Università degli Studi di Pavia
Dipartimento di Fisica

DOTTORATO DI RICERCA IN FISICA – XXVII CICLO

Analysis of Radiation Effects in Cancer Patients: from the Macro-scale of Organ Symptoms to the Micro-scale of Cell Death

Alessandro Cicchetti

Submitted to the Graduate School of Physics in partial fulfillment of the requirements for the degree of

DOTTORE DI RICERCA IN FISICA

DOCTOR OF PHILOSOPHY IN PHYSICS

at the

University of Pavia

Supervisors: Prof. Francesca Ballarini
Dr. Tiziana Rancati

Cover: Organ at risks affected by radiotherapy for prostate cancer

Analysis of Radiation Effects in Cancer Patients: from the Macro-scale of Organ Symptoms to the Micro-scale of Cell Death

Alessandro Cicchetti
PhD thesis - University of Pavia
Pavia, Italy, September 2019

Contents

Introduction

The origin of radiation biology should be attached to the discovery of X-rays made by Röntgen in December 1895 [1]. Indeed, the scientific community did not need a large time to understand that the "invisible rays" could lead to harmful effects when they were applied in the field of medicine. Their power was so evident that less than one year after the discovery several university hospitals were already equipped with their own Crooke's tube for "skiography" [1]. At that time patients were exposed to X-rays mainly for the treatment of diseases such as lupus and hypertrichosis, but also for the detection of foreign bodies embedded in the tissues, of fractures, of hepatic, renal and vescical calculi and even for the diagnosis and treatment of soft cancer [2] (V. Despeignes performed in 1896 the first documented anticancer radiation treatment of a gastric carcinoma [3]) . Surgeons were very excited for this new tool; however, like a large portion of scientific discoveries which benefit the human race on the whole, also X-rays bring in their train a certain amount of evil. Nausea, bilious attack, redness, vesicles, irritative eczema, exfoliation of the epidermis and purulent discharge were the main symptoms reported in several papers during the follow-up of the patients exposed to X-rays. It is not a coincidence that in April 1898 [4], the Roengten Society appointed a "Commitee on X-ray Injuries". A brilliant article on this topic was published in 1911 by the radiologist Leon Bouchacourt, entitled "About the sensitivity of different individuals and, for a given individual, of the different part of the body: what should one think about the treatment of hypertrichosis by radiotherapy?" [5]. He pointed out the individual differences in the occurrence of early skin reactions, noting also some "bystander effects" in the nonirradiated neighboring areas. The "idiosyncrasy", then replaced by the term radiosensitivity, was already known in the French academia. Infact, in a congress of electrophysiology organized in Lyon in 1906, Dr. Bergonie summarized the ongoing debates by saying that there were "two error types" that may affect the medical application of X-rays:
a) "the uncertaintes in the assessment of radiation dose"
b) "the differences in the sensitivity of the patients".

[1] Picture of a substance invisible to the eye on account of its covering, transferred to a negative by means of X-rays

Bergonie strongly believed that the last notion was due to "hereditary or acquired predisposition" [6]. In the same year he also defined the law of Bergonie and Tribondeau, which asserts that "radiosensitivity of cells is linked to their proliferation rate" [7]. The conclusion was that no "pseudo-harmlessness dose" can be defined, since every organ has its radio-resistance and every person has its own sensitivity to X-rays. Under totally different circumstances, the Radium Girls' saga [8] held an important place in the history of the field of health physics, convincing lawyers and the American government of the dangerousness of radiation exposure. Briefly, the Radium Girls were female factory workers who contracted radiation poisoning from painting watch dials with self-luminous paint. The women in each facility had been told the paint was harmless, and subsequently ingested deadly amounts of radium after being instructed to "point" their brushes on their lips in order to give them a thin point. Many of the women later began to suffer from anemia, bone fractures and necrosis of the jaw, a condition now known as radium jaw (see Figure 1). Similar side effects were found in people who did a massive use of Radithor, a radium water guaranteed by the manufacturer to contain $2\mu Ci$ of radium. The most famous example was Eben Byers, a well-known industrialist, who died of radiation poisoning from the magic water [9]. Other decades were needed

