons among patients very difficult. One solution, coming from the image-processing field, is the image registration process. It consists of transforming one image (by applying several types of deformation) to make it consistent or comparable with a reference image. This process is particularly complex because it requires each patient's image to be mapped on the coordinate space of a reference image. One alternative consists of mapping the CT scan images and then propagating the same transformation to the dose distribution images. It is important to note that CT scan images come from physical devices, whereas dose distribution images are generated from a treatment planning system. This thesis will assume that images have already been spatially normalized.

1.5 Objectives of the thesis

In the context of external beam radiotherapy for prostate cancer, the objectives of this thesis are:

- To find predictors of bladder and rectal complications from the analysis of clinical data,
- To apply new methodologies for improving toxicity prediction,
- To develop new statistical approaches that enable the exploitation of the treatment information and patient characteristics.

1.6 Modeling approaches developed in the thesis to predict bladder and rectal complications

Different modeling approaches can be used with the aim to predict rectal and bladder complications. In this thesis we follow a modeling pathway that can be characterized in terms of model inputs and type (Figure 1.9).

Concerning the input information, three different levels can be considered:

1. Patient parameters: it is the information regarding the characteristics of the patient such as age, medical history (history of surgery, diabetes, etc.) and tumor characteristics (T stage, Gleason score, PSA).
2. Treatment parameters: it consists of the different treatment parameters such as the irradiation technique (3D-CRT, IMRT, IGRT), average and maximum doses to different clinical target volumes and the DVH. Indeed, as it was presented above, most classic NTCP models use the DVH to compute a complication probability, and the DVH itself is considered as a reduction of the 3D dose distribution. We found, for example, that variables such as V72 (the volume of the organ receiving at least 72 Gy) or D25 (the dose delivered to at least 25% of the organ volume), can be extracted from the DVH. In this level no spatial relationships are considered.
3. 3D dose distribution: it can be regarded as a functional mapping that assigns to each point of the patient's anatomy a value of dose in Grays.

Concerning the models, we start using classical approaches, such as logistic regression, Kaplain-Meier risk estimation, and Cox proportional hazard regression, that exploit patient parameters and dosimetric parameters. Then, we use a random forest approach (RF) to improve toxicity prediction using both patient parameters and DVH. Then, we considered the 3D dose distribution and we apply two different methods: a tensor population value decomposition (PVD) that enable to look for dose patterns correlated to side effects and a nonparametric mixed-effects (NPME) model that allows for anatomical regions correlated to rectal toxicity to be highlighted. Figure 1.9 summarizes this information.

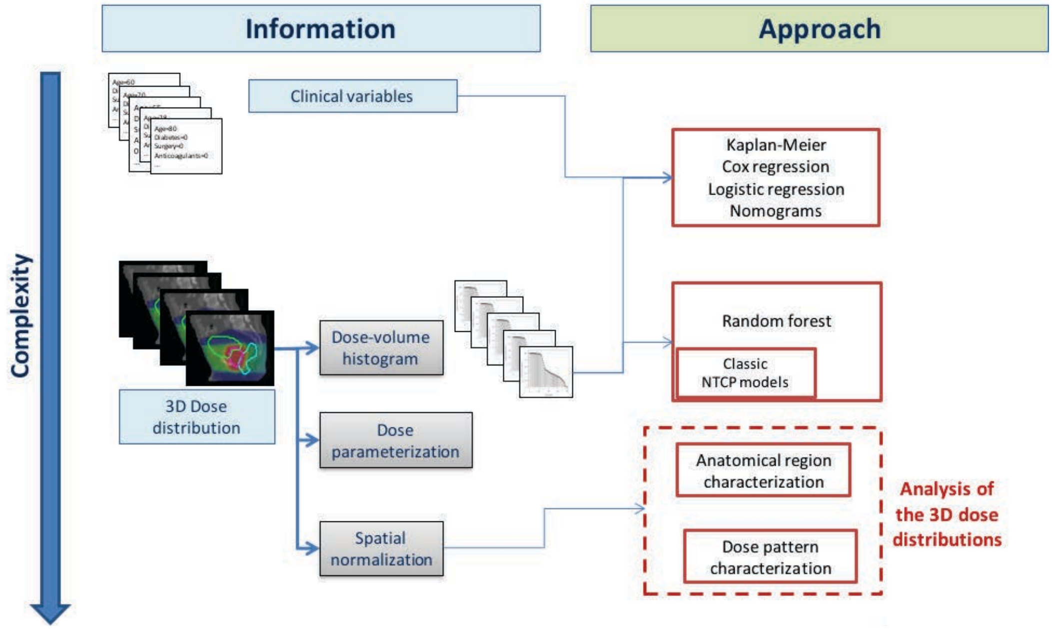


Figure 1.9: Data inputs and models considered in the thesis

Bibliography

[1] INCa, “La situation du cancer en France en 2011. collection rapports et syntheses edite par l’INCa,” INCa, Boulonge-Billancourt, Tech. Rep., October 2011.

[2] M. Pineros Petersen, Ed., *Anuario Estadistico 2009.* Bogota D.C., Colombia: Instituto Nacional de Cancerologia, 2009, vol. 7.

[3] Instituto Nacional de Cancerologia, “Cifras del cancer en Colombia,” Web site http://www.cancer.gov.co/, November 2011.

[4] A. Heidenreich, P. J. Bastian, J. Bellmunt, M. Bolla, S. Joniau, T. van der Kwast, M. Mason, V. Matveev, T. Wiegel, F. Zattoni, and N. J. o. R. O. Mottet, “{EAU} guidelines on prostate cancer. part 1: Screening, diagnosis, and local treatment with curative intentâupdate 2013,” *European Urology*, vol. 65, no. 1, pp. 124 – 137, 2014.

[5] M. Peyromaure, P. Beuzeboc, L. Salomon, P. Richaud, P. Coloby, B. Malavaud, X. Rebillard, P. Rischmann, A. Villers, and M. Soulie, “Le dépistage du cancer de la prostate en 2009: mise au point du comité de cancérologie de l’association française d’urologie,” *Progrès en urologie*, vol. 20, no. 1, pp. 17–23, 2010.

[6] H. Scher, J. Isaacs, and Z. MJ, “Prostate cancer,” *Clinical Oncology. 2nd ed. New York, NY: Churchill Livingstone*, pp. 1823–1884, 2000.

[7] O. H. Behars, D. E. Henson, R. V. P. Hutter, and M. H. Myers, *American joint committee on cancer: Manual for staging of cancer.* Philadelphia: Lippincott, 1992.

[8] L. Lemaître, A. Villers, D. Mouton, and P. Puech, "Echographie et biopsies de prostate," *Journal de Radiologie*, vol. 87, no. 2, pp. 201–209, 2006.

[9] A. W. Partin, L. A. Mangold, D. M. Lamm, P. C. Walsh, J. I. Epstein, and J. D. Pearson, "Contemporary update of prostate cancer staging nomograms (partin tables) for the new millennium," *Urology*, vol. 58, no. 6, pp. 843–848, 2001.

[10] F. Cornud, A. Villers, P. Mongiat-Artus, X. Rebillard, and M. Soulie, "Imagerie par résonance magnétique et cancer de la prostate," *Progrès en urologie*, vol. 18, no. 10, pp. 621–633, 2008.

[11] A. V. D'Amico, R. Whittington, S. B. Malkowicz, D. Schultz, K. Blank, G. A. Broderick, J. E. Tomaszewski, A. A. Renshaw, I. Kaplan, C. J. Beard, and A. Wein, "Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer," *JAMA: The Journal of the American Medical Association*, vol. 280, no. 11, pp. 969–974, 1998.

[12] G. Aus, C. Abbou, M. Bolla, A. Heidenreich, H.-P. Schmid, H. van Poppel, J. Wolff, and F. Zattoni, "EAU guidelines on prostate cancer," *European Urology*, vol. 48, no. 4, pp. 546 – 551, 2005.

[13] A. Magheli and A. L. Burnett, "Erectile dysfunction following prostatectomy: prevention and treatment," *Nat Rev Urol*, vol. 6, no. 8, pp. 415–427, August 2009.

[14] E. H. Lambert, K. Bolte, P. Masson, and A. E. Katz, "Focal cryosurgery: Encouraging health outcomes for unifocal prostate cancer," *Urology*, vol. 69, no. 6, pp. 1117 – 1120, 2007.

[15] C. Hill and G. ter Haar, "High intensity focused ultrasound–potential for cancer treatment," *British Journal of Radiology*, vol. 68, pp. 1296–1303, 1995.

[16] L. Poissonnier, F. Murat, A. Belot, R. Bouvier, M. Rabilloud, O. Rouviere, J. Chapelon, and A. Gelet, "Locally recurrent prostatic adenocarcinoma after exclusive radiotherapy: results of high intensity focused ultrasound." *Progrès en urologie: journal de l'Association française d'urologie et de la Société française d'urologie*, vol. 18, no. 4, p. 223, 2008.

[17] J. Jegu, B. Tretarre, M. Velten, A.-V. Guizard, A. Danzon, A. Buemi, M. Colonna, A.-M. Kadi-Hanifi, O. Ganry, F. Molinie *et al.*, "Le cancer de la prostate en france en 2001: état des pratiques et facteurs associés à la réalisation d'une prostatectomie totale," *Progrès en urologie*, vol. 20, no. 1, pp. 56–64, 2010.

[18] INCa, "Situation de la radiothérapie en 2011," Juin 2012.

[19] B. Hårdemark, A. Liander, H. Rehbinder, and J. Löf, "Direct machine parameter optimization with RayMachine® in Pinnacle3®," *RaySearch White Paper. Stockholm, Sweden: RaySearch Laboratories AB*, 2003.

[20] R. Cheung, S. L. Tucker, L. Dong, and D. Kuban, "Dose-response for biochemical control among high-risk prostate cancer patients after external beam radiotherapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 56, no. 5, pp. 1234–1240, 2003.

[21] R. Cheung, S. L. Tucker, A. K. Lee, R. de Crevoisier, L. Dong, A. Kamat, L. Pisters, and D. Kuban, "Dose–response characteristics of low-and intermediate-risk prostate cancer treated with external beam radiotherapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 61, no. 4, pp. 993–1002, 2005.

[22] C. Fiorino, R. Valdagni, T. Rancati, and G. Sanguineti, "Doseâvolume effects for normal tissues in external radiotherapy: Pelvis," *Radiotherapy and Oncology*, vol. 93, no. 2, pp. 153 – 167, 2009.

[23] W. Schaake, E. M. Wiegman, M. de Groot, H. P. van der Laan, C. P. van der Schans, A. C. van den Bergh, and J. A. Langendijk, "The impact of gastrointestinal and genitourinary toxicity on health related quality of life among irradiated prostate cancer patients," *Radiotherapy and Oncology*, vol. 110, no. 2, pp. 284 – 290, 2014.

[24] L. J. Boersma, M. van den Brink, A. M. Bruce, T. Shouman, L. Gras, A. Velde, and J. V. Lebesque, "Estimation of the incidence of late bladder and rectum complications after high-dose (70â78 gy) conformal radiotherapy for prostate cancer, using doseâvolume histograms," *International Journal of Radiation Oncology*Biology*Physics*, vol. 41, no. 1, pp. 83 – 92, 1998.

[25] H. P. van der Laan, A. van den Bergh, C. Schilstra, R. Vlasman, H. Meertens, and J. A. Langendijk, "Grading-system-dependent volume effects for late radiation-induced rectal toxicity after curative radiotherapy for prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 70, no. 4, pp. 1138 – 1145, 2008.

[26] S. M. Bentzen, L. S. Constine, J. O. Deasy, A. Eisbruch, A. Jackson, L. B. Marks, R. K. Ten Haken, and E. D. Yorke, "Quantitative analyses of normal tissue effects in the clinic (QUANTEC): An introduction to the scientific issues," *International Journal of Radiation Oncology*Biology*Physics*, vol. 76, no. 3, Supplement, pp. S3 – S9, 2010, quantitative Analyses of Normal Tissue Effects in the Clinic.

[27] J. M. Michalski, H. Gay, A. Jackson, S. L. Tucker, and J. O. Deasy, "Radiation dose-volume effects in radiation-induced rectal injury," *International Journal of Radiation Oncology*Biology*Physics*, vol. 76, no. 3, Supplement 1, pp. S123–S129, 2010, quantitative Analyses of Normal Tissue Effects in the Clinic.

[28] A. N. Viswanathan, E. D. Yorke, L. B. Marks, P. J. Eifel, and W. U. Shipley, “Radiation doseâvolume effects of the urinary bladder,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 76, no. 3, Supplement, pp. S116 – S122, 2010, quantitative Analyses of Normal Tissue Effects in the Clinic.

[29] K.-R. Trott, W. Doerr, A. Facoetti, J. Hopewell, J. Langendijk, P. van Luijk, A. Ottolenghi, and V. Smyth, “Biological mechanisms of normal tissue damage: Importance for the design of {NTCP} models,” *Radiotherapy and Oncology*, vol. 105, no. 1, pp. 79 – 85, 2012.

[30] J. T. Lyman, “Complication probability as assessed from dose-volume histograms,” *Radiation Research*, vol. 104, no. 2, pp. S13–S19, Nov. 1985.

[31] P. Rubin, A. Cooper, and T. L. Philips, Eds., ser. Set R.T.I: Radiation Oncology. Chicago: American College of Radiology, 1975.

[32] J. T. Lyman and A. B. Wolbarst, “Optimization of radiation therapy, iii: A method of assessing complication probabilities from dose-volume histograms,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 13, no. 1, pp. 103–109, 1987.

[33] G. J. Kutcher and C. Burman, “Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method gerald,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 16, no. 6, pp. 1623–1630, 1989.

[34] T. E. Schultheiss, C. G. Orton, and R. A. Peck, “Models in radiotherapy: Volume effects,” *Medical Physics*, vol. 10, no. 4, 1983.

[35] A. Dritschilo, J. T. Chaffey, W. D. Bloomer, and A. Marck, “The complication probability factor: a method for selection of radiation treatment plans,” *Br J Radiol*, vol. 51, no. 605, pp. 370–374, 1978.

[36] A. B. Wolbarst, E. S. Sternick, B. H. Curran, and A. Dritschilo, “Optimized radiotherapy treatment planning using the complication probability factor (cpf),” *International Journal of Radiation Oncology*Biology*Physics*, vol. 6, no. 6, pp. 723–728, 1980.

[37] G. J. Kutcher, C. Burman, L. Brewster, M. Goitein, and R. Mohan, “Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 21, pp. 137–146, 1991.

[38] P. Källman, A. Ågren, and A. Brahme, “Tumor and normal tissue responses to fractioned non-uniform dose delivery,” *International Journal of Radiation Biology*, vol. 62, no. 2, pp. 249–262, 1992.

[39] A. Niemierko and M. Goitein, “Modeling of normal tissue response to radiation: the critical volume model,” *International Journal of Radiation Oncology Biology Physics*, vol. 25, pp. 135–145, 1993.

[40] A. Jackson, G. J. Kutcher, and E. Yorke, “Probability of radiation-induced complications for normal tissues with parallel architecture subject to non-uniform irradiation.” *Journal of Medical Physics*, vol. 20, pp. 613–625, 1993.

[41] A. B. Wolbarst, L. Chin, and G. Svensson, “Optimization of radiation therapy: integral-response of a model biological system,” *International Journal of Radiation Oncology Biology Physics*, vol. 8, pp. 1761–1769, 1982.

[42] D. J. Venzon and S. H. Moolgavkar, “A method for computing profile-likelihood-based confidence intervals,” *Applied Statistics*, pp. 87–94, 1988.

[43] M. Söhn, M. Alber, and D. Yan, “Principal component analysis-based pattern analysis of doseâvolume histograms and influence on rectal toxicity,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 69, no. 1, pp. 230 – 239, 2007.

[44] C. Fiorino, C. Cozzarini, V. Vavassori, G. Sanguineti, C. Bianchi, G. M. Cattaneo, F. Foppiano, A. Magli, and A. Piazzolla, “Relationships between {DVHs} and late rectal bleeding after radiotherapy for prostate cancer: analysis of a large group of patients pooled from three institutions,” *Radiotherapy and Oncology*, vol. 64, no. 1, pp. 1 – 12, 2002.

[45] C. Fiorino, V. Vavassori, G. Sanguineti, C. Bianchi, G. M. Cattaneo, A. Piazzolla, and C. Cozzarini, “Rectum contouring variability in patients treated for prostate cancer: impact on rectum doseâvolume histograms and normal tissue complication probability,” *Radiotherapy and Oncology*, vol. 63, no. 3, pp. 249 – 255, 2002.

[46] S. Marzi, G. Arcangeli, B. Saracino, M. G. Petrongari, V. Bruzzaniti, G. Iaccarino, V. Landoni, A. Soriani, and M. Benassi, “Relationships between rectal wall doseâvolume constraints and radiobiologic indices of toxicity for patients with prostate cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 68, no. 1, pp. 41 – 49, 2007.

[47] V. A. Benk, J. A. Adams, W. U. Shipley, M. M. Urie, P. L. McManus, J. T. Efird, C. G. Willett, and M. Goitein, “Late rectal bleeding following combined X-ray and proton high dose irradiation for patients with stages T3-T4 prostate carcinoma,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 26, no. 3, pp. 551 – 557, 1993.

[48] T. Rancati, C. Fiorino, G. Gagliardi, G. M. Cattaneo, G. Sanguineti, V. Casanova Borca, C. Cozzarini, G. Fellin, F. Foppiano, G. Girelli, L. Menegotti, A. Piazzolla, V. Vavassori, and R. Valdagni, “Fitting late rectal bleeding data using different NTCP models: results from an italian multi-centric study (airopros0101),” *Radiotherapy and Oncology*, vol. 73, no. 1, pp. 21 – 32, 2004.

[49] C. Kupchak, J. Battista, and J. Van Dyk, “Experience-driven dose-volume histogram maps of NTCP risk as an aid for radiation treatment plan selection and optimization,” *Medical Physics*, vol. 35, no. 1, pp. 333–343, 2008.

[50] J. A. Hatton, P. B. Greer, C. Tang, P. Wright, A. Capp, S. Gupta, J. Parker, C. Wratten, and J. W. Denham, “Does the planning doseâvolume histogram represent treatment doses in image-guided prostate radiation therapy? assessment with cone-beam computerised tomography scans,” *Radiotherapy and Oncology*, vol. 98, no. 2, pp. 162 – 168, 2011.

[51] G. Defraene, L. Van den Bergh, A. Al-Mamgani, K. Haustermans, W. Heemsbergen, F. Van den Heuvel, and J. V. Lebesque, “The benefits of including clinical factors in rectal normal tissue complication probability modeling after radiotherapy for prostate cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 82, no. 3, pp. 1233 – 1242, 2012.

[52] T. Rancati, C. Fiorino, G. Fellin, V. Vavassori, E. Cagna, V. Casanova Borca, G. Girelli, L. Menegotti, A. F. Monti, F. Tortoreto, S. Delle Canne, and R. Valdagni, “Inclusion of clinical risk factors into {NTCP} modelling of late rectal toxicity after high dose radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, vol. 100, no. 1, pp. 124 – 130, 2011.

[53] L. A. Dawson, M. Biersack, G. Lockwood, A. Eisbruch, T. S. Lawrence, and R. K. Ten Haken, “Use of principal component analysis to evaluate the partial organ tolerance of normal tissues to radiation,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 62, no. 3, pp. 829 – 837, 2005.

[54] S. L. Gulliford, S. Webb, C. G. Rowbottom, D. W. Corne, and D. P. Dearnaley, “Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate,” *Radiotherapy and Oncology*, vol. 71, no. 1, pp. 3 – 12, 2004.

[55] S. Tomatis, T. Rancati, C. Fiorino, V. Vavassori, G. Fellin, E. Cagna, F. Mauro, G. Girelli, A. Monti, M. Baccolini, G. Naldi, C. Bianchi, L. Menegotti, M. Pasquino, M. Stasi, and R. Valdagni, “Late rectal bleeding after 3d-crt for prostate cancer: development of a neural-network-based predictive model,” *Physics in Medicine and Biology*, vol. 57, no. 5, p. 1399, 2012.

[56] R. Valdagni, T. Rancati, C. Fiorino, G. Fellin, A. Magli, M. Baccolini, C. Bianchi, E. Cagna, C. Greco, F. A. Mauro, A. F. Monti, F. Munoz, M. Stasi, P. Franzone, and V. Vavassori, “Development of a set of nomograms to predict acute lower gastrointestinal toxicity for prostate cancer 3d-crt,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 71, no. 4, pp. 1065 – 1073, 2008.

[57] S. L. Tucker, L. Dong, W. R. Bosch, J. Michalski, K. Winter, R. Mohan, J. A. Purdy, D. Kuban, A. K. Lee, M. R. Cheung, H. D. Thames, and J. D. Cox, “Late rectal toxicity on {RTOG} 94-06: Analysis using a mixture lyman model,”

*International Journal of Radiation Oncology*Biology*Physics*, vol. 78, no. 4, pp. 1253 – 1260, 2010.

[58] C. S. Mayo, C. Zankowski, M. Herman, R. Miller, K. Olivier, A. Vincent, and J. Suominen, “A method to vectorize the dose distribution, the dose volume histogram and create a dose vector histogram,” *Medical Physics*, vol. 40, no. 1, pp. –, 2013.

[59] S. L. Tucker, M. Zhang, L. Dong, R. Mohan, D. Kuban, and H. D. Thames, “Cluster model analysis of late rectal bleeding after {IMRT} of prostate cancer: A case-control study,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 64, no. 4, pp. 1255 – 1264, 2006.

[60] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, vol. 54, no. 21, p. 6535, 2009.

[61] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Modeling late rectal toxicities based on a parameterized representation of the 3d dose distribution,” *Physics in Medicine and Biology*, vol. 56, no. 7, p. 2103, 2011.

[62] W. D. Heemsbergen, A. Al-Mamgani, M. G. Witte, M. Van Herk, F. J. Pos, and J. V. Lebesque, “Urinary obstruction in prostate cancer patients from the dutch trial (68 gy vs. 78 gy): Relationships with local dose, acute effects, and baseline characteristics,” *International Journal of Radiation Oncology* Biology* Physics*, vol. 78, no. 1, pp. 19–25, 2010.

[63] M. G. Witte, W. D. Heemsbergen, R. Bohoslavsky, F. J. Pos, A. Al-Mamgani, J. V. Lebesque, and M. van Herk, “Relating dose outside the prostate with freedom from failure in the Dutch trial 68 gy vs. 78 gy,” *International Journal of Radiation Oncology* Biology* Physics*, vol. 77, no. 1, pp. 131–138, 2010.

Part II

Models based on patient parameters, dosimetry parameters and dose-volume histograms

Chapter 2

Classical modeling of bladder and rectal toxicity

In this chapter, a classical statistical modeling approach is used to characterize bladder and rectal toxicity. Using Kaplan-Meier estimation, the cumulative risks of bladder and rectal toxicity (and the associated symptoms) are described. With logistic regression we identify patient and treatment parameters impacting the risks of bladder and rectal toxicity. The main objective is to propose predictive models to help clinicians to decide about a treatment and to inform patients about treatment risks. The contribution of this chapter is the description of the main toxicities following prostate cancer radiotherapy rather than the developing new methodologies for toxicity prediction. From the point of view of the input data, only patient parameters and general treatment parameters were used. Dose-volume histogram (DVH) and dose distribution were not taken into account. The chapter is divided in two sections. Section 2.1 focuses on bladder toxicity and its associated symptoms whereas Section 2.2 is the rectal counterpart.

2.1 Nomograms to predict late urinary toxicity following prostate cancer radiotherapy

The content of this section is the result of a collaboration with doctor Romain Mathieu and was presented as part of his thesis in medicine (*thése d'exercice*) at Université de Rennes 1 [1] and published in the World Journal of Urology [2].

2.1.1 Introduction

Radical prostatectomy (RP) and radiotherapy (RT) are cornerstones of localized prostate cancer treatment, leading to relatively similar results in terms of local control [3]. However, the side effects of both are different, mainly concerning the urinary adverse events. If the intensity of this toxicity after RT is relatively well-reported in the literature, the description of the symptoms corresponding to this toxicity is often limited. Moreover, the patient and/or treatment factors related to each of the side effects are not well known. Their identification is crucial. These factors could be used to generate urinary toxicity predictive tools (like nomograms), to guide the physician in deciding the treatment and to inform the patient, in this context of different therapeutic alternatives. To identify which radiation parameters increase toxicity is essential in understand how to decrease toxicity, in particular due to new highly-conformal radiotherapy techniques, such as intensity-modulated radiation therapy IMRT and image-guided radiation therapy (IGRT). These new techniques allow for an increase of the dose in the prostate considering the strong dose-effect relationship for local control [4], while limiting the dose in the bladder and the rectum. Their part in decreasing urinary toxicity has not, however, been clearly shown. If the relationship between dose-volume and toxicity has been consistently demonstrated for the rectum, it remains unclear for the bladder [5, 6]. Moreover, hypofractionated RT may be more efficient in eradicating prostate adenocarcinoma, but the impact of such modified fractionation on toxicity has not been well-established. If nomograms have already been published to predict rectal bleeding and the risk of fecal incontinence [7, 8], to our knowledge no tools have as yet been developed to predict urinary toxicity. We thus analyzed a large group of patients having received prostate cancer RT to accurately quantify and describe late urinary toxicity, identify related risk factors and propose predictive nomograms.

2.1.2 Material and methods

2.1.2.1 Patient inclusion criteria

Records from 965 patients who received definitive radiotherapy for localized prostate adenocarcinoma were analyzed. Data were prospectively collected from 470 patients treated in 17 French institutions within two randomized studies: GETUG06 (comparing 70 Gy to 80 Gy) [9] and STIC-IGRT (testing IGRT) [10], and retrospectively from 623 patients treated in two of them. All patients had a biopsy-proven adenocarcinoma of the prostate. Pretreatment workup included a CT scan and a bone scan.

2.1.2.2 Patient and tumor characteristics

The following data were extracted from each randomized database: age, medical and surgical history (prior abdominal surgery, prior transurethral resection of prostate, anticoagulant treatment, diabetes, HTA, coronary insufficiency) and tumor characteristics (T stage, Gleason score, pretreatment PSA). Patients were staged by digital rectal examination according to the 1992 American Joint Committee on Cancer staging system [11]. Patients were classified into the three prognostic risk groups defined by D'Amico [3]. Patient and tumor characteristics of the prospective cohort are presented in Table 2.1.

2.1.2.3 Treatment characteristics

The target volume comprised the prostate only in the low risk group (16%), the prostate and the seminal vesicles in the other risk groups. The pelvic lymph nodes were not treated in the two randomized studies but may have been treated for high risk patients of the two institutions (not treated in the randomized study) (9%). Total median total dose to the prostate was 70 Gy (ranging from 65 Gy to 80 Gy), the seminal vesicles receiving 46 Gy, and the pelvic lymph nodes also 46 Gy, if treated. Dose per fraction was 2 Gy/day, 5 fractions/week for 69% of patients, or 2.5 Gy/day, 4 fractions/week for 31% of patients.

The radiation technique was 3D conformational for the vast majority of patients (85%) and 2D for 15% of the patients, depending on the treatment period. IMRT and IGRT have been more recently used in patients receiving 80 Gy. Among the 41% of the patients having received 80 Gy, the technique was a standard 3D conformal RT for 63%, IMRT only for 18% and IMRT combined with IGRT in 19%.

The 3D radiation technique was carried out following the French GETUG group recommendations, as previously reported [9]. Patients underwent simulation and treatment in the supine position. Target volume and organs at risk (bladder, rectum and femoral heads) were delineated on CT slices. A bladder wall was generated with a thickness of 7 mm from the external manually-delineated bladder contour. The planning target volume (PTV) was calculated by adding a 10-mm margin in all directions except in the posterior where a 5-mm margin was considered. The dose-volume histogram had to respect the GETUG constraints [9]. The following bladder dosimetric data were analyzed: volume of the bladder wall, Dmax (maximal dose received in the bladder), D25 (minimal dose received in 25% of the bladder wall) and D50 (minimal dose received in 50% of the bladder wall).

Androgen deprivation therapy was given to 23% of the patients, all presenting high-risk cancer. Details of treatment characteristics in the prospective are presented in Table 2.1.

2.1.2.4 Follow up and toxicity grading

All patients were evaluated every three months in the first year and every six months thereafter. Late urinary toxicity was defined as events occurring more than six months after the beginning of RT. To determine the severity and incidence of main late urinary complaints, records were prospectively extracted from standardized questionnaires or retrospectively from physicians' reports, taken at each follow-up visit. Urinary complaints were classified according to the LENT/SOMA morbidity scoring system into four categories of symptoms: urinary frequency, dysuria, incontinence and hematuria (Table 2.2). The analyses were performed for late urinary toxicity and for each of the symptoms, all being considered if $\geq$ grade 2.

2.1.2.5 Statistical analysis

Non-parametric tests were used to compare the distribution of the parameters between different groups of treatment. The Kaplan-Meier method was used to calculate cumulative incidences of late urinary toxicity events ($\geq$ grade 2) and corresponding symptoms. The differences between the survival curves were assessed using the log-rank test.

Patient characteristics		*Prescribed dose and fractionation*	
Number of patients	965	65 Gy	
Mean age, year (range)	68 (45-83)	2.5 Gy. 4/week	15%
Diabetes^a	7%	70 Gy	
Anticoagulant treatment^b	21%	2.5 Gy. 4/week	16%
Prior abdominal or pelvic surgery	34%	2 Gy. 5/week	28%
Prior transuretral resection of prostate	6%	80 Gy	
Hypertension	19%	2 Gy. 4/week	41%
Coronary insufficiency	9%	*Target volume*	
Tumor characteristics		Prostate only	16%
PSA, ng/ml (range)	15 (0-133)	Prostate + SV	75%
Gleason Score		Prostate + SV + PLN	9%
<7	53%	*Dosimetric parameters*	
7	38%	Bladder (wall)^c	
>7	9%	Volume (cc) ± SD	70.7 ± 39.5
T stage		Dmax Gy ± SD	75.8 ± 4.7
T1	25%	D25 (Gy) ± SD	64.8 ± 11.6
T2	62%	D50 (Gy) ± SD	43.1 ± 15.2
T3	13%	*Target*	
Prognostic group (D'Amico)		PTV prostate and SV (cc) ± SD	234.5 ± 60.9
Low	18%	PTV prostate only (cc) ± SD	174.2 ± 55.6
Intermediate	51%	V95 (%)	93.1 ± 10.8
High	31%	Androgen deprivation	23%
Treatment characteristics			
Radiotherapy technique			
2D Technique	15%		
"Standard" 3D conformational (without IMRT)	66%		
IMRT (without IGRT)	7%		
IGRT (with IMRT)	12%		

Table 2.1: SD: standard deviation; D$_{max}$: maximum dose; D25: minimal dose received in 25% of the bladder wall volume; D50: minimal dose received in 50% of the bladder wall volume; PTV: planning target volume; V95: volume in the prostate PTV (in %) receiving 95% of the prescribed dose; PLN: pelvic lymph nodes

[a] Type 1 or 2

[b] Vitamin K antagonist and antiplatelet drugs

[c] Thickness of the bladder wall = 7 mm

	Grade 1	Grade 2	Grade 3	Grade 4
Subjective				
Dysuria	Occasional and minimal	Intermittent and tolerable	Persistent and intense	Refractory and excruciating
Decrease stream	Occasionally weak	Intermittent	Persistent bu incomplete obstruction	Complete obstruction
Frequency	3-4-h intervals (6-8/day)	2-3-h intervals (9-12/day)	1-2-h intervals (13-24/day	Hourly (>24/day)
Hematuria	Occasional	Intermittent	Persistent with clot	Refractory
Incontinence	<Weekly episodes	>Daily episodes	Pads/undergarments /day	Refractory
Management				
Dysuria and decreased stream	Occasional, non-narcotic	Regular non-narcotic >1/day self-cahteterization	Regular narcotic Dilatation or TUR >1/day self-catheterization	Permanent catheter, surgical intervention
Frequency		Occasional antispasmodic	Regular narcotic	Cystectomy
Hematuria	Alkalization iron therapy	Single transfusion or cauterization	Frequent transfusions or coagulations	Surgical intervention
Incontinence	Occasional use of incontinence pads	Intermittent use of incontinence pads	Regular use of incontinence pads or self	Catheterization permanent catheter

Table 2.2: LENT-SOMA grading scale (urinary symptoms)

The impact of the following parameters on late urinary toxicity ($\geq$ grade 2) was tested at the 5-years mark in the prospective cohort:

- Patient parameters: age, diabetes (types 1 and 2), anticoagulant treatment (vitamin K antagonist or antiplatelet drug), prior abdominal or pelvic surgery, prior transurethral resection of prostate, hypertension, coronary insufficiency;
- Tumor parameters: Gleason Score, T stage, prognostic group (D'amico);
- Treatment parameters: RT technique (2D technique, 3DConformational technique, with or without IMRT/IGRT), total dose and fractionation, target volume, dosimetric bladder parameters (volume of the bladder wall, maximal dose (Dmax), D25, D50) and androgen deprivation.

There were no statistically significant differences in the toxicity risk between prospective and retrospective patients. Relationships between late urinary toxicity and patient, tumor or treatment parameters were first analyzed using Cox proportional hazard regression at univariate level. Multivariate analyses, including covariates statistically significant in univariate analysis, were carried out using the Cox proportional hazards model. The 5-year late urinary toxicity events were analyzed using logistic regression at univariate and multivariate levels. A p-value ≤ 0.05 was considered statistically significant. Nomograms to predict 5-year late urinary toxicity were built up according to the logistic model. To assess nomogram performance, a nonparametric fit of the predicted probability as regards the actual observed probability was made for each nomogram. The analyses were performed using the SPSS V18 (Chicago, IL) and R with the *rms* package [12].

2.1.3 Results

The median follow-up was 61 months (range 6-140).

2.1.3.1 Late urinary toxicity: global quantification and symptom description

Among the 965 patients, 183 events of late urinary toxicity grade 2 or greater were reported. Among them, only 14% were toxicity grade 3 or 4. Ninety-two (50%) corresponded to an increase in urinary frequency, 36 (20%) to dysuria, and 36 (26%) to hematuria. Only seven consisted of urinary incontinence grade 2 or greater. The 5-year an 10-year rates of grade 2 or higher urinary toxicity, urinary frequency, hematuria, dysuria and urinary incontinence were 15% (95% CI 12-18%) and 24% (95% CI 19-29%), 10% (95% CI 12-18%) and 15% (95% CI 11-19%), 5% (95% CI 4-6%) and 8% (95% CI 5-11%), 3% (95% CI 2-4%) and 8% (95% CI 4-12%), and 1% (95% CI 0-2%) and 2% (95% CI 0-4%), respectively. Figure 2.1 presents cumulative incidence of global late urinary toxicity and the corresponding symptoms ($\geq$grade 2). The 5- and 10-year rates of grade 3 or higher global urinary toxicity were 3% (95% CI 2-4%) and 7% (95% CI 5-9%).

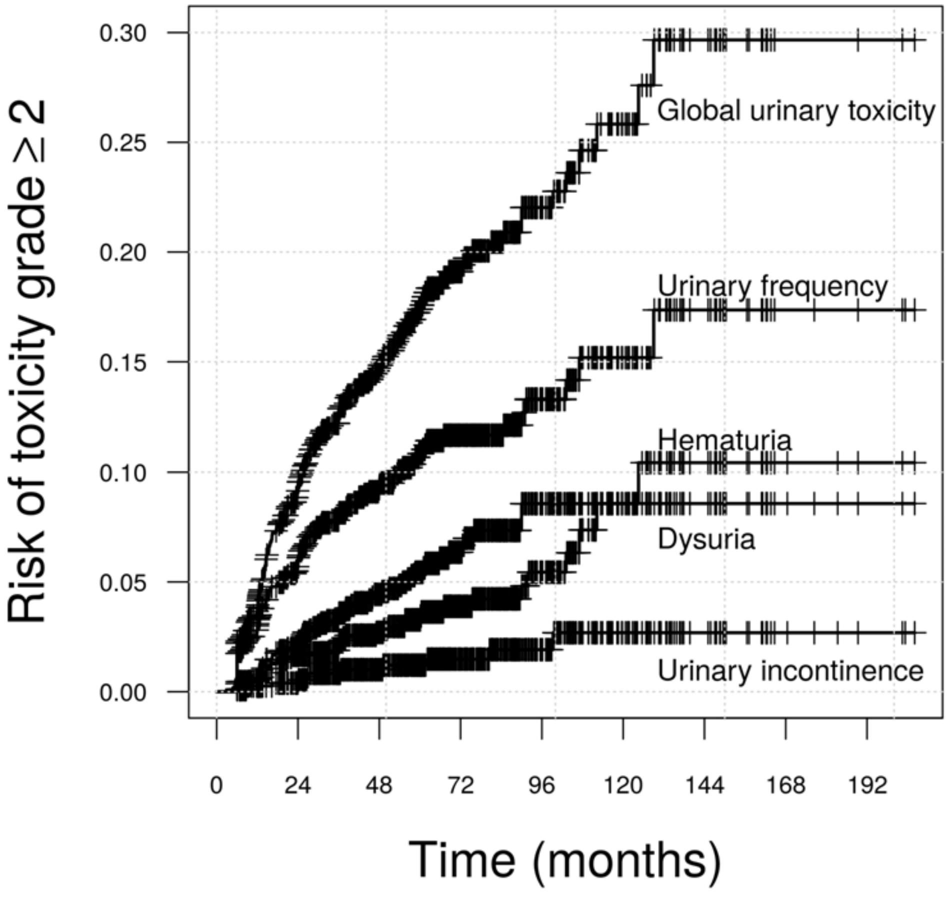


Figure 2.1: Incidence of global and by symptoms late urinary toxicity (≥ grade 2) according to LENT/SOMA classification

Factors	Late urinary toxicity		Urinary frequency	
	RR (95% CI)	p value	RR (95% CI)	p value
Anticoagulant treatment	2.35 (1.33-4.14)	<0.01	-	-
Total dose	1.09 (1.05-1.14)	<0.01	1.07 (1.02-1.13)	0.01
Diabetes	-	-	4.00 (1.42-11.27)	0.01
D25	1.03 (1.00-1.06)	0.04	-	-
Dmax	1.10 (1.04-1.17)	<0.01	-	-
Age	1.06 (1.01-1.11)	0.02	-	-
	Hematuria		Dysuria	
	RR (95% CI)	p value	RR (95% CI)	p value
Anticoagulant treatment	2.89 (1.29-6.46)	0.01	-	-
Total dose	-	-	1.10 (1.02-1.17)	0.01
Diabetes	-	-	-	-
D25	-	-	-	-
Dmax	-	-	-	-
Age	-	-	-	-

Table 2.3: Patient, tumor and treatment factors related to 5-year risk of late urinary toxicity and corresponding symptoms (multivariate regression logistic analysis)

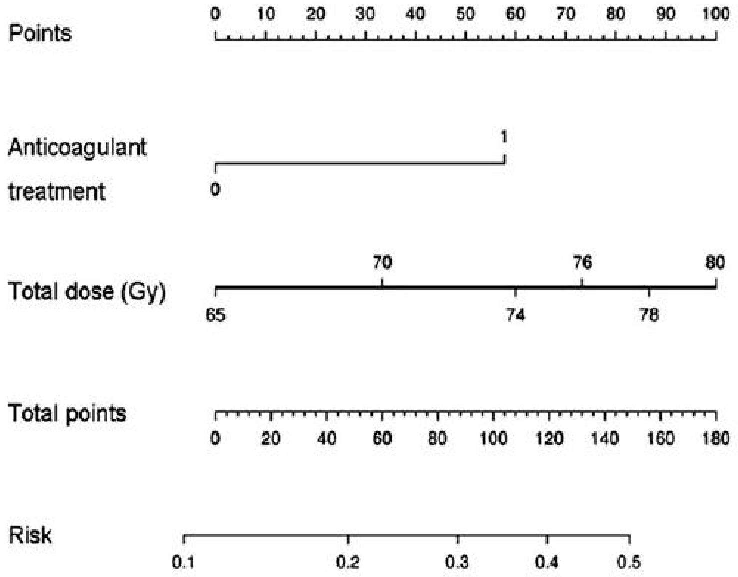
2.1.3.2 Nomogram to predict five-year late urinary toxicity

In multivariate analysis, the following preplanning parameters were found to be significantly and positively associated with the 5-year risk of urinary toxicity: anticoagulant treatment (RR=2.35), total dose (RR=1.09), age (RR=1.06), D25 (RR=1.03), and Dmax (RR=1.1) received by the bladder (Table 2.3). Nomogram including pretreatment factors to predict 5-year risk of global late urinary toxicity (and its calibration plot) is presented in Figure 2.2.

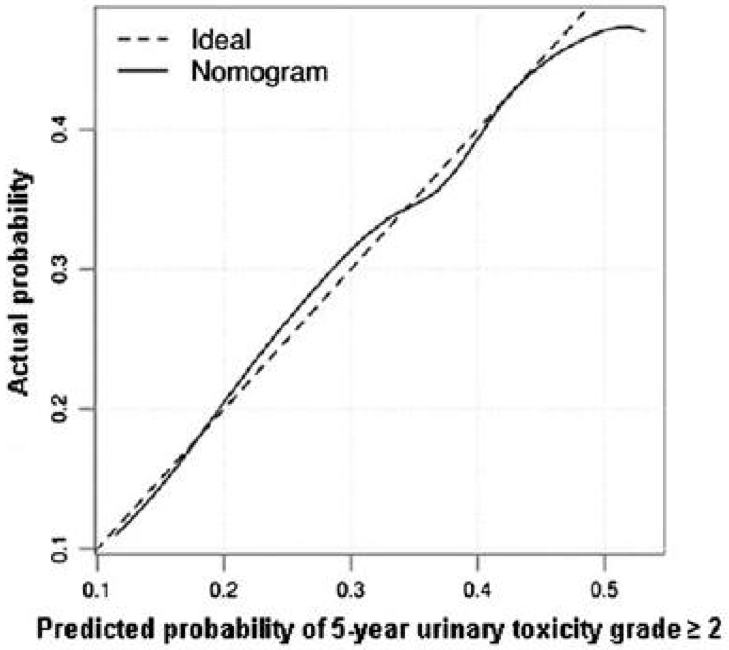
The 5-year risk of urinary frequency was related to total dose (RR=1.07) and diabetes (RR=4). For dysuria, the total dose was the only significant factor (RR=1.1) (Table 2.3). Figures 2.3 and 2.4 present nomograms to predict the 5-year risk of these urinary symptoms. The 5-year risk of hematuria was significantly increased by anticoagulant treatment (RR=2.9).

2.1.4 Discussion

We showed that the incidence of late urinary toxicity symptoms continuously increases after RT, reaching a rate of 24% and 7% at 10 years, for more than grade 2 and grade 3 urinary toxicity, respectively. These rates appear relatively similar to those previously observed after RT [9]. Urinary toxicity events may occur late after RT, later than those observed for late gastro-intestinal toxicity which generally reaches a plateau at three years after RT [13]. A longer follow-up is consequently required to properly estimate late urinary toxicity [13]. Comparing the risk of urinary toxicity after different treatments

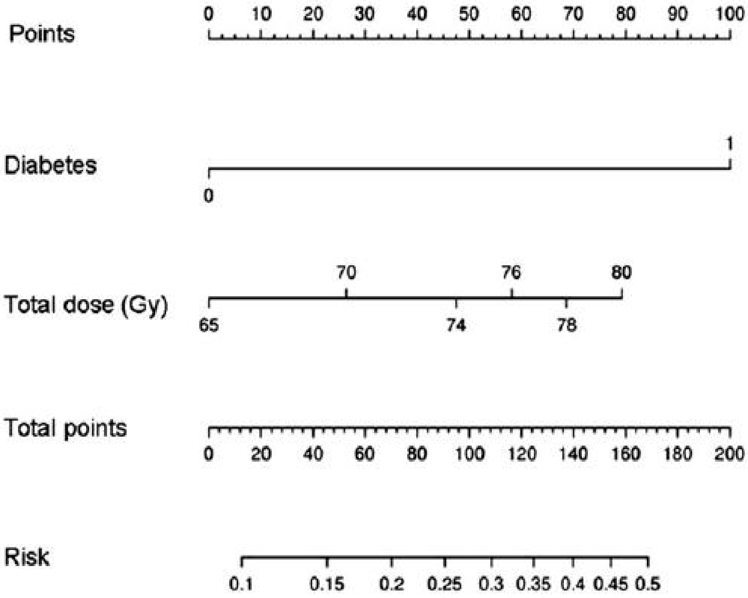


(a) Nomogram

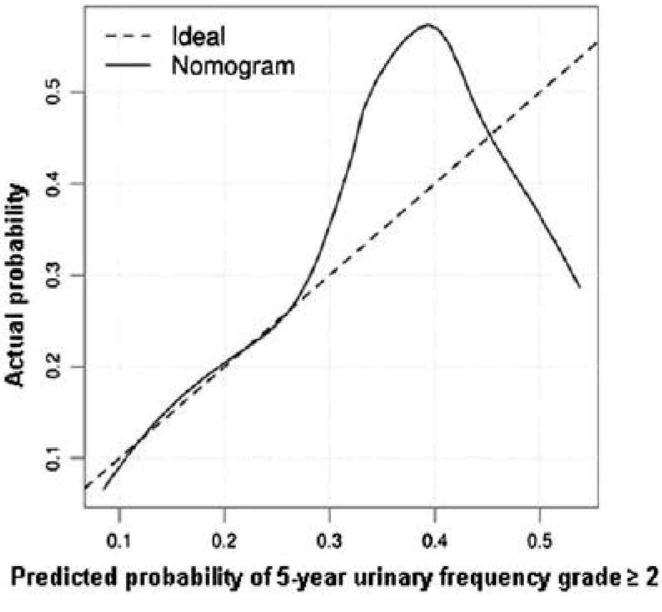


(b) Calibration plot

Figure 2.2: Five-year risk of global late urinary toxicity grade ≥ 2: nomogram and calibration plot. Calibration plot assessing the nomogram performance by a nonparametric fit of the predicted probability versus the actual observed probability

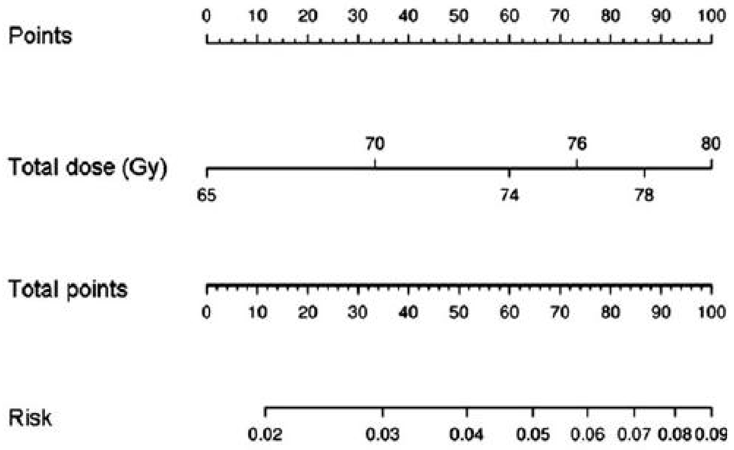


(a) Nomogram

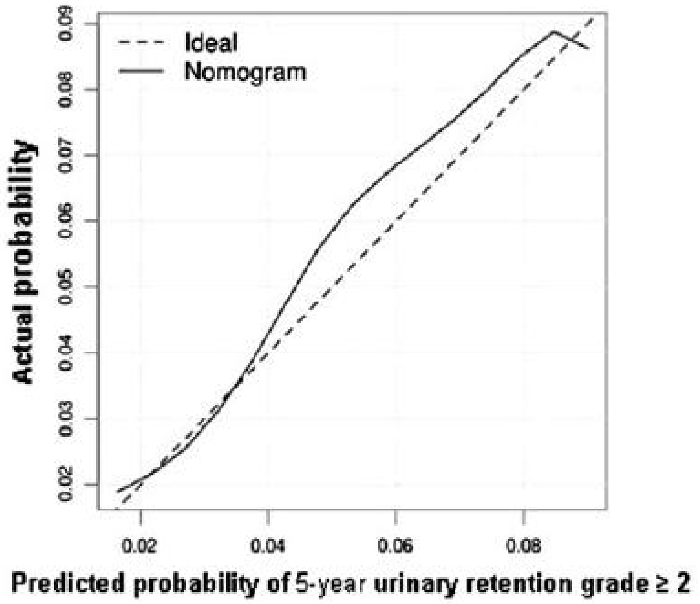


(b) Calibration plot

Figure 2.3: Five-year risk of urinary frequency grade ≥ 2: nomogram and calibration plot. Calibration plot assessing the nomogram performance by a nonparametric fit of the predicted probability versus the actual observed probability



(a) Nomogram



(b) Calibration plot

Figure 2.4: Five-year risk of dysuria grade ≥2: nomogram and calibration plot. Calibration plot assessing the nomogram performance by a nonparametric fit of the predicted probability versus the actual observed probability

should therefore carefully consider the same follow-up. Late urinary toxicity symptoms are mainly characterized by urinary frequency (50% of all events) and, to a lesser extent, by urinary retention (dysuria) and hematuria. Incontinence is very rare (1% at 5 years). We identified dose escalation and hormonal treatment as factors increasing the risk of late urinary toxicity by multivariate analysis. Both parameters represent cornerstones of intermediate and high risk prostate cancer treatment [14]. In intermediate risk prostate cancer, association of androgen deprivation to radiotherapy can be discussed according the total dose. Thus, any predictive model addressing toxicity according these factors could represent a user-friendly tool for treatment decision. Androgen deprivation has been previously associated with late toxicity [15, 16]. The reason would be a decrease in prostate volume, in the meantime between planning and irradiation, responsible for discrepancy between the dose was planned and delivered. Another explication would be biologic transformation through vascular or repair alteration [17]. However, its implication in late urinary toxicity is still controversial [17, 18]. We found a strong dose-effect relationship in urinary toxicity. Most of the randomized studies comparing a "standard" dose (68 to 70 Gy) to a higher one (76 to 80 Gy) did not demonstrate such a significant increase in late urinary toxicity [19, 20, 21, 22]. However, in a large group of patients who received a dose escalation with 3D-CRT/IMRT, Zelefsky et al reported a significant increase of genitourinary (GU) toxicities after 10 years in patients who received higher doses (20). More recently, the GETUG 06 randomized trial reported such an increase, when escalating the dose from 70 Gy to 80 Gy [9]. These differences can be explained by several reasons. The first one is the lack of follow-up. Indeed, the initial MDACC report that compared 78Gy to 70Gy did not show a significant difference in late GI toxicity, while it was found with a longer follow-up [19]. Secondly, studies might be different in terms of their treatment scheme (dose, target volume, technique), patient characteristics and grading scale. Finally, urinary toxicity might be more related to patient risk factors than dose parameters. The lack of correlation between dose distribution (dose-volume histogram) and urinary toxicity may also be due to the large bladder volume variation occurring at the planning stage and at the different fractions, so that the planned dose distribution does not represent the actual delivered dose to the bladder. IMRT and IGRT aim at increasing local control by allowing dose escalation, while reducing toxicity by sparing normal surrounding tissues. Compared to "standard" 3D conformal technique, IMRT clearly reduces the risk of long-term rectal toxicity and bowel dysfunction [23, 24, 25]. However, IMRT fails to decrease late urinary toxicity in most studies, as in the present one. Late urinary toxicity could even be partly due to prostatic urethra lesions [26], and modern techniques such as IMRT combined with IGRT still do not allow for the preservation of this area. A recent non-randomized study of Zelefsky et al reported however that, with a median follow-up of 2.8 years and a high dose (86.4 Gy) delivered to the prostate by IMRT, patients treated with IGRT (using fiducials) experienced significantly less urinary toxicity than non-IGRT treated patients [27]. One limitation of our study is the difficulty of validation of our nomogram with the retrospective cohort along all ranges. Complication grading evaluation was based on physicians report in the retrospective cohort. Such evaluation could have underestimated the significance of urinary toxicity. it could partly explain discrepancies between calibration plot with both cohorts. Thus, we believe performance of this nomogram should be also validated using external data from prospective studies.

2.1.5 Conclusions

We were able to identify several parameters increasing the risk of urinary toxicity after prostate cancer radiotherapy. The first nomogram to predict global late urinary toxicity was generated, resulting in new tool for patient management and treatment decision, particularly between RT and surgery.

2.2 Nomograms to predict late rectal toxicity following prostate cancer radiotherapy

The contents of this section is the result of a collaboration with doctor Jean-Bernard Delobel and was presented as part of his thesis in medicine (*thése d'exercice*) at Université de Rennes 1 [28] and it has been submitted for publication. With respect to Section 2.1, the patient cohort was enlarged and only patients treated by 3D conformal radiotherapy were considered for the analysis.

2.2.1 Introduction

The main limiting side-effects for prostate cancer radiotherapy (RT) escalation are rectal and bladder toxicities, with a negative impact on patient quality of life [29]. The symptoms of late radio-induced rectal toxicity include rectal bleeding. diarrhea, mucoid discharge, urgency, tenesmus, rectal pain and fecal incontinence. The incidence and severity of rectal toxicity after prostate cancer RT vary considerably among different studies, mainly depending on patient and radiation characteristics, such as dose and irradiated rectal volume. Yet the literature rarely provides precise descriptions of the rectal toxicity symptoms. Moreover, the use of a reliable rectal toxicity predictive tool appears crucial for both informing the patient and guiding physicians to decide upon a particular prostate cancer treatment in a context where equally efficient, alternative therapies such as radical prostatectomy are available. The early identification of the patients at higher risk of complications could also be useful to propose preventive treatments.

The clear correlation between dose and tumor control justifies dose-escalation strategies (> 76 Gy) in prostate cancer, which also inevitably leads to increase the doses delivered to normal surrounding tissues [30, 16]. In order to reduce these doses, new irradiation techniques such as intensity-modulated radiation therapy (IMRT) and image-guided radiation therapy (IGRT), have been developed. By modulating the intensity of photons within the treatment beam, IMRT enables a highly conformal dose to be delivered to the area within the concave shape of the prostate target, while limiting the dose to the convex shape of the rectum [31]. This is expected to decrease rectal toxicity. However, prostate position may significantly change (by up to 2cm) within the pelvis between fractions, due in part to rectal filling variations. This type of non-corrected movements can therefore increase the risk of under-dosing the prostate and over-dosing the rectum, exposing the patient to increased risks of both tumor recurrence and rectal toxicity. By means of either intra-prostatic fiducial markers or on-board CT imaging (CBCT), IGRT aims to detect and correct prostate displacements prior to each treatment fraction. The benefits of employing these novel combined RT techniques in reducing rectal toxicity risk have yet, however, to be clearly assessed. This study was conducted involving a large-scale population of patients treated with RT for prostate cancer, with the following objectives:

- to quantify the incidence and severity of acute and late radio-induced rectal toxicity,

- to identify predictors of late radio-induced rectal toxicity, incorporating both patient and treatment characteristics, such as novel RT techniques (IMRT and IGRT) and dose escalation, namely increasing the total dose (up to 80 Gy) and dose per fraction ($>$ 2Gy, in hypofractionated RT schedule),
- to build up nomograms to assess the risk of late radio-induced rectal toxicity for a given patient in routine practice.

2.2.2 Materials and methods

2.2.2.1 Patient and tumor characteristics

A total of 972 patients receiving definitive radiotherapy for localized prostate adenocarcinoma were included in this study. Data was prospectively compiled for 487 patients (50%) treated in 17 French institutions between 2000 and 2012 in two randomized trials, namely GETUG 06 (comparing 70Gy with 80Gy) [9] and STIC-IGRT (testing IGRT) [10, 32], and retrospectively recorded for 485 patients treated in two of these institutions (used only to assess the impact of dose fractionation). All patients exhibited a biopsy-proven prostate adenocarcinoma. Pretreatment evaluation included clinical history, physical examination, laboratory studies, CT scan, and bone scan. Patients were classified into prognostic risk groups according to pretreatment PSA levels, clinical stage (T1-4), and Gleason score, as described by D'Amico [3]. Patient and tumor characteristics are provided in Table 2.4.

2.2.2.2 Radiotherapy description

All patients received 3D conformational RT (3DCRT), carried out in following the French GETUG group recommendations [33]. Each patient underwent simulation and treatment in supine position. Target volume and organs at risk, namely the bladder, rectum, and femoral heads, were manually delineated on 3 mm to 5 mm thickness CT slices. The planning target volume (PTV) was calculated as macroscopic tumor volume + margins to take into account the risk of microscopic involvement, patient set-up error, and anatomical variations. It included the prostate + seminal vesicles, with a 10 mm additional margin in each space direction, except in the posterior where it was reduced to 5 mm so as to spare the rectum. The seminal vesicles were spared in cases of low-risk tumor. The pelvic lymph nodes were not treated. The rectum was manually delineated from 2 cm above to 2 cm below the prostate and the seminal vesicles. The rectal wall was generated with a thickness of 5 mm from the external manually-delineated rectal contour.

The mean total dose delivered to the prostate was 74 Gy, ranging from 70 to 80 Gy, compared to a 46 Gy mean dose delivered to the seminal vesicles. A standard fractionation scheme was followed, administering 2 Gy per day, five fractions per week, to 71% of patients. In order to assess the impact of a moderate hypofractionated scheme, we also included a series of 401 patients treated in a single institution, all having received a total dose of 70 Gy over 7 weeks, either at 2 Gy per fraction with five fractions per week (standard fractionation) or at 2.5 Gy per fraction with four fractions per week (hypofractionation). IMRT and IGRT were given to 20% and 14% of all patients, respectively. IGRT was always used in combination with IMRT, and given to patients having received 80 Gy to the prostate. For this patient subgroup, the choice of the RT technique could be classified into three increasing complexity levels: standard 3DCRT for 56%, IMRT only for 15%, and IMRT combined with IGRT for 29%. All details of treatment techniques are presented in Table 2.4.

Patient characteristics		
Number of patients		972
Age (year), mean (range)		69 (45-83)
Diabetes[a] (%)		9
Anticoagulant treatment[b] (%)		24
Hypertension (%)		21
Coronary insufficiency (%)		9
Prior abdominal or pelvic surgery (%)		34
Androgen deprivation (%)		11
Tumor characteristics		
PSA (ng/ml), mean range		13 (0.3-84)
T stage (%)	T1	33
	T2	57
	T3	10
Prognostic group (%)	Low	16
	Intermediate	58
	High	26
Treatment characteristics		
Radiotherapy technique		
Standard 3D Conformational RT (%)		80
IMRT (without IGRT) (%)		6
IGRT (with IMRT) (%)		14
Total dose and fractionation		
70 Gy (%)	2.5 Gy. 4/week	29
	2 Gy. 5/week	29
80Gy (%)	2 Gy. 5/week	42
Target volume		
Prostate only (%)		16
Prostate + SV (%)		84

Table 2.4: Patient, tumor and treatment characteristics

[a] Type 1 or 2

[b] Vitamin K antagonist and antiplatelet drugs

2.2.2.3 Follow-up and toxicity grading

Patients were weekly assessed during treatment, then every 3 months over 1 year, and every 6 months thereafter. To determine the severity and incidence of principal late rectal complaints, patient records were prospectively collected from standardized questionnaires or retrospectively from physicians' medical files, filled at each follow-up visit. Acute toxicity was defined as adverse events occurring both during treatment, and within 3 months of treatment completion, recorded according to RTOG [34] and CTCAE V3.0 [35] toxicity grading (Table 2.6). Late toxicity was defined as adverse events occurring at least 6 months after RT initiation. Late rectal complaints were classified according to different symptoms, such as rectal bleeding, proctitis (urgency, tenesmus), diarrhea, and fecal incontinence, in compliance with a modified LENT-SOMA morbidity scoring system [34] (Table 2.5).

Symptoms	Grade 1	Grade 2	Grade 3	Grade 4
Rectal hemorrhage	Mild, intervention (other than iron supplements) not indicated	Two or fewer coagulations for bleeding	Three or more coagulations or any transfusion	Life threatening bleeding
Diarrhea	2-4 per day	5-8 per day	>8 per day	Refractory diarrhea
Proctitis (urgency/tenesmus)	Occasional urgency and/or occasional mild tenesmus	Intermittent urgency and/or intermittent tolerable tenesmus	Persistent urgency and/or persistent and intense tenesmus	Refractory and intolerable
Fecal incontinence	Occasional use of incontinence pads	Intermittent use of incontinence pads	Persistent use of incontinence pads	Surgical intervention Permanent colostomy

Table 2.5: Gastro-intestinal toxicity grading scale (modified LENT-SOMA) [35]

2.2.2.4 Statistical analysis

The Kaplan-Meier method was used to calculate the cumulative risk of Grade ≥ 2 rectal toxicity. The impact of the following parameters on acute and late radio-induced rectal toxicity was assessed: patient-related parameters, such as age, diabetes mellitus, anticoagulant treatment (vitamin K antagonist or antiplatelet drug), hypertension, coronary insufficiency, and history of abdominal surgery; tumor-related parameters, such as T-stage and risk group; radiation-related parameters, such as total dose, dose per fraction, IMRT, and IGRT; androgen deprivation. Logistic regression was used to assess the impact of the different parameters on acute and late toxicity at different follow-up time points, ranging from 3 to 5 years. Cox proportional hazard regression was employed to assess the impact of these parameters on the risk of late toxicity over time following treatment. Both methods provided the basis for univariate and multivariate analyses.

Differences between survival curves were assessed using the log-rank test. We employed only prospective data to analyze acute toxicity and to assess the risk of late toxicity. We conducted two investigations regarding the impact of the parameters on the risk of late toxicity: firstly using prospective data only, and secondly using all patient series to test the impact of moderate hypofractionation (2Gy vs. 2.5Gy/fraction).

Nomograms to predict rectal toxicity were build up based on the logistic regression model. To do this, the patient cohort was split into training (70% of patients) and validation (30% of patients) subgroups. The logistic regression model parameters were estimated using the training group, then applied to predict the complication probability of patients in the validation subgroup. The plot of actual versus predicted probability fit was used to assess the nomogram accuracy. The analyses were performed using the SPSS V18 (Chicago, IL) and R, by means of the rms package [12].

2.2.3 Results

Median follow-up was 60 months, ranging from 6 to 235 months.

2.2.3.1 Quantification of acute radio-induced rectal toxicity

Overall, 35.9% of patients exhibited Grade 1 radio-induced rectal toxicity during radiotherapy, while maximum acute radiation rectal toxicity was recorded as Grade 2 and 3 in 20.7 and 1.5% of patients, respectively. The primary types of acute toxicity were diarrhea and urgency/tenesmus, with Grade ≥ 2 events affecting 8.3% and 5% of patients, respectively. No patient experienced Grade ≥ 2 acute rectal bleeding or fecal incontinence. Details of acute radiation rectal toxicity, both overall [34] and discriminated by symptoms [35], in cases of IMRT combined with IGRT are presented in Table 2.6.

Acute toxicity	Grade 1	Grade 2	Grade 3
Diarrhea	19.3%	0.8%	0.0%
Proctitis	13.4%	5.0%	0.0%
Rectal bleeding	4.2%	0.0%	0.0%
Fecal incontinence	1.7%	0.0%	0.0%
Overall toxicity	38.6%	0.0%	0.0%

Table 2.6: Acute radio-induced rectal toxicity after prostate IMRT combined with IGRT (CTCAE v3.0)

2.2.3.2 Quantification of late radio-induced rectal toxicity

Figure 2.5 displays cumulative risks of Grade ≥ 2 late rectal toxicities, both overall and classified by symptom. At 3- and 5-year follow-up, these rates were 14.6% (95% CI: 11.2-18.0%) and 17.4% (95% CI: 13.6-20.2%) for overall toxicity, 8.7% (95% CI: 6.1-11.3%) and 10.7% (95% CI: 7.7-13.7) for rectal bleeding, 2.9% (95% CI: 1.3-4.5%) and

3.9% (CI: 2.3-5.5) for diarrhea, 2.6% (95% CI: 1.0-4.2%) and 2.9% (95% CI: 1.6-4.5) for urgency/tenesmus, 0.40% (95% CI: 0-1.0%), and 0.40% (95% CI: 0-1.0%) for fecal incontinence, respectively. Median time for Grade $\geq$ 2 late rectal side effects was 22 months, and 84.4% of events occurred within the first 3 years of treatment initiation. The 3- and 5-year Grade 3 overall late radio-induced rectal toxicity rates were 3.2% (95% CI: 1.6-4.8%) and 3.5% (95% CI: 1.7-5.3%), respectively. No patient exhibited Grade 4 toxicity.

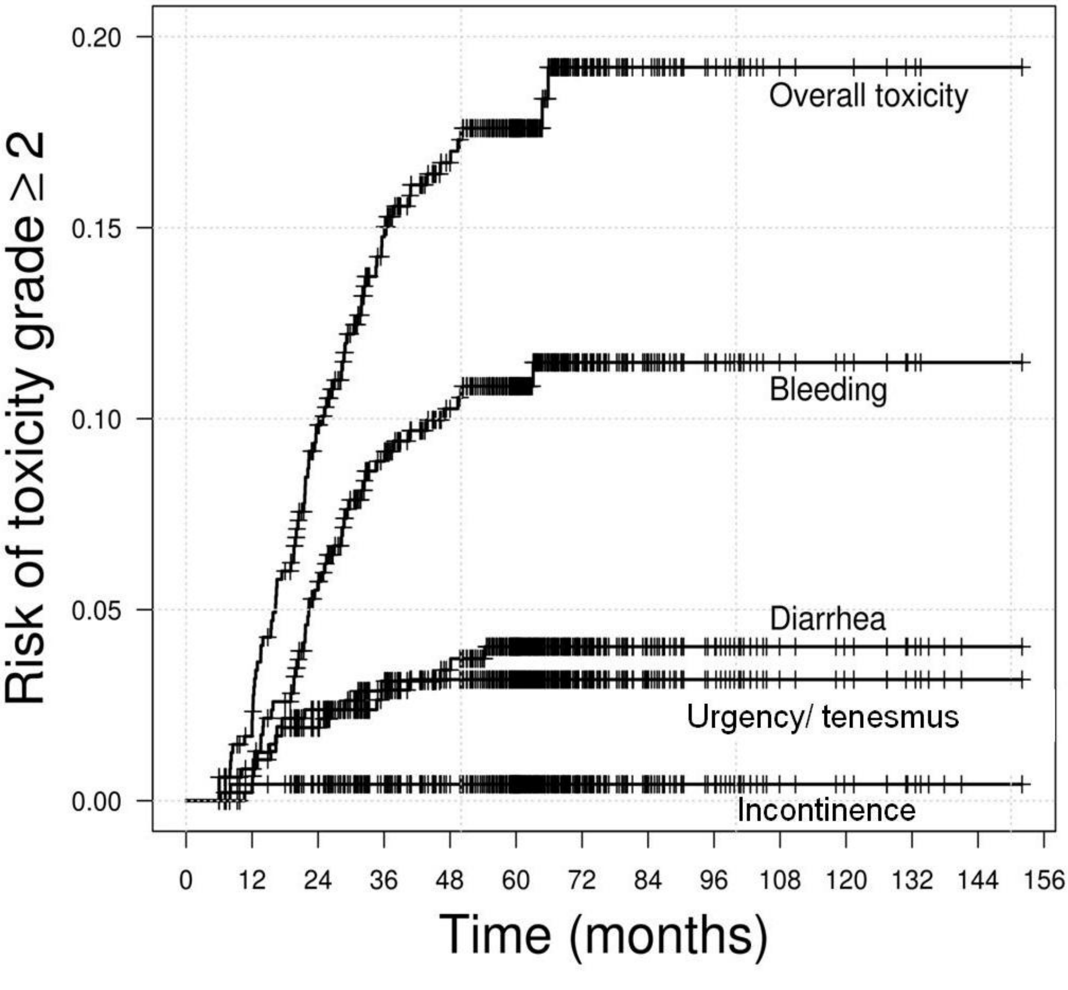


Figure 2.5: Risk of late rectal toxicity (Grade $\geq$2), overall and by symptoms (modified SOMA-LENT classification)

2.2.3.3 Factors impacting on the risk of acute radio-induced rectal toxicity

Univariate analysis revealed a significant increase in acute Grade $\geq$ 2 radio-induced rectal toxicity risk with total dose (p=0.05), along with a decrease with the combination of IMRT and IGRT ($p<0.0001$) and androgen deprivation ($p=0.05$). Multivariate analysis confirmed that the use of both IMRT and IGRT significantly decreased the risk of acute Grade $\geq$2 radio-induced rectal toxicity (RR=0.18, $p<0.0001$). When receiving both IGRT and IMRT, 6.7% of patients exhibited toxicity, compared to 16.7% of the patients treated with IMRT only, and 28.5% of those treated with 3D-CRT only. No association was observed between acute radio-induced rectal toxicity risk and abdominal surgery history, anticoagulant treatment, or diabetes mellitus.

2.2.3.4 Factors impacting on the risk of late radio-induced rectal toxicity and corresponding nomograms

The following results correspond to the multivariate analysis for overall Grade ≥ 2 late rectal toxicity, presented first, and rectal bleeding, presented second. No significant predictors were found for the other GI (gastro-intestinal) symptoms. The risk of Grade ≥ 2 overall late rectal toxicity was significantly increased by total dose (RR=1.7, $p=0.046$), and decreased by the combination of IMRT and IGRT (RR=0.25, $p=0.003$), demonstrated in the Cox analysis and only using the prospective cohort. The impact of total dose and RT techniques on the risk of Grade ≥ 2 late rectal complications by time (Kaplan-Meier) is shown in Figures 2.6 and 2.7. The 3-year risk of late rectal toxicity was significantly increased by total dose (RR=1.95, $p=0.03$), and decreased by both the use of IMRT and combined IMRT and IGRT, compared to the risk resulting from a “standard” 3DRT technique (RR=0.85 and RR=0.16, respectively, $p=0.01$). When delivering high doses (80Gy) to the prostate with IGRT and IMRT, the 3-year risk of toxicity was 4.5% (95% CI: 0.1-8.9%).

Using the entire cohort (prospective and retrospective patients) to assess the impact of fractionation modification, the 3-year risk of late rectal toxicity was found to be increased by total dose (RR=2.4, $p=0.002$) and dose per fraction (RR=3.3, $p=0.03$), while being reduced by the use of IMRT and combined IMRT and IGRT (RR=0.6 and RR=0.18, respectively, $p=0.02$). The impact of fractionation on the risk of late rectal toxicity when delivering 70 Gy over 7 weeks is presented in Figure 2.8. The corresponding nomogram and calibration plot (validation cohort) are shown in Figure 2.9. When incorporating acute toxicity (grade ≥ 2) events into the Cox analysis, these events were also predictor of late toxicity ($p=0.0001$, RR=3.3), together with the total dose ($p=0.043$, RR=1.7) and IGRT ($p=0.053$, RR=0.39).

The 5-year risk of Grade ≥ 1 rectal bleeding was increased by total dose (RR=2.3, $p=0.01$), dose per fraction (RR=4.2, $p=0.023$), and history of angina (RR=2.3, $p=0.046$). The corresponding nomogram and calibration plot (validation cohort) are shown in Figure 2.10. The 4-year risk of Grade ≥ 2 rectal bleeding was increased by total dose (RR=2.4, $p=0.04$) and history of angina (RR=4.9, $p=0.04$). The corresponding nomogram and calibration plot (validation cohort) are shown in Figure 2.11.

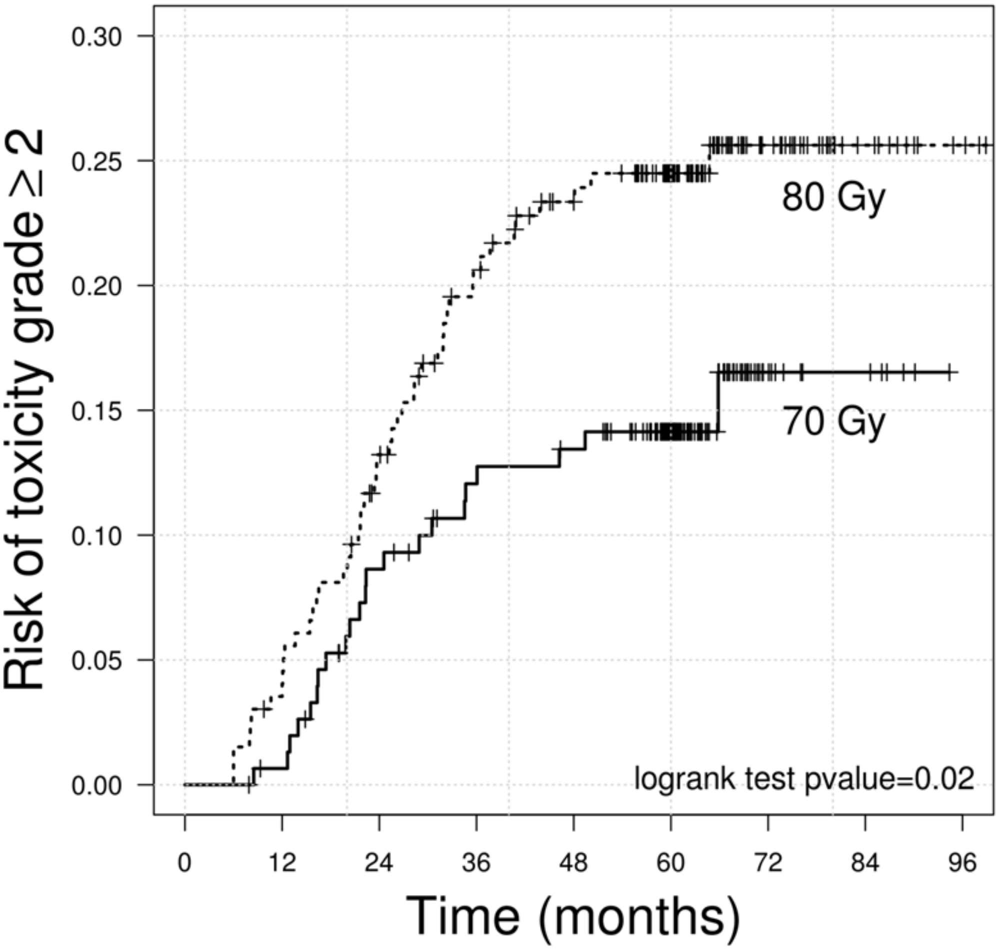


Figure 2.6: Impact of total dose (70Gy vs. 80Gy) on the risk of late rectal toxicity (Grade ≥ 2)

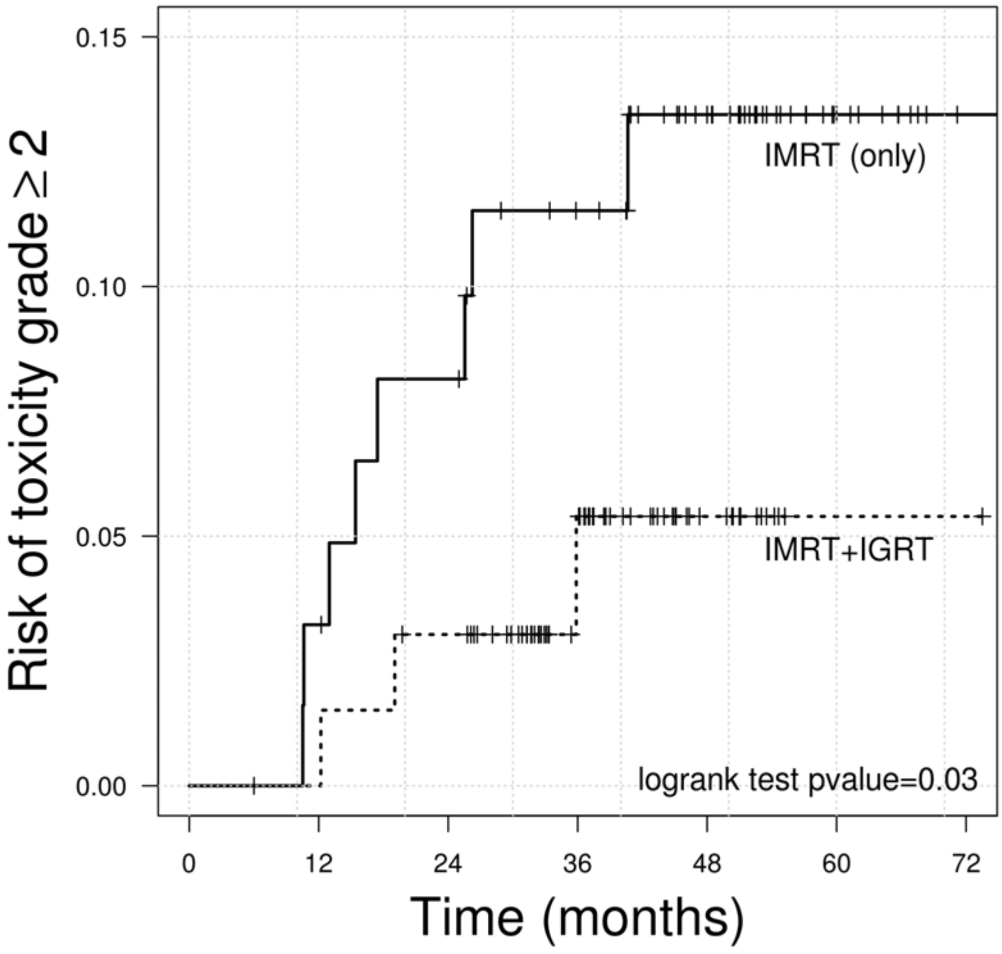


Figure 2.7: Impact of IGRT on the risk of late rectal toxicity (Grade ≥2) in case of high-dose IMRT (80Gy)

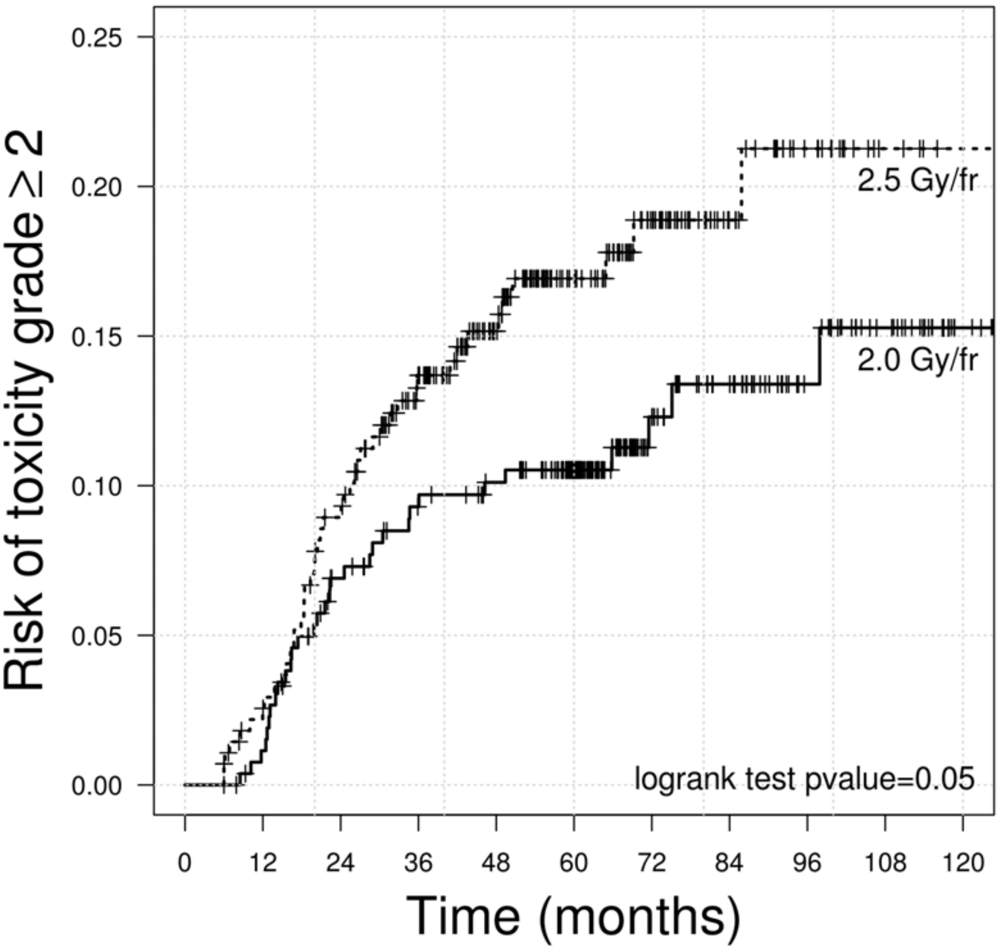
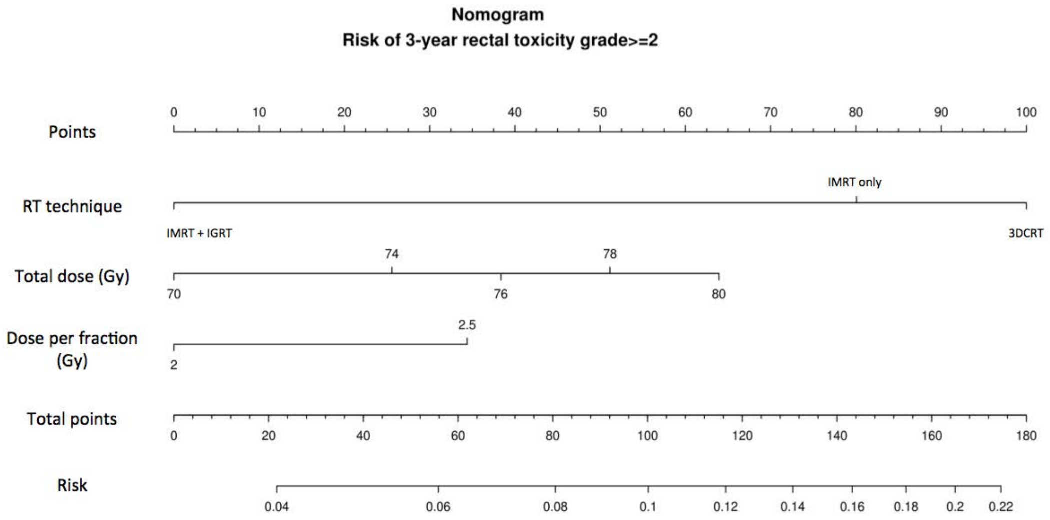
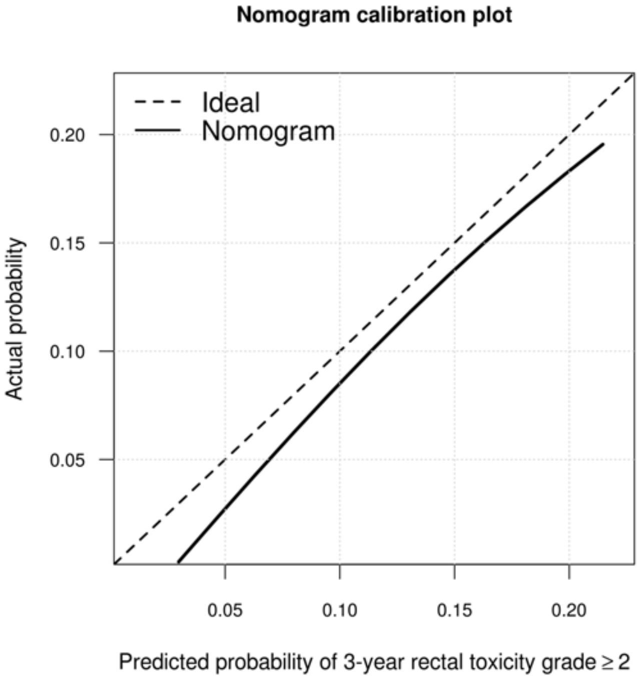


Figure 2.8: Impact of dose fractionation (2Gy vs. 2.5Gy) on the risk of late rectal toxicity (Grade $\geq$2) when delivering a total dose of 70Gy in 7 week

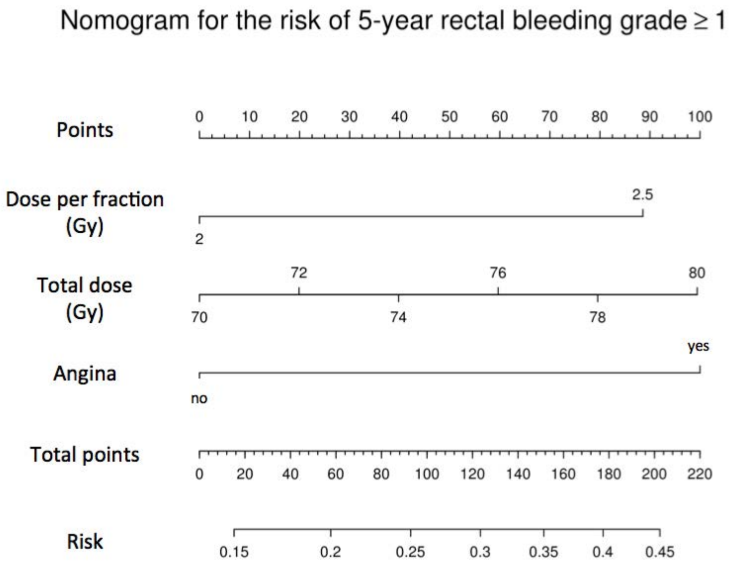


(a) Nomogram

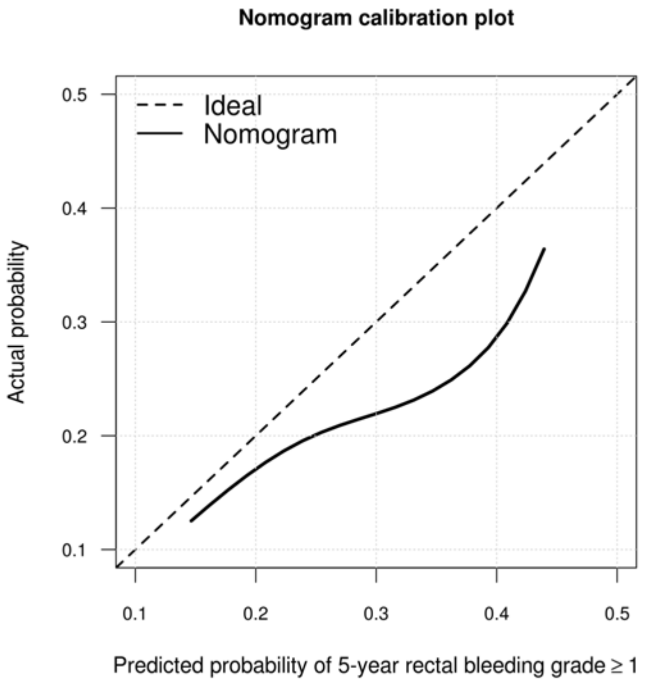


(b) Calibration plot

Figure 2.9: Nomogram and calibration plot (validation cohort) for the 3-year risk of Grade ≥2 overall late rectal toxicity

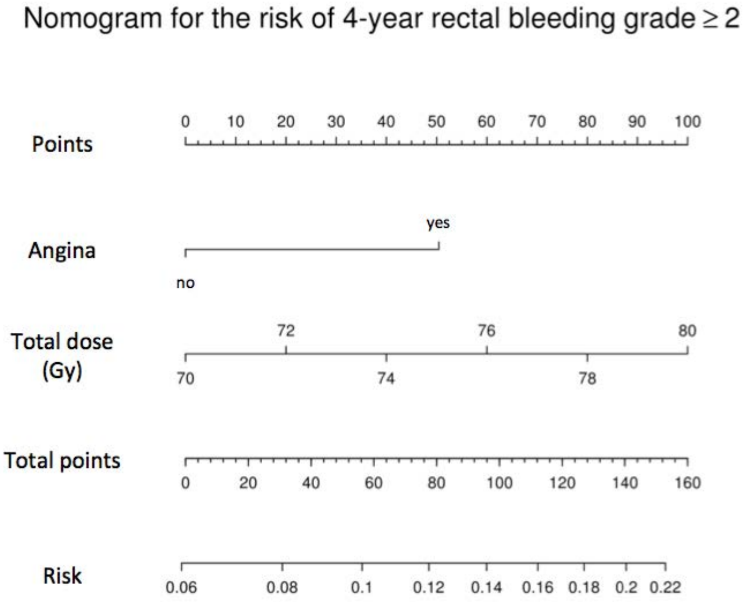


(a) Nomogram

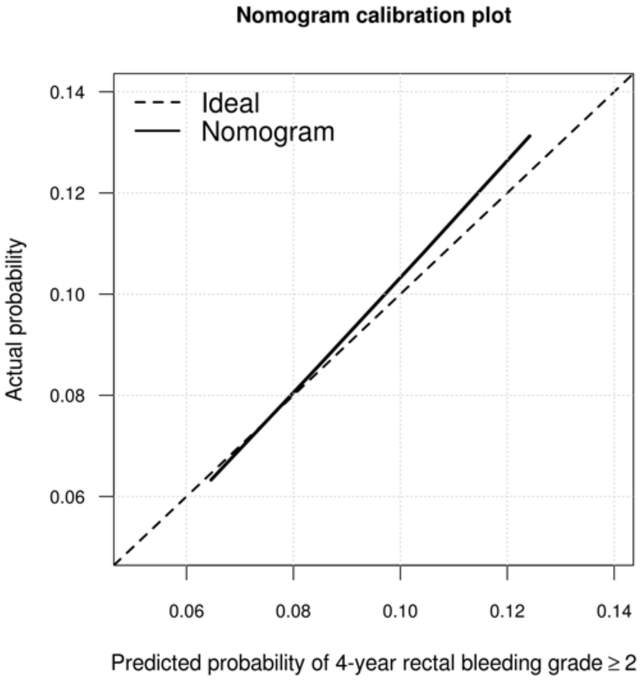


(b) Calibration plot

Figure 2.10: Nomogram and calibration plot (validation cohort) for the 5-year risk of Grade ≥1 rectal bleeding



(a) Nomogram



(b) Calibration plot

Figure 2.11: Nomogram and calibration plot (validation cohort) for the 4-year risk of Grade ≥2 rectal bleeding

2.2.4 Discussion

Our study analyzed the risk of radio-induced rectal toxicities in a large cohort of patients receiving radiotherapy for prostate cancer. The most frequently occurring symptoms of acute and late toxicity were diarrhea and rectal bleeding, respectively. Our findings confirmed the dose-effect correlation with radio-induced rectal toxicity side-effect occurrence, and demonstrated that combining novel radiation techniques dramatically decreases the risk of both acute and late radio-induced rectal toxicity. Even when increasing the dose targeting the prostate to 80 Gy, the risks of Grade ≥ 2 acute or late toxicity amounted to less than 10%. The nomograms we created for the study enabled the individual prediction of overall late rectal toxicity and rectal bleeding risks, taking into account both technique types and patient risk factors, such as history of angina.

Unlike the urinary toxicity risk, which increases linearly following treatment, the risk of rectal toxicity plateaued at the 3-year mark in our study, in line with previous publications [5]. Our data suggests the patient can thus be reassured that, once past this point in time, they are at low risk of developing new radio-induced rectal toxicity events. Rectal bleeding was the principal side-effect observed, while diarrhea and urgency/tenesmus were much less common, and fecal incontinence was a rare occurrence. Given that the pelvic lymph nodes were not included in the target volume, the small bowel was not irradiated, potentially accounting for the low incidence of diarrhea. The observed GI toxicities were therefore mainly associated with the dose received by the rectum and the anal canal. Other studies, however, observed slightly higher rates of fecal incontinence requiring pads, with 3-year rates reported at around 3 to 9% [5, 7, 36].

Several risk factors for rectal bleeding have been identified, the first being total dose. Increasing total dose to the prostate from 8 to 10Gy significantly increased the risk of overall late rectal toxicity and rectal bleeding from around 8 to 10% in our series (Figure 2.6), as was the case in most of the randomized studies [30, 16, 21], and in a meta-analysis involving 2,812 patients [37].

Reactive oxygen species produced by the ionizing radiation induce replicative and apoptotic death of both epithelial and microvascular endothelial cells of the mucosa, leading to ischemia, fibrosis and the development of neovascular lesions. The intestinal response to radiation injury depends also on several complex pathogenic processes involving the cell differentiation and the extracellular matrix, within the concept of “coordinated multicellular response” [38, 39]. Incidence and severity of chronic enteritis may also be related to radio-induced compromise of host defense responses to luminal pathogens or antigens [40]. Moreover, the pathophysiology of the “overall” late rectal toxicity depends clearly on each symptom. Since bleeding is caused by telangiectasias, functional symptoms (urgency/tenesmus) are due to changes in the rectal compliance and incontinence is due to loss of sphincter function. From a radiobiological point of view, all the dosimetric analyses indicated a prevalently serial-like behavior to rectal bleeding incidences [5], which demonstrates that bleeding is primarily accounted for the highest dose delivered to the anterior rectal wall, rather than to the high “average” dose delivered to the full rectum. This fact is also strongly supported by endoscopic telangiectasia findings, revealing telangiectasia increase from the posterior to the anterior rectum wall, with Grade 3 telangiectasia limited to the high dose region at the anterior rectum wall [41]. The practice of increasing the dose per fraction, while decreasing the total number of fractions, and consequently the treatment duration (i.e. hypofractionation), is also matter of debate in prostate radiotherapy. The benefits should be patient comfort and cost reduction, while maintaining the same tumor control and without increasing toxicity. We found that slightly increasing the dose per fraction (+0.5Gy) also moderately increased the risk of radio-induced rectal toxicity, as the same total dose was delivered. Three randomized studies [42, 43, 44] testing the benefits of moderate hypofractionation (2.4Gy/fr, 2.7Gy/fr, and 3.1Gy/fr), while also decreasing the total dose, revealed the same GI toxicity risk as with standard fractionation, although the risk of urinary toxicity was significantly increased in two studies [42, 43].

Our study revealed that the combination of IMRT and IGRT had a great impact on decreasing the risk of both acute and late radio-induced rectal toxicity, particularly in cases where high doses were delivered to the prostate (Figure 2.7). Though the benefits of IMRT have been well documented as decreasing the dose to the rectum, and therefore GI toxicity [31, 45], the benefits of IGRT have not yet been assessed. Moreover, at the best of our knowledge, no randomized study has been designed to compare between a “3D conformal technique” and IMRT, or indeed between one of the previously mentioned techniques and IGRT. Zelefsky et al recently published a retrospective study involving 376 patients, all treated to a high dose with IMRT (86.4 Gy), with half receiving IGRT [27]. With a limited median follow-up of 2.8 years, their findings revealed a significant reduction in late urinary toxicity among IGRT-treated patients, with the incidence of Grade $\geq$ 2 rectal toxicity being very low for both groups (1.0 and 1.6%, respectively; p=0.81).