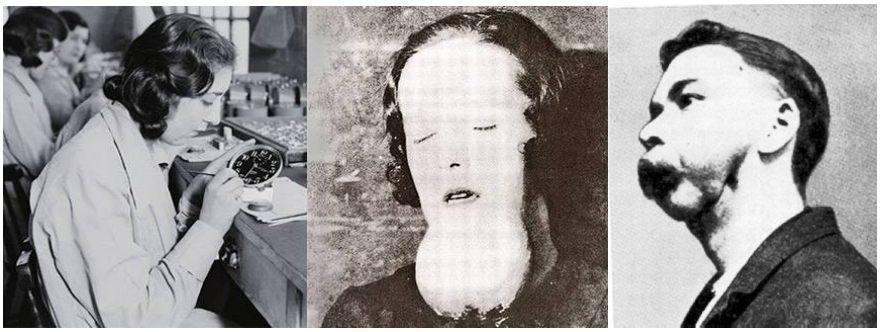


Figure 1: *From left to right: a radium girl painting clocks; example of radiation toxicity on the jaw; Eben Byers after prolonged use of Radithor. Images from references [8, 9]*

to the biologists to discover the constituents of the cell and, particularly, to determine into the cell nucleus the presence of the double-helix structue of the Deoxyribonucleic acid (DNA)[2] [10]. A molecule composed of two chains that coil around each other to form a double helix carrying genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. Few years later, in 1956 [11], the clonogenic assay was

[2]DNA was isolatd for the first time at the University of Tübingen by Miescher, who called it nuclein (now nucleic acid). It was the 1869. To understand the DNA strucuture a longer period of time was needed. Specifically, it was the 1952 when the diffraction photograph of the B form of DNA was taken by Rosalind Franklin. Data derived from this photograph were instrumental in allowing James Watson and Francis Crick to construct their Nobel Prize-winning model for DNA.

defined to assess the radiosensitivity of mammalian and animal cell lines to radiation. Further details on this technique and cell survival curves are presented in the thesis, but it is worth noting that an increasing interest in radiobiology was given to the understanding of the processes of cell death. From this point of view, the first question to answer was to define which structure of the cell is most responsible for cell death. A couple of studies by Warters [12, 13] were fundamental to solve this open question. This team of research was able to irradiate the cell locally. Basically, they irradiated three structures of the cell: the membrane, the cytoplasm and the nucleus. High doses could be given to the first two units without causing cell death. On the other hand, the nucleus resulted to play a crucial role for cell killing (see Table 1). Particularly, the

Table 1: *The toxicity of radioisotopes depends upon their sub-cellular distribution. Particularly, for each of these three treatments a dose has been chosen that gives 50% cell killing in Chinese hamster ovary (CHO) cells. The absorbed radiation doses to the nucleus, cytoplasm or membranes have then been calculated. 3H-Thymidine was bound to DNA and 125I-concanavalin to cell membranes. It is the nuclear dose that is constant and thus correlates with cell killing, not the cytoplasmic or membrane doses. Table modified from [12].*

	Radiation dose to part of the Cell (Gy)		
Radiation source/type	Nucleus	Cytoplasm	Membrane
X-Rays	3.3	3.3	3.3
^{3}H-thymidine	3.8	0.27	0.01
^{1}25I-Concanavalin	4.1	24.7	516.7

researchers were able to proof that when one or two alpha particles hit the nucleus (using small polonium needles), cell death resulted. The reason for that was explained by the lethal damage inflicted to the nucleus DNA[3]. Because of its role, cells have developed a series of processes for preserving DNA structure from the continuous internal and external attack. Therefore, a system of DNA damage response has been developed for detecting and repairing bases, single-strand breaks, double-strand breaks and other damages. Every mechanism is a coordinated system within which a group of highly interrelated pathways act. The system can be divided in two parts: the DNA damage *sensors* and the *effectors* of damage response. The former are proteins that constantly survey the genome for the presence of damage. The latter, still mediated by proteins, can induce cell death, checkpoint activation or DNA repair. Physics and biology have walked different paths to increase the efficacy of treatment. The first one develops new technologies for diagnosing the active volume of the tumour and increasing the conformity of the dose to the target volume and, thus, sparing as much as possible the healthy tissues. Biologists build up new assays for improving the knowledge on the radiobiological pathways and the

[3]The number of DNA lesions per cell detected immediately after such a dose is approximately: base damage (>1,000), single-strand breaks ($\sim$ 1000), double-strand breaks ($\sim$ 40).

sensitivity of the different tumour cells, but they also spend many efforts in targeted cancer therapies.