Our multivariate analysis found angina to be the only patient-related predictor of rectal bleeding. Other studies have reported on the impact of other factors on this side-effect, such as history of abdominal or pelvic surgery [46] and androgen deprivation [47], also enabling nomograms to be generated [48]. Our findings revealed that the intensity of acute radio-induced rectal toxicity reactions was strongly correlated with the risk of late radio-induced rectal toxicity, in line with numerous other publications [16, 49, 50]. Given that the treatment has already been completed, the interest of such a factor is however relatively limited.

In order to assess the impact of moderate hypofractionation, we combined both retrospective and prospective data, all prospective patients having undergone standard fractionation. This retrospective data was, nevertheless, informative, since it was taken from a single institution, which irradiated all prostate cancers using the same 3D conformal technique, to the same total dose (70Gy), and within the same time period (7 weeks). Fractionation was, however, set at either 2Gy/fraction (five fractions/week), or 2.5Gy/fraction (four fractions/week), for practical reasons. Despite this, our full analysis was conducted in two stages, either excluding (for the vast majority of the results) or including this retrospective data. We did not include planning parameters such as dose-volume histograms in the analyses, as the study aim was to identify predictors and generate nomograms prior to planning, as a tool to guide oncologists in their decision-making process with respect to treatment options, such as choosing between radical prostatectomy and radiotherapy. Our study did not include the search of biological markers of GI toxicity. Indeed, faecal excretion of human DNA and calprotectin have been shown to increase during pelvic treatment and could be considered therefore as biomarkers of intestinal toxicity [51]. On the other hand, if blood Citrulline was found to decrease during pelvic radiotherapy, its variation was not correlated with mucosal toxicity [52]. Genetic predisposition for late rectal bleeding was also reported, involving single nucleotide polymorphisms in SOD2, XRCC1 and XRCC3 [53] or LIG4, MLH1 and ERCC2 [54] or reduced expression of AKR1B1, BAZB1, LSM7 and NUTD1 [55]. The combination of genetic susceptibility, clinical and dosimetric data should finally strongly increase radio-induced rectal injury prediction.

2.2.5 Conclusion

This study demonstrated that acute and late radio-induced rectal toxicity were significantly decreased by the combination of novel RT techniques such as IMRT and IGRT, which should thus be considered standard therapy in cases of high-dose prostate cancer radiotherapy. While the nomograms we generated require validation with external data, they remain simple, helpful tools aiding in treatment decision and patient information. Furthermore, an optimal predictive tool in routine prostate cancer practice is meant to be able to incorporate all GI, urinary, and sexual toxicities, as well as cancer control probabilities.

We have thus characterized in this chapter the risks of bladder and rectal toxicities and their associated symptoms. To do this, patient parameters and other dosimetry parameters than the DVH were used. Parameters impacting the risk of toxicity were identified using logistic regression, which is a parametric model. The next chapter will introduce a nonparametric model, namely a random forest model, which uses both DVH and patient parameters to predict rectal toxicity following prostate cancer radiotherapy.

Bibliography

[1] R. Mathieu and R. de Crevoisier, “Analyse de la toxicité urinaire tardive après radiothérapie externe pour adénocarcinome prostatique,” Ph.D. dissertation, Universiteé de Rennes 1. Université europeéenne de Bretagne., Rennes, France, 2012.

[2] R. Mathieu, J. D. Ospina Arango, V. Beckendorf, J.-B. Delobel, T. Messai, C. Chira, A. Bossi, E. Prisé, S. Guerif, J.-M. Simon, B. Dubray, J. Zhu, J.-L. Lagrange, P. Pommier, K. Gnep, O. Acosta, and R. de Crevoisier, “Nomograms to predict late urinary toxicity after prostate cancer radiotherapy,” *World Journal of Urology*, pp. 1–9, 2013.

[3] A. V. D’Amico, R. Whittington, S. Malkowicz, D. Schultz, K. Blank, G. A. Broderick, J. E. Tomaszewski, A. A. Renshaw, I. Kaplan, C. J. Beard, and A. Wein, “Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for linically localized prostate cancer,” *JAMA*, vol. 280, no. 11, pp. 969–974, 1998.

[4] R. Cheung, S. L. Tucker, A. K. Lee, R. de Crevoisier, L. Dong, A. Kamat, L. Pisters, and D. Kuban, “Doseâresponse characteristics of low- and intermediate-risk prostate cancer treated with external beam radiotherapy,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 61, no. 4, pp. 993 – 1002, 2005.

[5] C. Fiorino, R. Valdagni, T. Rancati, and G. Sanguineti, “Doseâvolume effects for normal tissues in external radiotherapy: Pelvis,” *Radiotherapy and Oncology*, vol. 93, no. 2, pp. 153 – 167, 2009.

[6] L. Budäus, M. Bolla, A. Bossi, C. Cozzarini, J. Crook, A. Widmark, and T. Wiegel, “Functional outcomes and complications following radiation therapy for prostate cancer: A critical analysis of the literature,” *European Urology*, vol. 61, no. 1, pp. 112 – 127, 2012.

[7] R. Valdagni, M. W. Kattan, T. Rancati, C. Yu, V. Vavassori, G. Fellin, E. Cagna, P. Gabriele, F. A. Mauro, M. Baccolini, C. Bianchi, L. Menegotti, A. F. Monti, M. Stasi, M. O. Giganti, and C. Fiorino, “Is it time to tailor the prediction of radio-induced toxicity in prostate cancer patients? building the first set of nomograms for late rectal syndrome,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 82, no. 5, pp. 1957 – 1966, 2012.

[8] R. Valdagni, T. Rancati, and C. Fiorino, “Predictive models of toxicity with external radiotherapy for prostate cancer,” *Cancer*, vol. 115, no. S13, pp. 3141–3149, 2009.

[9] V. Beckendorf, S. Guerif, E. Le Prisé, J.-M. Cosset, A. Bougnoux, B. Chauvet, N. Salem, O. Chapet, S. Bourdain, J.-M. Bachaud, P. Maingon, J.-M. Hannoun-Levi, L. Malissard, J.-M. Simon, P. Pommier, M. Hay, B. Dubray, J.-L.

Lagrange, E. Luporsi, and P. Bey, “70 gy versus 80 gy in localized prostate cancer: 5-year results ofÂ getug 06 randomized trial,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 80, no. 4, pp. 1056 – 1063, 2011.

[10] R. de Crevoisier, P. Pommier, J. Bachaud, G. Crehange, C. Boutry, B. Chauvet, T. Nguyen, A. Laplanche, M. Aubelle, and J. Lagrange, “Image-guided radiation therapy (igrt) in prostate cancer: Preliminary results in prostate registration and acute toxicity of a randomized study,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 75, no. 3, Supplement, pp. S99 –, 2009, proceedings of the American Society for Radiation Oncology 51st Annual Meeting 51st Annual Meeting of the American Society for Radiation Oncology.

[11] O. H. Behars, D. E. Henson, R. V. P. Hutter, and M. H. Myers, *American joint committee on cancer: Manual for staging of cancer.* Philadelphia: Lippincott, 1992.

[12] F. E. J. Harrell, *rms: Regression Modeling Strategies*, 2013, r package version 4.0-0.

[13] B. G. Gardner, A. L. Zietman, W. U. Shipley, U. E. Skowronski, and P. McManus, “Late normal tissue sequelae in the second decade after high dose radiation therapy with combined photons and conformal protons for locally advanced prostate cancer,” *The Journal of Urology*, vol. 167, no. 1, pp. 123 – 126, 2002.

[14] A. Heidenreich, J. Bellmunt, M. Bolla, S. Joniau, M. Mason, V. Matveev, N. Mottet, H.-P. Schmid, T. van der Kwast, T. Wiegel, and F. Zattoni, “{EAU} guidelines on prostate cancer. part 1: Screening, diagnosis, and treatment of clinically localised disease,” *European Urology*, vol. 59, no. 1, pp. 61 – 71, 2011.

[15] T. E. Schultheiss, W. Lee, M. A. Hunt, A. L. Hanlon, R. S. Peter, and G. E. Hanks, “Late {GI} and {GU} complications in the treatment of prostate cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 37, no. 1, pp. 3 – 11, 1997.

[16] S. T. A. Peeters, W. D. Heemsbergen, W. L. J. van Putten, A. Slot, H. Tabak, J. W. Mens, J. V. Lebesque, and P. C. Koper, “Acute and late complications after radiotherapy for prostate cancer: Results of a multicenter randomized trial comparing 68 gy to 78 gy,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 61, no. 4, pp. 1019 – 1034, 2005.

[17] M. Liu, T. Pickles, A. Agranovich, E. Berthelet, G. Duncan, M. Keyes, W. Kwan, M. McKenzie, J. Morris, H. Pai, S. Tyldesley, and J. Wu, “Impact of neoadjuvant androgen ablation and other factors on late toxicity after external beam prostate radiotherapy,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 58, no. 1, pp. 59 – 67, 2004.

[18] O. Cahlon, M. J. Zelefsky, A. Shippy, H. Chan, Z. Fuks, Y. Yamada, M. Hunt, S. Greenstein, and H. Amols, “Ultra-high dose (86.4 gy) {IMRT} for localized prostate cancer: Toxicity and biochemical outcomes,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 71, no. 2, pp. 330 – 337, 2008.

[19] A. Pollack, G. K. Zagars, G. Starkschall, J. A. Antolak, J. J. Lee, E. Huang, A. C. von Eschenbach, D. A. Kuban, and I. Rosen, “Prostate cancer radiation dose response: results of the m. d. anderson phase {III} randomized trial,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 53, no. 5, pp. 1097 – 1105, 2002.

[20] A. L. Zietman, M. L. DeSilvio, J. D. Slater, C. J. Rossi, D. W. Miller, J. A. Adams, and W. U. Shipley, “Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: A randomized controlled trial,” *JAMA*, vol. 294, no. 10, pp. 1233–1239, 2005.

[21] D. P. Dearnaley, M. R. Sydes, J. D. Graham, E. G. Aird, D. Bottomley, R. A. Cowan, R. A. Huddart, C. C. Jose, J. H. L. Matthews, J. Millar, A. R. Moore, R. C. Morgan, J. M. Russell, C. D. Scrase, R. J. Stephens, I. Syndikus, and M. K. B. Parmar, “Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the {MRC} {RT01} randomised controlled trial,” *The Lancet Oncology*, vol. 8, no. 6, pp. 475 – 487, 2007.

[22] A. Al-Mamgani, W. L. van Putten, W. D. Heemsbergen, G. J. van Leenders, A. Slot, M. F. Dielwart, L. Incrocci, and J. V. Lebesque, “Update of dutch multicenter dose-escalation trial of radiotherapy for localized prostate cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 72, no. 4, pp. 980 – 988, 2008.

[23] M. J. Zelefsky, E. J. Levin, M. Hunt, Y. Yamada, A. M. Shippy, A. Jackson, and H. I. Amols, “Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer,” *International Journal of Radiation Oncology*Biology*Physics*, vol. 70, no. 4, pp. 1124 – 1129, 2008.

[24] S. Namiki, S. Ishidoya, A. Ito, T. Tochigi, I. Numata, K. Narazaki, S. Yamada, Y. Takai, and Y. Arai, “Five-year follow-up of health-related quality of life after intensity-modulated radiation therapy for prostate cancer,” *Japanese Journal of Clinical Oncology*, vol. 39, no. 11, pp. 732–738, 2009.

[25] Z. A. Alicikus, Y. Yamada, Z. Zhang, X. Pei, M. Hunt, M. Kollmeier, B. Cox, and M. J. Zelefsky, “Ten-year outcomes of high-dose, intensity-modulated radiotherapy for localized prostate cancer,” *Cancer*, vol. 117, no. 7, pp. 1429–1437, 2011.

[26] K. Wallner, J. Roy, and L. Harrison, “Dosimetry guidelines to minimize urethral and rectal morbidity following transperineal I-125 prostate brachytherapy,”

*International Journal of Radiation Oncology*Biology*Physics*, vol. 32, no. 2, pp. 465 – 471, 1995.

[27] M. J. Zelefsky, M. Kollmeier, B. Cox, A. Fidaleo, D. Sperling, X. Pei, B. Carver, J. Coleman, M. Lovelock, and M. Hunt, "Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 84, no. 1, pp. 125 – 129, 2012.

[28] J.-B. Delobel and R. de Crevoisier, "Prédiction de la toxicité digestive après radiothérapie prostatique," Ph.D. dissertation, Universiteé de Rennes 1. Université européene de Bretagne., [S.l.], 2013.

[29] B. Gami, K. Harrington, P. Blake, D. Dearnaley, D. Tait, J. Davies, A. R. Norman, and H. J. N. Andreyev, "How patients manage gastrointestinal symptoms after pelvic radiotherapy," *Alimentary Pharmacology & Therapeutics*, vol. 18, no. 10, pp. 987–994, 2003.

[30] D. A. Kuban, S. L. Tucker, L. Dong, G. Starkschall, E. H. Huang, M. R. Cheung, A. K. Lee, and A. Pollack, "Long-term results of the m. d. anderson randomized dose-escalation trial for prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 70, no. 1, pp. 67 – 74, 2008.

[31] O. Cahlon, M. Hunt, and M. J. Zelefsky, "Intensity-modulated radiation therapy: Supportive data for prostate cancer," *Seminars in Radiation Oncology*, vol. 18, no. 1, pp. 48 – 57, 2008, advances in the Management of Prostate Cancer.

[32] L. Perrier, M. Morelle, P. Pommier, J.-L. Lagrange, A. Laplanche, P. Dudouet, S. Supiot, B. Chauvet, T.-D. Nguyen, G. Créhange, V. Beckendorf, F. Pene, X. Muracciole, J.-M. Bachaud, E. Le Prisé, and R. de Crevoisier, "Cost of prostate image-guided radiation therapy: Results of a randomized trial," *Radiotherapy and Oncology*, vol. 106, no. 1, pp. 50 – 58, 2013.

[33] V. Beckendorf, S. Guérif, E. Le Prisé, J. M. Cosset, O. Lefloch, B. Chauvet, N. Salem, O. Chapet, S. Bourdin, J. M. Bachaud, P. Maingon, J.-L. Lagrange, L. Malissard, J.-M. Simon, P. Pommier, M. H. Hay, B. Dubray, E. Luporsi, and P. Bey, "The {GETUG} 70 gy vs. 80 gy randomized trial for localized prostate cancer: Feasibility and acute toxicity," *International Journal of Radiation Oncology*Biology*Physics*, vol. 60, no. 4, pp. 1056 – 1065, 2004.

[34] J. D. Cox, J. Stetz, and T. F. Pajak, "Toxicity criteria of the radiation therapy oncology group (RTOG) and the European organization for research and treatment of cancer (EORTC)," *International Journal of Radiation Oncology* Biology* Physics*, vol. 31, no. 5, pp. 1341–1346, 1995.

[35] *Cancer Therapy Evaluation Program, Common Terminology Criteria for Adverse Events, Version 3.0.*, DCTD, NCI,NIH, DHHS Std., August 2006.

[36] C. Fiorino, T. Rancati, G. Fellin, V. Vavassori, E. Cagna, V. Casanova Borca, G. Girelli, L. Menegotti, A. F. Monti, F. Tortoreto, S. Delle Canne, and R. Valdagni, "Late fecal incontinence after high-dose radiotherapy for prostate cancer: Better prediction using longitudinal definitions," *International Journal of Radiation Oncology*Biology*Physics*, vol. 83, no. 1, pp. 38 – 45, 2012.

[37] G. A. Viani, E. J. Stefano, and S. L. Afonso, "Higher-than-conventional radiation doses in localized prostate cancer treatment: A meta-analysis of randomized, controlled trials," *International Journal of Radiation Oncology*Biology*Physics*, vol. 74, no. 5, pp. 1405 – 1418, 2009.

[38] J. Kountouras and C. Zavos, "Recent advances in the management of radiation colitis," *World Journal of Gastroenterology*, vol. 14, no. 48, pp. 7289–7301, December 2008.

[39] M. C. Vozenin-Brotons, "Tissue toxicity induced by ionizing radiation to the normal intestine: Understanding the pathophysiological mechanisms to improve the medical management," *World Journal of Gastroenterology*, vol. 13, no. 22, pp. 3031–3032, 2007.

[40] MaCnaughton, "Review article: new insights into the pathogenesis of radiation-induced intestinal dysfunction," *Alimentary Pharmacology & Therapeutics*, vol. 14, no. 5, pp. 523–528, 2000.

[41] S. Wachter, N. Gerstner, G. Goldner, R. Pötzi, A. Wambersie, and R. Pötter, "Endoscopic scoring of late rectal mucosal damage after conformal radiotherapy for prostatic carcinoma," *Radiotherapy and Oncology*, vol. 54, no. 1, pp. 11 – 19, 2000.

[42] D. A. Kuban, G. M. Nogueras-Gonzalez, L. Hamblin, A. K. Lee, S. Choi, S. J. Frank, Q. N. Nguyen, K. E. Hoffman, S. E. McGuire, and M. F. Munsell, "Preliminary report of a randomized dose escalation trial for prostate cancer using hypofractionation," *International Journal of Radiation Oncology* Biology* Physics*, vol. 78, no. 3, pp. S58–S59, 2010.

[43] A. Pollack, G. Walker, M. Buyyounouski, E. Horwitz, R. Price, S. Feigenberg, A. Konski, R. Stoyanova, and C. Ma, "Five year results of a randomized external beam radiotherapy hypofractionation trial for prostate cancer," *International Journal of Radiation Oncology * Biology * Physics*, vol. 81, no. 10, 2011.

[44] G. Arcangeli, B. Saracino, S. Gomellini, M. G. Petrongari, S. Arcangeli, S. Sentinelli, S. Marzi, V. Landoni, J. Fowler, and L. Strigari, "A prospective phase {III} randomized trial of hypofractionation versus conventional fractionation in patients with high-risk prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 78, no. 1, pp. 11 – 18, 2010.

[45] D. E. Spratt, X. Pei, J. Yamada, M. A. Kollmeier, B. Cox, and M. J. Zelefsky, "Long-term survival and toxicity in patients treated with high-dose intensity modulated radiation therapy for localized prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 85, no. 3, pp. 686 – 692, 2013.

[46] R. Valdagni, V. Vavassori, T. Rancati, G. Fellin, M. Baccolini, C. Bianchi, E. Cagna, P. Gabriele, F. Mauro, L. Menegotti, A. F. Monti, M. Stasi, and C. Fiorino, "Increasing the risk of late rectal bleeding after high-dose radiotherapy for prostate cancer: The case of previous abdominal surgery. results from a prospective trial," *Radiotherapy and Oncology*, vol. 103, no. 2, pp. 252 – 255, 2012.

[47] S. J. Feigenberg, A. L. Hanlon, E. M. Horwitz, R. G. Uzzo, D. Eisenberg, and A. Pollack, "Long-term androgen deprivation increases grade 2 and higher late morbidity in prostate cancer patients treated with three-dimensional conformal radiation therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 62, no. 2, pp. 397 – 405, 2005.

[48] R. Valdagni, T. Rancati, C. Fiorino, G. Fellin, A. Magli, M. Baccolini, C. Bianchi, E. Cagna, C. Greco, F. A. Mauro, A. F. Monti, F. Munoz, M. Stasi, P. Franzone, and V. Vavassori, "Development of a set of nomograms to predict acute lower gastrointestinal toxicity for prostate cancer 3d-CRT," *International Journal of Radiation Oncology*Biology*Physics*, vol. 71, no. 4, pp. 1065 – 1073, 2008.

[49] C. Vargas, A. Martinez, L. L. Kestin, D. Yan, I. Grills, D. S. Brabbins, D. M. Lockman, J. Liang, G. S. Gustafson, P. Y. Chen, F. A. Vicini, and J. W. Wong, "Dose-volume analysis of predictors for chronic rectal toxicity after treatment of prostate cancer with adaptive image-guided radiotherapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 62, no. 5, pp. 1297 – 1308, 2005.

[50] M. J. Zelefsky, D. Cowen, Z. Fuks, M. Shike, C. Burman, A. Jackson, E. S. Venkatramen, and S. A. Leibel, "Long term tolerance of high dose three-dimensional conformal radiotherapy in patients with localized prostate carcinoma," *Cancer*, vol. 85, no. 11, pp. 2460–2468, 1999.

[51] E. Varela, M. Antolin, F. Guarner, R. Verges, J. Giralt, and J.-R. Malagelada, "Faecal dna and calprotectin as biomarkers of acute intestinal toxicity in patients undergoing pelvic radiotherapy," *Alimentary Pharmacology & Therapeutics*, vol. 30, no. 2, pp. 175–185, 2009.

[52] L. Wedlake, C. McGough, C. Hackett, K. Thomas, P. Blake, K. Harrington, D. Tait, V. Khoo, D. Dearnaley, and H. J. N. Andreyev, "Can biological markers act as non-invasive, sensitive indicators of radiation-induced effects in the gastrointestinal mucosa?" *Alimentary Pharmacology & Therapeutics*, vol. 27, no. 10, pp. 980–987, 2008.

[53] R. J. Burri, R. G. Stock, J. A. Cesaretti, D. P. Atencio, S. Peters, C. A. Peters, G. Fan, N. N. Stone, H. Ostrer, and B. S. Rosenstein, "Association of single

nucleotide polymorphisms in SOD2, XRCC1 and XRCC3 with susceptibility for the development of adverse effects resulting from radiotherapy for prostate cancer," *Radiation Research*, vol. 170, no. 1, pp. 49–59, Jul. 2008.

[54] S. Damaraju, D. Murray, J. Dufour, D. Carandang, S. Myrehaug, G. Fallone, C. Field, R. Greiner, J. Hanson, C. E. Cass, and M. Parliament, "Association of dna repair and steroid metabolism gene polymorphisms with clinical late toxicity in patients treated with conformal radiotherapy for prostate cancer," *Clinical Cancer Research*, vol. 12, no. 8, pp. 2545–2554, 2006.

[55] R. Valdagni, T. Rancati, M. Ghilotti, C. Cozzarini, V. Vavassori, G. Fellin, C. Fiorino, G. Girelli, S. Barra, N. Zaffaroni, M. Alessandro, Pierotti, and M. Gariboldi, "To bleed or not to bleed. a prediction based on individual gene profiling combined with doseâvolume histogram shapes in prostate cancer patients undergoing three-dimensional conformal radiation therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 74, no. 5, pp. 1431 – 1440, 2009.

Chapter 3

Random forests to predict rectal toxicity following prostate cancer radiotherapy

This chapter presents the first methodological contribution of this thesis. It consist of the implementation of a random forest model to predict rectal toxicity following prostate cancer radiotherapy. In contrast with Chapter 2, DVH and patient parameters are simultaneously used to predict toxicity following prostate cancer radiotherapy. The content of this chapter has been accepted for publication in the International Journal of Radiation Oncology*Biology*Physics [1].

3.1 Introduction

In the context of dose escalation in prostate cancer radiotherapy, limiting toxicities are rectal and bladder complications. In order to assess the risk of toxicity, normal tissue complication probability (NTCP) models have been proposed by Lyman in the context of uniform radiation therapy, and subsequently adapted to 3D-conformal radiation [2]. These models operate in two different steps: 1) dose-volume histogram (DVH) reduction; and 2) probability-mapping. The DVH is first reduced to a single representative measure (i.e., equivalent uniform dose (EUD) [3]), before being converted into a probability of complication by a mathematical model conditioned to certain parameters. These parameters are estimated from empirical data [4]. One of the difficulties in improving toxicity prediction is not only including dose parameters, but also non-dosimetric parameters (such as clinical history, tumor characteristics or combined treatment) within the predictive models. One solution is to stratify the population and then estimate the model parameters at each stratum or to add extra parameters to "classic" NTCP models [5, 6], but at the expense of statistical power. Thus, each subgroup does not always meet the necessary number of patients to estimate the model parameters, or the estimation procedure becomes more unreliable.

Other approaches have been attempted. Using principal component-analysis, the most relevant DVH bins are extracted and then used to predict late toxicity [7]. Artificial neural networks (ANN) predict toxicity using additional inputs, other than DVH [8, 9], but their performance still needs to be compared to classic NTCP models. Based on all available information (DVH and patient parameters), one of the main issues is therefore identifying models that improve toxicity prediction.

The random forests (RF) method is a machine-learning technique proposed by Breiman [10] which intensively uses intuitive classification and regression trees (CART) [11]. Used for classification and regression, RF does not require *a priori* information concerning the relationship between input and output variables. Indeed, it provides a framework to assess the importance of input variables at predicting the outcome. RF presents several advantages over other predictive methods. In contrast to other machine-learning methods such as ANN, RF does not overfit data, as was demonstrated by Breiman [10]. RF can deal with data sets containing more variables than observations. It enables the implementation of variable-selection procedures based on their impact on the prediction of the outcome. To our knowledge, RF has been used for prostate cancer staging (11), but never to predict irradiation-induced toxicity.

With the aim of improving rectal toxicity prediction following prostate cancer radiotherapy, the objectives of this work are, firstly, to propose a random forest normal tissue complication probability (RF-NTCP) method to predict late rectal toxicity following prostate cancer radiotherapy and, secondly, to compare this RF approach with the classic Lyman-Kutcher-Burman (LKB) NTCP and logistic regression models.

3.2 Materials and methods

3.2.1 Patients and treatment

Records from 471 patients who received definitive radiotherapy for localized prostate adenocarcinoma were analysed. Data was prospectively collected from 295 patients treated in seventeen French institutions under two randomized studies: GETUG 06 (comparing 70 Gy to 80 Gy) [12] and STIC-IGRT (testing image-guide radiotherapy (IGRT)) [13], and retrospectively from 176 patients treated in two of them. Ethics committee approval was obtained and all prospective patients gave informed consent according to the current revision of the Helsinki Declaration. All patients had a biopsy-proven prostate adenocarcinoma. Pre-treatment workup included a CT scan and a bone scan. The mean age was 67.6 years (ranging from 45 to 82 years). Prior abdominal surgery was reported in 29.9% of patients, anticoagulant treatment in 16.8% and diabetes in 7.5%. The T stages according to the 1992 American Joint Committee on Cancer Staging System [14] were T1 in 25% of the patients, T2 in 63% and T3 in 12%. Prognostic risk groups defined by D'Amico [15] were low in 13% of the patients, intermediate in 57% and high in 30%.

The target volume comprised the prostate only in the low-risk group, and the prostate and seminal vesicles (SV) in the other risk groups. The pelvic lymph nodes were not treated. The mean prescribed dose to the prostate was 76.6 Gy (from 70 to 80 Gy) at 2 Gy per fraction, the seminal vesicles receiving 46 Gy. All of the patients received 3D conformational RT. Intensity-modulated RT (IMRT) and image-guided RT (IGRT) were given to 63% and 25% of patients, respectively. In the case of standard non-IGRT technique, orthogonal portal imaging was verified on Days 1, 2 and 3 and weekly thereafter. Prostate IGRT was performed using either cone beam CT (CBCT) or implanted fiducial markers within the prostate, either daily (for 50% of patients) or weekly (for 50% of patients). The 3D radiation technique was carried out following the French GETUG group recommendations, as previously reported [12]. Patients underwent simulation and treatment in the supine position. Target volume and organs at risk (bladder, rectum and femoral heads) were delineated on CT slices. According to the GTUG group, the rectal wall was generated with a thickness of 5 mm from the external manually-delineated rectal contour. The rectal length was defined as 1 cm below the planning target volume (PTV). The PTV was calculated by adding a 10 mm margin in all directions except for the posterior, where a 5 mm margin was considered. The median rectal volume, CTV1 (prostate and SV), CTV2 (prostate only), PTV1 (prostate and SV) and PTV2 (prostate only) were: 45.6 cc (from 24.0 to 111.4 cc), 74.5 cc (from 10.4 to 251.3 cc), 64.3 cc (from 7.4 to 185.9 cc), 222.3 cc (from 90.3 to 507.2 cc) and 167.0 cc (from 77.9 to 440.6 cc), respectively. The rectal DVH had to respect GETUG recommendations [12], being $V72<25\%$ with the maximum dose (within 1.8 cc) <76 Gy. The full rectal DVH was analyzed with a bin-step of one Gray. The median rectal V72 was 8.27% (from 0 to 35.2%) and the median maximum dose was 76 Gy (from 62 to 81 Gy). Androgen deprivation therapy was given to 21.9% of the patients.

3.2.2 Follow-up and toxicity grading

All patients were evaluated every three months in the first year and every six months thereafter. Late rectal toxicity was defined as an event occurring more than six months after the beginning of RT. To determine the severity and incidence of late rectal toxicity, records were prospectively extracted from standardized questionnaires or retrospectively from physicians' reports taken at each follow-up visit. Rectal toxicity was classified according to the SOMA/LENT morbidity scoring system. Patients with a history of hemorrhoids were not scored as Grade 1 bleeding. Analyses were performed for overall late rectal toxicity ($\geq$ grade 2) and for rectal bleeding ($\geq$ grade 1 and $\geq$ grade 2). The 5-year risk of rectal toxicity rates were assessed by the Kaplan-Meier method (471 patients). For a given patient only the highest toxicity was considered. For estimating the models only patients with a follow-up longer than five years were selected (261 patients). Among these 261 patients, 13 and 4 received IMRT only and IMRT combined with IGRT, respectively.

3.2.3 Logistic regression to identify significant parameters and estimate the risk of rectal toxicity

Standard logistic regression was used to estimate the risk of toxicity, including statistically-significant parameters found in multivariate analysis. The following parameters were tested in univariate and multivariate analysis as potentially impacting on the risk of 5-year rectal toxicity: patient characteristics (age, history of abdominal surgery, diabetes and anticoagulant treatment), tumor characteristics (T stage and D'Amico stage), DVH parameters (CTV, PTV, rectal volume, DVH bins) and other treatment parameters (IMRT and androgen deprivation). The impact of IGRT could not be assessed at the 5-year stage due to a lack of follow-up with this technique. A logistic regression was also fit with the parameters found by the RF-NTCP model.

3.2.4 Random forest NTCP model

The RF is based on classification and regression trees (CART), aggregating the different predictions coming from independent trees [10]. The mechanism to build a random forest can be summarized in three steps: 1) generating new databases by re-sampling with replacement from the original database; 2) building a classification tree on each one of the new databases; 3) finally, aggregating the predictions given by each classification tree. The details of this method are presented in Appendix A. As each tree is a set of if-then rules, the forest can be stored in software and once a RF-NTCP model has been trained for a specific toxicity endpoint, the model can be used to predict the NTCP for a new patient and the trained RF-NTCP can be also transferred via software to different centers.

Three RF were trained to predict three different 5-year rectal toxicity endpoints according to the SOMA/LENT classification: Grade$\geq$2 overall rectal toxicity, Grade$\geq$1 and Grade$\geq$2 rectal bleeding. In order to train a RF to predict 5-year Grade$\geq$1 rectal bleeding, the database was randomly split into a training cohort consisting of 75% of patients and a validation cohort consisting of 25% of patients. Firstly, a RF was trained on the training cohort using only the DVH, for comparison with the standard DVH-based NTCP model. The resulting RF contained only the most important DVH bins. A second RF was then trained using the selected DVH bins and additional clinical parameters (age, history of diabetes, anticoagulant treatment and previous abdominal surgery). The resulting RF contained the previously-selected DVH bins and the clinical parameters that contribute to improve prediction, measured in terms of the out-of-bag error and the Gini index (described in Appendix A). The resulting RF was tested using the validation cohort and this process was carried out 1000 times, to allow for the calculation of average area under the curve AUC values on the left-out samples. This was also done for LKB-NCTP and logistic regression models. An additional test was performed including the LKB-estimated NTCP as an input variable within the RF-NTCP model. For 5-year Grade$\geq$2 overall rectal toxicity and Grade$\geq$2 rectal bleeding, the low number of events did not enable us to split the cohort. For 5-year Grade$\geq$1 rectal bleeding, a RF was also trained using prospective patients only and then validated using retrospective patients. The number of trees, q, was set to 20,000 and the rest of the function parameters were set to the default values of the *randomForest* function in R [16]. The importance of the different variables was assessed using the variable importance measures described in Appendix A.

3.2.5 LKB model

The LKB model has been widely described in the literature [2, 17]. The model parameters are: the dose that yields a toxicity rate of 50%, TD_{50}; the volume effect n; and the steepness m of the dose response at TD_{50}. The model equations are shown in (3.1) to (3.3),

$$NTCP\left(DVH_i|TD_{50}, n, m\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{\frac{-x^2}{2}} dx, \tag{3.1}$$

$$t = \frac{D_{max,i} - TD_{50}\left(v\right)}{m \times TD_{50}\left(v\right)}, \tag{3.2}$$

$$TD_{50}\left(v\right) = TD_{50} \times EUD^{-1} \text{ and } EUD = \left(\sum_i v_i d_i^{1/n}\right)^n, \tag{3.3}$$

where $D_{max,i}$ is the maximum dose to the rectum for the i-th patient. In order to integrate clinical variables to the LKB-NTCP model, we used a modified version of this model, as in previous studies [5, 6]. The TD_{50} is defined as a function of clinical parameters $x_1, \ldots, x_n$. We used TD'_{50} instead of TD_{50} in equation (3.3) as defined in equation (3.4),

$$TD'_{50} = TD_{50} + \beta_1 x_1 + \cdots + \beta_n x_n. \tag{3.4}$$

Using this parameterization, TD_{50} is the dose that yields a toxicity rate of 50% under standard conditions (when clinical parameters equal zero). The coefficients $\beta_1, \ldots, \beta_n$ reflect the impact of the clinical variables on the dose that causes a toxicity rate of 50%. For a binary variable (which is either zero or one), a negative value for its corresponding coefficient means that a lower dose causes a toxicity rate of 50%. For a continuous variable, the mean was subtracted in order to have a straightforward interpretation of the associated coefficient. The model was estimated using the maximum likelihood method described in Section 1.4.5.

3.2.6 Assessment and comparison of the predicted capabilities of the different models

In order to assess the predictive capabilities of all the different NTCP models (LKB-NTCP, logistic regression and RF-NTCP), the receiver operator characteristic (ROC) curve was used. The area under the curve (AUC) was used as a measure of the performance of the different NTCP models, as previously described [18]. A nonparametric (Wilcoxon) test was performed to assess whether the mean AUC value obtained with RF-NTCP was higher than those obtained using LKB-NTCP and logistic regression models in Grade≥1 5-year rectal bleeding prediction. All of the calculations were performed using the R environment and the *randomForest* R package [16].

3.3 Results

The median follow-up was 62 months (6-155). The 5-year Grade≥2 overall rectal toxicity, Grade≥1 and Grade≥2 rectal bleeding rates were: 16% (95% CI: 13-20%); 25% (95% CI: 20-29%); and 10% (95% CI: 7-13%), respectively. For these three endpoints, 261 patients had a follow-up longer than 60 months, with 45, 65 and 29 toxicity events for each of the toxicity endpoints, respectively.

3.3.1 Significant parameters and risk of rectal toxicity by logistic regression

The univariate analysis found the following factors to be significantly associated ($p<0.05$) with a higher risk of 5-year rectal toxicity (Grade≥2 overall toxicity, Grade≥1 and Grade≥2 rectal bleeding): the total prescribed dose to the prostate, the maximum rectal dose and the rectal DVH bins from V70 to V74. In addition, V75 was found to be associated with Grade≥1 and Grade≥2 5-year rectal bleeding and V69 with Grade≥1 5-year rectal bleeding. In multivariate logistic regression, the maximum dose in the rectum was the only significant parameter for all rectal toxicity endpoints.

3.3.2 Random forests prediction

Significant predictive capabilities were obtained for all endpoints using the RF-NTCP model, including both training and validation cohorts, as shown in Tables 3.1-3.3. The AUC used to measure the performance of RF models ranges from 0.66 to 0.76, depending on the toxicity endpoint. The AUC of the RF-NTCP model trained using prospective patients was 0.66 and the corresponding AUC calculated from predictions made for retrospective patients was 0.65.

Several DVH points (from V21 to V73, V73 and V70 being the most important) were identified as relevant parameters in predicting all aforementioned toxicity events. In the case of rectal bleeding Grade≥1 and Grade≥2, age and anticoagulant treatment were also identified as relevant parameters.

Figure 3.6 shows the mean decrease accuracy and the mean decrease Gini index (c.f. Appendix A), the variable importance measures used in predicting 5-year Grade≥1 rectal bleeding using a RF-NTCP with age and anticoagulants as additional parameters.

Figures 3.1 to 3.3 present three different instances of trees taken from the RF-NTCP model to predict the risk of 5-year Grade≥1 rectal bleeding. Using just a single tree is nonsense, as each tree was grown using a bootstrap sample of the original dataset, but visualizing these threes could help to understand what a RF is made of. These three examples illustrate the diversity (differences among trees) of a RF.

3.3.3 LKB NTCP model prediction

NTCP parameters (TD_{50}, n and m) have been estimated for the LKB-NTCP model. Depending on the toxicity endpoint, TD_{50}, n and m ranged from 70.41 to 82.14 Gy, 0.0030 to 0.06 and 0.09 to 0.12, respectively. For the modified LKB model (integrating

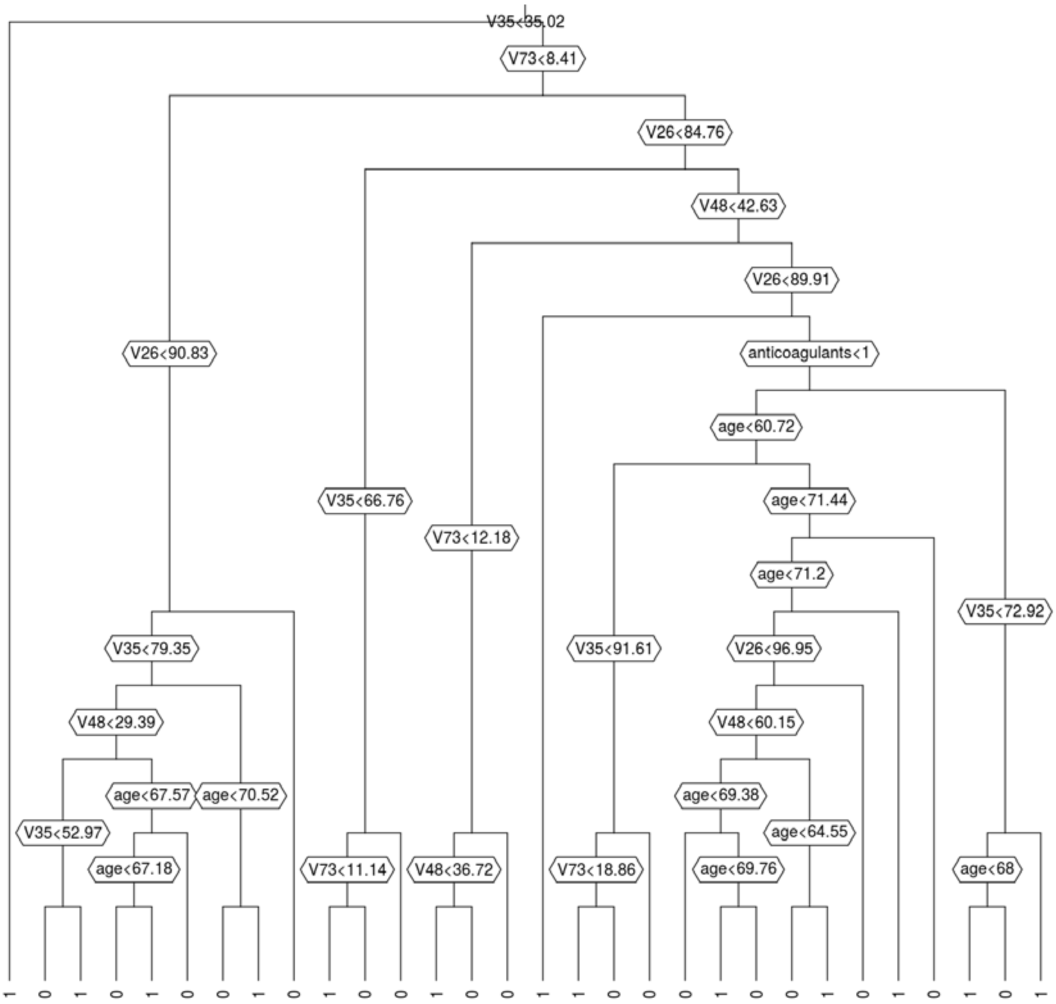


Figure 3.1: Example of a tree within a RF-NTCP model to predict the risk of 5-year Grade≥1 rectal bleeding (1 stands for complication and 0 for non complication)

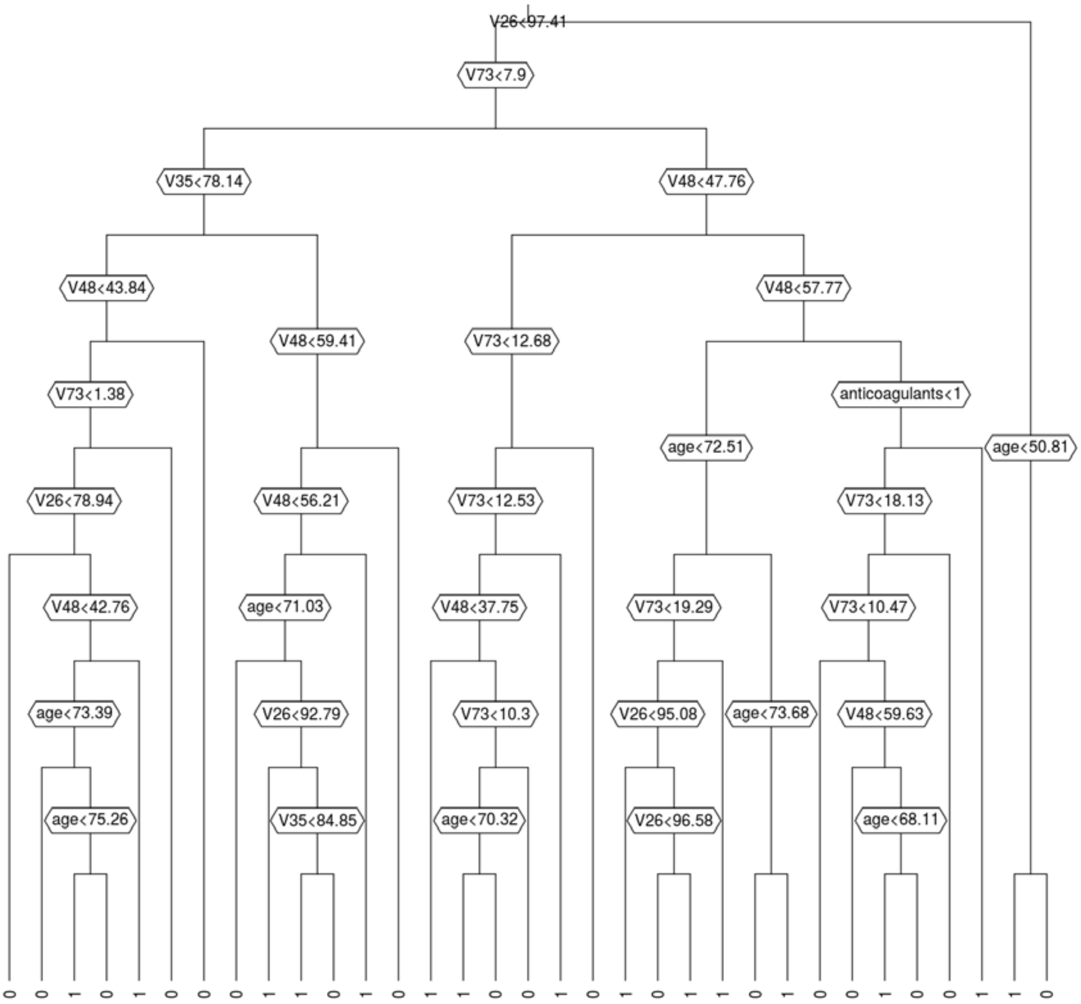


Figure 3.2: Example of a tree within a RF-NTCP model to predict the risk of 5-year Grade≥1 rectal bleeding (1 stands for complication and 0 for non complication)

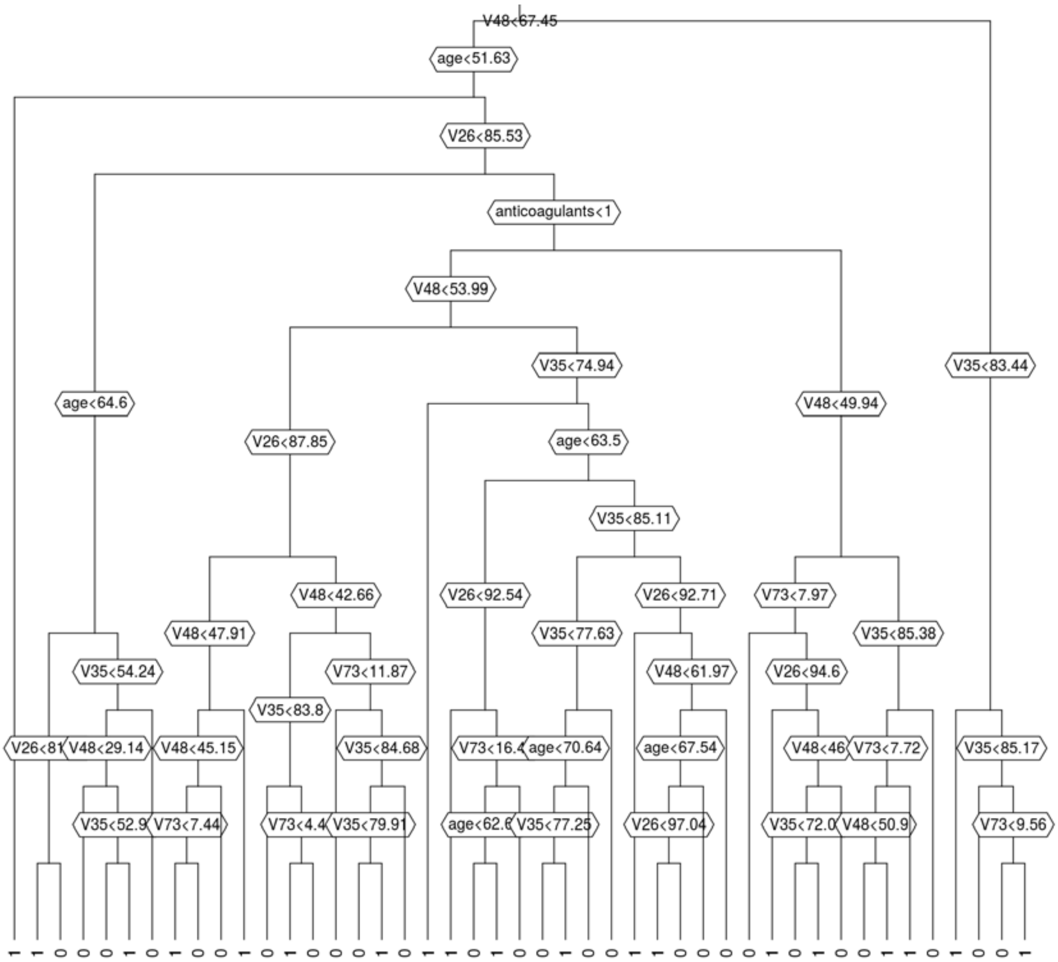


Figure 3.3: Example of a tree within a RF-NTCP model to predict the risk of 5-year Grade≥1 rectal bleeding (1 stands for complication and 0 for non complication)

clinical variables), negative values were obtained for the coefficients associated to anticoagulants in the case of Grade≥1 and Grade≥2 rectal bleeding, meaning that this is a risk factor. For age, the coefficient estimate is not statistically meaningful (the 95% confidence interval includes 0). Both age and anticoagulant treatment were added to the LKB-NTCP model in order to perform a fair comparison (using the same parameters) between RF-NTCP and the modified LKB model.

3.3.4 Comparison of random forest with LKB-NTCP models and logistic regression

Figures 3.4 and 3.5 show the ROC curves and the respective AUC of the different complication models used to predict 5-year Grade≥1 rectal bleeding, with training and validation data, respectively. Table 3.1 shows parameter estimates (for the LKB-NTCP and logistic regression models), selected variables (for the RF-NTCP model) and AUC values (training and validation) for 5-year Grade≥1 rectal bleeding. The use of covariates improved AUC values for both RF and LKB-NTCP models. Tables 3.2 and 3.3 show parameter estimates and selected variables for 5-year Grade≥2 rectal bleeding and rectal toxicity. In total, AUC values calculated using the RF-NTCP model with only DVH, (ranging from 0.66 to 0.74) are always higher than those calculated using the standard LKB-NTCP or logistic regression models (ranging from 0.62 to 0.71). The AUC values calculated using the RF-NTCP model integrating clinical variables are also higher (ranging from 0.66 to 0.76) than those calculated for the modified LKB-NTCP model (ranging from 0.64 to 0.69). The mean AUC values calculated using left-out samples of 5-year Grade≥1 rectal bleeding were 0.67 for the RF-NTCP model, 0.64 for the LKB-NTCP model (p-value<0.001) and 0.62 for the logistic regression model (p-value<0.001). The inclusion in the RF-NTCP model of the predicted probabilities generated by the LKB-NTCP model did not improve performance in the RF-NTCP prediction.

3.4 Discussion

The aim of our work was to evaluate the interest of RF-NTCP model in predicting late rectal toxicity following prostate cancer radiotherapy. To assess the predictive capabilities of the RF-NTCP model and to compare its performance to "classic" NTCP models (LKB), AUC values were calculated from ROC curves of each model. A "modified" LKB-NTCP model has also been used to integrate clinical variables to the LKB model [5, 6]. For each toxicity endpoint, a RF-NTCP model was trained using firstly only the DVH, leading to a higher AUC than that of the LKB model (ranging from 0.66 to 0.74 for the RF-NTCP model and from 0.62 to 0.65 for the LKB-NTCP model). Then, in the case of Grade≥1 and Grade≥2 rectal bleeding, two additional variables were added to the RF-NTCP model, namely age and anticoagulant treatment. These same variables were added to a modified LKB-NTCP model for comparison, and again the AUC values obtained for the RF-NTCP model were higher (ranging from 0.66 to 0.76 for the RF-NTCP model and from 0.64 to 0.69 for the LKB-NTCP model). In

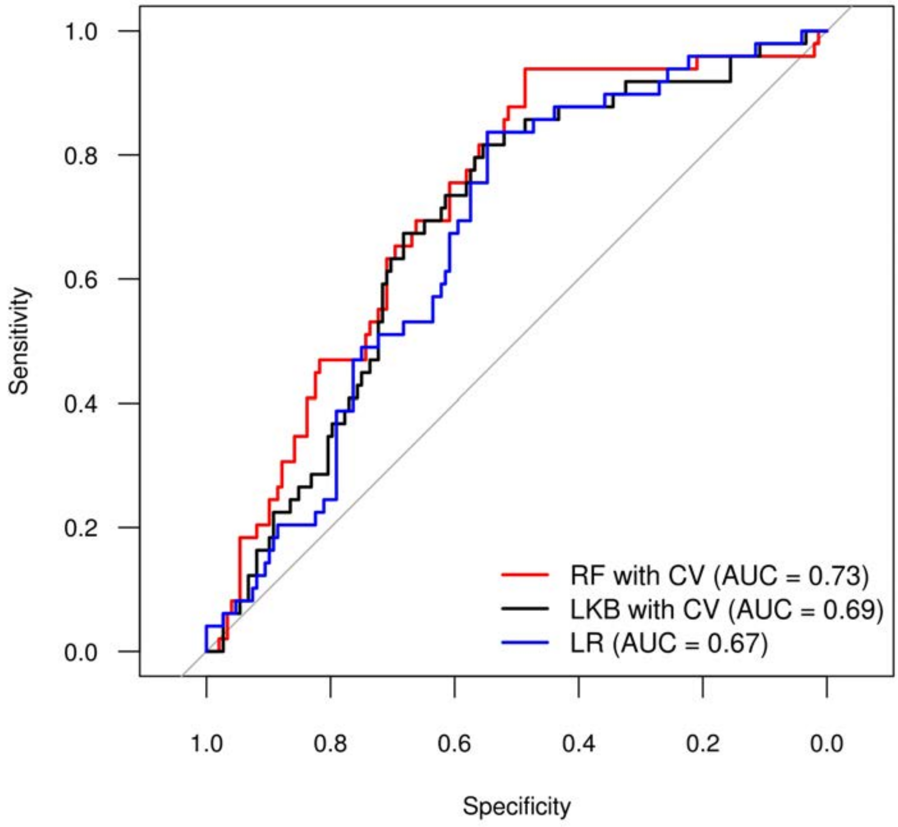


Figure 3.4: ROC and AUC in predicting 5-year Grade≥1 rectal bleeding for RF and LKB models using clinical variables (CV) (training cohort only)

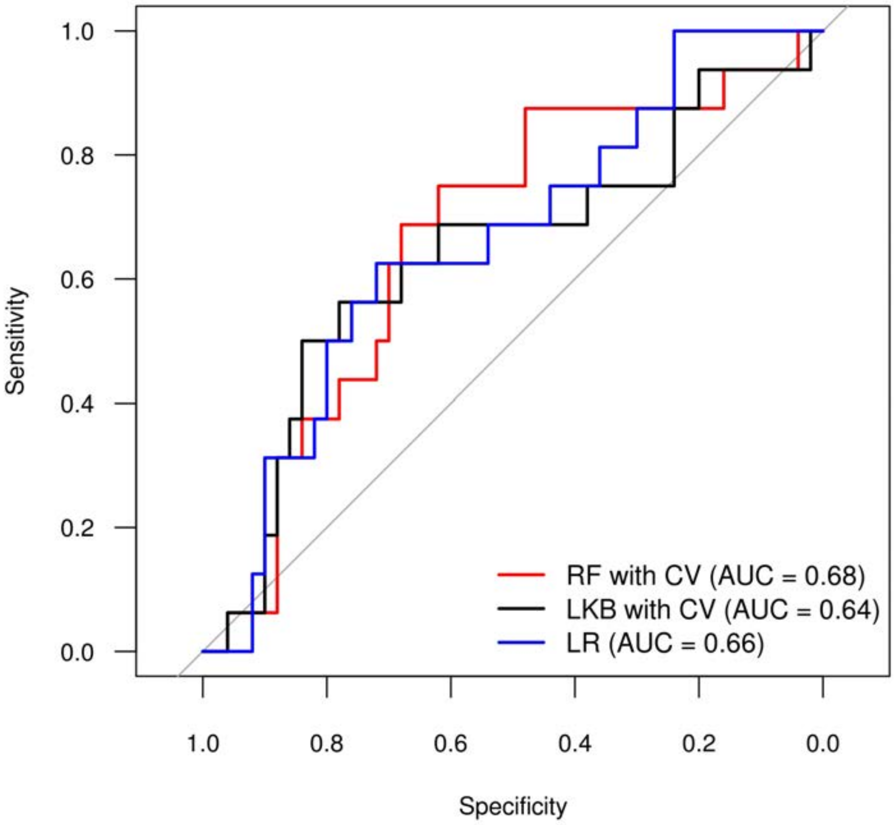


Figure 3.5: ROC and AUC in predicting 5-year Grade≥1 rectal bleeding for RF and LKB models using clinical variables (CV) (validation cohort only)

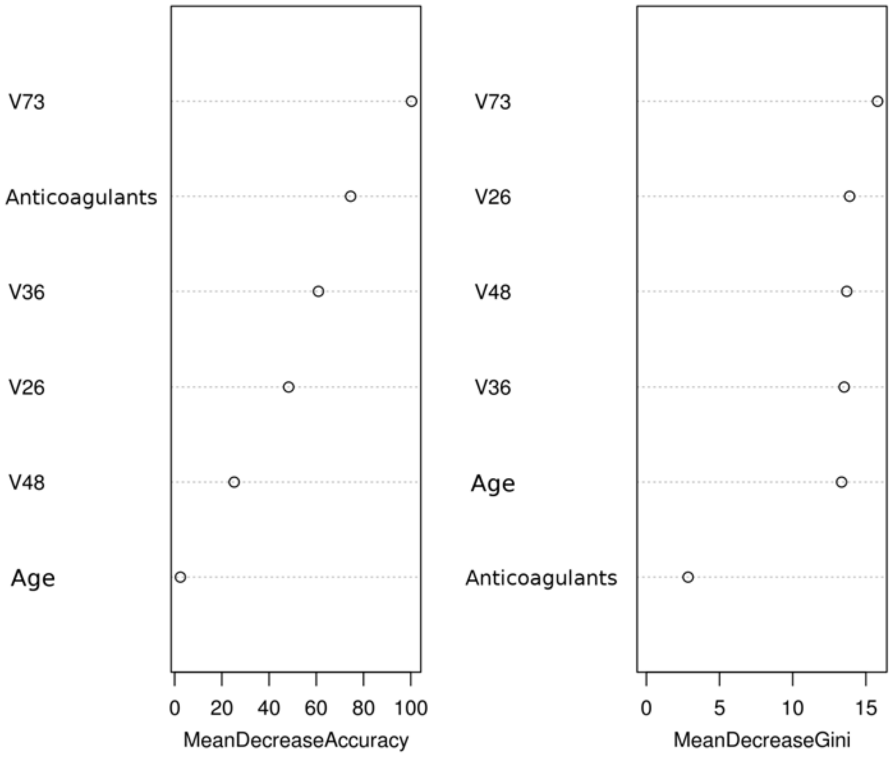


Figure 3.6: Variable-importance measures for a RF-NTCP model with clinical variables to predict 5-year Grade≥1 rectal bleeding. The mean decrease accuracy measures the predictive capabilities of a variable whereas the Gini index measures the discriminant capabilities of a variable during the construction of the RF.

Model		Parameters					AUC
RF-NTCP	DVH only	V73, V36, V26, V48*					Training 0.70 Validation 0.66
	Integrating DVH and clinical variables	V73, anticoagulants, V36, V26, V48, age*					Training 0.73 Validation 0.68
LKB-NTCP		TD_50	n	m	Others	LLH	
	Standard DVH only (1)	79.11 (77.24 - 81.12	0.003 (0.0003 - 0.0118)	0.12 (0.09 - 0.16	-	139.88	Training 0.65 Validation 0.64
	Modified integrating DVH and clinical variables (4,5)	70.41 (69.07, 71.83)	0.06 (0.04, 0.07)	0.09 (0.07, 0.13)	Age: 0.07 (-0.14 - 0.29) Anticoagulants: -1.92 (-4.04 - 0.40)	103.29	Training 0.69 Validation 0.64
Logistic regression	DVH only	V73: 0.05 (0.01, 0.09); V48: 0.01 (-0.03, 0.05) V36: 0.00 (-0.06, 0.06); V26: -0.02 (-0.07, 0.02)				105.81	Training 0.65 Validation 0.61
	Integrating DVH and clinical variables	V73: 0.05 (0.01, 0.09); V48: 0.03 (-0.01, 0.07); V36: 0.00 (-0.06, 0.06); V26: -0.02 (-0.07, 0.02); Age: -0.02 (-0.08, 0.04) Anticoagulants: 0.75 (-0.05, 1.5)				103.95	Training 0.69 Validation 0.65

Table 3.1: NTCP models for Grade≥1 rectal bleeding prediction. RF: Random Forest; LKB: Lyman-Kutcher-Burman; LLH: negative of the log-likelihood function; AUC: area under the curve from the receiving operating characteristic curve; *: The parameters retained by the RF-NTCP model are presented in order of decreasing importance; (xx - xx): 95% confidence interval; Training: results obtained on the 75% of patients; Validation: results obtained on the 25% of patients.

the case of Grade≥1 rectal bleeding, the results obtained on a training cohort (with or without additional clinical variables) were confirmed on a validation cohort. We also confirmed that the AUC value obtained using a RF-NTCP model trained with only prospective patients was close to the AUC value calculated during the validation step using only retrospective patients, thereby demonstrating the transportability of the proposed method. The RF-NTCP model therefore appears to be a strong competitor of "classic" NTCP models in predicting late rectal toxicity. Amongst the retained RF-NTCP parameters, the high dose appears to be the most important variable that increases the risk of rectal toxicity. Amongst clinical variables, our RF-NTCP model identified age and anticoagulant treatment as increasing the risk of rectal bleeding.

Previous studies have shown that different "classic" NTCP models (Logit-EUD, Poisson-EUD, serial reduction model, LKB) have similar predictive capabilities [19, 20]. In our work we used the LKB model, which was also found to be predictive concerning the three rectal toxicity endpoints (AUC>0.62). In the case of 5-year Grade≥2 rectal bleeding, for the LKB model our parameter estimates fall into the range of previous works reviewed by Fiorino et al. [21], with estimated values for TD_{50} ranging from 78.4 to 97.7 Gy, for n ranging from 0.085 to 0.24 and for m ranging from 0.10 to 0.27. Our results for the LKB model are also consistent with Tucker et al. [22], where a mixture LKB model was fit to predict Grade≥2 rectal toxicity. Our estimates for n, however, are lower for all endpoints, showing strong serial behavior for late rectal toxicity. Moreover, the slight differences in NTCP parameter values among different studies

Model		Parameters					AUC
RF-NTCP	DVH only	V21, V40, V62, V70*					0.74
	Integrating DVH and clinical variables	V21, V40, V62, V70, anticoagulants, age*					0.76
LKB-NTCP		TD_50	n	m	Others	LLH	
	Standard DVH only (1)	82.14 (79.96, 84.56)	0.0173 (0.0060, 0.0335)	0.12 (0.10, 0.14)		87.61	0.63
	Modified integrating DVH and clinical variables (4,5)	78.42 (76.22, 82.21)	0.03 (0.02, 0.05)	0.10 (0.08, 0.12)	Age: 0.10 (-0.19 - 0.04) Anticoagulants: -2.76 (-5.33 - 0.18)	85.9	0.64
Logistic regression	DVH only	V70: 0.09 (0.03, 0.15); V62: 0.09 (0.01, 0.17) V40: 0.03 (0.00, 0.07); V21: 0.01 (-0.07, 0.09)				83.47	0.71
	Integrating DVH and clinical variables	V70: 0.11 (0.05, 0.17); V62: -0.09 (-0.17, -0.01) V40: 0.04 (-0.02, 0.10); V21: 0.01 (-0.07, 0.09) age: -0.01 (-0.09, 0.07); anticoagulants: 1.00 (0.00, 2.00)				81.38	0.74

Table 3.2: NTCP models for Grade≥2 rectal bleeding prediction. RF: Random Forest; LKB: Lyman-Kutcher-Burman; LLH: negative of the log-likelihood function; AUC: area under the curve from the receiving operating characteristic curve; *: The parameters retained by the RF-NTCP model are presented in order of decreasing importance; (xx - xx): 95% confidence interval.

Model	Parameters				AUC
RF-NTCP	V21, V55, V48, V73*				0.67
LKB-NTCP	TD_{50}	n	m	LLH	
	80.46 (78.56, 82.52)	0.0103 (0.003, 0.0221)	0.12 (0.10, 0.15)	115.27	0.62
Logistic regression	V73: 0.06 (0.02, 0.10); V55: -0.12 (-0.22, 0.02) V48: 0.09 (0.00, 0.19); V21: 0.01 (-0.05, 0.07)			112.67	0.68

Table 3.3: NTCP models for Overall Grade≥2 rectal toxicity prediction. RF: Random Forest; LKB: Lyman-Kutcher-Burman; LLH: negative of the log-likelihood function; AUC: area under the curve from the receiving operating characteristic curve; *: The parameters retained by the RF-NTCP model are presented in order of decreasing importance; (xx - xx): 95% confidence interval.

can be explained by the shape of the DVH, which is related to the conformity of the technique (relating to the use of IMRT) and the recommended dose-volume constraints (e.g. RTOG and GETUG), the dose calculation algorithm, the delineation guidelines and to the volume being considered (e.g., rectal wall or whole rectum). Had the number of patients been higher, it may have been informative to model patients receiving IMRT and IGRT separately. Interestingly, the logistic regression model fit with the same parameters found by the RF-NTCP model exhibits a better performance in terms of AUC than the LKB-NTCP model. This suggests that RF can be used to find the most predictive parameters (DVH bins or patient characteristics) and then a parametric model can be fit using those parameters.

Different rectal complications are likely to be produced by different rectal pathogenic mechanisms and may therefore be related to different dose and patient-specific parameters [23]. Both representative DVH points as well as patient features can be therefore combined to improve toxicity prediction. In the case of rectal bleeding, our RF-NTCP model selected DVH points as well as age and anticoagulants as additional parameters.

Indeed, the impact of age on rectal toxicity has already been demonstrated in another study [24]. The impact of anticoagulant treatment on rectal bleeding has also been reported [25, 26]. Both RF-NTCP and logistic regression found the highest dose to have a significant impact on rectal toxicity. Moreover, in the case of 5-year Grade$\geq$1 rectal bleeding prediction, V73 is the most important variable, according to the two variable importance measures used (mean decrease accuracy and Gini index). The anticoagulant treatment is the second most important variable involved in the prediction of this toxicity endpoint (from the mean decrease accuracy).

The RF-NTCP model could be useful at different levels on a daily basis. Once a RF-NTCP model has been trained, the model can be stored in software (such as R) and be shared between different institutions to predict the toxicity risk for any given patient. As for EUD [27] or “classic” NTCP models, the RF-NTCP model may be used in the scope of treatment planning, for treatment comparison or as a constraint during the inverse planning step of IMRT.

One limitation of this study is the use of retrospectively-collected data (37%). The use of retrospective data was indeed justified to increase the number of patients allowing for meaningful results to be obtained. However, the cohort could be split into a training and validation data sets for the RF-NTCP model trained to predict 5-year rectal bleeding. The superiority of RF-NTCP over the LKB model, in terms of the average AUC, was confirmed on the validation data across 1000 random splits of the data set.

To conclude, the proposed RF-NTCP framework appears a promising tool in predicting late rectal toxicity following prostate cancer radiotherapy. The RF-NTCP model is a method that can potentially identify a large amount of predictive variables, including both DVH and patient parameters. The interest of RF-NTCP models concerning other radio-induced toxicities could also be evaluated.

* * *

At this point, we have explored the use of parametric and nonparametric models to predict toxicity without taking into account the spatial dose effects. The next part of the thesis introduces two methodologies to perform population analysis using 3D images. The purpose is to characterize spatial dose patterns and anatomical regions implied in toxicity following prostate cancer radiotherapy.

Bibliography

[1] J. D. Ospina, J. Zhu, C. Chira, A. Bossi, J. B. Delobel, V. Beckendorf, B. Dubray, J. L. Lagrange, J. C. Correa, A. Simon, O. Acosta, and R. de Crevoisier, "Random forests to predict rectal toxicity following prostate cancer radiotherapy," *International Journal of Radiation Oncology*Biology*Physics*, 2014.

[2] J. T. Lyman, "Complication probability as assessed from dose-volume histograms," *Radiation Research*, vol. 104, no. 2s, pp. S13–S19, November 1985.

[3] A. Niemierko, "Reporting and analyzing dose distributions: A concept of equivalent uniform dose," *Medical Physics*, vol. 24, no. 1, pp. 103–110, 1997.

[4] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. Munzenrider, B. Shank, L. Solin, and M. Wesson, "Tolerance of normal tissue to therapeutic irradiation," *International Journal of Radiation Oncology*Biology*Physics*, vol. 21, no. 1, pp. 109 – 122, 1991, three-Dimensional Photon Treatment Planning Report of the Collaborative Working Group on the Evaluation of Treatment Planning for External Photon Beam Radiotherapy.

[5] G. Defraene, L. Van den Bergh, A. Al-Mamgani, K. Haustermans, W. Heemsbergen, F. Van den Heuvel, and J. V. Lebesque, "The benefits of including clinical factors in rectal normal tissue complication probability modeling after radiotherapy for prostate cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 82, no. 3, pp. 1233 – 1242, 2012.

[6] T. Rancati, C. Fiorino, G. Fellin, V. Vavassori, E. Cagna, V. Casanova Borca, G. Girelli, L. Menegotti, A. F. Monti, F. Tortoreto, S. Delle Canne, and R. Valdagni, "Inclusion of clinical risk factors into {NTCP} modelling of late rectal toxicity after high dose radiotherapy for prostate cancer," *Radiotherapy and Oncology*, vol. 100, no. 1, pp. 124 – 130, 2011.

[7] L. A. Dawson, M. Biersack, G. Lockwood, A. Eisbruch, T. S. Lawrence, and R. K. Ten Haken, "Use of principal component analysis to evaluate the partial organ tolerance of normal tissues to radiation," *International Journal of Radiation Oncology*Biology*Physics*, vol. 62, no. 3, pp. 829 – 837, 2005.

[8] S. L. Gulliford, S. Webb, C. G. Rowbottom, D. W. Corne, and D. P. Dearnaley, "Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate," *Radiotherapy and Oncology*, vol. 71, no. 1, pp. 3 – 12, 2004.

[9] S. Tomatis, T. Rancati, C. Fiorino, V. Vavassori, G. Fellin, E. Cagna, F. Mauro, G. Girelli, A. Monti, M. Baccolini, G. Naldi, C. Bianchi, L. Menegotti, M. Pasquino, M. Stasi, and R. Valdagni, "Late rectal bleeding after 3d-crt for prostate cancer: development of a neural-network-based predictive model," *Physics in Medicine and Biology*, vol. 57, no. 5, p. 1399, 2012.

[10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[11] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees.* Monterey, CA: Wadsworth & Brooks, 1984.

[12] V. Beckendorf, S. Guerif, E. Le Prisé, J.-M. Cosset, A. Bougnoux, B. Chauvet, N. Salem, O. Chapet, S. Bourdain, J.-M. Bachaud, P. Maingon, J.-M. Hannoun-Levi, L. Malissard, J.-M. Simon, P. Pommier, M. Hay, B. Dubray, J.-L. Lagrange, E. Luporsi, and P. Bey, "70 gy versus 80 gy in localized prostate cancer: 5-year results ofÂ getug 06 randomized trial," *International Journal of Radiation Oncology*Biology*Physics*, vol. 80, no. 4, pp. 1056 – 1063, 2011.

[13] R. de Crevoisier, P. Pommier, J. Bachaud, G. Crehange, C. Boutry, B. Chauvet, T. Nguyen, A. Laplanche, M. Aubelle, and J. Lagrange, "Image-guided radiation therapy (igrt) in prostate cancer: Preliminary results in prostate registration and acute toxicity of a randomized study," *International Journal of Radiation Oncology*Biology*Physics*, vol. 75, no. 3, Supplement, pp. S99 –, 2009, proceedings of the American Society for Radiation Oncology 51st Annual Meeting 51st Annual Meeting of the American Society for Radiation Oncology.

[14] O. H. Behars, D. E. Henson, R. V. P. Hutter, and M. H. Myers, *American joint committee on cancer: Manual for staging of cancer.* Philadelphia: Lippincott, 1992.

[15] A. V. D'Amico, R. Whittington, S. Malkowicz, D. Schultz, K. Blank, G. A. Broderick, J. E. Tomaszewski, A. A. Renshaw, I. Kaplan, C. J. Beard, and A. Wein, "Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for linically localized prostate cancer," *JAMA*, vol. 280, no. 11, pp. 969–974, 1998.

[16] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 2, pp. 18–22, 2002.

[17] G. J. Kutcher, C. Burman, L. Brewster, M. Goitein, and R. Mohan, "Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations," *International Journal of Radiation Oncology*Biology*Physics*, vol. 21, no. 1, pp. 137 – 146, 1991, three-Dimensional Photon Treatment Planning Report of the Collaborative Working Group on the Evaluation of Treatment Planning for External Photon Beam Radiotherapy.

[18] C. Fiorino, T. Rancati, G. Fellin, V. Vavassori, E. Cagna, V. Casanova Borca, G. Girelli, L. Menegotti, A. F. Monti, F. Tortoreto, S. Delle Canne, and R. Valdagni, "Late fecal incontinence after high-dose radiotherapy for prostate cancer: Better prediction using longitudinal definitions," *International Journal of Radiation Oncology*Biology*Physics*, vol. 83, no. 1, pp. 38 – 45, 2012.

[19] T. Rancati, C. Fiorino, G. Gagliardi, G. M. Cattaneo, G. Sanguineti, V. Casanova Borca, C. Cozzarini, G. Fellin, F. Foppiano, G. Girelli, L. Menegotti, A. Piazzolla, V. Vavassori, and R. Valdagni, "Fitting late rectal bleeding data using different NTCP models: results from an italian multi-centric study (airopros0101)," *Radiotherapy and Oncology*, vol. 73, no. 1, pp. 21 – 32, 2004.

[20] M. Söhn, D. Yan, J. Liang, E. Meldolesi, C. Vargas, and M. Alber, "Incidence of late rectal bleeding in high-dose conformal radiotherapy of prostate cancer using equivalent uniform dose-based and dose-volume-based normal tissue complication probability models," *International Journal of Radiation Oncology*Biology*Physics*, vol. 67, no. 4, pp. 1066 – 1073, 2007.

[21] C. Fiorino, R. Valdagni, T. Rancati, and G. Sanguineti, "Dose-volume effects for normal tissues in external radiotherapy: Pelvis," *Radiotherapy and Oncology*, vol. 93, no. 2, pp. 153 – 167, 2009.

[22] S. L. Tucker, L. Dong, W. R. Bosch, J. Michalski, K. Winter, R. Mohan, J. A. Purdy, D. Kuban, A. K. Lee, M. R. Cheung, H. D. Thames, and J. D. Cox, "Late rectal toxicity on {RTOG} 94-06: Analysis using a mixture lyman model," *International Journal of Radiation Oncology*Biology*Physics*, vol. 78, no. 4, pp. 1253 – 1260, 2010.

[23] K.-R. Trott, W. Doerr, A. Facoetti, J. Hopewell, J. Langendijk, P. van Luijk, A. Ottolenghi, and V. Smyth, "Biological mechanisms of normal tissue damage: Importance for the design of {NTCP} models," *Radiotherapy and Oncology*, vol. 105, no. 1, pp. 79 – 85, 2012.

[24] M. Liu, V. Moiseenko, A. Agranovich, A. Karvat, W. Kwan, Z. H. Saleh, A. A. Apte, and J. O. Deasy, "Normal tissue complication probability (NTCP) modeling of late rectal bleeding following external beam radiotherapy for prostate cancer: A test of the quantec-recommended NTCP model," *Acta Oncologica*, vol. 49, no. 7, pp. 1040–1044, 2010, pMID: 20831493.

[25] A. V. D'Amico, J. Manola, E. McMahon, M. Loffredo, L. Lopes, J. Ching, M. Albert, M. Hurwitz, W. W. Suh, T. A. Vivenzio, and C. Beard, "A prospective evaluation of rectal bleeding after dose-escalated three-dimensional conformal radiation therapy using an intrarectal balloon for prostate gland localization and immobilization," *Urology*, vol. 67, no. 4, pp. 780 – 784, 2006.

[26] K. S. Choe, A. B. Jani, and S. L. Liauw, "External beam radiotherapy for prostate cancer patients on anticoagulation therapy: How significant is the bleeding toxicity?" *International Journal of Radiation Oncology*Biology*Physics*, vol. 76, no. 3, pp. 755 – 760, 2010.

[27] Q. Wu, R. Mohan, A. Niemierko, and R. Schmidt-Ullrich, "Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose,"

*International Journal of Radiation Oncology*Biology*Physics*, vol. 52, no. 1, pp. 224 – 235, 2002.

Part III

Analysis of the 3D dose distribution for a better understanding of its implication in rectal toxicity

Chapter 4

A tensor-based population value decomposition to explain rectal toxicity following prostate cancer radiotherapy

This chapter is devoted to develop a method to perform population analysis using 3D images. There are two main contributions: firstly, we present an extension of a 2D population value decomposition (PVD) technique to the tensor case and, secondly, we applied this tensor PVD to characterize a dose pattern possibly correlated to rectal toxicity. The input data of this chapter consists of the 3D dose distribution and toxicity outcome of a cohort of patients treated from prostate cancer with radiotherapy. The contents of this chapter have been presented in the International Conference on Medical Image Computing and Computer-Assisted Intervention held in Nagoya, Japan in September 2013 [1].

4.1 Introduction

Radiotherapy is one of the prescribed treatments for prostate cancer. Its objective is to deliver high doses of radiation to the tumor cells while sparing neighboring organs, often called organs at risk (OAR). The prediction of normal tissue complications has traditionally been addressed using dose-volume histograms (DVH) [2] or DSH, which are reductions of 3D dose distribution received by the organs. Although many studies have shown a correlation between dose, volume and rectal toxicity [3, 4], they lack spatial accuracy and are not able to correlate treatment outcome with a specific dose pattern. One of the main reasons why it is not possible is is that normal tissue complication probability (NTCP) models mainly use dosimetry information via DVH, which does not preserve spatial information. It is a well-known fact in radiotherapy that similar DVHs may come from different dose distributions.

In the context of intensity-modulate radiotherapy (IMRT), more treatments adapted to the individual could be proposed by including new constraints during the treatment planning step. Thus, the identification of anatomical regions correlated with toxicity is crucial to provide new recommendations for treatment planning. Some attempts have been made to introduce the notion of spatiality [5], but within the dose-volume space. Buettner et al have shown that late complications in the rectum are not only related to volumetric aspects of the dose, but particularly to the shape of the dose distribution [6]. More recently, spatial considerations were incorporated by parameterizing the 3D dose patterns [7, 8]. In this way, by selecting a limited set of predictive features, their method outperforms classical models based on DVH/DSH. These approaches still lie in the reduction of feature dimensionality by fitting analytic functions to each dose shape. A principal component analysis (PCA) is developed on the registered dose matrices (expressed as row vectors) to obtain the principal modes of variation [9]. These modes of variation are then used to classify patients according to a specific toxicity. The main issue of this work is the difficulty in interpreting results. Even if some principal components are related to a toxicity event, a spatial dose pattern is still hard to identify.

It would be desirable to have a method that correctly identifies dose patterns. To accomplish this, we propose to generalise and use the population value decomposition (PVD) technique proposed by Crainiceaunu et al [10] for 2D image analysis. We propose a tensor-based generalization of this PVD combined with a method to carry out statistical analysis that mimics voxel-based morphometry (VBM) methods [11, 12]. We show that the direct application of these methods, as proposed in [13], is difficult mainly due to normality assumptions for test statistics. In the end, we show that it is possible to exhibit dose patterns associated with some types of toxicity following prostate cancer radiotherapy treatment.

4.2 Tensor-based population value decomposition

The tensor-based population value decomposition presented herein is a generalization of the PVD presented in [10]. In this former reference, a PVD method is developed using a possibly-truncated singular value decomposition (SVD) along with a PCA over the individuals' left- and right-singular eigen-vector matrices. This method enables compressing, resulting in a more efficient representation of the subjects.

4.2.1 Matrix-based population value decomposition

Consider a population consisting of N 2D images, $\mathbf{Y}_i \in \mathbb{R}^{F \times G}$, $i = 1, 2, \ldots, N$. The objective is to express $\mathbf{Y}_i$ as in equation (4.1),

$$\mathbf{Y}_i = \mathbf{U}\mathbf{R}_i\mathbf{W}, \tag{4.1}$$

where $\mathbf{U}$ and $\mathbf{W}$ are common to the whole population and $\mathbf{R}_i$ is specific to the i-th subject. The steps depicted in [10] to obtain the individual and population matrices are, as follows:

1. Obtain the SVD of each image, $\mathbf{Y}_i = \mathbf{S}_i\mathbf{V}_i\mathbf{D}_i^T$.

2. Create the concatenated resulting matrices, $\mathbf{S} = [\mathbf{S}_1|\dots|\mathbf{S}_N]$ and $\mathbf{D} = [\mathbf{D}_1|\dots|\mathbf{D}_N]$.

3. Perform a PCA over $\mathbf{S}$ to obtain the eigen-vector matrix $\mathbf{U}$ of the covariance matrix $\mathbf{SS^T}$ such as $\mathbf{S} = \mathbf{U}\left(\mathbf{U}^T\mathbf{S}\right)$.

4. Perform a PCA over $\mathbf{D}$ to obtain the eigen-vector matrix $\mathbf{W}$ of the covariance matrix $\mathbf{DD^T}$ such as $\mathbf{D} = \mathbf{W}\left(\mathbf{W}^T\mathbf{D}\right)$.

5. Express $\mathbf{Y}_i$ in terms of the population matrices as

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{S}_i\mathbf{V}_i\mathbf{D}_i^T \\ &= \mathbf{U}\{\left(\mathbf{U}^T\mathbf{S}_i\right)\mathbf{V}_i\left(\mathbf{D}_i^T\mathbf{W}\right)\}\mathbf{W}^T \end{aligned} \tag{4.2}$$

6. Define $\mathbf{R}_i = \left(\mathbf{U}^T\mathbf{S}_i\right)\mathbf{V}_i\left(\mathbf{D}_i^T\mathbf{W}\right)$ to obtain the decomposition $\mathbf{Y}_i = \mathbf{U}\mathbf{R}_i\mathbf{W}$.

The rationale behind this procedure is that the space spanned by $\mathbf{S}_i$ (or $\mathbf{D}_i$) columns contains subject-specific left (or right) singular-vectors that, although not equal, should be similar to each other. Thus, applying a PCA should allow for the calculation of the principal modes of variation of the left (or right) singular-vectors.

In addition, a more efficient representation could be obtained as a previous step to population analysis. This decomposition allows for two different compression levels:

1. At the individual level, some $\mathbf{S}_i$ and $\mathbf{D}_i$ columns could be discarded, taking into account some reconstruction error, leading to the approximation $\mathbf{Y}_i \approx \mathbf{S}_{L_i}\mathbf{V}_{L_iJ_i}\mathbf{D}_{J_i}^T$, where $\mathbf{S}_{L_i}$ is the matrix resulting of taking only the first L_i columns of $\mathbf{S}_i$. $\mathbf{D}_{J_i}$ is analogous with respect to $\mathbf{D}_i$ and $\mathbf{V}_{L_iJ_i}$ consists only of the first L_i columns and the first J_i rows of $\mathbf{V}_i$.

2. At the population level, it is possible to take only the first A columns of $\mathbf{U}$ (or the first B columns of $\mathbf{W}$), considering for example the explained percentage of variance, to build up the projection matrix $\mathbf{U}$ (or $\mathbf{W}$). The same reduction is possible when the PCA is performed over $\mathbf{S}_L = [\mathbf{S}_{L_1}|\dots|\mathbf{S}_{L_N}]$ and $\mathbf{D}_J = [\mathbf{D}_{J_1}|\dots|\mathbf{D}_{J_N}]$, with $L = \sum_{i=1}^N L_i$ and $J = \sum_{i=1}^N J_i$.

As one can notice, the resulting individual specific matrices $\mathbf{R}_i$ are not diagonal. These matrices can be regarded as projections onto a subspace of $\mathbb{R}^{A\times B}$, with $A \leq F$ and $B \leq G$.

4.2.2 Extension to 3D images

The following approach is based on the fact that a 3D image can be regarded as a third-order tensor. The SVD step of the 2D-PVD described above is emulated using a HOSVD [14]. The population consists of 3D images, $\underline{\mathbf{Y}}_i \in \mathbb{R}^{A\times B\times C}$, $i = 1, 2, \dots, N$, regarded as tensors.

The procedure, to obtain a 3D-PVD, was implemented as follows:

1. Obtain the HOSVD of each image, as in equation (4.3),

$$\underline{\mathbf{Y}}_i = \underline{\mathbf{V}}_i \times_1 \mathbf{S}_i^1 \times_2 \mathbf{S}_i^2 \times_3 \cdots \times_M \mathbf{S}_i^M, i = 1, \ldots, N, \tag{4.3}$$

   where $\underline{\mathbf{V}}_i$ is the core tensor and $\mathbf{S}_i^1, \ldots, \mathbf{S}_i^M$ are the corresponding factor matrices. For a 3D image, $M = 3$.

2. Create the M matrices, $\mathbf{S}^k = \left[\mathbf{S}_1^k | \ldots | \mathbf{S}_N^k\right]$, $k = 1, \ldots, M$.

3. Perform a PCA over each $\mathbf{S}^k$ to obtain the projection matrices $\mathbf{P}^k$ such as $\mathbf{S}^k = \mathbf{P}^k \left(\mathbf{P}^k\right)^T \mathbf{S}^k$, $k = 1, \ldots, M$.

4. Express each individual's image as in equation (4.4),

$$\underline{\mathbf{Y}}_i = \underline{\mathbf{\Lambda}}_i \times_1 \mathbf{P}^1 \times_2 \cdots \times_M \mathbf{P}^M, \tag{4.4}$$

   where $\underline{\mathbf{\Lambda}}_i = \underline{\mathbf{V}}_i \times_1 \left(\mathbf{P}^1\right)^T \mathbf{S}_i^1 \times_2 \cdots \times_M \left(\mathbf{P}^M\right)^T \mathbf{S}_i^M$.

As can be seen from equation (4.4), the image corresponding to the i-th individual can be expressed in terms of a specific-individual core tensor, namely $\underline{\mathbf{\Lambda}}_i$, and M population factor matrices. In addition, as in the case of 2D images, two compression levels are allowed:

1. At the individual level, each image can be expressed as in equation (4.5),

$$\underline{\mathbf{Y}}_i \approx \underline{\hat{\mathbf{V}}}_i \times_1 \mathbf{S}_{L_i^1}^1 \times_2 \mathbf{S}_{L_i^2}^2 \times_3 \cdots \times_M \mathbf{S}_{L_i^M}^M, \tag{4.5}$$

   where $\mathbf{S}_{L_i^k}^k$ is a matrix consisting of the the first L_i^k columns of $\mathbf{S}_i^k$, and $\underline{\hat{\mathbf{V}}}_i \in \mathbb{R}^{L_i^1 \times \cdots \times L_i^M}$ is the truncated core tensor.

2. At the population level, the $\mathbf{P}^k$ matrices can be truncated by taking only the first Q_k columns, leading to the reduced matrix $\mathbf{P}^{Q_k}$. The same reduction is possible when the PCA is carried out on $\mathbf{S}^{L_k} = \left[\mathbf{S}_{L_1^k}^k | \ldots | \mathbf{S}_{L_N^k}^k\right]$.

Using both levels of compression, the i-th individual's image can be approximated as in equation (4.6),

$$\underline{\mathbf{Y}}_i \approx \underline{\mathbf{\Lambda}}_i \times_1 \mathbf{P}^{Q_1} \times_2 \cdots \times_M \mathbf{P}^{Q_M}, \tag{4.6}$$

where $\underline{\mathbf{\Lambda}}_i = \underline{\hat{\mathbf{V}}}_i \times_1 \left(\mathbf{P}^{Q_1}\right)^T \mathbf{S}_{L_i^1}^1 \times_2 \cdots \times_M \left(\mathbf{P}^{Q_M}\right)^T \mathbf{S}_{L_i^M}^M$ is the i-th individual's core tensor and $\mathbf{P}^{Q_1}, \ldots, \mathbf{P}^{Q_M}$ are the population factor matrices.

4.2.3 Population analysis

Once each individual has been rewritten as in equation (4.6), statistical analysis can be carried out using the reduced individuals representation, $\underline{\boldsymbol{\Lambda}}_i$, $i = 1, \ldots, N$. If the population can be split into two mutually-exclusive groups, say G_1 and G_2, an element-by-element comparison of $\underline{\boldsymbol{\Lambda}}$ could help to highlight the differences among both groups. This comparison can be done using a two-sample t-test. The resulting p-values could then be used to decide which components are different among the two groups. Let $\underline{\boldsymbol{\Lambda}}^{(G_1)}$ and $\underline{\boldsymbol{\Lambda}}^{(G_2)}$ be the typical core tensors of groups G_1 and G_2, respectively, with components $\lambda^{(G_1)}_{mnp}$ and $\lambda^{(G_1)}_{mnp}$. In addition, let $\underline{\mathbf{Pv}}$ be the resulting p-value tensor, with entries pv_{mnp}. In this work, the components of $\underline{\boldsymbol{\Lambda}}^{(G_1)}$ and $\underline{\boldsymbol{\Lambda}}^{(G_2)}$ are calculated as $\lambda^{(G_j)}_{mnp} = \frac{1}{|G_j|}\sum\limits_{i \in G_j} \lambda^{i}_{mnp}$, $j = 1, 2$ if $pv_{mnp} < p_{tr}$ and $\lambda^{(G_j)}_{mnp} = \frac{1}{|G_1 \cup G_2|}\sum\limits_{i \in G_1 \cup G_2} \lambda^{i}_{mnp}$ otherwise, where p_{tr} is some specified threshold. It means that, for the components where no statistical difference is found, reconstruction is performed using the information of the whole population. For those components where a statistical difference is accepted, component reconstruction is performed with the information coming from either group.

Using this approach, the typical images for each group can be reconstructed as shown in equation (4.7),

$$\underline{\mathbf{Y}}^{(G_j)} = \underline{\boldsymbol{\Lambda}}^{(G_j)} \times_1 \mathbf{P}^{Q_1} \times_2 \cdots \times_M \mathbf{P}^{Q_M}, j = 1, 2. \tag{4.7}$$

4.3 Rectal bleeding study

4.3.1 Data

A total of 63 patients treated for localized prostate cancer with IMRT were included in the study. The treatment planning system used was Pinnacle V7.4 (Philips Medical System). The total prescribed dose was 46 Gy to the seminal vesicles delivered in 4.6 weeks, and 80 Gy to the prostate delivered in 8 weeks, with a standard fractionation of 2 Gy/fraction. The methods of patient positioning, CT acquisition, volume delineations and dose constraints relied upon and complied with GETUG06 recommendations, as described in [15]. The constraints concerning the rectal wall were maximal dose $\leq$76 Gy and V72 Gy$\leq$25% (the volume received by 72% of the rectum wall volume). The CT images consisted of 135 slices whose size in the axial plane was 512 × 512 pixels, with 1mm image resolution and 2mm slice thickness. The median follow-up period was 38 months, with a minimum of 24 months for all patients. Rectal toxicity events were prospectively collected and scored according to the common terminology criteria for adverse events, CTCAE, version 3.0. The endpoint of the study was 2-year Grade ≥ 1 rectal bleeding, excluding acute toxicity. Patients with a history of hemorrhoids were not allowed to be scored as Grade 1 bleeding. In total, 12 patients presented at least a Grade 1 rectal bleeding event, which occurred between 6 and 24 months following treatment.

4.3.2 Image processing

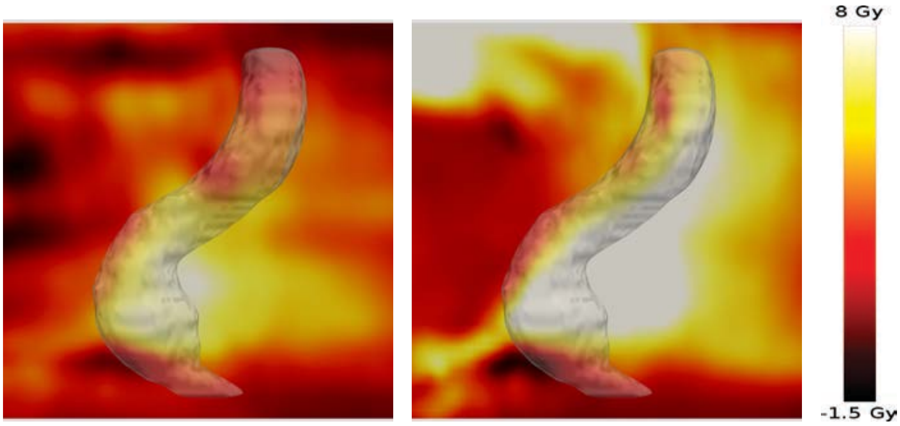
Patients' planned CTs and dose distributions were non-rigidly registered with the demons algorithm [16] on a single coordinate system by combining the CT and organ delineations as explained in [9]. The patients' CTs were first registered and then the same transformation was applied to the dose distributions in order to obtain the mapped dose distributions on a single template.

The dose images were cropped in the axial plane to be $87px \times 87px$ and only 51 slices were considered. Each dose was decomposed using HOSVD, as shown in equation (4.3), and no compression was allowed for the individual or population levels, as the precise identification of dose patterns remained the principal focus of the study.

Once all patients' core tensors were computed, an element-wise comparison was performed in the core tensor space using a two-sample t-test, comparing those patients that presented at least a Grade 1 rectal bleeding event against those that did not. A p-value lower than 0.05 was considered statistically significant. A "typical" core tensor for each group was calculated. Each element of the typical core tensors was computed as an average of the corresponding element across the population. For those elements that were found to be statistically different, the average was computed using only the values of the corresponding group. The average was computed for the remaining elements using the values of both groups. The "typical" dose image for each group was reconstructed using the corresponding "typical" core tensor and population factor matrices. The difference between the two "typical" dose distributions was used to find regions correlated with rectal bleeding Grade ≥ 1.

4.3.3 Results

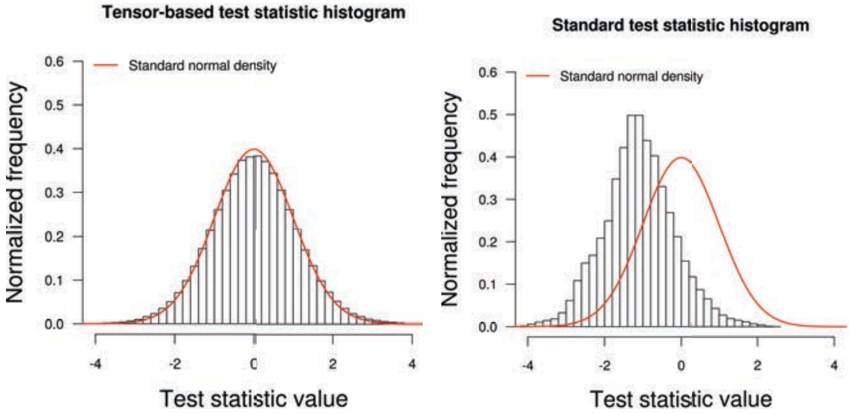
Figure 4.1(a) shows the dose difference between a "typical" patient presenting rectal bleeding and the "typical" patient without sings of rectal bleeding, obtained by the proposed method. This result suggests an excess of dose (around 6 Gy) on patients presenting rectal bleeding Grade ≥ 1 in the area of the posterior wall. For the sake of comparison, the difference between the means of the patients presenting rectal bleeding and those that did not is presented in Figure 4.1(b). The normality assumption was checked using a Shapiro-Wilk test in the core tensor space with a 95% confidence level. Only 22% of the elements rejected the normality assumption. This test was also performed in the native image space and the normality assumption was rejected by 67% of the voxels. Conversely, in the tensor case, no difference was detected in the elements who rejected the normality assumption. Figure 4.2 illustrates the problem of normality assumptions and the validity of performing a two-sample t-test in the native space by demonstrating the histograms of the test statistics for both cases: the tensor case and the native image space. If the normality assumption is valid, it is expected that the normalized histograms look like a standard normal distribution. This is the case for the tensor-based test statistics but not for those computed in the native image space.