In this framework, this thesis work was devoted to the analysis of radiation damage at different levels, from organs and tissues down to single cells. In the first chapter the concept of therapeutic ratio is presented. Its formulation established the importance of understanding the radiation dose tolerance of the healthy tissues surrounding the tumor region, which is the main limitation of the radiotherapy treatment. The 4 Rs of radiotherapy are also discussed in this part of the thesis.
The second chapter describes the evolution of the methods used to investigate the radiation-damage to the normal tissues, from the development of human sequelae to the analysis of cell death and also to gene sequencing, going through the use of imaging modalities and tissues-on-a-chip.
The following three Chapters (3,4,5) are focused on the analysis performed during the three years of PhD project: Chapter 3 describes the development and validation of NTCP models based on patient-reported questionnaire in prostate cancer patients treated with radiotherapy; Chapter 4 contains results on tissue investigations including also a preliminary test of a tissue simulation; Chapter 5 is based on the analysis of the direct effect of radiation on muscle cell lines and their DNA, including also Montecarlo simulation of cell death.

Chapter 1

The Concept of Therapeutic Ratio

The success of radiotherapy in eradicating tumours depends chiefly on the total radiation dose. It is a fundamental principle that clonogenic tumour cell kill increases with increasing radiation dose, in tumour cell lines in the laboratory, in experimental animal tumours, and in clinical diseases in patients. However, as turnout dose is increased, so the incidence and severity of normal tissue complications also rises. The goal of radiotherapy is to deliver high doses of ionizing radiation to eradicate tumour cells, while at the same time minimizing the risks of damaging the surrounding normal tissue [14].
Radiotherapy outcomes are usually characterized by two indices: *Tumour Control Probability* (TCP), which is the probability of the extinction of clonogenic tumour cells after radiotherapy, and *Normal Tissue Complication Probability* (NTCP), which is the probability of healthy normal tissue injury. What limits clinical radiotherapy doses is the tolerance of normal tissues surrounding the tumour that fall within the treatment volume, particularly those tissues that are late-reacting. One of the key components of modern radiation oncology research is to predict treatment outcomes during treatment planning or during a fractionated course of therapy to personalize prescription and optimize response. Outcome models can also inform clinicians when weighting different treatment options with their patients or adapting radiotherapy fractionation subject to patient specific variables.

The first dose response data in patients were reported for skin cancer by Miescher in 1934. Using this data and data on the formation of teleangiectasia, Holthusen [15] constructed the first radiation dose response curves (Figure 1.1) for both TCP and NTCP, respectively. In that trial a sigmoid (S-shape) relationship was observed for both observations: the incidence is to zero when no dose is given to the tissues; as radiation dose is increased, there will be a tendency for tumor response to increase, and the same is also true for damage to normal tissues; finally, at very large doses the incidence tends to 100%. The central objective of optimizing radiotherapy (RT) is therefore to obtain a dif-

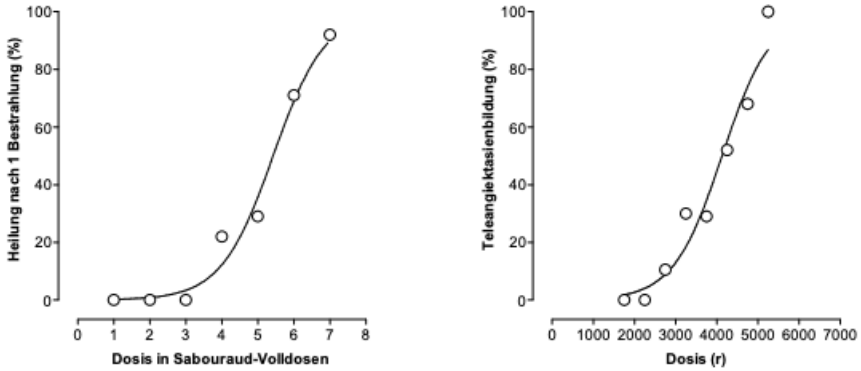


Figure 1.1: *Tumor control rate after single dose exposure (left). Formation of teleangiectasia (right) after fractionated radiotherapy with 280 to 300 r per fraction. Redrawn from [15].*

ferential effect of radiation on tumor and normal tissue. Any discussion of the possible benefit of a change in treatment strategy must always consider simultaneously the effects on tumour response and on normal-tissue damage. This ratio is known in the literature as *therapeutic index*[1] and should be optimized to give the best clinical outcome.