(a) Tensor-based PVD difference of means.

(b) Simple difference of means.

Figure 4.1: "Typical" toxic dose distribution minus "typical" non-toxic dose distribution. The rectum of the template patient in the sagital plane is overlaid.



(a) Tensor-based test statistic histogram.

(b) Native space test statistic histogram.

Figure 4.2: Histograms of the test statistics used in the t-test.

4.4 Conclusions

This chapter presented a method to perform population analysis on high order data ($M > 2$) and its application on the study of rectal bleeding following prostate cancer radiotherapy. The presented method does not depend on the type of toxicity and it can be used to investigate other dose patterns correlated with different side effects, even in different organs (such as the bladder). The main contribution of the tensor-based PVD was to enable the use of a classical test statistic to perform group comparison through the construction of the "typical" individual for both groups. The results obtained should be carefully interpreted, as dose excess zone depicted in Figure 4.1(a) could depend on the selected template for registering all the images but also on the registration algorithm. Future work will include the segmentation of this zone and the study of the associated DVH.

The direct application of the t-test on dose images is not advised as the normality assumption is rejected by most of the voxels. This is not the case in the core tensor space, however, highlighting the fact that the zones where both groups are statistically different in the native image space is a subject of current research.

Future work will also include a tensor-based PVD of each group and will attempt to reveal inter-group differences by the comparison of population matrices. The effect of applying compression at the individual or population level and the validity of the method are also a matter for future study.

We have tackled the problem of characterizing dose patterns possibly correlated with toxicity. Future work should be devoted to understand how these dose patterns are produced or, in different words, how the treatment setup (machine configuration) influences the apparition of these patterns.

On the other hand, the question of the localization of the anatomical region affected by this pattern remains to be answered. Next chapter will present another methodology to perform population analysis which aims at identifying anatomical regions implied in toxicity following prostate cancer radiotherapy.

Bibliography

[1] J. D. Ospina, F. Commandeur, R. Ríos, G. Dréan, J. C. Correa, A. Simon, P. Haigron, R. de Crevoisier, and O. Acosta, "A tensor-based population value decomposition to explain rectal toxicity after prostate cancer radiotherapy," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 387–394.

[2] J. Y. Ting, X. Wu, J. A. Fiedler, C. Yang, M. L. Watzich, and A. Markoe, "Dose-volume histograms for bladder and rectum." *International Journal of Radiation Oncology*Biology*Physics*, vol. 38, no. 5, pp. 1105–1111, Jul 1997.

[3] M. Sohn, M. Alber, and D. Yan, "Principal component analysis-based pattern analysis of dose-volume histograms and influence on rectal toxicity." *International Journal of Radiation Oncology*Biology*Physics*, vol. 69, no. 1, pp. 230–239, Sep 2007.

[4] C. Fiorino, V. Vavassori, G. Sanguineti, C. Bianchi, G. M. Cattaneo, A. Piazzolla, and C. Cozzarini, "Rectum contouring variability in patients treated for prostate cancer: impact on rectum dose-volume histograms and normal tissue complication probability." *Radiotherapy and Oncology*, vol. 63, no. 3, pp. 249–255, Jun 2002.

[5] C. Kupchak, J. Battista, and J. Van Dyk, "Experience-driven dose-volume histogram maps of NTCP risk as an aid for radiation treatment plan selection and optimization." *Medical Physics*, vol. 35, no. 1, pp. 333–343, Jan 2008.

[6] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, "Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the mrc rt01 trial (isrctn 47772397)," *Physics in Medicine and Biology*, vol. 54, no. 21, p. 6535, 2009.

[7] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, "Modeling late rectal toxicities based on a parameterized representation of the 3d dose distribution," *Physics in Medicine and Biology*, vol. 56, no. 7, p. 2103, 2011.

[8] R. Lee, E. K. Chan, R. Kosztyla, M. Liu, and V. Moiseenko, "Dose-distance metric that predicts late rectal bleeding in patients receiving radical prostate external-beam radiotherapy," *Physics in medicine and biology*, vol. 57, no. 24, p. 8297, 2012.

[9] B. Chen, O. Acosta, A. Kachenoura, J. Ospina, G. Dréan, A. Simon, J.-J. Bellanger, P. Haigron, and R. Crevoisier, "Spatial characterization and classification of rectal bleeding in prostate cancer radiotherapy with a voxel-based principal components analysis model for 3d dose distribution," in *Prostate Cancer Imaging. Image Analysis and Image-Guided Interventions*, ser. Lecture Notes in Computer Science, A. Madabhushi, J. Dowling, H. Huisman, and D. Barratt, Eds. Springer Berlin Heidelberg, 2011, vol. 6963, pp. 60–69.

[10] C. M. Crainiceanu, B. S. Caffo, S. Luo, V. M. Zipunnikov, and N. M. Punjabi, “Population value decomposition, a framework for the analysis of image populations,” *Journal of the American Statistical Association*, vol. 106, no. 495, pp. 775–790, 2011.

[11] K. Friston, A. Holmes, K. Worsley, J. Poline, C. Frith, and R. Frackowiak, “Statistical parametric maps in functional imaging: a general linear approach,” *Human brain mapping*, vol. 2, no. 4, pp. 189–210, 1994.

[12] J. Ashburner and K. Friston, “Nonlinear spatial normalizatioon using basis functions,” *in Human Brain Mapping*, vol. 7, pp. 254–266, 1999.

[13] O. Acosta, G. Drean, J. D. Ospina, A. Simon, P. Haigron, C. Lafond, and R. De Crevoisier, “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy,” *Physics in medicine and biology*, vol. 58, no. 8, p. 2581, 2013.

[14] L. de Lathauwer, B. de Moor, and J. Vandewalle, “A multilinear singular value decomposition,” *SIAM journal on Matrix Analysis and Applications*, vol. 21, pp. 1253–1278, 2000.

[15] V. Beckendorf, S. Guerif, E. Le Prisé, J.-M. Cosset, A. Bougnoux, B. Chauvet, N. Salem, O. Chapet, S. Bourdain, J.-M. Bachaud, P. Maingon, J.-M. Hannoun-Levi, L. Malissard, J.-M. Simon, P. Pommier, M. Hay, B. Dubray, J.-L. Lagrange, E. Luporsi, and P. Bey, “70 Gy versus 80 Gy in localized prostate cancer: 5-year results of GETUG 06 randomized trial,” *International Journal of Radiation Oncology* Biology* Physics*, vol. 80, no. 4, pp. 1056 – 1063, 2011.

[16] J.-P. Thirion, “Image matching as a diffusion process: an analogy with maxwell’s demons,” *Medical image analysis*, vol. 2, no. 3, pp. 243–260, 1998.

Chapter 5

Spatio-temporal nonparametric mixed-effects model for population analysis with 3D images

This chapter presents a method to perform population analysis using 3D images. With respect to the previous chapter, the objective is to find spatial locations and time points at where differences between two groups of individuals are statistically significant. The framework developed herein deals with spatio-temporal data. In this thesis, however, we present an application, namely a study of rectal bleeding following prostate cancer radiotherapy, which illustrates the cross-sectional case only. From the methodological point of view, the contribution of this chapter is the extension of a nonparametric mixed-effects model to the spatio-temporal case. We developed the formulas for model estimation and we study the properties of the estimators, namely bias, variance and asymptotic behavior which are presented in Appendix B.

Applications of the proposed method in the longitudinal case have been shown in two international conferences [1, 2]. An extended version of this chapter has been submitted for publication in a scientific journal.

5.1 Introduction

The problem of assessing differences between groups by comparing multi-modal images across a population is frequently found in biomedical applications. In order to assess differences among individuals, possibly due to a pathology or correlated to post-treated side effect, those comparisons may be performed either voxel- or image-wise. In image-wise approaches, a set of image features are extracted, for example, by segmenting the same anatomical region across the whole sample of individuals. These features are then used to characterize differences between groups.

An example of an image-wise approach can be found in [3], where the population and individuals' features are separated using matrix or tensor decompositions. Conversely, voxel-based methods provide a way to reveal regional changes between groups by locally computing the difference of a signal across a given population [4]. For voxel-wise comparisons, individual images must be previously co-registered and then voxel-wise statistical tests are performed to assess the differences between two groups or the effect of a covariate. Among these techniques, voxel-based morphometry (VBM) [5] in its standard form assesses the differences between groups by fitting a general linear model (GLM) to the available signal from all subjects at each voxel and relating it to different covariates such as age, gender, diagnosis, cognitive scores, etc. Thus, findings may be related to density changes of a given tissue such as, for example, gray matter (GM) in brain studies [6].

In order to compare similar spatial patterns, in the context of voxel-wise approaches, following spatial normalization, one of the issues deals with the generation of test statistics at voxel level and the correction of these statistics accounting for false positives and the multiple comparison problem. To reduce the effect of misregistration, most of the approaches smooth data, thereby increasing the overlap across individuals. This smoothing step is usually accomplished by convolving the image with a Gaussian kernel. However, this may reduce sensitivity in detecting differences and may lead to a less precise localization of regional differences, as it ignores the spatial correlation structure. Concerning voxel-wise hypothesis testing, it may not fulfill the normality assumptions leading to a misspecification of the testing statistics and its distributions under the null hypothesis. Some solutions have been proposed to tackle the multiple comparison problem, such as the Bonferroni correction, the false discovery rate [7], permutation testing [8] or Gaussian random fields [9].

There are two sources of variability in population analysis: intra- and inter-subject variability. Intra-subject variability is related to a subject's characteristics whereas inter-subject variability refers to more general population characteristics. In order to decrease uncertainty in the estimation of population features, intra-subject variability must be controlled. This is usually accomplished by taking several measures for each individual. A mixed-effects model can then be used voxel-wise to separate population effects from individual effects [10, 11, 12].

In this paper, we propose an approach for voxel-wise population analysis of 3D images based on a nonparametric mixed-effects models. Our choice is motivated by three aspects:

1. Intra-subject variability can be handled at voxel level in both spatial and time domains;
2. The assumption of an *a priori* model is not required;
3. The proposed method generalizes the voxel-wise approach by using a generic (and possibly nonlinear) model instead of the GLM to generate test statistics.

As opposed to classical voxel-based approaches, our proposal exploits intra-individual spatial correlation at each voxel location, leading to a separation of: i) underlying average population behavior; ii) individual specificities and iii) the effect of any given variable. This is not trivial, as intra-individual variability could lead to an over estimation of the variance of an estimator of a population feature, in this way masking the effect of this feature.

Our choice of nonparametric framework is justified by its flexibility and general range of application. Nonparametric methods are a broad class of techniques for data modeling (probability distributions, regression models, etc.), by using the data itself to produce a data description with no intermediation of an *a priori* functional relationship between input and output variables. Image processing applications can be found in [13].

The nonparametric mixed-effects (NPME) model was first proposed in [14] and the mixed-effects (ME) model has been used in VBM to handle intra-subject variability over time [10, 11, 12]. The formulation presented herein stands on the theoretical framework of Hoover et al [14] and Wu and Zhang [15] and on further extensions performed by Ospina et al [16, 1], with the difference being that this present work extends the NPME model to handle spatio-temporal data which enables longitudinal analysis using 3D images at any given time point for each individual. The present approach allows for intra-class variability to be handled and takes advantage of the smoothing step to characterize this intra-class variability in both spatial and time domains. Other approaches model spatio-temporal data, but they do not account for individual effects [17, 18]. To the best of our knowledge, the herein-proposed spatio-temporal modeling procedure has never been applied to 3D-volume images.

The next section introduces the model, including the distributional assumptions and the estimation strategy based on a local-likelihood approach. An application is presented in Section 5.3 to show how our method can be used in a cross-sectional application. The planned dose distributions of two groups of patients are compared using our NPME model to study the local-dose effect relationship in patients treated with external beam radiotherapy for prostate cancer. The chapter ends with our conclusions and proposed future work.

5.2 Statistical model: nonparametric approach

Let's consider a collection of n independent subjects, whose images were acquired or processed following the same protocol, thereby representing the same physical quantity. For each subject i, n_i different images at several time points $t_{i,1}, \ldots, t_{i,n_i}$ are available. The images are supposed to be spatially-normalized to a single coordinate system and therefore re-sampled into the same lattice $D \subseteq \mathbb{R}^3$ containing d points. Let's assume that the response (image intensity), for the i-th subject at time t and point $\mathbf{x}$ is $f_i\left(\mathbf{x}, \mathbf{t}\right)$. Following [14] an approximative model for $f_i\left(\mathbf{x}, \mathbf{t}\right)$ would be

$$f_i\left(\mathbf{x}, t\right) = m\left(\mathbf{x}, t\right) + e_i\left(\mathbf{x}, t\right) + \varepsilon_i\left(\mathbf{x}, t\right), \tag{5.1}$$

where $m\left(\mathbf{x}, t\right)$ is the mean value for the whole population at time t and voxel location $\mathbf{x}$, the function $e_i\left(\mathbf{x}, t\right)$ is the i-th subject deviation from the mean and $\varepsilon_i\left(\mathbf{x}, t\right)$ is a random error. The random term $e_i\left(\mathbf{x}, t\right)$ is included in the model to handle intra-individual variability and it is assumed to be independent of $\varepsilon_i\left(\mathbf{x}, t\right)$. Because this kind of model is composed of population and subject features, it is referred to in the literature as a *mixed-effects model.* In the nonparametric approach, $m(\mathbf{x},t)$ and $e_i(\mathbf{x},t)$ are unknown functions to be locally estimated.

If the n independent subjects can be separated into two different groups (T and its complement T^C), one goal is therefore to identify the differences between the two groups at each spatial point $\mathbf{x}$ and time point t. The model (5.1) could then be extended as follows

$$f_i(\mathbf{x},t) = m(\mathbf{x},t) + g(\mathbf{x},t)\,\phi_i + e_i(\mathbf{x},t) + \varepsilon_i(\mathbf{x},t)\,, \tag{5.2}$$

where $m(\mathbf{x},t)$, $g(\mathbf{x},t)$ and $e_i(\mathbf{x},t)$ are again unknown functions and $\varepsilon_i(\mathbf{x},t)$ is a random error. The binary variable ϕ_i equals one if the i-th individual belongs to the T group and zero otherwise. The term $m(\mathbf{x},t) + g(\mathbf{x},t)$ represents the expected value at point $\mathbf{x}$ and time t for a subject from the T group whereas $m(\mathbf{x},t)$ represents the same for a subject from the T^C group. Furthermore, $e_i(\mathbf{x},t)$ is an individual feature meaning to be the deviation of the i-th individual from the corresponding mean of its group and it is assumed to be random from the population perspective and independent of $\varepsilon_i(\mathbf{x},t)$.

Let $\mathrm{E}[\cdot]$, $\mathrm{Var}(\cdot)$ and $\mathrm{Cov}(\cdot,\cdot)$ denote expectation, variance and covariance operators. For the two different voxels, $\mathbf{x}_1$ and $\mathbf{x}_2$, and two different time point, t_1 and t_2, the distributional assumptions are:

1. $\mathrm{E}[f_i(\mathbf{x}_1,t_1)\,|e_i(\mathbf{x}_1,t_1)] = m(\mathbf{x}_1,t_1) + g(\mathbf{x}_1,t_1)\,\phi_i + e_i(\mathbf{x}_1,t_1)$;
2. $\mathrm{E}[f_i(\mathbf{x}_1,t_1)] = m(\mathbf{x}_1,t_1) + g(\mathbf{x}_1,t_1)\,\phi_i$;
3. $\varepsilon_i(\mathbf{x}_1,t_1)$ and $\varepsilon_j(\mathbf{x}_2,t_2)$ are independent for all i, j, $\mathbf{x}_1$ and $\mathbf{x}_2$ and t_1 and t_2;
4. $e_i(\mathbf{x}_1,t_1)$ and $\varepsilon_j(\mathbf{x}_2,t_2)$ are independent for all i, j, $\mathbf{x}_1$, $\mathbf{x}_2$, t_1 and t_2;
5. $e_i(\mathbf{x}_1,t_1)$ and $e_j(\mathbf{x}_2,t_2)$ are independent for all $i \neq j$, $\mathbf{x}_1$, $\mathbf{x}_2$, t_1 and t_2.

From these assumptions we find that:

1. $\mathrm{Cov}(f_i(\mathbf{x}_1,t_1), f_j(\mathbf{x}_2,t_2)) = 0$ if $i \neq j$;
2. $\mathrm{Cov}(f_i(\mathbf{x}_1,t_1), f_i(\mathbf{x}_2,t_2)) = \mathrm{Cov}(e_i(\mathbf{x}_1,t_1), e_i(\mathbf{x}_2,t_2))$;
3. $\mathrm{Var}(f_i(\mathbf{x}_1,t_1)) = \mathrm{Var}(e_i(\mathbf{x}_1,t_1)) + \mathrm{Var}(\varepsilon_i(\mathbf{x}_1,t_1))$.

The equation (5.2) presents an additive model where the terms are unknown functions to be estimated. The functions $m(\cdot,\cdot)$ and $g(\cdot,\cdot)$ are population characteristics, whereas $e_i(\cdot,\cdot)$, $i = 1,\ldots,n$, represent subject features. The goal here is therefore to estimate the population functions describing the behavior of the subjects in both spatial and time domains.

5.2.1 Model estimation

The idea behind this estimation procedure is to use a Taylor expansion of the functions $m(\cdot,\cdot)$, $g(\cdot,\cdot)$ and $e_i(\cdot,\cdot)$, then to take only the first components of these expansions and to estimate them using a linear mixed-effects regression model.

Assume that all functions in (5.2) admit, for example, a first order approximation around $\mathbf{x_0}$ and t_0 such as $f_i(\mathbf{x},t)$ can be approximated as follows

$$\begin{aligned} f_i(\mathbf{x},t) &\approx m(\mathbf{x}_0,t_0) + \frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}-\mathbf{x}_0) + \frac{\partial m(\mathbf{x}_0,t_0)}{\partial t}(t-t_0) + g(\mathbf{x}_0,t_0)\phi_i \\ &\quad + \frac{\partial g^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}-\mathbf{x}_0)\phi_i + \frac{\partial g(\mathbf{x}_0,t_0)}{\partial t}(t-t_0)\phi_i + e_i(\mathbf{x}_0,t_0) \\ &\quad + \frac{\partial e_i^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}-\mathbf{x}_0) + \frac{\partial e_i(\mathbf{x}_0,t_0)}{\partial t}(t-t_0) + \varepsilon_i(\mathbf{x},t). \end{aligned} \tag{5.3}$$

Now let $\boldsymbol{\beta}_m$, $\boldsymbol{\beta}_g$ and $\mathbf{b_i}$, vectors such that $\boldsymbol{\beta}_m = \left[m(\mathbf{x}_0,t_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}\frac{\partial m(\mathbf{x}_0,t_0)}{\partial t}\right]^T$, $\boldsymbol{\beta}_g = \left[g(\mathbf{x}_0,t_0)\frac{\partial g^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}\frac{\partial g(\mathbf{x}_0,t_0)}{\partial t}\right]^T$ and $\mathbf{b_i} = \left[e_i(\mathbf{x}_0,t_0)\frac{\partial e_i^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}\frac{\partial e_i(\mathbf{x}_0,t_0)}{\partial t}\right]^T$.

Then, if we define $\Delta_{\mathbf{x},t} = \left[1(\mathbf{x}-\mathbf{x}_0)^T(t-t_0)\right]^T$, $f_i(\mathbf{x},t)$ from (5.2) can be approximated as in (5.4),

$$f_i(\mathbf{x},t) \approx \Delta_{\mathbf{x},t}^T\boldsymbol{\beta}_m + \phi_i\Delta_{\mathbf{x},t}^T\boldsymbol{\beta}_g + \Delta_{\mathbf{x},t}^T\mathbf{b}_i + \varepsilon_i(\mathbf{x},t). \tag{5.4}$$

We can also assume that a zero order approximation holds,

$$f_i(\mathbf{x},t) \approx m(\mathbf{x}_0,t_0) + g(\mathbf{x}_0,t_0)\phi_i + e_i(\mathbf{x}_0,t_0) + \varepsilon_i(\mathbf{x},t), \tag{5.5}$$

with $\boldsymbol{\beta}_m = m(\mathbf{x}_0,t_0)$, $\boldsymbol{\beta}_g = g(\mathbf{x}_0,t_0)$ and $\mathbf{b_i} = e_i(\mathbf{x}_0,t_0)$ and $\Delta_{\mathbf{x},t} = 1$.

Let us denote the d image voxels as $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_d$ and by $t_{i_1},\ldots,t_{in_i}$, the n_i different time points for which images for the i-th subject are available. For each subject, the response vector around $(\mathbf{x}_j,t_0)$ is defined as $\mathbf{Y}_i = [f_i(\mathbf{x}_1,t_{i_1})\ldots f_i(\mathbf{x}_d,t_{i_1})\ldots f_i(\mathbf{x}_1,t_{in_i})\ldots f_i(\mathbf{x}_d,t_{in_i})]^T$, that in matrix notation can be expressed as in (5.6),

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{X}^*\mathbf{b}_i + \boldsymbol{\varepsilon}_i, \tag{5.6}$$

with $\mathbf{X}^* = \left[\Delta_{j1,t_{i_1}}\ldots\Delta_{jd,t_{i_1}}\ldots\Delta_{j1,t_{in_i}}\ldots\Delta_{jd,t_{in_i}}\right]^T$, $\mathbf{X}_i = \left[\begin{array}{cc}\mathbf{X}^* & \phi_i\mathbf{X}^*\end{array}\right]$, $\boldsymbol{\beta} = \left[\begin{array}{cc}\boldsymbol{\beta}_m^T & \boldsymbol{\beta}_g^T\end{array}\right]^T$ and $\Delta_{jk,t_{i_r}} = \left[1(\mathbf{x}_k-\mathbf{x}_j)^T(t_{i_r}-t_0)\right]^T$.

Moreover, by concatenating the vectors and matrices as $\mathbf{Y} = \left[\begin{array}{ccc}\mathbf{Y}_1^T & \ldots & \mathbf{Y}_n^T\end{array}\right]^T$, $\mathbf{X} = \left[\mathbf{X}_1{}^T\ldots\mathbf{X}_n{}^T\right]^T$ $\mathbf{Z} = \operatorname{diag}(\mathbf{X},\ldots,\mathbf{X})$ and $\mathbf{b} = \left[\begin{array}{ccc}\mathbf{b}_1^T & \ldots & \mathbf{b}_n^T\end{array}\right]^T$, the model for all individuals can be expressed as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}. \tag{5.7}$$

The vector $\boldsymbol{\varepsilon}$ is the concatenation of the $\boldsymbol{\varepsilon}_i$s vectors, which contain the errors for all individuals at each voxel. The model (5.7) has two random terms: $\mathbf{b}$, which represents intra-individual variability, and the error vector $\boldsymbol{\varepsilon}$, which models inter-individual variability. This model has the form of a linear mixed-effects model (LME) and can be estimated using maximum likelihood (ML) or restricted maximum likelihood (RML) as it is shown in [19]. The estimate of $\boldsymbol{\beta}$ and the predictor for $\mathbf{b}$, say $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{b}}$ respectively, contain the estimates of $m(\mathbf{x}_j, t_0)$, $g(\mathbf{x}_j, t_0)$ and $e_i(\mathbf{x}_j, t_0)$, $i = 1, \ldots, n$.

When assessing subjects at voxel $\mathbf{x}_k$ and at a given time t_0, the contribution of information from other voxels and time points of individuals should be weighted in order to improve local approximation. The information from a voxel close to $\mathbf{x}_k$, coming from an image acquired at time t, close to t_0, should have a bigger weight with respect to another voxel placed far from $\mathbf{x}_k$, corresponding to an image acquired at time t_0, far from t. This weighting is done using a kernel function, $k_{x,t}(\cdot,\cdot) = \frac{1}{h_x h_t} k_x\left(\frac{\cdot}{h_x}\right) k_t\left(\frac{\cdot}{h_t}\right)$. The h_x and h_t parameters are bandwidths for spatial and time kernels, $k_x\left(\frac{\cdot}{h_x}\right)$ and $k_t\left(\frac{\cdot}{h_t}\right)$, respectively. Using this function, we can define a diagonal matrix $\mathbf{K}_i$ containing the weights for all components of the vector $\mathbf{Y}_i$ from (5.6). The total weight matrix is defined as a diagonal matrix, $\mathbf{K}_h = \operatorname{diag}(\mathbf{K}_1, \ldots, \mathbf{K}_n)$.

The model to be estimated is the one from (5.7), but pre-multiplying both sides by $\mathbf{K}_h^{1/2}$. Assuming that $\mathbf{b}$ and $\boldsymbol{\varepsilon}$ are both independent Gaussian vectors with covariance matrices $\mathbf{D}$ and $\mathbf{R}$, respectively, the local joint-weighted log-likelihood can be written as

$$l(\boldsymbol{\beta}, \mathbf{b}|\mathbf{Y}) = -\frac{1}{2}\left(\mathbf{r}^{\mathbf{T}}\mathbf{V}_h\mathbf{r} + \mathbf{b}^T\mathbf{D}^{-1}\mathbf{b}\right) + C, \tag{5.8}$$

with $\mathbf{r} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b}$, $\mathbf{V}_h = \mathbf{K}_h^{1/2}\mathbf{R}^{-1}\mathbf{K}_h^{1/2}$, $C = -\frac{1}{2}\log|\mathbf{R}| - \frac{1}{2}\log|\mathbf{D}| - \frac{N_0}{2}\log(2\pi)$ and $N_0 = d \times \sum_{i=1}^{n} n_i$. By minimizing (5.8), the local polynomial ME (LP ME) estimates are obtained as in [20, 15]. This minimization should be carried out for each voxel $\mathbf{x}_j$ and each time point t_0, in order to have the corresponding estimates $\hat{m}(\mathbf{x}_j, t_0)$ and $\hat{g}(\mathbf{x}_j, t_0)$. The proposed method can be viewed as a generalization of previous works on voxel-based difference-testing. For example, using a zero order approximation for $f_i(\cdot,\cdot)$, as well as assuming that $e_i(\cdot,\cdot)$ is always zero for all individuals (meaning that there is no individual effect) and conveniently choosing the kernel function, it can be shown that $g(\mathbf{x}_j, t_0)$ is the difference, at $\mathbf{x}_j$, between the means of individuals belonging to the T group and of those which do not, at time $t = t_0$. Notice that data from all individuals are used. Even if for each group a different mean function is computed, the population and intra-individual variances are calculated using all available data. The formulas concerning the estimators for the fixed and random effects and their corresponding variances are provided in Appendix B.2.

5.2.2 Hypothesis testing

The purpose of comparing individuals is to test, both, voxel- and time-wise, if both groups are statistically different. In this work, the estimates divided by their corresponding standard errors are used as test statistics to assess the significance of the estimates. Following standard procedures [20, 21] one could show that the null distribution of these statistics is asymptotically standard normal (see Appendix B.4). Dividing $\hat{\beta}_1^g$ (the estimate of $g\left(\mathbf{x}_j, t_0\right)$) by its standard deviation (say $sd\left[\hat{\beta}_1^g\right]$, obtained via the fitting procedure under these hypothesis) results in Z-scores, allowing for p-values to be computed.

If there is only one image available for each individual, and if in (5.2) $e_i\left(\mathbf{x}, t\right) = 0$ for $i = 1, \ldots, n$, meaning that there are no individual effects, the model becomes a simple nonparametric model. This model could be estimated voxel-wise with a general linear model, (5.7) thus becoming $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. In addition, when h_t is big enough, the time effect disappears from the model. With the kernel function carefully selected, it is possible to show that using a local likelihood approach is equivalent to using a linear model voxel-wise on smoothed images (see Appendix B.1). This implies that when no individual effects are considered, the local likelihood approach is equivalent to performing a linear regression on smoothed images, as in classical VBM [5].

5.3 Application

A scenario with real data was devised to illustrate the application of the proposed method. An application on data coming from a study of rectal bleeding following prostate cancer treatment with external beam radiotherapy, which demonstrates how the model can be used in a cross-sectional study and how intra-individual variability is handled in the spatial domain to reveal anatomical regions probably related to rectal bleeding following prostate cancer radiotherapy. The pipeline followed in this applications is depicted in Figure 5.1. The inputs of our method are 3D images and their covariates. These covariates can be, in the case of medical images, patient characteristics such as the age or treatment outcome. The 3D images must be spatially-normalized. This process is problem-dependent and different methods for image registration may be used accordingly. Once images are spatially-normalized, a nonparametric mixed-effects model is fitted voxel-wise. The fitting of this model produces 3D volumes of estimators of the model and of their standard errors. With this information, 3D volumes of statistics are generated to test the different hypotheses of interest. The general procedure is depicted in Algorithm 1. If the model must be estimated at time points $t_0, t_1, \ldots, t_{k-1}, t_k$ and spatial locations (voxels) $\mathbf{x}_1, \ldots, \mathbf{x}_d$, then $d \times (k+1)$ linear mixed-effects regressions are required. The estimation of our proposed model was implemented using R software [22, 23, 24]. The statistical significance was evaluated using standard theory (Wald test) for results concerning the estimation of $g\left(\mathbf{x}, t\right)$. Test statistics were generated voxel- and time-wise as $Z\left(\mathbf{x}, t\right) = \frac{\hat{g}(\mathbf{x},t)}{sd[\hat{g}(\mathbf{x},t)]}$, which under the null hypothesis H_0: $g\left(\mathbf{x}, t\right) = 0$ could be approximated to a standard normal distribution, allowing for p-values to be computed.

Intuitively, one would presume that if image resolution is good enough a low order approximation would lead to good results. In both applications, we obtained the same results using a zero-order approximation and a first-order approximation. When estimating the model for an specific voxel only the adjacent voxels were considered, which

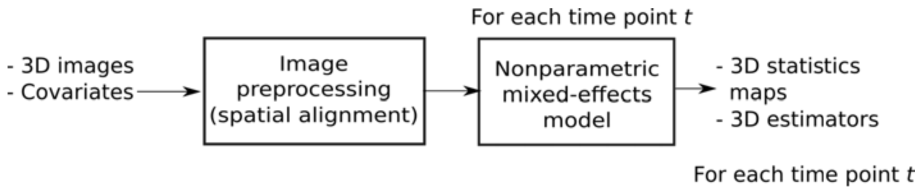


Figure 5.1: General VBM flowchart. We propose to use a nonparametric mixed-effect model voxel-wise to compare between groups.

means that $27 = 3^3$ voxels are simultaneously considered to estimate the model at one voxel (at a specific time t). This implies that for estimating the model at one voxel, $27 \times N$ observations are used in the regression model, where $N = \sum_{i=1}^{n} n_i$ with n_i the number of available images for the i-th individual and n the total number of subjects. The number of voxels d is then the required number of regressions to fully estimate the model at some specific time point.

Algorithm 1 Pseudo code for the implementation of the spatio-temporal NPME model

```
$t_0, t_1, \ldots, t_{k-1}, t_k$: different time points at which the model will be estimated.
$\mathbf{x}_{s_1}, \ldots, \mathbf{x}_{s_d}$: different voxels at which the model will be estimated.
Build up the matrices $\mathbf{Y}$, $\mathbf{X}$ and $\mathbf{Z}$ of equation (5.7).
for i=0:k do
  Calculate the weighting matrix $K_h$
  for j=1:d do
    Minimize the local log-likelihood of equation (5.8) to obtain the model coefficients and their standard errors.
  end for
end for
Compute the test statistics and the associated p-values under the null hypothesis that there is no difference among groups.
```

5.3.1 Study of local dose and rectal toxicity relationship in prostate cancer treatment with radiotherapy

In prostate cancer radiotherapy, the aim is to deliver high doses of radiation to the prostate while sparing normal tissues, namely the bladder and rectum. Predicting side-effects following prostate cancer radiotherapy is usually carried out via a reduction in 3D dose distribution, namely the dose-volume histogram (DVH) and a family of logistic models known as normal tissue complication probability (NTCP) models. As these NTCP models use a reduction of the 3D dose distribution, they lack spatial accuracy. The goal of this voxel-wise analysis was to investigate the relationship between the treatment plan and the side-effects produced by a harmful over-irradiation of the neighboring organs at risk, specifically the rectum. We selected 63 prostate cancer patients treated with external radiotherapy in the same institution between July 2006 and June 2010.

Clinical outcomes (rectal bleeding, within a two-year follow-up) and 3D dose distributions were available. For each patient, the dose was computed on 512×512×256 pelvic computed tomography (CT) scans, using manual delineations, according to standard clinical protocol. Rectal toxicity was scored according to the common terminology criteria for adverse events (CTCAE) version 3.0 and the endpoint of this study was two-year grade$\geq$1 rectal bleeding. The image pre-processing step was performed as follows: We first mapped the organs to a common template using a hybrid organ/intensity non-rigid registration method, allowing for the alignment of barycentres and neighboring structures across the population; then, using the obtained transformations, the dose matrices were mapped on the same template, as described in [25]. This preprocessing step is depicted in Figure 5.2. Once we have aligned the dose distributions, two groups were constituted, namely patients who presented rectal bleeding grade$\geq$1 within a two-year period following treatment and those which did not (12 vs. 51 patients, respectively). Thus, according to (5.2), the planned dose distribution for the individual i, along the rectum, was modeled by the term $f_i(\mathbf{x})$ (without the time variable). All analyses were restricted to the rectum, as the registration method was designed for accuracy within the rectum and the rectal wall. The underlying mean of each of the groups at voxel $\mathbf{x}$ was $m(\mathbf{x}) + g(\mathbf{x})$ for patients presenting rectal bleeding and $m(\mathbf{x})$ for those that did not, respectively. The individual specific deviation from the mean behavior was $e_i(\mathbf{x})$, $i = 1, \ldots, 63$. This special case is described in the Appendix B.2.1.2. As the side effects are likely to be related to over-irradiation, the alternative hypothesis in this application was H_a: $g(\mathbf{x}) > 0$. The kernel bandwidth in the spatial domain was fixed to $h_x = 1.5$ mm.

5.3.1.1 Results

Figure 5.3 presents two different views of the region identified by the proposed method. In this region, significant dose differences appear between both groups, suggesting over-irradiation on patients who presented rectal bleeding. For the sake of clarity the prostate, bladder and rectum were overlaid, but only the dose within the rectum and rectal wall was analyzed. Using a threshold of 0.01 for the p-values, one zone of over-irradiation was identified beside the anterior rectal wall.

5.4 Conclusions and discussion

In this chapter, we presented a new method to perform voxel-wise comparisons with 3D images. This proposed method can be seen as a generalization of classical VBM approaches and it extends the capabilities of the current methods. It exploits intra-individual spatio-temporal correlation at each voxel location, leading to a separation of population features and individual features. We applied the methodology on clinical data, to show that the method can be used to find spatial patterns related to, for example, side-effects following prostate cancer radiotherapy.It is important to notice that even when there is only one image available for each patient, the method still allows for the characterization of the intra-subject variability at voxel level.

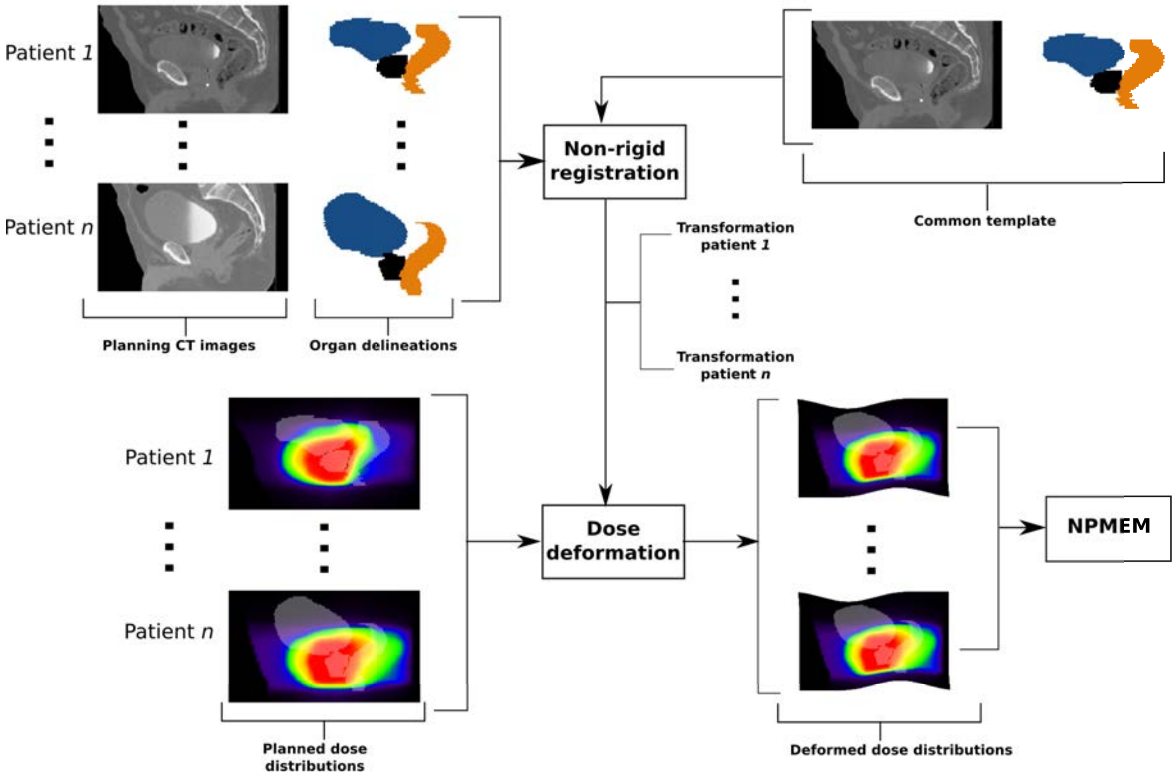
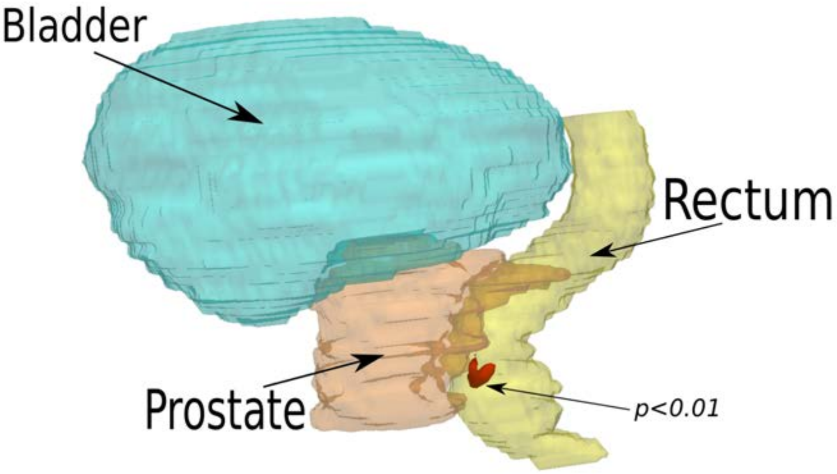
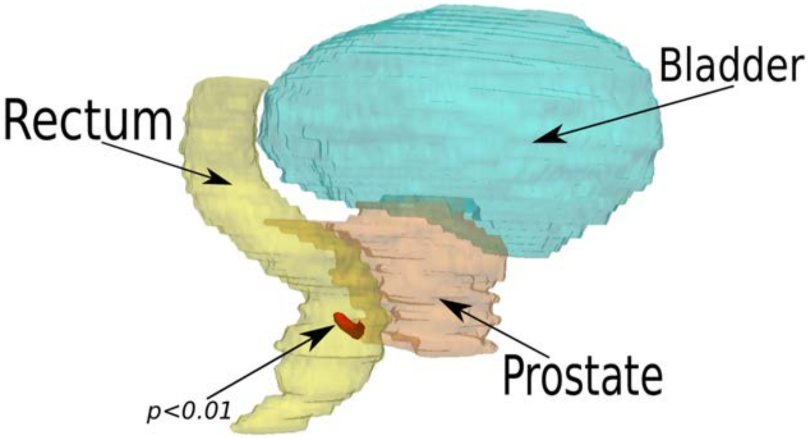


Figure 5.2: Preprocessing step in the rectal toxicity study. For each patient we have a CT scan, the manual organ delineations and planified dose distribution. All patient CT scans were non-rigidly registered towards the same common template using CT scans and organ delineations. Then, the obtained transformations were used to propagate the planned dose distributions which are analyzed using the nonparametric mixed-effects (NPME) model.



(a) Rectal bleeding A



(b) Rectal bleeding B

Figure 5.3: Zones found statistically over irradiated (p-value<0.01) on patients presenting rectal bleeding.

The local likelihood approach used to estimate the nonparametric model requires the definition of bandwidths to perform the weighting process. These parameters have been shown to affect results and some methods, based on cross-validation to find the optimal bandwidth, can be found in the literature. However, in the context of medical images, with information from millions of voxels to be processed, these methods are difficult to implement and new strategies must be developed. In the context of dose-image comparison in the prostate, as dose images are smooth, we set $h_x = 1.5$ mm.

5.4.1 Future work: bandwidth selection

There are several approaches found in the literature to chose the optimal bandwidth. Some authors suggest using different bandwidths to estimate fixed and random effects [26]. For such a purpose, bandwidth selection strategies based on cross-validation can be used [15, pp. 87]. However, they are computationally expensive to be applied in voxel-wise models and efficient methods should be developed in order to use a cross-validation bandwidth selection strategy. For instance, the optimal bandwidth could be selected as the one that minimizes the leave-one-subject-out cross-validation error shown in (5.9),

$$CV\left(h_x, h_t\right) = \frac{1}{N_0} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{k=1}^{d} \left(\hat{f}_i^{(h_x,h_t)} \left(\mathbf{x}_k, t_{i_j}\right) - y_{\mathbf{x}_k, t_{i_j}} \right)^2, \tag{5.9}$$

where $\hat{f}_i^{(h_x,h_t)} \left(\mathbf{x}_k, t_{ij}\right) = \hat{m}^{(h_x,h_t)} \left(\mathbf{x}_k, t_{i_j}\right) + \hat{g}^{(h_x,h_t)} \left(\mathbf{x}_k, t_{i_j}\right)$ is the unconditional estimate of $y_{\mathbf{x}_{s_k}, t_{ij}}$ (the observed value for the i-th individual at voxel $\mathbf{x}_k$) using a model estimated without the i-th individual and the bandwidths h_x and h_t. By choosing h_x and h_t in this way, the conclusions will be supported on the fitted model that best represents the population. Note that kernel bandwidths h_x and h_t can be adapted at each voxel $\mathbf{x}$ and time point t. However, this method, even if accurate, is intensively CPU-consuming. An alternative to cross-validation based procedures is to use the marginal and conditional Aikake's information criteria (mAIC and cAIC, respectively). Fang [27] demonstrates that mAIC is asymptotically equivalent to the leave-one-subject-out cross-validation error and cAIC to the leave-one-observation-out cross validation. Instead of minimizing (5.9), we could minimize

$$mAIC\left(h_x, h_t\right) = \left(\mathbf{Y} - \mathbf{X}\hat{\beta}_{h_x,h_t}\right)^T \Omega_h \left(\mathbf{Y} - \mathbf{X}\hat{\beta}_{h_x,h_t}\right), \tag{5.10}$$

where β_{h_x,h_t} and Ω_h are the parameter vector and the covariance matrix given in Appendix B.2. The advantage of minimizing (5.10) over minimizing (5.9) is that for each couple (h_x, h_t) only a model is estimated whereas minimizing (5.9) requires to estimate n models if we dispose of n subjects. Future work includes locally adapted bandwidths.

We conclude our work on the dose distribution with our spatio-temporal NPME model. Using the dose distribution, we developed two different and complementary methodologies to find dose patterns and anatomical regions implied in toxicity following prostate cancer radiotherapy. While the dose pattern can be used to understand how the ballistics is implied in the toxicity , the anatomical regions could be used to define new constraints during the treatment planning step in IMRT.

Bibliography

[1] J. Ospina, O. Acosta, R. de Crevoisier, J. Correa, and P. Haigron, “A non parametric mixed-effect model for population analysis: Application to Alzheimer’s disease data,” in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, May 2012, pp. 1124–1127.

[2] J. D. Ospina, P. Benquet, F. Commandeur, R. de Crevoisier, J. C. Correa, and O. Acosta, “A mixed effects spatio temporal model to assess the joint effect of ageing and Alzheimerâs disease in gray matter volume,” in *MICCAI-NIBAD’12: Workshop on Novel Imaging Biomarkers for Alzheimer’s Disease and Related Disorders*, Nice, France, 2012, pp. 190–7.

[3] C. M. Craniceanu, B. S. Caffo, S. Luo, V. M. Zipunnikov, and N. M. Punjabi, “Population value decomposition, a framework for the analysis of image populations,” *Journal of the American Statistical Association*, vol. 106, no. 495, pp. 775–790, 2011.

[4] K. Friston, A. Holmes, K. Worsley, J. Poline, C. Frith, and R. Frackowiak, “Statistical parametric maps in functional imaging: a general linear approach,” *Human brain mapping*, vol. 2, no. 4, pp. 189–210, 1994.

[5] J. Ashburner and K. Friston, “Voxel-based morphometry-the methods,” *Neuroimage*, vol. 11, no. 6 Pt 1, pp. 805–821, Jun 2000.

[6] Y. Qiang, Z. Ling, and C. Qin, “Voxel-based morphometric study of brain structure in alzheimer’s disease.” *Sichuan Da Xue Xue Bao Yi Xue Ban (Journal of Sichuan University. Medical science edition)*, vol. 39, no. 3, pp. 496–499, May 2008.

[7] C. R. Genovese, N. A. Lazar, and T. Nichols, “Thresholding of statistical maps in functional neuroimaging using the false discovery rate.” *Neuroimage*, vol. 15, no. 4, pp. 870–878, Apr 2002.

[8] T. E. Nichols and A. P. Holmes, “Nonparametric permutation tests for functional neuroimaging: a primer with examples,” *Human Brain Mapping*, vol. 15, no. 1, pp. 1–25, Jan 2002.

[9] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin, “A three-dimensional statistical analysis for CBF activation studies in human brain,” *Journal of Cerebral Blood Flow & Metabolism*, vol. 12, no. 6, pp. 900–918, Nov 1992.

[10] K. J. Friston, K. E. Stephan, T. E. Lund, A. Morcom, and S. Kiebel, “Mixed-effects and fMRI studies,” *NeuroImage*, vol. 24, pp. 244–252, 2005.

[11] J. Mumford and T. Nichols, “Modeling and inference of multisubject fMRI data,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 25, no. 2, pp. 42 –51, march-april 2006.

[12] A. Roche, S. Mériaux, M. Keller, and B. Thirion, "Mixed-effects statistics for group analysis in fMRI: A nonparametric maximum likelihood approach," *Neuroimage*, vol. 38, pp. 501–510, 2007.

[13] Y.-Q. Song, Z. Liu, J.-M. Chen, F. Zhu, and C.-H. Xie, "Medical image segmentation based on non-parametric mixture models with spatial information," *Signal, Image and Video Processing*, vol. 6, no. 4, pp. 569–578, 2012.

[14] D. R. Hoover, J. Rice, C. Wu, and L. P. Yang, "Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data," *Biometrika*, vol. 85, pp. 809–822, 1998.

[15] H. Wu and J.-T. Zhang, *Nonparametric Regression Methods for Longitudinal Data Analysis.* New Jersey: John Wiley & Sons, Inc., 2006.

[16] J. D. Ospina, O. Acosta, G. Dréan, G. Cazoulat, A. Simon, J. C. Correa, P. Haigron, and R. de Crevoisier, "Spatial nonparametric mixed-effects model with spatial-varying coefficients for analysis of populations," in *Machine Learning in Medical Imaging*, ser. Lecture Notes in Computer Science, K. Suzuki, F. Wang, D. Shen, and P. Yan, Eds. Springer Berlin Heidelberg, 2011, vol. 7009, pp. 142–150.

[17] A. E. Gelfand, H.-J. Kim, C. F. Sirmans, and S. Banerjee, "Spatial modeling with spatially varying coefficient processes," *Journal of the American Statistical Association*, vol. 98, no. 462, pp. 387–396, 2003.

[18] A. E. Gelfand, S. Banerjee, and D. Gamerman, "Spatial process modelling for univariate and multivariate dynamic spatial data," *Environmetrics*, vol. 16, no. 5, pp. 465–479, 2005.

[19] J. C. Pinheiro and D. M. Bates, *Mixed-Effects models in S and S-Plus*, 1st ed. Springer, 2000.

[20] H. Wu and J.-T. Zhang, "Local polynomial mixed-effects models for longitudinal data," *Journal of the American Statistical Association*, vol. 97, no. 459, pp. 883–897, 2002.

[21] T. L. Lai and M. Shih, "Nonparametric estimation in nonlinear mixed effects models," *Biometrika*, vol. 90, no. 1, pp. 1–13, 2003.

[22] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2011, ISBN 3-900051-07-0.

[23] D. Bates, M. Maechler, and B. Bolker, *lme4: Linear mixed-effects models using S4 classes*, 2011, r package version 0.999375-42.

[24] C. Bordier, M. Dojat, and P. Lafaye de Micheaux, “Temporal and spatial independent component analysis for fMRI data sets embedded in the AnalyzeFMRI R package,” *Journal of Statistical Software*, vol. 44, no. 9, pp. 1–24, 2011.

[25] O. Acosta, G. Dréan, J. D. Ospina, A. Simon, P. Haigron, C. Lafond, and R. de Crevoisier, “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy,” *Physics in Medicine and Biology*, vol. 58, no. 8, p. 2581, 2013.

[26] J. A. Rice and B. W. Silverman, “Estimating the mean and covariance structure nonparametrically when the data are curves.” *Journal of Royal Statistical Society, Series B*, vol. 57, pp. 253–259, 1991.

[27] Y. Fang, “Asymptotic equivalence between cross-validation and aikake information criteria in mixed-effects models,” *Journal of Data Science*, vol. 9, pp. 15–21, 2011.

Conclusion and perspectives

Conclusion and perspectives

Contributions of the thesis

We presented different modeling strategies aiming at improving toxicity prediction following prostate cancer radiotherapy. We have characterized cumulative risk of bladder and rectal toxicity and identified different predictors of these two toxicities and their associated symptoms.

Using a RF-NTCP model, we integrated clinical variables and treatment parameters into toxicity prediction. Beyond the statistical significance, our RF-NTCP modeling framework used the impact of each variable on toxicity prediction. The pertinence of this approach was demonstrated by the AUC values obtained for the RF-NTCP model, which were always higher than those obtained by the classic LKB-NTCP and logistic regression models.

In the context of 3D dose analysis, two different methods were developed. We generalized a 2D population value decomposition (PVD) technique to a tensorial framework. This tensor-based PVD was applied to find a pattern of dose possibly correlated to rectal toxicity. Although our application is specific to rectal toxicity characterization, the developed methodology is quite general and can be applied to characterize differences between groups in 3D image population analysis. The other contribution to 3D image population analysis was the development of a spatio-temporal nonparametric mixed-effects model that is able to handle intra-subject variability in both spatial and time domains. This method was applied in a cross-sectional study to analyze rectal bleeding following external beam radiotherapy for prostate cancer, but it also extents the voxel-wise morphometry framework and has application in other domains, as for example in modeling the progression of neurodegenerative diseases such as Alzheimer's disease. For this nonparametric model, we provided the formulas for the estimators and the asymptotic results that allow for hypothesis testing.

Predictive model comparison

Different modeling approaches have been developed in this thesis with the aim of improving toxicity prediction. Each one of the modeling approaches presents advantages and disadvantages with respect to the information they require, the number of patient needed, the model flexibility and algorithmic implementation. The path we followed started from classical binary logistic regression and Kaplan-Meier risk estimation, and continued throughout nonparametric models, namely random forest, to end up with nonparametric models for 3D image population analysis. As information complexity increased, the complexity of the proposed models also did. One of our objectives was to highlight anatomical regions and dose patterns correlated to rectal toxicity. However, when working on the 3D dose distribution, we found the difficulty of including patient parameters due to a restricted number of available patients. Table 5.1 summarizes the main advantages and disadvantages of the different approaches considered in the thesis.

Models	**Advantages**	**Disadvantages**
Classical regression modeling	- Easy to interpret - Fast implementation	- Makes strong hypothesis about the existing relationships between output and input variables -Lacks of spatial accuracy
Random forest normal tissue complication probability model	- Integration of DHV and patient parameters - The importance of a variable is assessed in terms of the impact on toxicity prediction	- Requires an important number of patients to be trained - Lacks of spatial accuracy - Lacks of and odds ratio-like measure for the selected variables
3D population value decompostion	- It allows for image compression - The analysis are performed in a different space, where the original spatial relationships do not remain	- Lacks of the capacity to integrate patient characteristics - Its applicability is restricted to cross-sectional studies.
Spatio temporal nonparametric mixed-effects model	- Handling of intra-individual variability at voxel-level - Non *a priori* assumption about the relationships among variables.	- It is time and computing consuming - Requires an additional step to select the bandwidth values

Table 5.1: Predictive model comparison: main advantages and disadvantages of the different approaches considered in the thesis

Limitations of the work

One limitation to develop and apply statistical methodologies is the data availability. To develop our models, we used both prospective and retrospective data. The use of retrospectively collected data was justified to increase the statistical power and to be able to implement nonparametric methods such as random forest. The data collection process was by itself a challenge in this work. Different data sets were combined and two medical doctors worked on the scoring of toxicity events using a standard scale. The heterogeneity in the data may be therefore a source of uncertainty.

Another limitation of our work is the use of information from the planning step. Indeed, it has been shown that planned dose distribution and actual delivered dose may significantly differ for normal tissues due to anatomical variations. This divergence may be at the origin of serious uncertainties in toxicity prediction. One, almost unrealistic, solution could have been to analyze per-treatment images, such as CBCT.

The applications of the proposed methods to predict bladder toxicity was relatively limited. We haven chosen to illustrate the feasibility of our methodological developments mainly on rectal toxicity because, as we focused more in methodological rather than clinical aspects, the comprehensive literature in rectal toxicity gave us a framework to compare our work. This is not the case in bladder toxicity. Another reason was that the 3D data used in this thesis was the outcome of another thesis in image registration, simultaneously developed in our Laboratory. In that thesis, different algorithms for inter-individual co-registration of delivered-dose-to-the-rectum images were developed. When new registration methods for the images of the dose received by the bladder will be developed, the methods proposed in this present thesis will be implemented.

In the case of the proposed 3D approaches, the pattern of dose and the anatomical region found in this thesis need to be validated using an independent data set.

Another issue is the joint modeling of the impact of the dose distribution shape, clinical factors and patient characteristics. A recurrent criticism to classic NTCP models is that they lack of spatial accuracy and the capacity to integrate dosimetric and patient parameters. The use of a RF-NTCP model answer to this issue. However, the proposed 3D approaches aiming at be spatially accurate lack of the capacity to integrate a significant number of patient characteristics, and this must be matter of further research.

We lack of tools to validate the likelihood of the dose pattern found using the tensorial population value decomposition technique and future work in this area should be developed. Another issue of the proposed methods to perform 3D image population analysis is the assumption that a good inter-individual spatial alignment is possible. The impact of the registration methods needs to be studied.

Perspectives

From a methodological point of view, there is still room for improvement in the approaches proposed in this thesis. For instance, highly correlated variables have been shown to affect the performance of random forest. Our RF-NTCP framework lacks of an automatic variable selection procedure to deal with the highly correlated DVH bins. Another possibility with RF-NTCP would be the development of RF-NTCP models for specific symptoms, which requires bigger data bases.

Our tensor-based population value decomposition technique could be exploited to perform classification. Using the information embedded in the core tensors it would be possible to define metrics allowing for distance computation and using these distances to perform clustering. Another possibility is to develop optimization algorithms to avoid the patterns correlated to side-effects during treatment planning.

The spatio-temporal nonparametric mixed-effects model provides a flexible framework to perform voxel-wise analysis. It extends the already available voxel-base morphometry methods. However, it lacks of an automatic procedure to select the kernel bandwidths. We believe that this can be done using the asymptotic equivalences between cross-validation criteria and marginal Aikake's information criteria.

Future work should include a joint analysis of the dose pattern and anatomical regions implied in toxicity following prostate cancer radiotherapy as well as their use in the clinic, as for example in inverse planning.

On the other hand, if toxicity is explained by patients and dosimetric parameters, the risk of toxicity is likely also related to an individual radiosensibility. Current works in radiobiological laboratories aim to identify biological tests (such as apoptosis in lymphocyte or DNA repairs in fibroblasts) to assess such individual radiosensibility. These biological parameters could be also integrated as another input in our predictive models

From a clinical point of view, the proposed methodologies can be applied to study other radio-induced toxicities for other tumor localizations, within the frame of radiotherapy prospective trials.

Conclusion

In this thesis, we applied different modeling approaches to improve the prediction of toxicity events following external beam radiotherapy for prostate cancer. We implemented classic models to predict toxicity which allowed for the risks of bladder and rectal toxicities. We identified, however, some drawbacks of these classic approaches, namely i) the difficulty to include in a single model both dosimetric and patient parameters and ii) the lack of spatial accuracy. Our random forest normal tissue complication probability model answers to the first issue. Indeed, we have demonstrated that the RF-NTCP was able to predict rectal toxicity and that it was a strong competitor for classic NTCP models. Our tensorial value decomposition and spatio-temporal nonparametric mixed-effects models answer to the second issue, demonstrating that it was possible to find dose patterns and anatomical regions correlated to rectal toxicity.

Our 3D approaches can still be improved. Further research aiming at integrating patient parameters to the 3D image analysis should be conducted.

Appendices

List of publications

International journal papers

- Ospina J.D., Zhu J., Chira C., Bossi A., Delobel J.B., Beckendorf V., Dubray B., Lagrange J.L., Correa J.C., Simon A., Acosta O., and de Crevoisier R., “Random forests to predict rectal toxicity following prostate cancer radiotherapy,” International Journal of Radiation Oncology*Biology*Physics, Accepted for publication on April, 14th 2014.
- Mathieu, R., J. D. Ospina, V. Beckendorf, J. B. Delobel, T. Messai, C. Chira, A. Bossi et al. “Nomograms to predict late urinary toxicity after prostate cancer radiotherapy.” World journal of urology, 1-9(2013).
- Acosta, O., Dréan, G., Ospina, J. D., Simon, A., Haigron, P., Lafond, C., and de Crevoisier, R. “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy.” Physics in medicine and biology, 58(8), 2581 (2013).

Peer-reviewed conference papers

- Ospina, J. D., Commandeur, F., Rìos, R., Dr éan, G., Correa, J.C., Simon, A., Haigron, P., de Crevoisier, R., and Acosta, O. “A Tensor-Based Population Value Decomposition to Explain Rectal Toxicity after Prostate Cancer Radiotherapy.” In Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013, pp. 387-394. Springer Berlin Heidelberg, 2013.
- Dréan, G., Acosta, O., Ospina,, J.D. Voisin, C., Rigaud, B., Simon, A., Haigron, P., and de Crevoisier, R. “How to identify rectal sub-regions likely involved in rectal bleeding in prostate cancer radiotherapy?” In IX International Seminar on Medical Information Processing and Analysis, pp. 89220L-89220L. International Society for Optics and Photonics, 2013.
- Ospina J.D., Acosta O, de Crevoisier R., Correa JC, Haigron P, “A non parametric Mixed-Effect Model for population Analysis: Application to Alzheimer’s Disease Data” in The IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, SPAIN. IEEE ISBI p. 1124-27

- Oscar, A., Dréan, G., Ospina, J.D., Simon, A., Haigron, P., and de Crevoisier, R. “Explaining relationships between local dose and rectal toxicity in prostate cancer radiotherapy with voxel-based population analysis.” In MICCAI-workshop on Image-Guidance and Multimodal Dose Planning in Radiation Therapy, pp. 88-95. 2012.

- Ospina, J. D., Benquet, P., Correa, J. C., Acosta, O. “A mixed-effects spatio-temporal model to assess the joint effects of ageing and Alzheimerâs disease in gray matter volumen”. MICCAI 2012 Workshop on Novel Imaging Biomarkers for Alzheimer’s Disease and Related Disorders, Nice, France, 2012

- Ospina, J.D., Acosta, O., Dréan, G., Cazoulat, G., Simon, A., Correa, J.C, Haigron, P. and de Crevoisier, R. “Spatial nonparametric mixed-effects model with spatial-varying coefficients for analysis of populations” in Machine Learning in Medical imaging Workshop, MICCAI 2011, Toronto, CA, Sep 2011, Lecture Notes in Computer Science, Volume 7009/2011, 142-150.

- Chen, B., Acosta, O., Kachenoura, A., Ospina, J.D., Dréan, G., Simon, A., Bellanger, J.J., Haigron, P. and de Crevoisier, R. “Spatial characterization and classification of rectal bleeding in prostate cancer radiotherapy with a voxel-based Principal Components Analysis model for 3D dose distribution,” in Prostate Cancer Imaging Workshop, MICCAI 2011, Toronto, CA, Sep 2011, Lecture Notes in Computer Science, Volume 6963/2011, 60-69.

Communications in international conferences

- Ospina, J., J. Zhu, E. Le Pris é, A. Bossi, P. Haigron, K. Gnep, V. Beckendorf, O. Acosta, A. Simon, and R. De Crevoisier. “Random Forest are Strong Competitors of Published NTCP Models for Rectal and Bladder Toxicity Prediction.” International Journal of Radiation Oncology* Biology* Physics 84, no. 3 (2012): S387. *Proceedings of the ASTRO 54th meeting, Boston, USA. Poster*

- Ospina, J. D., J. Zhu, E. Le Pris é, A. Bossi, T. Messai, K. Gnep, V. Beckendorf, O. Acosta, A. Simon, and R. de Crevoisier. OC-0477 RANDOM FOREST VERSUS PUBLISHED NTCP MODELS FOR RECTAL TOXICITY PREDICTION. Radiotherapy and Oncology 103 (2012): S191-S192. *Proceedings of the ESTRO 31th meeting, Barcelona, Spain. Oral presentation*

- de Crevoisier, R., Messai, T., Wibault, P., Bridier, A., Blanchard, P., Ospina, J.D., Bakkour, M., and Bossi,A.. “7005 POSTER DISCUSSION PSA Measurement at the Fifth Week of Radiotherapy Is an Independent Predictor of Failure in Intermediate Risk Prostate Cancer Patients.” European Journal of Cancer 47 (2011): S485.

Submitted

- Ospina, J.D., Correa, J.C., Commandeur, F., Dréan, G., de Crevoisier, R., and Acosta, O. “Spatio-temporal nonparametric mixed-effects model for population analysis with 3D images.”
- Delobel, J.B, Ospina, J.D., Beckendorf, V., Chira C., Zhu, J., Bossi, A., Messai, T., and de Crevosier, R. “Nomograms to predict radiation proctopathy following prostate cancer radiotherapy.”

Collaborations

- Roman-Jimenez, G., Ospina, J.D., Leseur, J., Devillers, A., Castelli, J., Simon, A., Terve, P., Acosta, O., and de Crevoisier, R. “Investigating the contribution of pre- and per-treatment 18F-FDG PET-CT segmentation methodologies for post-treatment tumor recurrence prediction in cervical cancer.” IRBM, Volume 34, Issues 4â5, November 2013, Pages 274-277

Appendix A

Random forest description

A.1 Random forest

Random forest calculation is divided into three steps: 1) Generating q new databases by re-sampling with replacement from the original database; 2) building a classification tree for each of the new q databases; finally, 3) aggregating the predictions given by each classification tree.

The first step is usually carried out using bootstrap (described in [1]). A bootstrap sample is a sample with replacement taken from the original database. For example, if a database with five patients is available, namely $DB = \{Pt_1, Pt_2, Pt_3, Pt_4, Pt_5\}$, five bootstrap samples could be $DB^{(1)} = \{Pt_2, Pt_2, Pt_1, Pt_1, Pt_1\}$, $DB^{(2)} = \{Pt_4, Pt_1, Pt_4, Pt_4, Pt_3\}$, $DB^{(3)} = \{Pt_3, Pt_4, Pt_5, Pt_5, Pt_3\}$, $DB^{(4)} = \{Pt_3, Pt_1, Pt_4, Pt_4, Pt_3\}$ and $DB^{(5)} = \{Pt_4, Pt_3, Pt_2, Pt_3, Pt_2\}$. In this case, $q = 5$. Of course, this is only an example and a bigger database is required in order to use the bootstrap methodology.

The methodology to generate a classification tree could be modified if the analyst wants to add more randomness during the model-building process. In traditional classification, a tree is created by partitioning the input variable space, which is usually referred to as *growing* a tree. For example, if the input space consists of three continuous variables, V_1, V_2 and V_3, to set up the k-th node of the tree, a threshold for each input variable is calculated, namely v_{01}^k, v_{02}^k and v_{03}^k. Using theses thresholds, if-then classification rules are set. When individuals are classed in two categories, namely $[0, 1]$ these rules are of the type "if $V_i > v_{0i}^k$ then the individual is classed as 0, otherwise as 1." The thresholds v_{0i}^k are chosen in such a way that they minimize some impurity measure, such as the Gini index [2], that may be interpreted as a classification error. The threshold with the lowest impurity measure becomes the k-th node of the tree. The same procedure is followed within each partition until a stopping criterion is met, mostly associated with a minimum number of observations within each partition. At this point, the tree has reached its maximal size and should be pruned to avoid over fitting. In RF with random inputs, a subset of inputs is randomly chosen to generate each node. This introduces another source of randomness. In addition, no pruning is performed in

RF. These last two features give two characteristics to the tree ensemble: 1) each tree is different from the others; and 2) each tree is a good predictor on its training data, both characteristics essential for good results when aggregating individual outcomes.

The aggregation step depends on the type of problem, i.e. classification or regression. For regression, the outcome of the RF could be defined as the average of all tree outcomes, whereas concerning classification the RF outcome could be defined as the most voted category. Moreover, the portion of trees voting each category could be used as an estimate of the probabilities that a given individual belongs to each class, which in our work would be interpreted as a probability of normal tissue complication.

A.2 Random forest quality assessment

The quality of a RF can be measured using the out-of-bag (OOB) error. It consists of predicting each individual's outcome using only the trees where that individual was not present during the training step. In the illustrative example above, the OOB error for the Pt_2 individual should be calculated using the predictions from the trees built using $DB^{(2)}$, $DB^{(3)}$ and $DB^{(4)}$. For a given individual, it is expected that approximately 37% of the trees do not make use of it during training. The OOB subsample associated with $DB^{(i)}$ is defined as the set of patients that do not belong to it. For example, the OOB subsample associated to $DB^{(1)}$ is $DB^{(2)}_{OOB} = \{Pt_3, Pt_4, Pt_5\}$. The OOB error associated with $DB^{(2)}_{OOB}$ is the sum of the OOB errors obtained for patients Pt_3, Pt_4 and Pt_5.

A.3 Variable importance assessment

The logic behind measuring the importance of a variable can be stated as follows: if an important variable is disturbed, then predictions are expected to be worse, whereas disturbing a non-important variable should not have a significant impact on prediction. Variable importance assessment is then carried out using OOB errors. To assess the importance of variable V_j, first all OOB subsamples are obtained; secondly, the values of V_j are randomly permuted among all individuals, leading this to perturbed OOB subsamples denoted by $DB^{(i)}_{OOB,V_j}$; finally, the OOB error associated with each $DB^{(i)}_{OOB,V_j}$, namely $errorOOB^{(i)}_{V_j}$, is compared with its correspondent associated to $DB^{(i)}_{OOB}$, namely $errorOOB^{(i)}$. If the variable is important, then it is expected to have $errorOOB^{(i)}_{V_j} > errorOOB^{(i)}$. The importance of the variable V_j, $VI\,(V_j)$, can be thus defined as the average difference between OOB errors before and after permuting the values of the variable for all trees, as is shown in equation (A.1). It is thus called the mean decrease accuracy,

$$VI\,(V_j) = \frac{1}{q}\sum_{i=0}^{q}\left(errorOOB^{(i)}_{V_j} - errorOOB^{(i)}\right). \tag{A.1}$$

Another variable importance measure can be defined using the Gini index, which is used in classification trees to assess the impurity of a node. Suppose again that an individual can be classed in two categories, $[0, 1]$. For the k-th node, let p_{k0} and p_{k1} be the proportion of individuals classed as 0 and 1, respectively. The Gini index of the k-th node is calculated as in equation (A.2),

$$imp\left(k\right) = 1 - p_{k0}^2 - p_{k1}^2. \tag{A.2}$$

Note that when a node achieves perfect separation among the two classes, the Gini index equals 0. To split the k-th node, all the variables and several thresholds for each variable are tested. For instance, assuming that the k-th node is split using variable V_j and threshold v_{0j}^k, the Gini index for the left (L) and right (R) nodes can be calculated. Let's represent those by $imp\left(V_j, v_{0j}^k, L\right)$ and $imp\left(V_j, v_{0j}^k, R\right)$, respectively, and let $p_{(k,L),0}$ and $p_{(k,R),0}$ represent the portion of the data in the k-th node that lies on the left and on the right side, respectively, after the splitting. The Gini gain for variable V_j is then calculated as in equation (A.3),

$$\Delta imp\left(k, V_j\right) = imp\left(k\right) - p_{(k,L),0} imp\left(V_j, v_{0j}^k, L\right) - p_{(k,R),0} imp\left(V_j, v_{0j}^k, R\right). \tag{A.3}$$

Equation (A.3) is minimized to choose both the variable and its threshold. To calculate the importance of a variable during the construction of the tree, one can aggregate all the Gini gains associated to this variable. Thus, if N_j represents all the node splits using the variable V_j, the total Gini gain for variable V_j can be defined as in equation (A.4),

$$Gini\left(V_j\right) = \frac{1}{|N_j|} \sum_{i \in N_j} \Delta imp\left(i, V_j\right). \tag{A.4}$$

A high value for $Gini\left(V_j\right)$ indicates that variable V_j is important in discriminating the individuals. In a RF, it is possible to have the total Gini gain for each variable. Another measure of variable importance is defined in equation (A.5), which is called the mean decreasing Gini,

$$VI_G\left(V_j\right) = \frac{1}{q} \sum_{i=1}^{q} Gini_i\left(V_j\right), \tag{A.5}$$

where $Gini_i\left(V_j\right)$ is total Gini gain for variable V_j computed on $DB^{(i)}$. The former $VI_G\left(V_j\right)$ is to be used in classification rather than in regression problems. These two measures, $Gini\left(V_j\right)$ and $VI_G\left(V_j\right)$, help to quantify the performance of a variable in terms of prediction results and discrimination capabilities, respectively. These two variable importance measures will be used as they are furnished by the R implementation of RF (see [3]).

Bibliography

[1] B. Efron, “Computers and the theory of statistics: Thinking the unthinkable,” *SIAM Review*, vol. 21, no. 4, pp. 460–480, 1979.

[2] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees.* Monterey, CA: Wadsworth & Brooks, 1984.

[3] A. Liaw and M. Wiener, “Classification and regression by randomforest,” *R News*, vol. 2, pp. 18–22, 2002.

Appendix B

Results for the spatio-temporal nonparametric mixed-effects model

B.1 Smoothing beforehand is equivalent to using a local-likelihood

In this appendix, we demonstrate that under some assumptions the strategy of smoothing images by convolving them with a kernel function is equivalent to using a local-likelihood approach. In this way, we show that our nonparametric approach generalizes the modeling strategy of VBM methods.

B.1.1 Voxel-wise regression using smoothed images

Let image $Y'_{i,t} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ be a volume image, such that $Y'_{i,t} = \left(y'_{(jkl),t}\right)$, that is

$$Y'_{i,t} = \begin{bmatrix} y'_{111,i,t} & \cdots & y'_{1d_2 1,i,t} \\ \vdots & \ddots & \vdots \\ y'_{d_1 11,i,t} & \cdots & y'_{d_1 d_2 1,i,t} \end{bmatrix}, \ldots, \begin{bmatrix} y'_{11d_3,i,t} & \cdots & y'_{1d_1 d_3,i,t} \\ \vdots & \ddots & \vdots \\ y'_{d_1 1d_3,i,t} & \cdots & y'_{d_1 d_2 d_3,i,t} \end{bmatrix}.$$

Now assume that the image is convolved with a kernel function to obtain $Y_{i,t} = \left(y_{(jkl),t}\right)$ such that $y_{jkl,i,t} = \sum\limits_{p_1 p_2 p_3} y'_{p_1 p_2 p_3,i,t} \omega^{jkl}_{p_1 p_2 p_3}$ with $\sum\limits_{p_1 p_2 p_3} \omega^{(jkl)}_{p_1 p_2 p_3} = 1$ and $1 \leq p_1 \leq d_1$, $1 \leq p_2 \leq d_2$, $1 \leq p_3 \leq d_3$. In practice, $v^{jkl}_{p_1 p_2 p_3} = \frac{1}{h_x} k_x \left(\frac{\|(p_1,p_2,p_3)-(j,k,l)\|}{h_x} \right)$ and $\omega^{jkl}_{p_1 p_2 p_3} = \frac{v^{jkl}_{p_1 p_2 p_3}}{\sum_{p_1 p_2 p_3} v^{jkl}_{p_1 p_2 p_3}}$, resulting in

$$Y_{i,t} = \begin{bmatrix} y_{111,i,t} & \cdots & y_{1d_2 1,i,t} \\ \vdots & \ddots & \vdots \\ y_{d_1 11,i,t} & \cdots & y_{d_1 d_2 1,i,t} \end{bmatrix}, \ldots, \begin{bmatrix} y_{11d_3,i,t} & \cdots & y_{1d_1 d_3,i,t} \\ \vdots & \ddots & \vdots \\ y_{d_1 1d_3,i,t} & \cdots & y_{d_1 d_2 d_3,i,t} \end{bmatrix}. \tag{B.1}$$

Assume now that there is only one image for each individual, which we denote as Y_i. Let us define $\mathbf{y}_i = vec\left(Y_i\right)$ as the column vector resulting from staking all rows of all slices of image Y_i. We can fit the following linear model shown in (B.2) to each voxel $\mathbf{x}_0 = (j,k,l)$,

$$y_{jkl,i} = \mu_0^{(jkl)} + \phi_i \mu^{(jkl)} + \varepsilon_{jkl,i}. \tag{B.2}$$

This model can be written for the whole population (at voxel $\mathbf{x}_0 = (j,k,l)$) as in (B.3),

$$\begin{bmatrix} y_{jkl,1} \\ \vdots \\ y_{jkl,n} \end{bmatrix} = \begin{bmatrix} 1 & \phi_1 \\ \vdots & \vdots \\ 1 & \phi_n \end{bmatrix} \begin{bmatrix} \mu_0^{(jkl)} \\ \mu^{(jkl)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{jkl,1} \\ \vdots \\ \varepsilon_{jkl,n} \end{bmatrix}, \tag{B.3}$$

which is a model of the form $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$. Under the assumption that $\varepsilon \sim N_n\left(\mathbf{0}, \sigma^2 \mathbf{I}_{n\times n}\right)$, the maximum likelihood estimator of β is given by $\hat{\beta}^T = \left[\hat{\mu}_0^{(jkl)}, \hat{\mu}^{(jkl)}\right]^T = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}$. Thus, if we define $\phi = \sum_{i=1}^n \phi_i$, the ML estimators of $\mu_0^{(jkl)}$ and $\mu^{(jkl)}$ can be calculated as in (B.4),

$$\begin{aligned} \hat{\mu}_0^{(jkl)} &= \frac{1}{n-\phi}\left(\sum_{q=1}^n y_{jkl,q} - \sum_{q=1}^n \phi_q y_{jkl,q}\right), \\ \hat{\mu}^{(jkl)} &= \frac{1}{n-\phi}\left(\frac{n}{\phi}\sum_{q=1}^n \phi_q y_{jkl,q} - \sum_{q=1}^n y_{jkl,q}\right). \end{aligned} \tag{B.4}$$

These estimators are expressed in terms of smoothed images, but they could be re-written in terms of original images, as in (B.5) and (B.6),

$$\hat{\mu}_0^{(jkl)} = \frac{1}{n-\phi}\left(\sum_{q=1}^n \sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega_{p_1p_2p_3}^{jkl} - \sum_{q=1}^n \phi_q \sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega_{p_1p_2p_3}^{jkl}\right), \tag{B.5}$$

$$\hat{\mu}^{(jkl)} = \frac{1}{n-\phi}\left(\frac{n}{\phi}\sum_{q=1}^n \phi_q \sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega_{p_1p_2p_3}^{jkl} - \sum_{q=1}^n \sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega_{p_1p_2p_3}^{jkl}\right). \tag{B.6}$$

B.1.2 Local-likelihood approach

The non-smoothed images are considered in the local-likelihood approach. Assume that a version of (5.2) with no individual effects holds such that (B.7) is an approximation at voxel $\mathbf{x}_0 = (j,k,l)$,

$$f_i\left(\mathbf{x}_0\right) = m\left(\mathbf{x}_0\right) + g\left(\mathbf{x}_0\right)\phi_i + \varepsilon_i\left(\mathbf{x}_0\right), \tag{B.7}$$

which can be re-written as

$$y'_{jkl,i} = m^{(jkl)} + g^{(jkl)}\phi_i + \varepsilon_{jkl,i}. \tag{B.8}$$

For the i-individual, the model approximated around $\mathbf{x}_0 = (j,k,l)$ can be written as in (B.9),

$$\begin{bmatrix} y'_{111,i} \\ \vdots \\ y'_{d_1d_2d_3,i} \end{bmatrix} = \begin{bmatrix} 1 & \phi_i \\ \vdots & \vdots \\ 1 & \phi_i \end{bmatrix} \begin{bmatrix} m^{(jkl)} \\ g^{(jkl)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111,i} \\ \vdots \\ \varepsilon_{d_1d_2d_3,i} \end{bmatrix} \tag{B.9}$$

which is a model of the form $\mathbf{Y}_i = \mathbf{X}_i\beta + \varepsilon_i$.

Now, let $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_n \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ and $\beta = \left[m^{(jkl)}, g^{(jkl)},\right]^T$. In this way, the model for the whole population around $\mathbf{x}_0 = (j,k,l)$ can be written in the form $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$. In order to use the local-likelihood approach we must define a weighting matrix $\mathbf{K}^{(jkl)}$ to weight the contribution of each voxel when estimating $m^{(jkl)}$ and $g^{(jkl)}$.

Again, we can make $\omega^{jkl}_{p_1p_2p_3} = \frac{v^{jkl}_{p_1p_2p_3}}{\sum_{p_1p_2p_3} v^{jkl}_{p_1p_2p_3}}$, with $v^{jkl}_{p_1p_2p_3} = \frac{1}{h_x} k_{h_x}\left(\frac{\|(p_1p_2p_3)-(jkl)\|}{h_x}\right)$.

We thus can define the weight matrix as $\mathbf{K}^{(jkl)} = \operatorname{diag}\left(\mathbf{K}_i^{(jkl)}, \ldots, \mathbf{K}_n^{(jkl)}\right)$, where $\mathbf{K}_i^{(jkl)} = \operatorname{diag}\left(\omega_{111}^{(jkl)}, \ldots, \omega_{d_1d_2d_3}^{(jkl)}\right)$. Proceeding in this way the log-likelihood can be expressed as in (B.10),

$$l\left(\boldsymbol{\beta}|Y\right) = -\frac{1}{2}\mathbf{r}^T\mathbf{V}_h\mathbf{r} + C, \tag{B.10}$$

where $\mathbf{V}_h = \mathbf{K}^{(jkl)^{1/2}}\Sigma^{-1}\mathbf{K}^{(jkl)^{1/2}}$ and C is a constant that does not depend on $\boldsymbol{\beta}$.

Assuming that the errors are independent and normally-distributed, we have $\varepsilon \sim N_{nd}\left(\mathbf{0}, \sigma^2\mathbf{I}_{nd\times nd}\right)$ and the ML estimator of $\boldsymbol{\beta}$ around $\mathbf{x} = (j,k,l)$ is given by $\hat{\boldsymbol{\beta}}^{(jkl)} = \left(\mathbf{X}^T\mathbf{K}^{(jkl)}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{K}^{(jkl)}\mathbf{Y}$. Thus, the ML estimators of $m\left(\mathbf{x}\right)$ and $g\left(\mathbf{x}\right)$ around $\mathbf{x} = (j,k,l)$ can be obtained as in (B.11) and (B.12), respectively,

$$\hat{m}^{(jkl)} = \frac{1}{n-\phi}\left(\sum_{q=1}^{n}\sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega^{jkl}_{p_1p_2p_3} - \sum_{q=1}^{n}\phi_q\sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega^{jkl}_{p_1p_2p_3}\right), \tag{B.11}$$

$$\hat{g}^{(jkl)} = \frac{1}{n-\phi}\left(\frac{n}{\phi}\sum_{q=1}^{n}\phi_q\sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega^{jkl}_{p_1p_2p_3} - \sum_{q=1}^{n}\sum_{p_1p_2p_3} y'_{p_1p_2p_3,q}\omega^{jkl}_{p_1p_2p_3}\right). \tag{B.12}$$

As (B.11) and (B.12) are equal to (B.5) and (B.6), respectively, we conclude that smoothing the images first and then applying a GLM has the same effect as using a nonparametric model and a local-likelihood approach with only fixed effects.

B.2 Estimators

In this section, we provide the formulas for the estimators of $m\left(\mathbf{x}_0, t_0\right)$ and $g\left(\mathbf{x}_0, t_0\right)$, which are obtained by estimating $m\left(\mathbf{x}, t\right)$ and $g\left(\mathbf{x}, t\right)$ around $\mathbf{x}_0 = (j, k, l)$ and $t = t_0$, assuming that a first-order approximation is good enough. It means that for voxel $\mathbf{x} = (p_1, p_2, p_3)$ at time t the model can be expressed as $y'_{p_1 p_2 p_3, t} \approx m\left(\mathbf{x}_0, t_0\right) + \phi_i g\left(\mathbf{x}_0, t_0\right) + e_i\left(\mathbf{x}_0, t_0\right) + \varepsilon_i\left(\mathbf{x}, t\right)$. Let us assume that t_{i_j} with $j = 1, \ldots, n_i$, are different time points at which images for the i-th individual were acquired and that we have of $N = \sum_{i=1}^{n} n_i$ images. We also consider the vector $\mathbf{y}_{i,j} = \left[y'_{111,t_{i_j}} \cdots y'_{d_1 d_2 d_3, t_{i_j}}\right]^T$, which is the image corresponding to the i-th individual at time t_{i_j}. The vector with all images for the i-th individual is $\mathbf{y}_i = \left[\mathbf{y}_{i,1}^T \ldots \mathbf{y}_{i,n_i}^T\right]^T$ and the vector with all the images for the whole population is $\mathbf{Y} = \left[\mathbf{y}_1^T \ldots \mathbf{y}_n^T\right]^T$. Let $\mathbf{X}_{i,j} = \begin{bmatrix} 1 & \varphi_i \\ \vdots & \vdots \\ 1 & \varphi_i \end{bmatrix} \in \mathbb{R}^{d \times 2}$, $\mathbf{X}_i = \left[\mathbf{X}_{i,1}^T \ldots \mathbf{X}_{i,n_i}^T\right]^T \in \mathbb{R}^{n_i d \times 2}$ and $\mathbf{X} = \left[\mathbf{X}_1^T \ldots \mathbf{X}_n^T\right]^T \in \mathbb{R}^{dN \times 2}$, $\mathbf{Z}_{i,j} = [1 \ldots 1]^T \in \mathbb{R}^{d}$, $\mathbf{Z}_i = \left[\mathbf{Z}_{i,1}^T \ldots \mathbf{Z}_{i,n_i}^T\right]^T \in \mathbb{R}^{dn_i}$ and $\mathbf{Z} = \operatorname{diag}\left(\mathbf{Z}_1 \ldots \mathbf{Z}_n\right) \in \mathbb{R}^{dN \times n}$.

With the coefficient vectors given by $\beta = \left[m\left(\mathbf{x}_0, t_0\right) g\left(\mathbf{x}_0, t_0\right)\right]^T$ and $\mathbf{b} = \left[e_1\left(\mathbf{x}_0, t_0\right) \ldots e_n\left(\mathbf{x}_0, t_0\right)\right]^T$ the model for the whole population can be written as $\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{b} + \varepsilon$, where $\mathbf{b} \sim N\left(\mathbf{0}, \mathbf{R}\right)$ and $\varepsilon \sim N\left(\mathbf{0}, \mathbf{D}\right)$ with $\mathbf{R} \in \mathbb{R}^{dN \times dN}$ and $\mathbf{D} \in \mathbb{R}^{n \times n}$. The corresponding log-likelihood is shown in (5.8). Differentiating (5.8) with respect to β and $\mathbf{b}$ and solving for the fixed and random effects, we obtain the equation system presented in (B.13),

$$\begin{bmatrix} \mathbf{X}^T\mathbf{V}_h\mathbf{X} & \mathbf{X}^T\mathbf{V}_h\mathbf{Z} \\ \mathbf{Z}^T\mathbf{V}_h\mathbf{X} & \mathbf{Z}^T\mathbf{V}_h\mathbf{Z} + \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T\mathbf{V}_h\mathbf{Y} \\ \mathbf{Z}^T\mathbf{V}_h\mathbf{Y} \end{bmatrix}. \tag{B.13}$$

From (B.13), we can obtain the expressions for β and $\mathbf{b}$ as in (B.14) and (B.15),

$$\hat{\beta} = \left(\mathbf{X}^T\mathbf{\Omega}_h\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{\Omega}_h\mathbf{Y}, \tag{B.14}$$

$$\hat{\mathbf{b}} = \left(\mathbf{Z}^T\mathbf{V}_h\mathbf{Z} + \mathbf{D}^{-1}\right)^{-1}\mathbf{Z}^T\mathbf{V}_h\left(\mathbf{Y} - \mathbf{X}\hat{\beta}\right)\mathbf{Y}, \tag{B.15}$$

with $\mathbf{\Omega}_h = \left(\mathbf{V}_h^{-1} + \mathbf{Z}\mathbf{D}\mathbf{Z}^T\right)^{-1}$. The weighting matrix here is defined as

$$\mathbf{K}_h = \mathbf{K}^{(jkl,t_0)} = \operatorname{diag}\left(\mathbf{K}_i^{(jkl,t_0)}, \ldots, \mathbf{K}_n^{(jkl,t_0)}\right),$$

where $\mathbf{K}_i^{(jkl,t_0)} = \operatorname{diag}\left(\mathbf{K}_{i,1}^{(jkl,t_0)} \ldots, \mathbf{K}_{i,n_i}^{(jkl,t_0)}\right)$ and
$\mathbf{K}_{i,j}^{(jkl,t_0)} = \operatorname{diag}\left(\omega_{111,t_{i_j}}^{(jkl,t_0)}, \ldots, \omega_{d_1 d_2 d_3, t_{i_j}}^{(jkl,t_0)}\right)$.

Each weight $\omega_{p_1 p_2 p_3, t_{i_j}}^{(jkl,t_0)}$ is calculated as $\omega_{p_1 p_2 p_3, t_{i_j}}^{(jkl,t_0)} = \frac{v_{p_1 p_2 p_3, t_{i_j}}^{(jkl,t_0)}}{C}$, with

$$v_{p_1 p_2 p_3, t_{i_j}}^{(jkl,t_0)} = k_{h_x, h_t}\left(\left\|(p_1, p_2, p_3) - (j, k, l)\right\|, \left\|t_{i_j} - t_0\right\|\right).$$

The normalization constant C is chosen such as $trace\left(\mathbf{K}_h\right) = 1$ (the weights are constrained to sum one).

Under the assumption that $\mathbf{R} = \sigma^2\mathbf{I}_{dN\times dN}$ and $\mathbf{D} = \sigma_\varepsilon^2\mathbf{I}_{n\times n}$, we can provide the expressions for the estimators of $m\left(\mathbf{x}_0, t_0\right)$ and $g\left(\mathbf{x}_0, t_0\right)$. In order to simplify the notation we define $\omega_i = \sum\limits_{j=1}^{n_i}\sum\limits_{p_1,p_2,p_3}\omega_{p_1p_2p_3,t_{i_j}}^{(jkl,t_0)}$, $\alpha_i = \frac{\sigma^2}{\sigma^2+\sigma_\varepsilon^2\omega_i}$, $N_\omega = \sum\limits_{i=1}^{n}\alpha_i\omega_i$, $\phi_\omega = \sum\limits_{i=1}^{n}\varphi_i\alpha_i\omega_i$, and

$$\bar{y}_i = \sum_{j=1}^{n_i}\sum_{p_1,p_2,p_3} y_{p_1p_2p_3,t_{i_j}}\,\omega_{p_1p_2p_3,t_{i_j}}^{(jkl,t_0)}. \tag{B.16}$$

The term $\bar{y}_i$ represents the weighted average of the images of the i-th individual taken at time points t_{i_j}n $j = 1,\ldots,n_i$. With this notation, the local weighted-likelihood estimators of $m\left(\mathbf{x}_0, t_0\right)$ and $g\left(\mathbf{x}_0, t_0\right)$ are given in (B.17) and (B.18),

$$\hat{m}\left(\mathbf{x}_0, t_0\right) = \frac{1}{N_\omega - \phi_\omega}\left(\sum_{i=1}^{n}\alpha_i\bar{y}_i - \sum_{i=1}^{n}\varphi_i\alpha_i\bar{y}_i\right) \tag{B.17}$$

$$\hat{g}\left(\mathbf{x}_0, t_0\right) = \frac{1}{N_\omega - \phi_\omega}\left(\frac{N_\omega}{\phi_\omega}\sum_{i=1}^{n}\varphi_i\alpha_i\bar{y}_i - \sum_{i=1}^{n}\alpha_i\bar{y}_i\right). \tag{B.18}$$

The estimator of $\hat{g}\left(\mathbf{x}_0, t_0\right)$ represents the difference among the two groups at voxel $\mathbf{x}_0$ and time point t_0. The corresponding mean for individuals belonging to group T^C is given in eq. (B.19),

$$\hat{m}\left(\mathbf{x}_0, t_0\right) + \hat{g}\left(\mathbf{x}_0, t_0\right) = \frac{1}{\phi_\omega}\sum_{i=1}^{n}\phi_i\alpha_i\bar{y}_i. \tag{B.19}$$

As it can be noted, $\hat{m}\left(\mathbf{x}_0, t_0\right)$ and $\hat{m}\left(\mathbf{x}_0, t_0\right) + \phi_i\hat{g}\left(\mathbf{x}_0, t_0\right)$ both rely on different information except for the variance components, σ^2 and σ_ε^2. The local-likelihood predictor for $\hat{e}_i\left(\mathbf{x}_0, t_0\right)$ is given in eq (B.20),

$$\hat{e}_i\left(\mathbf{x}_0, t_0\right) = \sigma_\varepsilon^2\alpha_i\left(\bar{y}_i - \omega_i\left[\hat{m}\left(\mathbf{x}_0, t_0\right) + \hat{g}\left(\mathbf{x}_0, t_0\right)\right]\right). \tag{B.20}$$

For most applications, the values of σ^2 and σ_ε^2 are unknown. These can be obtained through an iterative estimation procedure, such as RML, leading to $\hat{\sigma}^2\left(\mathbf{x}_0, t_0\right)$ and $\hat{\sigma}_\varepsilon^2\left(\mathbf{x}_0, t_0\right)$. Thus, in practice we have a nonparametric local estimation of inter- and intra-individual variance components as a function of spatial location and time. To obtain the standard errors associated with $\hat{m}\left(\mathbf{x}_0, t_0\right)$ and $\hat{g}\left(\mathbf{x}_0, t_0\right)$, we must note that $Var\left(\mathbf{Y}\right) = \mathbf{R} + \mathbf{Z}\mathbf{D}\mathbf{Z}^T$. Thus, if we define $\mathbf{C}_h = \left(\mathbf{X}^T\boldsymbol{\Omega}_h\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Omega}_h$, Cov $\left(\hat{\beta}\right)$ is given by (B.21),

$$\begin{aligned}\operatorname{Cov}\left(\hat{\beta}\right) &= \mathbf{C}_h Var\left(\mathbf{Y}\right)\mathbf{C}_h^T \\ &= \mathbf{C}_h\left(\mathbf{R}+\mathbf{ZDZ}^T\right)\mathbf{C}_h^T. \qquad \text{(B.21)}\end{aligned}$$

Developing the expression given in (B.21) under the assumption that $\mathbf{R} = \sigma^2\mathbf{I}_{dN\times dN}$ and $\mathbf{D} = \sigma_\varepsilon^2\mathbf{I}_{n\times n}$ and that $trace\left(\mathbf{K}_h\right) = 1$, we obtain the expressions (B.22) and (B.23) for the variances of the fixed effects,

$$\operatorname{Var}\left(\hat{m}\left(\mathbf{x}_0, t_0\right)\right) = \frac{1}{\left(N_\omega - \phi_\omega\right)^2}\left(\tau - \tau_\phi\right), \qquad \text{(B.22)}$$

$$\operatorname{Var}\left(\hat{g}\left(\mathbf{x}_0, t_0\right)\right) = \frac{1}{\left(N_\omega - \phi_\omega\right)^2}\left(\tau - 2\frac{N_\omega}{\phi_\omega}\tau_\phi + \frac{N_\omega^2}{\phi_\omega^2}\tau_\phi\right), \qquad \text{(B.23)}$$

where $\tau = \sigma^2\omega^{(2)} + \sigma_\varepsilon^2\bar{\omega}^{(2)}$ and $\tau_\phi = \sigma^2\omega_\phi^{(2)} + \sigma_\varepsilon^2\bar{\omega}_\phi^{(2)}$, with
$\omega^{(2)} = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\sum_{p_1,p_2,p_3}\left(\omega_{p_1p_2p_3,t_{ij}}^{(jkl,t_0)}\right)^2$, $\bar{\omega}^{(2)} = \sum_{i=1}^{n}\left(\omega_i\right)^2$,
$\omega_\phi^{(2)} = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\sum_{p_1,p_2,p_3}\phi_i\left(\omega_{p_1p_2p_3,t_{ij}}^{(jkl,t_0)}\right)^2$ and $\bar{\omega}_\phi^{(2)} = \sum_{i=1}^{n}\phi_i\left(\omega_i\right)^2$.
The covariance between $\hat{m}\left(\mathbf{x}_0, t_0\right)$ and $\hat{g}\left(\mathbf{x}_0, t_0\right)$ is given in (B.24),

$$\begin{aligned}\operatorname{Cov}\left(\hat{m}\left(\mathbf{x}_0, t_0\right), \hat{g}\left(\mathbf{x}_0, t_0\right)\right) &= \frac{1}{\left(N_\omega - \phi_\omega\right)^2}\left(\tau_\phi - \tau\right) \\ &= -\operatorname{Var}\left(\hat{m}\left(\mathbf{x}_0, t_0\right)\right). \qquad \text{(B.24)}\end{aligned}$$

B.2.1 Two special cases

B.2.1.1 Non-group effect

Under the assumption that there is no group effect, the proposed model can still describe the population behavior. In that case, we assume that $g\left(\mathbf{x}, t\right) = 0$ and the estimator of the population mean $m\left(\mathbf{x}_0, t_0\right)$ at voxel $\mathbf{x}_0$ and time t_0 is given by (B.43),

$$\hat{m}\left(\mathbf{x}_0, t_0\right) = \frac{1}{N_\omega}\sum_{i=1}^{n}\alpha_i\bar{y}_i. \qquad \text{(B.25)}$$

B.2.1.2 Cross-sectional studies

In the cross-sectional case we dispose of n images for n individuals (so $n_i = 1$ for $i = 1, \ldots, n$. If we want to model the effect of the time variable on the population, we can still make use of (B.17) and ((B.18), with $n_i = 1$ for all individuals. If the time variable t is not relevant, (B.17) and (B.18) can be modified to obtain the estimators

of $m(\mathbf{x}_0)$ and $g(\mathbf{x}_0)$. As in Appendix (B.1.1), the weights are defined as a function of the distance between two different voxels. Using a similar notation as before, we define $\omega_i = \sum_{p_1,p_2,p_3} \omega^{(jkl)}_{p_1p_2p_3}$, $\alpha_i = \frac{\sigma^2}{\sigma^2+\sigma^2_\varepsilon \omega_i}$, $N_\omega = \sum_{i=1}^{n} \alpha_i\omega_i$, $\phi_\omega = \sum_{i=1}^{n} \varphi_i\alpha_i\omega_i$, and
$\bar{y}_i = \sum_{p_1,p_2,p_3} y_{p_1p_2p_3,i}\omega^{(jkl)}_{p_1p_2p_3}$.
With this notation, the estimators of $m(\mathbf{x}_0)$ and $g(\mathbf{x}_0)$ obtained from the local-likelihood approach can be computed as in (B.26) and (B.27),

$$\hat{m}(\mathbf{x}_0) = \frac{1}{N_\omega - \phi_\omega}\left(\sum_{i=1}^{n} \alpha_i\bar{y}_i - \sum_{i=1}^{n} \varphi_i\alpha_i\bar{y}_i\right), \tag{B.26}$$

$$\hat{g}(\mathbf{x}_0) = \frac{1}{N_\omega - \phi_\omega}\left(\frac{N_\omega}{\phi_\omega}\sum_{i} \varphi_i\alpha_i\bar{y}_i - \sum_{i} \alpha_i\bar{y}_i\right). \tag{B.27}$$

Generally, when the variance components σ^2 and σ^2_ε are unknown, we can estimate them as $\hat{\sigma}^2(\mathbf{x}_0)$ and $\hat{\sigma}^2_\varepsilon(\mathbf{x}_0)$ using RML. The practical importance of this is that even in a cross-sectional study, individual variability can be handled using spatial information.