The uncomplicated tumor control
The maximum dose to irradiate a tumor is driven by the dose to warranty a percentage below a threshold value for the damage to the organs at risk. All the concepts were already clear to Holthusen who proposed also a formula to optimize the tumor dose. The probability to achieve tumor control without complications can be calculated as follows:

$$UCP = TCP * (1 - NTCP) \tag{1.1}$$

The bell-shaped curve obtained was called the *Uncomplicated Tumor Control* (UCP). Once the optimum dose is established, further improvements in UCP can only be achieved by either moving the TCP curve to lower doses, or the NTCP curve to higher doses.

Radiosensitizers and radioprotectors: how to shift the NTCP and TCP Curve
Indeed, there are factors that can impact and shift the sigmoid curves. Radiosensitizers, which act on tumor, can shift the TCP curve to the left, while selective radioprotectors, agents that prevent the damage caused by radiation, are able to move the NTCP curve increasing the distance with the TCP model, as depicted in Figure 1.2. Many efforts have been done in the past to under-

[1] Or, equivalently, Therapeutix Window.

stand which are these factors and the mechanisms of their action. Further discussions will follow in the next sections.

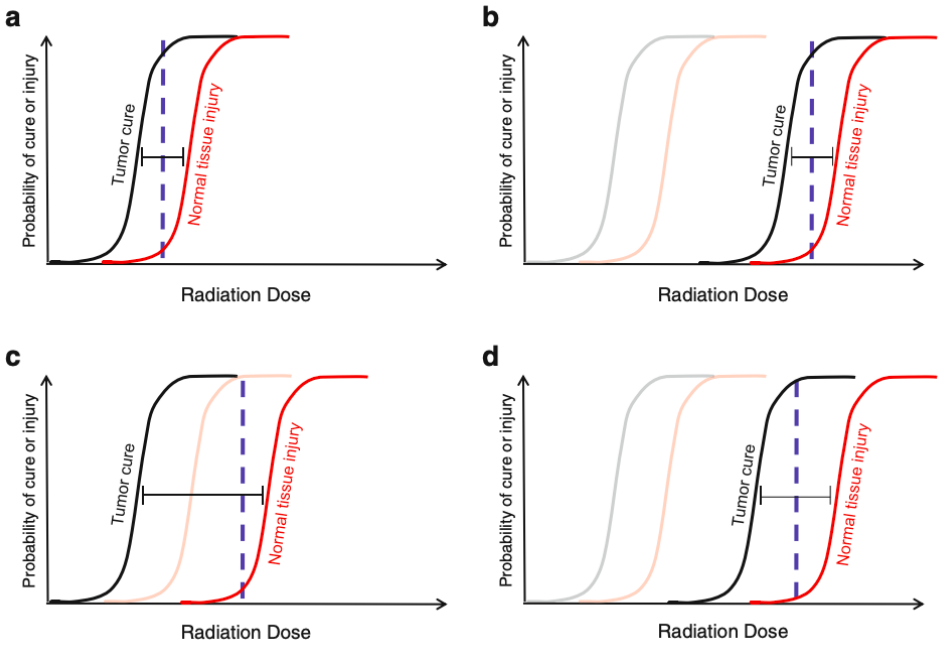


Figure 1.2: *Effects of radioprotectors on the therapeutic window . (a) The chosen treatment dose (blue vertical line) delivers a high chance of tumor cure (black) with a small chance of normal tissue injury (red). (b) A nonselective radioprotector indiscriminately protects tumor and normal tissue shifting both the tumor cure and normal tissue injury curve to the right. Consequently, the therapeutic window remains unchanged. Shifted curves are shown in red and black . The original curves are shown in gray and pale red . (c) A true selective radioprotector exclusively protects normal tissue and, thus, shifts only the normal tissue injury curve to the right. This affords a larger therapeutic window such that a higher dose can be given to achieve increased tumor cure with equal or less injury. A lower dose producing the same tumor control probability can also be given with less tissue injury. (d) Some nonselective radioprotectors can protect the normal tissue to a greater extent shifting the normal tissue curve further to the right than the tumor cure curve and resulting in an increased therapeutic window. Image from [16]*

Radiosensitive and radioresistant patients: how to change the slope of NTCP and TCP