B.3 Variance of volumes

For a given region D, it could be interesting to compute the following integral as a function of the time t: $V_{D,T}(t) = \int_{\mathbf{x}\in D}[m(\mathbf{x},t) + g(\mathbf{x},t)]\,d\mathbf{x}$ for a typical group T subject and $V_{D,T^C}(t) = \int_{\mathbf{x}\in D} m(\mathbf{x},t)\,d\mathbf{x}$ for a T^C subject. The estimators for these integrals are presented in eq. (B.28) and (B.29),

$$\hat{V}_{D,T}(t) = \sum_{\mathbf{x}_i\in D} \hat{m}(\mathbf{x}_i,t)\,\Delta\mathbf{x}, \tag{B.28}$$

$$\hat{V}_{D,T^C}(t) = \sum_{\mathbf{x}_i\in D} [\hat{m}(\mathbf{x}_i,t) + \hat{g}(\mathbf{x}_i,t)]\,\Delta\mathbf{x}, \tag{B.29}$$

where $\Delta\mathbf{x}$ is the volume of the voxels.

In order to derive confidence intervals for these estimators, we need to compute $\text{Var}\left(\hat{V}_{D,T}(t)\right)$ and $\text{Var}\left(\hat{V}_{D,T^C}(t)\right)$ as in (B.30) and (B.31),

$$\text{Var}\left(\sum_{\mathbf{x}_i\in D} \hat{m}(\mathbf{x}_i,t)\right) = \sum_{\mathbf{x}_i\in D} \text{Var}(\hat{m}(\mathbf{x}_i,t)) + \sum_{\mathbf{x}_i,\mathbf{x}_j\in D} \text{Cov}(\hat{m}(\mathbf{x}_i,t), \hat{m}(\mathbf{x}_j,t)), \tag{B.30}$$

$$
\begin{aligned}
&\operatorname{Var}\left(\sum_{\mathbf{x}_i \in D} \left(\hat{m}\left(\mathbf{x}_i, t\right) + \hat{g}\left(\mathbf{x}_i, t\right)\right)\right) = \sum_{\mathbf{x}_i \in D} \operatorname{Var}\left(\hat{m}\left(\mathbf{x}_i, t\right) + \hat{g}\left(\mathbf{x}_i, t\right)\right) \\
&+ \sum_{\mathbf{x}_i, \mathbf{x}_j \in D} \operatorname{Cov}\left(\hat{m}\left(\mathbf{x}_i, t\right) + \hat{g}\left(\mathbf{x}_i, t\right), \hat{m}\left(\mathbf{x}_j, t\right) + \hat{g}\left(\mathbf{x}_j, t\right)\right).
\end{aligned}
\tag{B.31}
$$

To compute these variances, we need some extra notation: now we will denote $\bar{y}_i$ by $\bar{y}_i\left(\mathbf{x}_0, t_0\right)$ in (B.16) to show the dependency (through the kernel function) of this weighted average on the spatial location ($\mathbf{x}_0 = (j, k, l,))$ and time point t_0. We will also write $\alpha_i\left(\mathbf{x}_0, t_0\right)$, $N_\omega\left(\mathbf{x}_0, t_0\right)$ and $\phi_\omega\left(\mathbf{x}_0, t_0\right)$. Because of the independence among subjects, we note that $\operatorname{Cov}\left(\bar{y}_i\left(\mathbf{x}_{i_1}, t_1\right), \bar{y}_j\left(\mathbf{x}_{i_2}, t_2\right)\right) = 0$ if $i \neq j$. With some algebra it is possible to derive an expression for $\operatorname{Cov}\left(\bar{y}_i\left(\mathbf{x}_{i_1}, t\right), \bar{y}_i\left(\mathbf{x}_{i_2}, t\right)\right)$ as in (B.32),

$$
\begin{aligned}
\operatorname{Cov}\left(\bar{y}_i\left(\mathbf{x}_{i_1}, t\right), \bar{y}_i\left(\mathbf{x}_{i_2}, t\right)\right) \quad = \quad & \sum_{j=1}^{n_i} \sum_{p} \omega_{p,t_{ij}}^{\left(\mathbf{x}_{i_1}, t\right)} \omega_{p,t_{ij}}^{\left(\mathbf{x}_{i_2}, t\right)} \left(\sigma^2\left(\mathbf{x}_p, t_{ij}\right) + \sigma_\varepsilon^2\left(\mathbf{x}_p, t_{ij}\right)\right) \\
& + \sum_{j_1 \neq j_2} \sum_{p \neq l} \omega_{p,t_{ij}}^{\left(\mathbf{x}_{i_1}, t\right)} \omega_{l,t_{ij}}^{\left(\mathbf{x}_{i_2}, t\right)} \operatorname{Cov}\left(e_i\left(\mathbf{x}_p, t_{ij_1}\right), e_i\left(\mathbf{x}_l, t_{ij_2}\right)\right),
\end{aligned}
\tag{B.32}
$$

where the sums are carried out over both voxels $\mathbf{x}_p$ and $\mathbf{x}_l$. The terms $\sigma^2\left(\mathbf{x}_p, t_{ij}\right)$ and $\sigma_\varepsilon^2\left(\mathbf{x}_p, t_{ij}\right)$ are estimated via RML and, as in [1], $\operatorname{Cov}\left(e_i\left(\mathbf{x}_p, t_{ij_1}\right), e_i\left(\mathbf{x}_l, t_{ij_2}\right)\right)$ can be estimated using the method of moments demonstrated in (B.33),

$$
\hat{\operatorname{Cov}}\left(e_i\left(\mathbf{x}_p, t_{ij_1}\right), e_i\left(\mathbf{x}_l, t_{ij_2}\right)\right) = \frac{1}{n-1} \sum_{i=1}^{n} \hat{e}_i\left(\mathbf{x}_p, t_{ij_1}\right) \times \hat{e}_i\left(\mathbf{x}_l, t_{ij_2}\right),
\tag{B.33}
$$

where $\hat{\operatorname{Cov}}\left(\cdot, \cdot\right)$ denotes the estimator of $\operatorname{Cov}\left(\cdot, \cdot\right)$. From (B.17) we can compute $\operatorname{Cov}\left(\hat{m}\left(\mathbf{x}_{i_1}, t\right), \hat{m}\left(\mathbf{x}_{i_2}, t\right)\right)$ as in (B.34),

$$
\begin{aligned}
&\operatorname{Cov}\left(\hat{m}\left(\mathbf{x}_{i_1}, t\right), \hat{m}\left(\mathbf{x}_{i_2}, t\right)\right) = \\
&\left(\frac{1}{\Psi_\omega\left(\mathbf{x}_{i_1}, t\right) \times \Psi_\omega\left(\mathbf{x}_{i_2}, t\right)}\right) \times \sum_{i=1}^{n} \left\{\left(1 - \phi_i\right) \operatorname{Cov}\left(\bar{y}_i\left(\mathbf{x}_{i_1}, t\right), \bar{y}_i\left(\mathbf{x}_{i_2}, t\right)\right) \times \alpha_i\left(\mathbf{x}_{i_1}, t\right) \alpha_i\left(\mathbf{x}_{i_2}, t\right)\right\}.
\end{aligned}
\tag{B.34}
$$

To simplify notation, we define $\Psi_\omega\left(\mathbf{x}, t\right) = N_\omega\left(\mathbf{x}, t\right) - \phi_\omega\left(\mathbf{x}, t\right)$. Using (B.19) it is possible to compute $\operatorname{Cov}\left(\hat{m}\left(\mathbf{x}_{i_1}, t\right) + \hat{g}\left(\mathbf{x}_{i_1}, t\right), \hat{m}\left(\mathbf{x}_{i_2}, t\right) + \hat{g}\left(\mathbf{x}_{i_2}, t\right)\right)$ as in (B.35),

$$\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_{i_1},t\right)+\hat{g}\left(\mathbf{x}_{i_1},t\right),\hat{m}\left(\mathbf{x}_{i_2},t\right)+\hat{g}\left(\mathbf{x}_{i_2},t\right)\right)=$$
$$\left(\frac{1}{\phi_\omega\left(\mathbf{x}_{i_1},t\right)\times\phi_\omega\left(\mathbf{x}_{i_2},t\right)}\right)\times\sum_{i=1}^{n}\phi_i\times\mathrm{Cov}\left(\bar{y}_i\left(\mathbf{x}_{i_1},t\right),\bar{y}_i\left(\mathbf{x}_{i_2},t\right)\right)\alpha_i\left(\mathbf{x}_{i_1},t\right)\alpha_i\left(\mathbf{x}_{i_2},t\right). \tag{B.35}$$

An alternative way to speed up the calculation of $\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_i,t\right),\hat{m}\left(\mathbf{x}_j,t\right)\right)$ is to assume that spatial correlation does not change significantly with time and that $\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_i,t\right),\hat{m}\left(\mathbf{x}_j,t\right)\right)\approx\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_i,t_1\right),\hat{m}\left(\mathbf{x}_j,t_2\right)\right)$ if $|t_1-t_2|$ is small. We can thus use a kernel estimation for $\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_i,t\right),\hat{m}\left(\mathbf{x}_j,t\right)\right)$ as in (B.36),

$$\mathrm{Cov}\left(\hat{m}\left(\mathbf{x}_i,t\right),\hat{m}\left(\mathbf{x}_j,t\right)\right)=$$
$$\left(\frac{1}{N_T-1}\right)\times\sum_{k=1}^{N_T}\omega_{t_k}\Big(\left(\hat{m}\left(\mathbf{x}_i,t_k\right)-\bar{m}\left(\mathbf{x}_i,\cdot\right)\right)\times\left(\hat{m}\left(\mathbf{x}_j,t_k\right)-\bar{m}\left(\mathbf{x}_j,\cdot\right)\right)\Big), \tag{B.36}$$

where t_k, $k=1,\dots,N_T$ are the different N_T time points at which model is estimated, ω_{t_k} is the weight assigned to time t_k when assessing the variance at time t and $\bar{m}\left(\mathbf{x}_i,\cdot\right)=\sum_{k=1}^{N_T}\omega_{t_k}\hat{m}\left(\mathbf{x}_i,t_k\right)$.

B.4 Asymptotic properties

We assume that the zero oder approximation holds,

$$f_i\left(\mathbf{x},t\right)\approx m\left(\mathbf{x}_0,t_0\right)+g\left(\mathbf{x}_0,t_0\right)\phi_i+e_i\left(\mathbf{x}_0,t_0\right)+\varepsilon_i\left(\mathbf{x},t\right). \tag{B.37}$$

B.4.1 Theorems and proofs

B.4.1.1 Assumptions

1. Images are acquired at time points t_{ij}, $i=1,\dots,n$, $j=1,\dots,n_i$ which are iid with density $f\left(\cdot\right)$ with continue first and second derivatives.

2. The fixed-effect curves, $m\left(\mathbf{x},t\right)$ and $m\left(\mathbf{x},t\right)+g\left(\mathbf{x},t\right)$, have first order continue partial derivatives with respect to $\mathbf{x}$ and to t.

3. For simplicity we assume that the covariance function $\mathrm{Cov}\left(e_i\left(\mathbf{x}_p,t_{ij_1}\right),e_i\left(\mathbf{x}_l,t_{ij_2}\right)\right)$ can be modeled as $\gamma_x\left(\mathbf{x}_p,\mathbf{x}_l\right)\gamma_t\left(t_{ij_1},t_{ij_2}\right)$, and both $\gamma_x\left(\cdot,\cdot\right)$ and $\gamma_t\left(\cdot,\cdot\right)$ admit at least a second order Taylor expansion.

4. The population variance $\sigma^2\left(\mathbf{x},t\right)$ is continuous in both arguments.

5. For the kernel function $k_{h_x,h_t}(\cdot,\cdot) = \frac{1}{h_x h_t} k_{\mathbf{x}}\left(\frac{\cdot}{h_x}\right) k_t\left(\frac{\cdot}{h_t}\right)$ we assume that $k_{\mathbf{x}}(\cdot) = k_t(\cdot)$ and that $k_t(\cdot)$ is a bounded symmetrical probability density function such as $k_t(0) > 0$, $\int k_t(v)dv = 1$, $\int v k_t(v)dv = 0$ and $Lim_{|v|\to\infty} v k_t(v) = 0$. We thereby define $B(k) = \int v^2 k_v(v)dv$ and $V(k) = \int k_t^2(v)dv$.

6. Image resolution: the image contains information for d voxels labeled $\mathbf{x}_1, \ldots, \mathbf{x}_p, \ldots, \mathbf{x}_d$. As the resolution increases ($d \to \infty$) we have that $h_x \to 0$ and $d \times h_x \to \infty$.

7. $\mathbf{D} = \{t_{1,1}, \ldots, t_{1,n_1}, \ldots, t_{n,1}, \ldots, t_{n,n_n}\}$ is the collection of all time points at which images are acquired.

8. $N_1 = \sum_{i=1}^{n}(1-\phi_i)n_i$ and $N_2 = \sum_{i=1}^{n}\phi_i n_i$ are the number of available images for groups T and T^C and $N_1^* = \sum_{i=1}^{n}(1-\phi_i)$ and $N_2^* = \sum_{i=1}^{n}\phi_i$ are the number of individuals in each of the groups.

9. We define $m_1^* = N_1^*\left(\sum_{i=1}^{n}(1-\phi_i)n_i^{-1}\right)^{-1}$ and $m_2^* = N_2^*\left(\sum_{i=1}^{n}\phi_i n_i^{-1}\right)^{-1}$

10. We assume that $h_x = O\left(\min\left\{(N_1 m_1^*)^{-2/5}, (N_2 m_2^*)^{-2/5}\right\}\right)$ and that $h_t = O\left(\min\left\{(N_1 m_1^*)^{-1/5}, (N_2 m_2^*)^{-1/5}\right\}\right)$.

11. As $N_1^* \to \infty$, $n_i = O\left(N_1^{*\delta}\right)$ with $1/4 < \delta < 3/2$ for $i \in T$ and as $N_2^* \to \infty$,$n_i = O\left(N_2^{*\delta}\right)$ for $i \in T^C$.

12. There is $C_K < \infty$ such as $|f_i(\mathbf{x},t) - m(\mathbf{x},t)| < C_K$, for all i, $\mathbf{x}$ and t.

B.4.1.2 Asymptotic properties

In this section we present the results that allow for the use of Z-scores as voxel-wise test statistics. Our results, derived under almost the same conditions and following the same demonstration techniques, coincide with Theorems 1 (a and b) and 2 (a) in [2] (Wu2002), with some modifications accounting for the spatial variable and the two groups.

The asymptotic bias and variance for $\hat{m}(\mathbf{x}_0,t_0)$ and $\hat{m}(\mathbf{x}_0,t_0) + \hat{g}(\mathbf{x}_0,t_0)$ are given by

$$\begin{aligned}\text{Bias}\left(\hat{m}(\mathbf{x}_0,t_0)|\mathbf{D}\right) &= \left\{\frac{C_2}{C} + \frac{h_t^2}{2}\left(\frac{f''(t_0)}{f(t_0)}\frac{C_2}{C}B(k_t) + 2\frac{f'(t_0)}{f(t_0)}\frac{\partial m(\mathbf{x}_0,t_0)}{\partial t}B(k_t)\right)\right\} \\ &\quad \times [1 + o_p(1)],\end{aligned} \tag{B.38}$$

$$\begin{aligned}\operatorname{Bias}\left(\hat{m}\left(\mathbf{x}_0,t_0\right)+\hat{g}\left(\mathbf{x}_0,t_0\right)|\mathbf{D}\right) &= \left[1+o_p\left(1\right)\right]\left\{\frac{C_2}{C}+\frac{h_t^2}{2}\left(\frac{f''\left(t_0\right)}{f\left(t_0\right)}\times\frac{C_2}{C}B\left(k_t\right)\right.\right.\\ &\left.\left.+2\frac{f'\left(t_0\right)}{f\left(t_0\right)}\left(\frac{\partial m\left(\mathbf{x}_0,t_0\right)}{\partial t}+\frac{\partial g\left(\mathbf{x}_0,t_0\right)}{\partial t}\right)\times B\left(k_t\right)\right)\right\}\times . \end{aligned} \tag{B.39}$$

$$\begin{aligned}\operatorname{Var}\left(\hat{m}\left(\mathbf{x}_0,t_0\right)|\mathbf{D}\right) &= \frac{\operatorname{Cov}\left(e_i\left(\mathbf{x}_0,t_0\right),e_i\left(\mathbf{x}_0,t_0\right)\right)}{N_1^*}+\left(\sigma^2\left(\mathbf{x}_0,t_0\right)+\sigma_\varepsilon^2\left(\mathbf{x}_0,t_0\right)\right)\times\\ &\times\frac{\frac{C^{(2)}}{C^2}V\left(k_t\right)}{N_1^*m_1^*h_tf\left(t_0\right)}+o_p\left(N_1^{*-1}+\left(N_1^*m_1^*h_t\right)^{-1}\right)\\ &= \frac{\operatorname{Cov}\left(e_i\left(\mathbf{x}_0,t_0\right),e_i\left(\mathbf{x}_0,t_0\right)\right)}{N_1^*}+o_p\left(N_1^{*-1}+\left(N_1^*m_1^*h_t\right)^{-1}\right).\end{aligned} \tag{B.40}$$

$$\begin{aligned}\operatorname{Var}\left(\hat{m}\left(\mathbf{x}_0,t_0\right)+\hat{g}\left(\mathbf{x}_0,t_0\right)|\mathbf{D}\right) &= \frac{\operatorname{Cov}\left(e_i\left(\mathbf{x}_0,t_0\right),e_i\left(\mathbf{x}_0,t_0\right)\right)}{N_2^*}\\ &+o_p\left(N_2^{*-1}+\left(N_2^*m_2^*h_t\right)^{-1}\right).\end{aligned} \tag{B.41}$$

The asymptotic distributions of $\hat{m}\left(\mathbf{x}_0,t_0\right)$ and $\hat{m}\left(\mathbf{x}_0,t_0\right)+\hat{g}\left(\mathbf{x}_0,t_0\right)$ are given by

$$N_1^{*1/2}\left(\hat{m}\left(\mathbf{x}_0,t_0\right)-m\left(\mathbf{x}_0,t_0\right)\right)\sim AN\left(0,\operatorname{Cov}\left(e_i\left(\mathbf{x}_0,t_0\right),e_i\left(\mathbf{x}_0,t_0\right)\right)\right) \tag{B.42}$$

and
$N_2^{*1/2}\left(\hat{m}\left(\mathbf{x}_0,t_0\right)+\hat{g}\left(\mathbf{x}_0,t_0\right)-m\left(\mathbf{x}_0,t_0\right)-g\left(\mathbf{x}_0,t_0\right)\right)$ is asymptotically normal with mean 0 and variance $\operatorname{Cov}\left(e_i\left(\mathbf{x}_0,t_0\right),e_i\left(\mathbf{x}_0,t_0\right)\right)$.
The constants C, $C^{(2)}$ and C_2 are given by $C=\sum_p\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)$, $C^{(2)}=\sum_p\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)$ and $C_2=\sum_p\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{\partial m^T\left(\mathbf{x}_0,t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_p-\mathbf{x}_0\right)$.

B.4.1.3 Asymptotic bias

We first calculate the asymptotic bias of $\hat{m}\left(\mathbf{x}_0,t_0\right)$, which estimator is given by

$$\begin{aligned}
\hat{m}\left(\mathbf{x}_0, t_0\right) &= \frac{1}{N_\omega - \Phi_\omega}\left(\sum_{i=1}^{n} \alpha_i \bar{y}_i - \sum_{i=1}^{n} \varphi_i \alpha_i \bar{y}_i\right) \\
&= \frac{1}{\sum_{i=1}^{n}\left(1-\phi_i\right)\frac{\sigma^2\omega_i}{\sigma^2+\sigma_\varepsilon^2\omega_i}} \sum_{i=1}^{n}\left(1-\phi_i\right)\frac{\sigma^2\bar{y}_i}{\sigma^2+\sigma_\varepsilon^2\omega_i} \\
&= \frac{1}{\sum_{i=1}^{n}\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2\omega_i}{\sigma^2+\sigma_\varepsilon^2\omega_i}} \sum_{i=1}^{n}\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2\bar{y}_i}{\sigma^2+\sigma_\varepsilon^2\omega_i} \qquad \text{(B.43)}
\end{aligned}$$

Now we define $A_i = \frac{\sigma_\varepsilon^2\omega_i}{\sigma^2}$ and $B_i = \frac{\sigma_\varepsilon^2\bar{y}_i}{\sigma^2}$. We have thus,

$$\frac{A_i}{1+A_i} = \frac{\sigma_\varepsilon^2\omega_i}{\sigma^2+\sigma_\varepsilon^2\omega_i},$$

and

$$\frac{B_i}{1+A_i} = \frac{\sigma_\varepsilon^2\bar{y}_i}{\sigma^2+\sigma_\varepsilon^2\omega_i},$$

and we can write $\hat{m}\left(\mathbf{x}_0, t_0\right)$ as

$$\begin{aligned}
\hat{m}\left(\mathbf{x}_0, t_0\right) &= \frac{1}{\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}} \sum_{i=1}^{n}\frac{\left(1-\phi_i\right)B_i}{1+A_i} \\
&= \frac{1}{\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}} \sum_{i=1}^{n}\frac{\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2}{\sigma^2}\sum\limits_{j=1}^{n_i}\sum\limits_{p_1,p_2,p_3} y_{p_1p_2p_3,t_{i_j}}\,\omega_{p_1p_2p_3,t_{i_j}}^{(jkl,t_0)}}{1+A_i} \\
&= \frac{1}{\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}} \sum_{i=1}^{n}\frac{\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2}{\sigma^2}\sum\limits_{j=1}^{n_i}\sum\limits_{p} y_{p,t_{i_j}}\,\omega_{p,t_{i_j}}^{(jkl,t_0)}}{1+A_i} \\
&= \frac{1}{\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}} \sum_{i=1}^{n}\frac{\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2}{\sigma^2}\sum\limits_{j=1}^{n_i}\sum\limits_{p}\left(y_{p,t_{i_j}} - m\left(\mathbf{x}_0, t_0\right) + m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{1+A_i} \\
&= m\left(\mathbf{x}_0, t_0\right) + \frac{1}{\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}} \sum_{i=1}^{n}\frac{\left(1-\phi_i\right)\frac{\sigma_\varepsilon^2}{\sigma^2}\sum\limits_{j=1}^{n_i}\sum\limits_{p}\left(y_{p,t_{i_j}} - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{1+A_i} \\
&= m\left(\mathbf{x}_0, t_0\right) + \sum_{i=1}^{n}\left(1-\phi_i\right) r_i \frac{\sum_{h=1}^{n_i}\sum_p\left(y_{p,t_{i_j}} - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{h=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}} \\
&= m\left(\mathbf{x}_0, t_0\right) + \sum_{i=1}^{n}\left(1-\phi_i\right) r_i Z_i, \qquad \text{(B.44)}
\end{aligned}$$

with $r_i = \frac{A_i}{1+A_i}\left(\sum_{i=1}^{n}\frac{(1-\phi_i)A_i}{1+A_i}\right)^{-1}$ and $Z_i = \frac{\sum_{j=1}^{n_i}\sum_p\left(y_{p,t_{ij}}-m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{ij}}^{(jkl,t_0)}}{\sum_{j=1}^{n_i}\sum_p\omega_{p,t_{ij}}^{(jkl,t_0)}}$. One can show that, $\omega_i = \sum_{j=1}^{n_i}\sum_p \omega_{p,t_{ij}}^{(jkl,t_0)} = E\left[\omega_i\right] + O_p\left(\left(\mathrm{Var}\left(\omega_i\right)\right)^{1/2}\right)$. We first compute

$$E\left[\omega_i\right] = E\left[\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{ij}}^{(jkl,t_0)}\right] = \sum_{j=1}^{n_i}\sum_p E\left[\omega_{p,t_{ij}}^{(jkl,t_0)}\right]. \tag{B.45}$$

Note that $E\left[\omega_{p,t_{ij}}^{(jkl,t_0)}\right] = E\left[\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right] = \frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)E\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right]$. This comes from the fact that the spatial points are fixed and not random. As only the time points near to t_0 have a significant weight, to compute the former expectation we can use a Taylor expansion of the density $f(\cdot)$ around t_0. Thus,

$$\begin{aligned}
E\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right] &= \int \frac{1}{h_t}k_t\left(\frac{t-t_0}{h_t}\right)f(t)\,dt \\
&= \int k_t(v)\,f(t_0+vh_t)\,dv \\
&= \int k_t(v)\left(f(t_0)+f'(t_0)\,vh_t+\frac{f''(t_0)}{2}v^2h_t^2\right)dv \\
&= f(t_0) + h_t^2\frac{f''(t_0)}{2}B(k) \\
&\approx f(t_0).
\end{aligned} \tag{B.46}$$

And the expected value of ω_i can be written as

$$\begin{aligned}
E\left[\omega_i\right] &= n_i f(t_0)\sum_p \frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right) \\
&= n_i f(t_0)\,C, \text{ with } C = \sum_p \frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right).
\end{aligned} \tag{B.47}$$

To calculate the variance of ω_i notice that, as time points are iid, we have

$$\begin{aligned}
\mathrm{Var}\left(\omega_i\right) &= \mathrm{Var}\left(\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{ij}}^{(jkl,t_0)}\right) \\
&= \sum_{j=1}^{n_i}\sum_{p_1}\sum_{p_2}\frac{1}{h_x^2}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\mathrm{Var}\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right].
\end{aligned} \tag{B.48}$$

To compute $\mathrm{Var}\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right]$ we calculate first

$$\begin{aligned} E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right] &= \frac{1}{h_t^2}\int k_t^2\left(\frac{t-t_0}{h_t}\right)f\left(t\right)dt \\ &= \frac{1}{h_t}\int k_t^2\left(v\right)\left(f\left(t_0\right)+f'\left(t_0\right)vh_t+\frac{f''\left(t_0\right)}{2}v^2h_t^2\right)dv \\ &\approx \frac{V\left(k_t\right)f\left(t_0\right)}{h_t}. \end{aligned} \tag{B.49}$$

Then, noticing that $h_t \to 0$, we can write

$$\begin{aligned} \mathrm{Var}\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right] &= E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right]-\left(E\left[\frac{1}{h_t}k_t\left(\frac{t_{ij}-t_0}{h_t}\right)\right]\right)^2 \\ &= \frac{V\left(k_t\right)f\left(t_0\right)}{h_t}-f^2\left(t_0\right) \\ &= \frac{V\left(k_t\right)f\left(t_0\right)-h_tf^2\left(t_0\right)}{h_t} \\ &\approx \frac{V\left(k_t\right)f\left(t_0\right)}{h_t} \end{aligned} \tag{B.50}$$

Now we can write $\mathrm{Var}\left(\omega_i\right)$ as

$$\mathrm{Var}\left(\omega_i\right) = n_iC^2\frac{V\left(k_t\right)f\left(t_0\right)}{h_t} = n_i^2O\left(\left(n_ih_t\right)^{-1}\right), \tag{B.51}$$

and ω_i can be approximated as

$$\begin{aligned} \omega_i &= \sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)} = E\left[\omega_i\right]+O_p\left(\left(\mathrm{Var}\left(\omega_i\right)\right)^{1/2}\right) \\ &= n_iCf\left(t_0\right)+O_p\left(\sqrt{n_i^2O\left(\left(n_ih_t\right)^{-1}\right)}\right) \\ &= n_iCf\left(t_0\right)\left[1+O_p\left(\left(n_ih_t\right)^{-1/2}\right)\right]. \end{aligned} \tag{B.52}$$

And then,

$$\begin{aligned} \frac{A_i}{1+A_i} &= 1-\frac{1}{1+A_i} = 1-\frac{1}{1+\frac{\sigma_\varepsilon^2}{\sigma^2}n_iCf\left(t_0\right)\left[1+O_p\left(\left(n_ih_t\right)^{-1/2}\right)\right]} \\ &= 1+O_p\left(n_i^{-1}\right), \end{aligned} \tag{B.53}$$

which leads to

$$\begin{aligned}\sum_{i=n}^{n}(1-\phi_i)\frac{A_i}{1+A_i} &= N_1^* + \sum_{i=n}^{n}(1-\phi_i)\,O_p\left(n_i^{-1}\right)\\ &= N_1^* + \frac{N_1^*}{N_1^*}\sum_{i=n}^{n}(1-\phi_i)\,O_p\left(n_i^{-1}\right)\\ &= N_1^* + N_1^* O_p\left(m_1^{*-1}\right) = N_1^*\left[1+O_p\left(m_1^{*-1}\right)\right], \qquad \text{(B.54)}\end{aligned}$$

and

$$\begin{aligned}r_i &= \frac{A_i}{1+A_i}\left(\sum_{i=1}^{n}\frac{(1-\phi_i)\,A_i}{1+A_i}\right)^{-1}\\ &= \frac{1+O_p\left(n_i^{-1}\right)}{N_1^*\left[1+O_p\left(m_1^{*-1}\right)\right]} = \frac{1}{N_1^*}\left[1+O_p\left(m_1^{*-1}\right)\right]. \qquad \text{(B.55)}\end{aligned}$$

On the other hand, let $Q_i = Z_i|\mathbf{D}$. We want to express Q_i as $E[Q_i] + O_p\left(\sqrt{\operatorname{Var}(Q_i)}\right)$. Notice that for individuals in group T,

$$E[Q_i] = \frac{\sum_{j=1}^{n_i}\sum_p\left(m(\mathbf{x}_p,t_{ij}) - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{n_i C f(t_0)\left[1+O_p\left((n_i h_t)^{-1/2}\right)\right]}. \qquad \text{(B.56)}$$

Now, let $R_i = \sum_{j=1}^{n_i}\sum_p\left(m(\mathbf{x}_p,t_{ij}) - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}$. Then,

$$E[R_i] = n_i\sum_p E\left[\left(m(\mathbf{x}_p,t) - m(\mathbf{x}_0,t_0)\right)\omega_{p,t}^{(jkl,t_0)}\right]. \qquad \text{(B.57)}$$

$$\begin{aligned}&E\left[\left(m(\mathbf{x}_p,t) - m(\mathbf{x}_0,t_0)\right)\omega_{p,t}^{(jkl,t_0)}\right] =\\ &\quad= \int\frac{\left(m(\mathbf{x}_p,t) - m(\mathbf{x}_0,t_0)\right)}{h_x h_t}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)k_t\left(\frac{t_{ij}-t_0}{h_x}\right)f(t)\,dt\\ &= \frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\left(\int k_t(v)\left(m(\mathbf{x}_p,t_0+vh_t) - m(\mathbf{x}_0,t_0)\right)f(t_0+vh_t)\,dv\right).\end{aligned}$$

$$\text{(B.58)}$$

To compute the last integral we use a second order Taylor expansion of $m(\mathbf{x}_p, t_0+vh_t)$ around $(\mathbf{x}_0,t_0)$ and we ignore high order terms. Then,

$$\begin{aligned}
&\int k_t(v)\left(m(\mathbf{x}_p, t_0 + vh_t) - m(\mathbf{x}_0, t_0)\right) f(t_0 + vh_t)\, dv = \\
&= \int k_t(v)\left(m(\mathbf{x}_0, t_0) + \frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0) + \frac{\partial m(\mathbf{x}_0, t_0)}{\partial t} v h_t - m(\mathbf{x}_0, t_0)\right) \times \\
&\qquad \times \left(f(t_0) + f'(t_0)\, v h_t + \frac{f''(t_0)}{2} v^2 h_t^2\right) dv \\
&= f(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0) + \frac{f''(t_0)}{2} B(k)\, h_t^2 \frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0) + \\
&\qquad + f'(t_0)\, B(k)\frac{\partial m(\mathbf{x}_0, t_0)}{\partial t} h_t^2 \\
&\qquad\qquad \approx f(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0). \qquad \text{(B.59)}
\end{aligned}$$

The expected value of $(m(\mathbf{x}_p, t) - m(\mathbf{x}_0, t_0))\,\omega_{p,t}^{(jkl,t_0)}$ is therefore given by

$$\begin{aligned}
E\left[(m(\mathbf{x}_p, t) - m(\mathbf{x}_0, t_0))\,\omega_{p,t}^{(jkl,t_0)}\right] &= \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p - \mathbf{x}_0\|}{h_x}\right)\left(f(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0)\right. \\
&\quad + h_t^2 B(k)\left[\frac{f''(t_0)}{2}\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0)\right. \\
&\quad \left.\left. + f'(t_0)\frac{\partial m(\mathbf{x}_0, t_0)}{\partial t}\right]\right) \\
&\approx \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p - \mathbf{x}_0\|}{h_x}\right) f(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0).
\end{aligned} \tag{B.60}$$

Then, $E[R_i]$ can be calculated as

$$\begin{aligned}
E[R_i] &= n_i f(t_0)\sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p - \mathbf{x}_0\|}{h_x}\right)\left(f(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0)\right. \\
&\quad \left. + h_t^2 B(k)\left[\frac{f''(t_0)}{2}\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0) + f'(t_0)\frac{\partial m(\mathbf{x}_0, t_0)}{\partial t}\right]\right) \\
&= n_i\left[f(t_0)\, C_2 + B(k)\, h_t^2\left(\frac{f''(t_0)}{2} C_2 + f'(t_0)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial t} C\right)\right] \\
&\approx n_i f(t_0)\, C_2, \text{ with } C_2 = \sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p - \mathbf{x}_0\|}{h_x}\right)\frac{\partial m^T(\mathbf{x}_0, t_0)}{\partial \mathbf{x}}(\mathbf{x}_p - \mathbf{x}_0).
\end{aligned} \tag{B.61}$$

To compute $\mathrm{Var}\left[R_i\right]$ we notice that

$$\begin{aligned}
&\mathrm{Var}\left[R_i\right] = \mathrm{Cov}\left(R_i, R_i\right) \\
&= \mathrm{Cov}\left(\sum_{j=1}^{n_i}\sum_{p}\left(m\left(\mathbf{x}_p, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}, \sum_{j=1}^{n_i}\sum_{p}\left(m\left(\mathbf{x}_p, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}\right) \\
&= \sum_{j=1}^{n_i}\sum_{p_1}\sum_{p_2}\mathrm{Cov}\left(\left(m\left(\mathbf{x}_{p_1}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p_1,t_{i_j}}^{(jkl,t_0)}, \left(m\left(\mathbf{x}_{p_2}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right) \\
&\qquad = n_i\sum_{p_1}\sum_{p_2}\left(\frac{1}{h_x^2}k_x\left(\frac{\left\|\mathbf{x}_{p_1} - \mathbf{x}_0\right\|}{h_x}\right)k_x\left(\frac{\left\|\mathbf{x}_{p_2} - \mathbf{x}_0\right\|}{h_x}\right)\times\right. \\
&\left.\mathrm{Cov}\left(\frac{\left(m\left(\mathbf{x}_{p_1}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right), \frac{\left(m\left(\mathbf{x}_{p_2}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right)\right)\right).
\end{aligned} \tag{B.62}$$

To calculate this former covariance we first compute

$$\begin{aligned}
&E\left[\frac{\left(m\left(\mathbf{x}_{p_1}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right)\times\frac{\left(m\left(\mathbf{x}_{p_2}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right)\right] = \\
&= \frac{1}{h_t}\int k_t^2\left(v\right)\left(m\left(\mathbf{x}_{p_1}, t_0 + vh_t\right) - m\left(\mathbf{x}_0, t_0\right)\right)\left(m\left(\mathbf{x}_{p_2}, t_0 + vh_t\right) - m\left(\mathbf{x}_0, t_0\right)\right)f\left(t_0 + vh_t\right)dv \\
&\quad = \frac{1}{h_t}\int\left[k_t^2\left(v\right)\left(\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_1} - \mathbf{x}_0\right) + \frac{\partial m\left(\mathbf{x}_0, t_0\right)}{\partial t}vh_t\right)\left(\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_2} - \mathbf{x}_0\right)\right.\right. \\
&\qquad\qquad \left.\left.+ \frac{\partial m\left(\mathbf{x}_0, t_0\right)}{\partial t}vh_t\right)\left(f\left(t_0\right) + f'\left(t_0\right)vh_t + \frac{f''\left(t_0\right)}{2}v^2h_t^2\right)\right]dv \\
&\qquad\qquad = \frac{V\left(k_t\right)}{h_t}\left[f\left(t_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_1} - \mathbf{x}_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_2} - \mathbf{x}_0\right)\right].
\end{aligned} \tag{B.63}$$

Then, using that $\mathrm{Cov}\left(X, Y\right) = E\left[XY\right] - E\left[X\right]E\left[Y\right]$, eq. (B.60) and that

$$E\left[\frac{\left(m\left(\mathbf{x}_{p_1}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right)\right] \approx V\left(k_t\right)f\left(t_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_p - \mathbf{x}_0\right)$$

we have that

$$\begin{aligned}
&\mathrm{Cov}\left(\frac{\left(m\left(\mathbf{x}_{p_1}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right), \frac{\left(m\left(\mathbf{x}_{p_2}, t_{ij}\right) - m\left(\mathbf{x}_0, t_0\right)\right)}{h_t}k_t\left(\frac{t_{ij} - t_0}{h_t}\right)\right) = \\
&\qquad = \frac{V\left(k_t\right)}{h_t}\left[f\left(t_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_1} - \mathbf{x}_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_2} - \mathbf{x}_0\right)\right] \\
&\qquad\quad - f^2\left(t_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_1} - \mathbf{x}_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_2} - \mathbf{x}_0\right) \\
&\qquad\quad \approx \frac{V\left(k_t\right)}{h_t}\left[f\left(t_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_1} - \mathbf{x}_0\right)\frac{\partial m^T\left(\mathbf{x}_0, t_0\right)}{\partial\mathbf{x}}\left(\mathbf{x}_{p_2} - \mathbf{x}_0\right)\right].
\end{aligned} \tag{B.64}$$

We can now calculate $\operatorname{Var}[R_i]$ as

$$\begin{aligned}\operatorname{Var}[R_i] = n_i \sum_{p_1}\sum_{p_2} &\left(\frac{1}{h_x^2} k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right) k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right) \times \right.\\ &\left.\frac{V(k_t)}{h_t}\left[f(t_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}_{p_1}-\mathbf{x}_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}_{p_2}-\mathbf{x}_0)\right]\right)\\ &= \frac{V(k_t)\,n_i}{h_t} f(t_0)\,C_2^2 = n_i^2\frac{V(k_t)}{n_i h_t} f(t_0)\,C_2^2 = n_i^2 O\left((n_i h_t)^{-1}\right). \end{aligned} \tag{B.65}$$

Thus, R_i can be approximated as follows

$$\begin{aligned} R_i &= \sum_{j=1}^{n_i}\sum_p \left(m(\mathbf{x}_p,t_{ij}) - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{ij}}^{(jkl,t_0)}\\ &= E[R_i] + O_p\left(\sqrt{\operatorname{Var}(R_i)}\right)\\ &= n_i\left[f(t_0)\,C_2 + B(k)\,h_t^2\left(\frac{f''(t_0)}{2}C_2 + f'(t_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}C\right)\right] + O_p\left(\sqrt{n_i^2 O\left((n_i h_t)^{-1}\right)}\right)\\ &= n_i\left[f(t_0)\,C_2 + B(k)\,h_t^2\left(\frac{f''(t_0)}{2}C_2 + f'(t_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}C\right)\right]\left[1 + O\left((n_i h_t)^{-1/2}\right)\right]\\ &\approx n_i f(t_0)\,C_2\left[1 + O\left((n_i h_t)^{-1/2}\right)\right]. \end{aligned} \tag{B.66}$$

Now, $E[Z_i|\mathbf{D}]$ can be written as

$$\begin{aligned} E[Z_i|\mathbf{D}] &= \frac{n_i\left[f(t_0)\,C_2 + B(k)\,h_t^2\left(\frac{f''(t_0)}{2}C_2 + f'(t_0)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}C\right)\right]\left[1 + O_p\left((n_i h_t)^{-1/2}\right)\right]}{n_i C f(t_0)\left[1 + O_p\left((n_i h_t)^{-1/2}\right)\right]}\\ &= \left[\frac{C_2}{C} + B(k)\,h_t^2\left(\frac{f''(t_0)}{2f(t_0)}\frac{C_2}{C} + \frac{f'(t_0)}{f(t_0)}\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}\right)\right]\left[1 + O_p\left((n_i h_t)^{-1/2}\right)\right]. \end{aligned} \tag{B.67}$$

From (B.44) and from the fact that $(n_i h_t)^{-1/2} = o(1)$, we can calculate the bias of $\hat{m}(\mathbf{x}_0,t_0)$ conditioned to $\mathbf{D}$ as

$$\begin{aligned} \operatorname{Bias}(\hat{m}(\mathbf{x}_0,t_0)|\mathbf{D}) &= E[\hat{m}(\mathbf{x}_0,t_0)|\mathbf{D}] - m(\mathbf{x}_0,t_0) = E\left[\sum_{i=1}^n (1-\phi_i)\,r_i Z_i|\mathbf{D}\right] =\\ = \sum_{i=1}^n &\left((1-\phi_i)\frac{1}{N_1^*}\left[1 + O_p\left(m_1^{*-1}\right)\right]\left[\frac{C_2}{C} + B(k)\,h_t^2\left(\frac{f''(t_0)}{2f(t_0)}\frac{C_2}{C} + \frac{f'(t_0)}{f(t_0)}\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}\right)\right]\right.\\ &\left.\times\left[1 + O_p\left((n_i h_t)^{-1/2}\right)\right]\right)\\ &= \left[\frac{C_2}{C} + B(k)\,h_t^2\left(\frac{f''(t_0)}{2f(t_0)}\frac{C_2}{C} + \frac{f'(t_0)}{f(t_0)}\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial t}\right)\right]\left[1 + o_p(1)\right]. \end{aligned} \tag{B.68}$$

Remark Notice that the term $\frac{C_2}{C} = \frac{\sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}_p-\mathbf{x}_0)}{\sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)}$ tends to $\sum_p \delta_{\mathbf{x}_0}(\mathbf{x}_p)\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}(\mathbf{x}_p-\mathbf{x}_0) = 0$, when $h_x \to \infty$, where $\delta_{\mathbf{x}_0}(\mathbf{x}_p)$ equals to 1 if $\mathbf{x}_p = \mathbf{x}_0$ and to 0 otherwise. In addition,

$$\left|\frac{C_2}{C}\right| \leq \left\|\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}\right\| \left\|\frac{\sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)(\mathbf{x}_p-\mathbf{x}_0)}{\sum_p \frac{1}{h_x} k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)}\right\| \tag{B.69}$$

$$\leq h_x \left\|\frac{\partial m^T(\mathbf{x}_0,t_0)}{\partial \mathbf{x}}\right\| \frac{\sum_p k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)}{\sum_p k_x\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)} = O(h_x) = o(1). \tag{B.70}$$

Our expression for the asymptotic bias depends on both h_t and h_x.

B.4.1.4 Asymptotic variance

Now we are going to derive the expression for the asymptotic variance. We shall start by $\mathrm{Var}(Z_i|\mathbf{D})$,

$$\begin{aligned}\mathrm{Var}(Z_i|\mathbf{D}) &= \mathrm{Var}\left(\frac{\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}}|\mathbf{D}\right)\\ &= \frac{\mathrm{Var}\left(\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}|\mathbf{D}\right)}{\left(\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2}.\end{aligned} \tag{B.71}$$

Using the properties of the variance of a sum we have

$$\begin{aligned}&\mathrm{Var}\left(\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m(\mathbf{x}_0,t_0)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}|\mathbf{D}\right) =\\ &\qquad = \sum_{j=1}^{n_i}\sum_p \mathrm{Var}\left(y_{p,t_{i_j}}\right)\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2 +\\ &\qquad + \sum\sum\sum\sum\nolimits_{j_1\neq j_2 \vee p_1\neq p_2} \omega_{p_1,t_{i_{j_1}}}^{(jkl,t_0)}\omega_{p_2,t_{i_{j_2}}}^{(jkl,t_0)}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right).\end{aligned} \tag{B.72}$$

We are now going to assume that $\mathrm{Var}\left(y_{p,t_{i_j}}\right) \approx \sigma^2(\mathbf{x}_0,t_0) + \sigma_\varepsilon^2(\mathbf{x}_0,t_0)$ for both voxels around $\mathbf{x}_0$ and time points around t_0. This implies that the terms in the first order Taylor expansion of $\sigma^2(\mathbf{x}_0,t_0) + \sigma_\varepsilon^2(\mathbf{x}_0,t_0)$ are negligible for voxels around $\mathbf{x}_0$ and time points around t_0. Under this assumption, the first term in the right side of the expression above can be expressed as

$$\sum_{j=1}^{n_i}\sum_p \mathrm{Var}\left(y_{p,t_{i_j}}\right)\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2 = \left(\sigma^2(\mathbf{x}_0,t_0) + \sigma_\varepsilon^2(\mathbf{x}_0,t_0)\right)\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2 \tag{B.73}$$

And now we can express $\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2$ as

$$\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2 = E\left[\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] + O_p\left(\sqrt{\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right)}\right). \tag{B.74}$$

This expectation can be computed as follows

$$E\left[\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] = \sum_{j=1}^{n_i}\sum_p E\left[\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] = n_i \sum_p E\left[\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right], \tag{B.75}$$

where

$$\begin{aligned} E\left[\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] = E\left[\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right] = \\ = \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right]. \end{aligned} \tag{B.76}$$

We have that

$$\begin{aligned} E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right] &= \frac{1}{h_t^2}\int k_t^2\left(\frac{t-t_0}{h_t}\right)f(t)\,dt \\ &= \frac{1}{h_t}\int k_t^2(v)\,f(t_0+vh_t)\,dv \\ &= \frac{1}{h_t}\int k_t^2(v)\left(f(t_0)+f'(t_0)\,vh_t+\frac{f''(t_0)}{2}v^2h_t^2\right)dv \\ &\approx \frac{f(t_0)\,V(k_t)}{h_t}, \end{aligned} \tag{B.77}$$

which leads to

$$E\left[\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] = \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{f(t_0)\,V(k_t)}{h_t}. \tag{B.78}$$

We can thus compute the expected value of $\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2$ as

$$\begin{aligned} E\left[\sum_{j=1}^{n_i}\sum_p \left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right] &= n_i\sum_p \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\right] \\ &= n_i\sum_p \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\frac{f(t_0)\,V(k_t)}{h_t} \\ &= \frac{n_iC^{(2)}f(t_0)\,V(k_t)}{h_t},\ \text{with } C^{(2)} = \sum_p \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right). \end{aligned} \tag{B.79}$$

Now, to compute the variance of $\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2$ we first notice that

$$\begin{aligned}\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_{p}\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right) &= \operatorname{Cov}\left(\sum_{j_1=1}^{n_i}\sum_{p_1}\left(\omega_{p_1,t_{i_{j_1}}}^{(jkl,t_0)}\right)^2,\sum_{j_2=1}^{n_i}\sum_{p_2}\left(\omega_{p_2,t_{i_{j_2}}}^{(jkl,t_0)}\right)^2\right)\\ &= \sum_{j=1}^{n_i}\sum_{p_1}\sum_{p_2}\operatorname{Cov}\left(\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2,\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right)\\ &= n_i\sum_{p_1}\sum_{p_2}\operatorname{Cov}\left(\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2,\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right). \qquad \text{(B.80)}\end{aligned}$$

To compute the covariance terms within the sum we calculate first

$$\begin{aligned}E\left[\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right] &= E\left[\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_t^4}k_t^4\left(\frac{t_{ij}-t_0}{h_x}\right)\right]\\ &= \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)E\left[\frac{1}{h_t^4}k_t^4\left(\frac{t_{ij}-t_0}{h_x}\right)\right]. \qquad \text{(B.81)}\end{aligned}$$

This requires us to calculate

$$\begin{aligned}E\left[\frac{1}{h_t^4}k_t^4\left(\frac{t_{ij}-t_0}{h_t}\right)\right] &= \frac{1}{h_t^4}\int k_t^4\left(\frac{t-t_0}{h_t}\right)f(t)\,dt\\ &= \frac{1}{h_t^3}\int k_t^4(v)\,f(t_0+vh_t)\,dv\\ &= \frac{1}{h_t^3}\int k_t^4(v)\left(f(t_0)+f'(t_0)\,vh_t+\frac{f''(t_0)}{2}v^2h_t^2\right)dv\\ &\approx \frac{f(t_0)\,V^{(2)}(k_t)}{h_t^3}, \text{ with } V^{(2)}(k_t)=\int k_t^4(v)\,dv. \qquad \text{(B.82)}\end{aligned}$$

Then, this former expectation can be written as

$$\begin{aligned}E\left[\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right] &=\\ &= \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{f(t_0)\,V^{(2)}(k_t)}{h_t^3}, \qquad \text{(B.83)}\end{aligned}$$

which allows for the covariance to be written as

$$\begin{aligned}
&\operatorname{Cov}\left(\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2,\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right)= \\
&\qquad = E\left[\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right]-E\left[\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2\right]E\left[\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right] \\
&\qquad = \frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{h_t^3} \\
&\qquad -\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{f\left(t_0\right)V\left(k_t\right)}{h_t}\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{f\left(t_0\right)V\left(k_t\right)}{h_t} \\
&\qquad = \frac{1}{h_x^4}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\left(\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{h_t^3}-\frac{f^2\left(t_0\right)V^2\left(k_t\right)}{h_t^2}\right) \\
&\qquad \approx \frac{1}{h_x^4}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{h_t^3}. \qquad \text{(B.84)}
\end{aligned}$$

Finally, for this part we can express $\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right)$ as

$$\begin{aligned}
\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right) &= \sum_{j=1}^{n_i}\sum_{p_1}\sum_{p_2}\operatorname{Cov}\left(\left(\omega_{p_1,t_{i_j}}^{(jkl,t_0)}\right)^2,\left(\omega_{p_2,t_{i_j}}^{(jkl,t_0)}\right)^2\right) \\
&= n_i\sum_{p_1}\sum_{p_2}\frac{1}{h_x^4}k_x^2\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)k_x^2\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{h_t^3} \\
&= n_i\left(C^{(2)}\right)^2\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{h_t^3} = \frac{n_i^2}{h_t^2}\left(C^{(2)}\right)^2\frac{f\left(t_0\right)V^{(2)}\left(k_t\right)}{n_ih_t} \\
&= \frac{n_i^2}{h_t^2}O\left(\left(n_ih_t\right)^{-1}\right). \qquad \text{(B.85)}
\end{aligned}$$

Then, an asymptotic approximation for the first term in the sum in (B.72) is given by

$$\begin{aligned}
\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2 &= E\left[\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right]+O_p\left(\sqrt{\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_p\left(\omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2\right)}\right) \\
&= \frac{n_iC^{(2)}f\left(t_0\right)V\left(k_t\right)}{h_t}+O_p\left(\sqrt{\frac{n_i^2}{h_t^2}O\left(\left(n_ih_t\right)^{-1}\right)}\right) \\
&= \frac{n_iC^{(2)}f\left(t_0\right)V\left(k_t\right)}{h_t}\left[1+O_p\left(\left(n_ih_t\right)^{-1/2}\right)\right]. \qquad \text{(B.86)}
\end{aligned}$$

For the second term in the right side in (B.72) we are going to use the same asymptotic equivalence to express

$$\begin{aligned}
&\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)=\\
&\qquad=E\left[\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]\\
&+O_p\left(\sqrt{\operatorname{Var}\left(\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right)}\right). \qquad \text{(B.87)}
\end{aligned}$$

In B.87 the $\sum\sum\sum\sum_{j_1\neq j_2\vee p_1\neq p_2}$ means double summation over all time points t_{j_1} and t_{j_2} such as $j_1\neq j_2$ and double summation over all voxels $\mathbf{x}_{p_1}$ and $\mathbf{x}_{p_2}$ such as $p_1\neq p_2$. We firstly compute the expectation as

$$\begin{aligned}
&E\left[\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]=\\
&\qquad=\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}E\left[\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\\
&\quad=\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)E\left[\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_1}}-t_0}{h_t}\right)\times\right.\\
&\qquad\qquad\left.\times\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_2}}-t_0}{h_t}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]\\
&\quad=\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\times\\
&\qquad\qquad\times E\left[\frac{1}{h_t}k_t\left(\frac{t_{i_{j_1}}-t_0}{h_t}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_2}}-t_0}{h_t}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]. \qquad \text{(B.88)}
\end{aligned}$$

For this expectation there are two possibilities, either $j_1\neq j_2$ or $j_1=j_2$. If $j_1\neq j_2$, as time points are independent we have

$$\begin{aligned}
&E\left[\frac{1}{h_t}k_t\left(\frac{t_{i_{j_1}}-t_0}{h_t}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_2}}-t_0}{h_t}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right]=\\
&\qquad=\int\int\frac{1}{h_t}k_t\left(\frac{t_1-t_0}{h_t}\right)\frac{1}{h_t}k_t\left(\frac{t_2-t_0}{h_t}\right)\gamma_t\left(t_1,t_2\right)f\left(t_1\right)f\left(t_2\right)dt_1dt_2\\
&\quad=\int\int k_t\left(v_1\right)k_t\left(v_2\right)\gamma_t\left(t_0+h_tv_1,t_0+h_tv_2\right)f\left(t_0+h_tv_1\right)f\left(t_0+h_tv_2\right)dv_1dv_2\\
&\quad=\int\int\left[k_t\left(v_1\right)k_t\left(v_2\right)\left(\gamma_t\left(t_0,t_0\right)+\frac{\partial\gamma_t\left(t_0,t_0\right)}{\partial t_1}h_tv_1+\frac{\partial\gamma_t\left(t_0,t_0\right)}{\partial t_2}h_tv_2\right)\right.\\
&\left.\left(f\left(t_0\right)+f'\left(t_0\right)v_1h_t+\frac{f''\left(t_0\right)}{2}v_1^2h_t^2\right)\left(f\left(t_0\right)+f'\left(t_0\right)v_2h_t+\frac{f''\left(t_0\right)}{2}v_2^2h_t^2\right)\right]dv_1dv_2\\
&\qquad\approx f^2\left(t_0\right)\gamma_t\left(t_0,t_0\right)+h_t^2f''\left(t_0\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)B\left(k_t\right). \qquad \text{(B.89)}
\end{aligned}$$

If $j_1 = j_2$,

$$
\begin{aligned}
E\Bigg[\frac{1}{h_t}k_t\left(\frac{t_{i_{j_1}}-t_0}{h_t}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_2}}-t_0}{h_t}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Bigg] &= \\
&= \int \frac{1}{h_t^2}k_t^2\left(\frac{t-t_0}{h_t}\right)\gamma_t(t,t)\,f(t)\,dt \\
&= \frac{1}{h_t}\int k_t(v)\,\gamma_t(t_0+h_tv,t_0+h_tv)\,f(t_0+h_tv)\,dv \\
&= \frac{1}{h_t}\int \Bigg[k_t(v)\left(\gamma_t(t_0,t_0)+\frac{\partial\gamma_t(t_0,t_0)}{\partial t_1}h_tv+\frac{\partial\gamma_t(t_0,t_0)}{\partial t_2}h_tv\right) \\
&\qquad \left(f(t_0)+f'(t_0)\,v_1h_t+\frac{f''(t_0)}{2}v_1^2h_t^2\right)\Bigg]dv \\
&= \frac{V(k_t)\,f(t_0)\,\gamma_t(t_0,t_0)}{h_t}.
\end{aligned}
\tag{B.90}
$$

Now the expectation in (B.87) can be expressed as

$$
\begin{aligned}
E\Big[\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Big] &= \\
E\Big[\sum\sum\sum\sum\nolimits_{j_1\neq j_2,p_1,p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Big] & \\
+E\Big[\sum\sum\sum\sum\nolimits_{j_1=j_2,p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Big]&.
\end{aligned}
\tag{B.91}
$$

In one hand we have (when $j_1 \neq j_2$),

$$
\begin{aligned}
E\Big[\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1,p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Big] &= \\
= n_i(n_i-1)\sum_{p_1}\sum_{p_2}\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,E\left[\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\right] & \\
= n_i(n_i-1)\sum_{p_1}\sum_{p_2}\Bigg\{\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right) & \\
E\Bigg[\frac{1}{h_t}k_t\left(\frac{t_{i_{j_1}}-t_0}{h_t}\right)\frac{1}{h_t}k_t\left(\frac{t_{i_{j_2}}-t_0}{h_t}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)\Bigg]\Bigg\} & \\
= n_i(n_i-1)\left(f^2(t_0)\,\gamma_t(t_0,t_0)+h_t^2f''(t_0)\,f(t_0)\,\gamma_t(t_0,t_0)\,B(k_t)\right) & \\
\times\sum_{p_1}\sum_{p_2}\Bigg\{\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\Bigg\} & \\
= n_i(n_i-1)\,\Gamma_1\times\left(f^2(t_0)\,\gamma_t(t_0,t_0)+h_t^2f''(t_0)\,f(t_0)\,\gamma_t(t_0,t_0)\,B(k_t)\right)&,
\end{aligned}
\tag{B.92}
$$

with $\Gamma_1 = \sum_{p_1}\sum_{p_2}\left\{\gamma_x(\mathbf{x}_{p_1},\mathbf{x}_{p_2})\,\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\right\}$. On the other hand we have (when $j_1 = j_2$),

$$
\begin{aligned}
&E\left[\sum\sum\sum\sum\nolimits_{j_1=j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{ij_1}}\omega^{(jkl,t_0)}_{p_2,t_{ij_2}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{ij_1},t_{ij_2}\right)\right]=\\
&\qquad= n_i\sum\sum\nolimits_{p_1\neq p_2}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)E\left[\omega^{(jkl,t_0)}_{p_1,t_{ij_1}}\omega^{(jkl,t_0)}_{p_2,t_{ij_2}}\gamma_t\left(t_{ij_1},t_{ij_2}\right)\right]\\
&\qquad= n_i\sum\sum\nolimits_{p_1\neq p_2}\left\{\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\right.\\
&\qquad\qquad\left.E\left[\frac{1}{h_t^2}k_t^2\left(\frac{t_{ij}-t_0}{h_t}\right)\gamma_t\left(t_{ij},t_{ij}\right)\right]\right\}\\
&\qquad= n_i\frac{V\left(k_t\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)}{h_t}\\
&\qquad\times\sum\sum\nolimits_{p_1\neq p_2}\left\{\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_1}-\mathbf{x}_0\|}{h_x}\right)\frac{1}{h_x}k_x\left(\frac{\|\mathbf{x}_{p_2}-\mathbf{x}_0\|}{h_x}\right)\right\}\\
&\qquad= n_i\frac{V\left(k_t\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)}{h_t}\left(\Gamma_1-\Gamma_2\right),
\end{aligned}
\tag{B.93}
$$

with $\Gamma_2=\sum_p\left\{\gamma_x\left(\mathbf{x}_p,\mathbf{x}_p\right)\frac{1}{h_x^2}k_x^2\left(\frac{\|\mathbf{x}_p-\mathbf{x}_0\|}{h_x}\right)\right\}$.

As for the expectation in (B.87), we compute the variance for both situations, namely $t_{ij_1}\neq t_{ij_2}$ or $t_{ij_1}=t_{ij_2}$. When $j_1\neq j_2$ carrying out similar calculations than before we can show that

$$
\operatorname{Var}\left[\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1,p_2}\omega^{(jkl,t_0)}_{p_1,t_{ij_1}}\omega^{(jkl,t_0)}_{p_2,t_{ij_2}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{ij_1},t_{ij_2}\right)\right]=n_i^4O\left(\left(n_ih_t\right)^{-2}\right).
\tag{B.94}
$$

On the other hand, for $j_1=j_2$ we have

$$
\operatorname{Var}\left[\sum\sum\sum\sum\nolimits_{j_1=j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{ij_1}}\omega^{(jkl,t_0)}_{p_2,t_{ij_2}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{ij_1},t_{ij_2}\right)\right]=n_i^4O\left(\left(n_ih_t\right)^{-3}\right).
\tag{B.95}
$$

From (B.92) and (B.87) we have that

$$
\begin{aligned}
&\sum\sum\sum\sum\nolimits_{j_1\neq j_2\vee p_1,p_2}\omega^{(jkl,t_0)}_{p_1,t_{ij_1}}\omega^{(jkl,t_0)}_{p_2,t_{ij_2}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{ij_1},t_{ij_2}\right)=\\
&n_i\left(n_i-1\right)\Gamma_1\times\left(f^2\left(t_0\right)\gamma_t\left(t_0,t_0\right)+h_t^2f''\left(t_0\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)B\left(k_t\right)\right)+O_p\left(\sqrt{n_i^4O\left(\left(n_ih_t\right)^{-2}\right)}\right)\\
&=n_i^2\Gamma_1\times\left(f^2\left(t_0\right)\gamma_t\left(t_0,t_0\right)+h_t^2f''\left(t_0\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)B\left(k_t\right)\right)\left[1+O_p\left(\left(n_ih_t\right)^{-1}\right)\right].
\end{aligned}
\tag{B.96}
$$

From (B.93) and (B.95) we have that

$$
\begin{aligned}
&\sum\sum\sum\sum\nolimits_{j_1=j_2\vee p_1\neq p_2}\omega^{(jkl,t_0)}_{p_1,t_{i_{j_1}}}\omega^{(jkl,t_0)}_{p_2,t_{i_{j_2}}}\gamma_x\left(\mathbf{x}_{p_1},\mathbf{x}_{p_2}\right)\gamma_t\left(t_{i_{j_1}},t_{i_{j_2}}\right)=\\
&\qquad= n_i\frac{V\left(k_t\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)}{h_t}\left(\Gamma_1-\Gamma_2\right)+O_p\left(\sqrt{n_i^4 O\left(\left(n_ih_t\right)^{-3}\right)}\right)\\
&\qquad= n_i\frac{V\left(k_t\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)}{h_t}\left(\Gamma_1-\Gamma_2\right)+n_i^2O_p\left(\left(n_ih_t\right)^{-3/2}\right)\\
&\qquad= n_i^2O_p\left(\left(n_ih_t\right)^{-1}\right)+n_i^2O_p\left(\left(n_ih_t\right)^{-3/2}\right)=n_i^2O_p\left(\left(n_ih_t\right)^{-1}\right).
\end{aligned}
\tag{B.97}
$$

Then,

$$
\begin{aligned}
&\operatorname{Var}\left(\sum_{j=1}^{n_i}\sum_{p}\left(y_{p,t_{i_j}}-m\left(\mathbf{x}_0,t_0\right)\right)\omega^{(jkl,t_0)}_{p,t_{i_j}}|\mathbf{D}\right)=\\
&\qquad=\left(\sigma^2\left(\mathbf{x}_0,t_0\right)+\sigma^2_{\varepsilon}\left(\mathbf{x}_0,t_0\right)\right)\frac{n_iC^{(2)}f\left(t_0\right)V\left(k_t\right)}{h_t}\left[1+O_p\left(\left(n_ih_t\right)^{-1/2}\right)\right]\\
&+n_i^2\Gamma_1\times\left(f^2\left(t_0\right)\gamma_t\left(t_0,t_0\right)+h_t^2f''\left(t_0\right)f\left(t_0\right)\gamma_t\left(t_0,t_0\right)B\left(k_t\right)\right)\left[1+O_p\left(\left(n_ih_t\right)^{-1}\right)\right]\\
&\qquad\qquad+n_i^2O_p\left(\left(n_ih_t\right)^{-1}\right)\\
&\qquad=\left(\sigma^2\left(\mathbf{x}_0,t_0\right)+\sigma^2_{\varepsilon}\left(\mathbf{x}_0,t_0\right)\right)\frac{n_iC^{(2)}f\left(t_0\right)V\left(k_t\right)}{h_t}\left[1+O_p\left(\left(n_ih_t\right)^{-1/2}\right)\right]\\
&\quad+n_i^2\Gamma_1\times\left(f^2\left(t_0\right)\gamma_t\left(t_0,t_0\right)\right)\left[1+O_p\left(\left(n_ih_t\right)^{-1}\right)\right]+n_i^2O_p\left(\left(n_ih_t\right)^{-1}\right).
\end{aligned}
\tag{B.98}
$$

Thus, $\mathrm{Var}\left(Z_i|\mathbf{D}\right)$ can be written as

$$
\begin{aligned}
&\mathrm{Var}\left(Z_i|\mathbf{D}\right) = \\
&\qquad = \frac{\mathrm{Var}\left(\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}|\mathbf{D}\right)}{\left(\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}\right)^2}. \\
&\qquad = \frac{\mathrm{Var}\left(\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m\left(\mathbf{x}_0, t_0\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}|\mathbf{D}\right)}{\left(n_i C f\left(t_0\right)\left[1 + O_p\left(\left(n_i h_t\right)^{-1/2}\right)\right]\right)^2} \\
&\qquad = \left(\sigma^2\left(\mathbf{x}_0, t_0\right) + \sigma_\varepsilon^2\left(\mathbf{x}_0, t_0\right)\right)\frac{C^{(2)} V\left(k_t\right)}{n_i h_t C^2 f\left(t_0\right)}\left[1 + O_p\left(\left(n_i h_t\right)^{-1/2}\right)\right] \\
&\qquad\quad + \frac{\Gamma_1 \times \gamma_t\left(t_0, t_0\right)}{C^2}\left[1 + O_p\left(\left(n_i h_t\right)^{-1/2}\right)\right] \\
&\qquad = \left(\sigma^2\left(\mathbf{x}_0, t_0\right) + \sigma_\varepsilon^2\left(\mathbf{x}_0, t_0\right)\right)\frac{C^{(2)} V\left(k_t\right)}{n_i h_t C^2 f\left(t_0\right)} + \frac{\Gamma_1 \times \gamma_t\left(t_0, t_0\right)}{C^2} + O_p\left(\left(n_i h_t\right)^{-1}\right) \\
&\qquad = \frac{\Gamma_1 \times \gamma_t\left(t_0, t_0\right)}{C^2} + O_p\left(\left(n_i h_t\right)^{-1}\right). \qquad \text{(B.99)}
\end{aligned}
$$

If $h_x \to 0$, then

$$
\begin{aligned}
\mathrm{Var}\left(Z_i|\mathbf{D}\right) &= \gamma_x\left(\mathbf{x}_0, \mathbf{x}_0\right) \times \gamma_t\left(t_0, t_0\right) + O_p\left(\left(n_i h_t\right)^{-1}\right) \\
&= \mathrm{Cov}\left(e_i\left(\mathbf{x}_0, t_0\right), e_i\left(\mathbf{x}_0, t_0\right)\right) + O_p\left(\left(n_i h_t\right)^{-1}\right) \qquad \text{(B.100)}
\end{aligned}
$$

From (B.44) we have that

$$
\begin{aligned}
\mathrm{Var}\left(\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right) &= \sum_{i=1}^{n}\left(1 - \phi_i\right) r_i^2\,\mathrm{Var}\left(Z_i|\mathbf{D}\right) \\
&= \sum_{i=1}^{n}\Bigg\{\left(1 - \phi_i\right)\left(\frac{1}{N_1^{*2}}\left[1 + O_p\left(m_1^{*-1}\right)\right]\right) \\
&\quad \times \left(\sigma^2\left(\mathbf{x}_0, t_0\right) + \sigma_\varepsilon^2\left(\mathbf{x}_0, t_0\right)\right)\frac{C^{(2)} V\left(k_t\right)}{n_i h_t C^2 f\left(t_0\right)} + \frac{\Gamma_1 \times \gamma_t\left(t_0, t_0\right)}{C^2} + O_p\left(\left(n_i h_t\right)^{-1}\right)\Bigg\} \\
&= \frac{\frac{\Gamma_1}{C^2} \times \gamma_t\left(t_0, t_0\right)}{N_1^*} + \left(\sigma^2\left(\mathbf{x}_0, t_0\right) + \sigma_\varepsilon^2\left(\mathbf{x}_0, t_0\right)\right)\frac{\frac{C^{(2)}}{C^2} V\left(k_t\right)}{N_1^* m_1^* h_t f\left(t_0\right)} \\
&\quad + o_p\left(N_1^{*-1} + \left(N_1^* m_1^* h_t\right)^{-1}\right) \\
&= \frac{\frac{\Gamma_1}{C^2} \times \gamma_t\left(t_0, t_0\right)}{N_1^*} + o_p\left(N_1^{*-1} + \left(N_1^* m_1^* h_t\right)^{-1}\right) \qquad \text{(B.101)}
\end{aligned}
$$

Remark Notice that when $h_x \to 0$, then $\mathrm{Var}\left(\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right) \to \frac{\mathrm{Cov}(e_i(\mathbf{x}_0,t_0),e_i(\mathbf{x}_0,t_0))}{N_1^*} + o_p\left(N_1^{*-1} + \left(N_1^* m_1^* h_t\right)^{-1}\right)$.

B.4.1.5 Asymptotic normality

Now, we are going to derive the asymptotic distribution for $\hat{m}\left(\mathbf{x}_0, t_0\right)$ under the previously established conditions. We shall start expressing

$$\begin{aligned}\hat{m}\left(\mathbf{x}_0, t_0\right) - E\left[\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right] &= \sum_{i=1}^{n}\left(1-\phi_i\right) r_i Z_i \\ &= \frac{1}{N_1^*}\sum_{i=1}^{n}\left(1-\phi_i\right)\left[1 + O_p\left(m^{*-1}\right)\right] Z_i.\end{aligned}$$

Using that $|f_i\left(\mathbf{x}, t\right) - m\left(\mathbf{x}, t\right)| < C_K$, for all i, $\mathbf{x}$ and t.

$$\begin{aligned}\hat{m}\left(\mathbf{x}_0, t_0\right) - E\left[\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right] &= \sum_{i=1}^{n}\left(1-\phi_i\right) r_i \frac{\sum_{h=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m\left(\mathbf{x}, t\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{h=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}} \\ &= \sum_{i=1}^{n}\left(1-\phi_i\right)\frac{1}{N_1^*}\left[1 + O_p\left(m_1^{*-1}\right)\right]\frac{\sum_{h=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m\left(\mathbf{x}, t\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{h=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}} \\ &= \frac{1}{N_1^*}\sum_{i=1}^{n}\left(1-\phi_i\right) W_i\left[1 + O_p\left(m_1^{*-1}\right)\right], \quad \text{(B.102)}\end{aligned}$$

with $W_i = \frac{\sum_{h=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m(\mathbf{x},t)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{h=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}}$. Then, under the condition that $|f_i\left(\mathbf{x}, t\right) - m\left(\mathbf{x}, t\right)|$ are uniformly bounded we have

$$\begin{aligned}|W_i| &= \left|\frac{\sum_{j=1}^{n_i}\sum_p \left(y_{p,t_{i_j}} - m\left(\mathbf{x}, t\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}}\right| \\ &= \left|\frac{\sum_{j=1}^{n_i}\sum_p \left(f_i\left(\mathbf{x}, t_{i_j}\right) - m\left(\mathbf{x}, t\right)\right)\omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}}\right| \\ &< \left|\frac{\sum_{j=1}^{n_i}\sum_p C_K \omega_{p,t_{i_j}}^{(jkl,t_0)}}{\sum_{j=1}^{n_i}\sum_p \omega_{p,t_{i_j}}^{(jkl,t_0)}}\right| = C_K. \quad \text{(B.103)}\end{aligned}$$

For individuals from group T, $E\left[W_i\right] = 0$ (as $E\left[y_{p,t_{i_j}}\right] = E\left[f_i\left(\mathbf{x}, t_{i_j}\right)\right] = m\left(\mathbf{x}, t\right)$) and $\operatorname{Var}\left(W_i\right) = \operatorname{Var}\left(Z_i\right) = \operatorname{Cov}\left(e_i\left(\mathbf{x}_0, t_0\right), e_i\left(\mathbf{x}_0, t_0\right)\right) + \left(\sigma^2\left(\mathbf{x}_0, t_0\right) + \sigma_\varepsilon^2\left(\mathbf{x}_0, t_0\right)\right)\frac{V(k_t)}{n_i h_t f(t_0)}\left[1 + o_p\left(1\right)\right]$. Moreover, the W_i are independent (as individuals are independent) and then, by the Lyapunov's central limit theorem we have

$$N_1^{*1/2}\left(\hat{m}\left(\mathbf{x}_0, t_0\right) - E\left[\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right]\right) \sim AN\left(0, \operatorname{Cov}\left(e_i\left(\mathbf{x}_0, t_0\right), e_i\left(\mathbf{x}_0, t_0\right)\right)\right). \quad \text{(B.104)}$$

From (B.68) we can see that $E\left[\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right] = \hat{m}\left(\mathbf{x}_0, t_0\right) + O_p\left(h_x\right) + O_p\left(h_t^2\right)$ and then $N_1^{*1/2}\left(\hat{m}\left(\mathbf{x}_0, t_0\right) - E\left[\hat{m}\left(\mathbf{x}_0, t_0\right)|\mathbf{D}\right]\right) = O_p\left(N_1^{*1/2}\left(h_x + h_t^2\right)\right) = O_p\left(N_1^{*-2(\delta-3/2)/5}\right) = o_p\left(1\right)$, therefore obtaining

$$N_1^{*1/2}\left(\hat{m}\left(\mathbf{x}_0, t_0\right) - m\left(\mathbf{x}_0, t_0\right)\right) \sim AN\left(0, \operatorname{Cov}\left(e_i\left(\mathbf{x}_0, t_0\right), e_i\left(\mathbf{x}_0, t_0\right)\right)\right). \quad \text{(B.105)}$$

Bibliography

[1] H. Wu and J.-T. Zhang, *Nonparametric Regression Methods for Longitudinal Data Analysis.* New Jersey: John Wiley & Sons, Inc., 2006.

[2] ——, "Local polynomial mixed-effects models for longitudinal data," *Journal of the American Statistical Association*, vol. 97, no. 459, pp. 883–897, 2002.

Summary

External beam radiotherapy (EBRT) is one of the cornerstones of prostate cancer treatment. The objectives of radiotherapy are, firstly, to deliver a high dose of radiation to the tumor (prostate and seminal vesicles) in order to achieve a maximal local control and, secondly, to spare the neighboring organs (mainly the rectum and the bladder) to avoid normal tissue complications. Both objectives are, however, in conflict and a compromise is needed to achieve an effective treatment and maintain a good quality of life after prostate EBRT.

Normal tissue complication probability (NTCP) models are then needed to assess the feasibility of the treatment and inform the patient about the risk of side effects, to derive dose-volume constraints and to compare different treatments. Proposed in the 1970's, the classic NTCP models can be modified to be adapted to both biological paradigm and treatment technique changes. The NTCP models have evolved from empirical models with parameters estimated via consensus to machine-learning methods trained on real data.

In the context of EBRT, the objectives of this thesis were to find predictors of bladder and rectal complications following treatment; to develop new NTCP models that allow for the integration of both dosimetric and patient parameters; to compare the predictive capabilities of these new models to the classic NTCP models and to develop new methodologies to identify dose patterns correlated to normal complications following EBRT for prostate cancer treatment. A large cohort of patient treated by conformal EBRT for prostate cancer under several prospective French clinical trials was used for the study.

In a first step, a traditional statistical regression approach was used to find predictors of bladder and rectal complications following EBRT. Using Kaplan-Meier nonparametric estimation, the incidence of the main genitourinary and gastrointestinal symptoms have been described. With another classical approach, namely logistic regression, some predictors of genitourinary and gastrointestinal complications were identified. The logistic regression models were then graphically represented to obtain nomograms, a graphical tool that enables clinicians to rapidly assess the complication risks associated with a treatment and to inform patients. This information can be used by patients and clinicians to select a treatment among several options (e.g. EBRT or radical prostatectomy).

In a second step, the difficulty of including both dosimetric and patient parameters in classical NTCP models was identified. Although this can be done by stratifying a population and then fitting a different model at each stratum, this stratification may

lead to a loss of statistical power as not all strata met the required number of patients to fit the models. Another strategy is adding new parameters to existing models, but this implies making biological assumptions which are difficult to confirm with treatment parameters and patient outcome data only. We thus proposed the use of random forest, a machine-learning technique, to predict the risk of complications following EBRT for prostate cancer. Random forest show a similar accuracy as the most accurate state-of-the-art machine-learning methods but without overfitting. Our random forest NTCP (RF-NTCP) model, which includes both clinical and patient parameters, was compared to traditional Lyman-Kutcher-Burman (LKB) NTCP and logistic regression models. The superiority of the RF-NTCP, assessed by the area under the curve (AUC) of the receiving operative characteristic (ROC) curve, was established.

In a third step, the 3D dose distribution was studied. A 2D population value decomposition (PVD) technique was extended to a tensorial framework to be applied on 3D volume image analysis. Using this tensorial PVD, a population analysis was carried out to find a pattern of dose possibly correlated to a normal tissue complication following EBRT. Also in the context of 3D image population analysis, a spatio-temporal nonparametric mixed-effects model was developed. This model was applied to find an anatomical region where the dose could be correlated to a normal tissue complication following EBRT.

In conclusion, the main contribution of our work is the development of new predictive models of rectal and bladder toxicities following EBRT and the demonstration that these models are strong competitors of the classic NTCP models.

Résumé

La prostate est une glande de l'appareil reproducteur masculin. Elle se situe en-dessous de la vessie et face au rectum. Le cancer de la prostate est le type de cancer le plus fréquent chez l'homme dans les pays occidentaux. En France, en 2011, il y a eu 71,000 nouveaux cas et 8,700 décès. La radiothérapie externe (EBRT en anglais pour *External Beam Radiotherapy*) est l'un des traitements référence du cancer de prostate. Les objectifs de la radiothérapie sont, premièrement, de délivrer une haute dose de radiations dans la cible tumorale (prostate et vésicules séminales) afin d'assurer un contrôle local de la maladie et, deuxièmement, d'épargner les organes à risque voisins (principalement le rectum et la vessie) afin de limiter les effets secondaires. Le traitement par radiothérapie nécessite donc un compromis entre ces deux objectifs. Une fois l'anatomie du patient caractérisée au moyen des images scanner, le plan de délivrance de la dose est optimisé afin de respecter les contraintes de dose maximale sur les organes à risque et de dose prescrite sur la prostate. Le résultat est une matrice de dose planifiée, la source d'information dosimétrique considérée dans cette thèse.

Des modèles de probabilité de complication des tissus sains (NTCP pour *Normal Tissue Complication Probability*) sont nécessaires pour (i) estimer les risques de présenter des effets secondaires au traitement, ainsi que pour (ii) définir des contraintes de dose par unité de volume ou encore (iii) comparer différents traitements. Proposés dans les années 1970, les modèles NTCP classiques peuvent être améliorés et adaptés au paradigme biologique ainsi qu'aux évolutions des techniques thérapeutiques. Ces premiers modèles empiriques avec estimation de paramètres par consensus ont évolué vers des méthodes d'apprentissage de machines entrainées sur données réelles. Les premiers modèles considéraient qu'une fraction du volume de l'organe à risque étudié était irradiée uniformément par une même dose (exprimée en énergie par unité de masse). Il y avait donc deux variables liées à la toxicité : la fraction du volume de l'organe irradiée et la dose. Cependant, avec l'apparition de la radiothérapie conformationnelle, les organes à risque n'étaient plus irradiés de façon uniforme. Il a fallu alors considérer l'organe à risque comme un ensemble de plusieurs sous-unités fonctionnelles liées entre elles selon un type d'architecture (en parallèle ou en série). La probabilité de complications de l'organe à risque est devenue une fonction de la probabilité de complication des sous-unités composant l'organe. L'enjeu pour les modèles NTCP est d'établir des liens entre des modèles de dose et la probabilité de présenter des effets secondaires, tout en prenant compte les caractéristiques des patients.

Dans le contexte de la radiothérapie externe, les objectifs de cette thèse étaient (i) d'identifier des paramètres prédictifs de complications rectales et vésicales secondaires au traitement; de (ii) développer de nouveaux modèles NTCP permettant l'intégration de paramètres dosimétriques et de paramètres propres aux patients; (iii) de comparer les capacités prédictives de ces nouveaux modèles à celles des modèles classiques et (iv) de développer de nouvelles méthodologies d'identification des modèles de dose corrélés à l'apparition de complications. Une importante base de données de patients traités par radiothérapie conformationnelle, construite à partir de plusieurs études cliniques prospectives françaises, a été utilisée pour ces travaux.

L'approche utilisée pour étudier la toxicité survenue suite au traitement du cancer de la prostate par radiothérapie commence par l'analyse de données relativement simples (l'histogramme dose-volume, la technique d'irradiation et les caractéristiques des patients) et finit avec l'analyse de la matrice de dose. Nous avons choisi d'illustrer la faisabilité des méthodologies proposées sur la toxicité rectale principalement, la littérature disponible étant plus riche pour ce type de toxicité.

Dans un premier temps, une approche classique de régression statistique a été utilisée pour estimer des paramètres prédictifs de complications rectales et vésicales. La fréquence des symptômes gastro-intestinaux et génito-urinaires a été décrite par une estimation non-paramétrique de Kaplan-Meier. Des prédicteurs de complications gastro-intestinales et génito-urinaires ont été identifiés via une autre approche classique : la régression logistique. Ces types de modèle peuvent être représentés graphiquement afin d'aider les médecins en pratique clinique pour informer les patients sur les possibles effets secondaires de la radiothérapie. Les nomogrammes en sont un exemple. Dans cette thèse, des nomogrammes ont été proposés et validés à partir des résultats obtenus par régression logistique.

Inclure les paramètres dosimétriques et les paramètres des patients dans les modèles NTCP classiques apparaît complexe. Bien que cela puisse être réalisé par stratification de la population et adaptation de modèles différents à chaque strate, ceci peut engendrer une perte de puissance statistique, du fait qu'une ou plusieurs strates puissent être représentées par un nombre insuffisant de patients. Une autre stratégie possible est d'ajouter de nouveaux paramètres aux modèles existants. Ceci implique cependant de faire des hypothèses biologiques difficilement vérifiables uniquement à l'aide de données thérapeutiques et de résultats de traitement. Nous avons proposé l'utilisation de la méthode d'apprentissage de machine des forêts aléatoires (RF en anglais pour *Random Forests*) pour estimer le risque de complications. La méthode de forêts aléatoires consiste à créer des arbres de décisions sur des échantillons (avec remise) de la base de données originale. Les forêts aléatoires ont deux caractéristiques : chaque arbre est un bon prédicteur sur la base ayant été utilisée pour son entraînement (précision) et les arbres sont différents entre eux (diversité). Les forêts aléatoires montrent une performance similaire à celles des méthodes d'apprentissage les plus performantes de la littérature, sans toutefois présenter l'inconvénient du sur-apprentissage. Les performances de ce modèle incluant des paramètres cliniques et patients, mesurées par l'aire sous la courbe d'opération caractéristique, surpassent celles du modèle NTCP de Lyman-Kutcher-Burman (LKB) et celles de la régression logistique. Cette méthode a aussi permis d'établir une hiérarchie entre les différents paramètres retenus par la forêt aléatoire au regard de leur impact sur la prédiction des effets secondaires.

Enfin, la dose 3D a été étudiée. Une méthode de décomposition en valeurs populationnelles (PVD en anglais pour *Population Value Decomposition*) en 2D a été généralisée au cas tensoriel et appliquée à l'analyse d'images 3D. Une étape préalable à l'application de cette méthode est la mise en correspondance spatiale des matrices de dose de tous le patients. Cette mise en correspondance a été l'objet d'une autre thèse qui s'est déroulée simultanément. En utilisant des méthodes de recalage d'image une transformation est obtenue pour chaque patient pour propager son anatomie sur l'anatomie d'un patient référent. Ensuite, la même transformation est appliquée à la matrice de dose de chaque patient. Notre PVD se décompose elle même en deux étapes : une première décomposition (en valeurs singulières en haute dimension) est faite sur la matrice de dose de chaque patient. Le résultat de cet étape est un tenseur (une image 3D) et trois matrices facteur (une pour chaque dimension) pour chaque patient. Ensuite, pour chaque dimension, on applique une décomposition (avec la méthode d'analyse en composantes principales) sur l'ensemble de matrices de tous les patients. Des composantes populationnelles sont crées à partir des résultats précédents. Ils vont permettre de représenter la matrice de dose de chaque patient dans un espace transformé. Dans cet espace, les différences entre les matrices de dose des patients présentant ou non des effets secondaires sont étudiées. Enfin, on reconstruit une distribution de dose typique d'un patient présentant un effet secondaire et une distribution de dose typique d'un patient ne présentant pas cet effet secondaire. La comparaison de ces deux distributions de dose permet de soulever des hypothèses sur le modèle (pattern en anglais) de dose qui entraîne l'apparition d'un effet secondaire.

Pour étudier la dose 3D nous avons aussi développé un modèle non-paramétrique à effets mixtes. Ce modèle permet d'estimer, voxel par voxel, la différence moyenne de la dose reçue par les patients qui présentent un effet secondaire et ceux ne le présentant pas. Comme la PVD, cette méthode requière que toutes les images soient dans un espace commun. Une fois toutes les anatomies dans le même espace, on applique notre méthode non-paramétrique dans la région d'intérêt, permettant ainsi de trouver une région d'intérêt où la différence de dose entre les deux groupes est susceptible d'être liée à l'apparition d'un effet secondaire. La méthode développée prend en compte la variabilité intra-patient et inter-patient pour pouvoir dégager des estimations fiables sur les caractéristiques populationnelles. Bien que la méthode soit appliquée dans le cadre de cette thèse où l'on a une seule matrice de dose par patient, la méthode peut être aussi généralisée au cas où on a plusieurs images par patient et la méthode devient alors spatio-temporelle. Les propriétés des estimateurs (bias, variance et loi asymptotique) ont été aussi étudiées et sont présentées en annexes. L'application de cette méthode à une analyse de population a été menée afin d'extraire un motif de dose corrélée à l'apparition de complications après EBRT. Nous avons donc développé un modèle non-paramétrique à effets mixtes spatio-temporels pour l'analyse de population d'images tridimensionnelles afin d'identifier une région anatomique dans laquelle la dose pourrait être corrélée à l'apparition d'effets secondaires.

Nous avons enfin analysé les points forts et les points faibles des méthodes proposées. La méthode de forêts aléatoires, bien qu'elle permette d'intégrer des paramètres dosimétriques et des paramétres patient, elle partage une des limitations caractéristiques des modèles NTCP classiques : le manque de précision spatiale. Aussi, nous n'avons pas de mesure facilement interprétable pour quantifier l'augmentation du risque de présenter un effet secondaire en fonction des variables retenues par la forêt aléatoire.

D'autre part, notre PVD permet d'étudier la dose dans un espace transformé sans considérer les relations spatiales existantes. Cependant, la méthode doit encore être adaptée afin de prendre en compte des paramètres patient.

Finalement, notre modèle non-paramétrique permet de contrôler la variabilité intra- et inter-patients au niveau du voxel et ne requière pas de fortes hypothèses sur les relations entre les différentes variables étudiées. Néanmoins, cette méthode demande d'important ressources informatiques. En outre, elle peut être considérée comme appartenant à la famille de méthodes de régression par lissage et nécessite une étape additionnelle d'estimation des paramètres de lissage optimaux.

En conclusion, la principale contribution de notre travail est le développement de nouveaux modèles prédictifs de toxicité rectale et vésicale après EBRT, sérieux concurrents pour les modèles NTCP classiques.

Resumen

La radioterapia externa es una de las piedras angulares del tratamiento del cáncer de próstata. Los objetivos de la radioterapia son, primero entregar altas dosis de radiación al tumor (próstata y vesículas seminales) para maximizar el control local y, segundo, evitar irradiar los órganos vecinos (principalmente el recto y la vejiga) para minimizar los efectos secundarios. Sin embargo, ambos objetivos están en conflicto y se necesita un compromiso para lograr un tratamiento efectivo y una buena calidad de vida después de la radioterapia externa.

Se requieren entonces modelos para estimar la probabilidad de complicación de los tejidos sanos (NTCP por *normal tissue complication probabilities*) para evaluar la factibilidad de un tratamiento e informar al paciente de los riesgos de efectos secundarios, para generar restricciones sobre la dosis durante el tratamiento y para comparar diferentes tratamientos. Propuestos en la década de 1970, los modelos NTCP pueden ser modificados para adaptarlos a los cambios en los paradigmas biológicos y a las técnicas de tratamiento. Los modelos NTCP han evolucionado desde modelos empíricos con parámetros ajustados a través del consenso clínico hasta métodos de aprendizaje de máquina entrenados con datos reales.

En el contexto de la radioterapia externa, los objetivos de esta tesis fueron encontrar factores predictivos de los efectos secundarios que siguen al tratamiento en el recto y la vejiga; desarrollar nuevos modelos NTCP que permitan la integración simultánea de parámetros dosimétricos y específicos de los pacientes; comparar las capacidades predictivas de estos nuevos modelos con los modelos NTCP clásicos; y desarrollar nuevas metodologías que para identificar patrones de dosis correlacionados con los efectos secundarios de la radioterapia externa. Para ello se empleó una base de datos con un importante número de pacientres provenientes de varios estudios prospectivos llevados a cabo en Francia.

En un primer momento se utilizó una aproximación de regresión estadística clásica para encontrar factores predictivos de los efectos secundarios localizados en el recto y la vejia. Utilizando la estimación de Kaplan-Meier se caracterizó la incidencia de los principales síntomas asociados a los efectos secundarios del recto y la vejiga. Usando regresión logística, otra herramienta clásica, se identificaron algunos factores predictivos de los efectos secundarios en el recto y la vejiga. Los modelos de regresión logística se representaron de manera gráfica a través de nomogramas, que son herramientas visuales para ayudar a los médicos a evaluar rápidamente los riesgos de efectos secundarios asociados a un tratamiento y para informar a los pacientes. Esta información puede ser

utilizada por médicos y pacientes para seleccionar un tratamiento entre varias opciones (por ejemplo entre radioterapia externa y prostateoctomía).

En un segundo momento se identificó la dificultad para incluir al mismo tiempo dentro de los modelos NTCP clásicos los parámetros específicos del paciente y los del tratamiento. Aunque esto se puede resolver estratificando una población de pacientes y estimando un modelo en cada estrato, dicha estratificación puede resultar en una pérdida de potencia estadística, ya que algunos estratos pueden no tiener el número de pacientes necesarios para estimar el modelo. Otra estrategia consiste en agregar parámetros a los módelos existentes, pero esto implica hacer hipótesis que en la práctica resultan difíciles de verificar solamente con información sobre parámetros del tratatamiento y el resultado observado en el paciente. Frente a esto, propusimos el uso de bosques aleatorios para predecir los efectos secundarios derivados de la radioterapia externa. Los bosques aleatorios son una técnica de aprendizaje de máquina cuyo desempeño compite con el de las técnicas de referencia en el estado del arte, y adicionalmente presentan la ventaja de no sobreajustar los datos. Nuestro modelo NTCP basado en bosques aleatorios (que denominamos RF-NTCP) incluye parámetros específicos del paciente y parámetros del tratamiento. El desempeño del RF-NTCP ha sido comparado con el modelo clásico de Lyman-Kutcher-Burman (LKB) y con la regresión logística. Utilizando el área bajo la curva (AUC por *area under the curve*) de la curva de operación característica (ROC por *receiving operating characteristic*) se pudo establecer la superioridad del RF-NTCP.

En un tercer momento se analizó la distribución 3D de la dosis. Se extendió al caso tensorial una técnica de descomposición en valores poblacionales (PVD por *population value decomposition*) propuesta para el análisis de imágenes 2D. Usando esta PVD se llevo a cabo un análisis poblacional que permitió identificar un patrón de dosis correlacionado con un efecto secundario derivado de la radioterapia externa. También, en el contexto del análisis de poblaciones de imágenes 3D, se desarrolló un modelo no paramétrico de efectos mixtos en variables espacial y temporal. Este modelo fue aplicado para encontrar una región anatómica relacionada con un efecto secundario de la radioterapia externa.

En conclusión, el aporte principal de este trabajo es el desarrollo de nuevos modelos predictivos para los efectos secundarios en el recto y la vejiga después del tratamiento del cáncer de próstata con radioterapia y la demostración de que estos modelos son fuertes competidores de los modelos NTCP clásicos.