Holthusen also interpreted the shape of the curves as expression of an underlying distribution of radiosensitivities (different slope for each patient) and concluded that, for statistical reasons, only observations of large samples of patients, but not case-reports, will contribute to radiotherapeutic progress. The concept is reasonable in terms of statistical approach, but it also highlighted the theoretical possibilities to understand which are the inner classes

of "radiosensitive" (sigmoid curve with a slope steeper than the average) and "radioresistant" (sigmoid curve with a slope less steep than the average) in between the treated patients, so who is more or less prone to reach the tumour control or develop side effects after radiotherapy[17]. Moreover, according to Holthusen the scale of cell sensitivity should correlate with the one at the patient level. To investigate this patient sensitivity to radiation, toxicity has been leveled to other complex diseases and analyses of genes are becoming more frequent in clinical trials, due to the viable technology of the *Next Generation Sequencing* (NGS). An impressive effort in the association of genomic features patterns, particularly *Single Nucleotide Polimorfisms* (SNPs), is the current aim of the *Genome-Wide Association Study* (GWAS). Indeed, the basic idea is to recognize a specific SNP fingerprint for each type of late toxicity after radiotherapy.

Real clinical practice

Cutoff values on UCP should be influenced by subjective features such as age and radiosensitivity, but also by objective characteristics such as the impact of that symptom on the *Quality Of Life* (QoL) of the patient or the chronicity of the dysfunction and also the recovery capability. TCP and NTCP used in the clinical practice describe the average trend within a population. This is mainly due to our lack of knowledge. Potentially, each patient has its own TCP and NTCP curves: the two curves would be well separated for radioresistant patients harbouring a radiosensitive tumour, conversely, they could be almost overlapping for a radiosensitive patient with a radioresistant tumour. The Uncomplicated Tumor Control dose should thus be dependent on the single patient/single tumour characteristics: a radioresistant patient could tolerate a higher amount of dose before reaching a threshold value of toxicity rate which is clinically unacceptable. In the current clinical practice, it is not possible to define such curves at the patient level, and Uncomplicated Tumor Control doses are defined at a population level, setting population-based acceptable toxicity rates. Thus, normal tissue tolerance is more a practical concept: the maximum dose that gives an acceptable incidence or probability of complications in a group of patients. What is an acceptable incidence depends, as we said, on a number of factors but what is usually done in clinical practice is to consider a threshold of 5-10% for moderate-severe toxicity and 15-20% for mild toxicity and, consequently, to fulfill one dose-volume constraint or a set of constraints for a specific organ.

From this point of view, the gold standard is the *Quantitative Analysis of Normal Tissue Effects in the Clinic* (QUANTEC), a special supplement of the International Journal of Radiation Oncology published in 2010 [18]. The manuscripts in the volume provided summaries of the dose/volume/outcome information for more than 15 organs. The organs discussed were selected because the authors believed that there were meaningful data to review, and a clinical need to better summarize the available dose/volume data for these

organs. Most of the available studies were related to conventionally fractionated conformal irradiation. Thus, hypofractionated or intensity-modulated approaches and imaging, biological and genetic factors were not taken into account in the review. Updates in all of these terms are needed with the aim of achieving a personalized treatment, where patient characteristics (comorbidity, radiosensitivity, tumour features) are considered for defining the prescription dose to the tumor.

1.1 Radiation Effects on the Tumour

Radiation effects on tumours under clinical as well as experimental conditions can be measured by different endpoints, including local tumour control, tumor regrowth delay and tumour regression. The aim of curative radiotherapy is, obviously, the local tumour control and it is also the preferable endpoint also for experimental investigations. A tumour is locally controlled when all its clonogenic cells (i.e. cells with the capacity to proliferate and to cause recurrence after radiotherapy) have been inactivated. The probability function for local control is radiation dose dependent and is related to the number of surviving clonogenic tumour cells. In these terms, radiotherapy is highly effective in killing clonogenic tumour cells. Moreover, the quantitative relationship between prescription dose, inactivation of clonogenic cells and local tumour control is well estabilished in the literature [19, 20, 21]. It has been demonstrated that the logarithm of the surviving clonogenic tumour cells decreases linearly with total radiation dose. If the radiation dose is high enough to sterilize all the active cells, then local control is achieved.
Munro and Gilbert also published a landmark paper in which they formulated the target-cell hypothesis of tumour control: "The objective of treating a tumour by radiotherapy is to damage every single potentially malignant cell to such an extent that it cannot continue to proliferate". From this assumption and the random nature of cell killing by radiation they derived a mathematical formula for the probability of tumour cure after irradiation of a "number of tumours each composed of N identical cells" [19]. More precisely, they showed that this probability depends only on the average number of surviving clonogens per tumour. Moreover, the relative frequencies of surviving clonogens follow a Poisson distribution (as many processes involving the counting of random events). Thus, when describing TCP, it is the probability of surviving clonogens in a tumour that is of interest. This is the zero-order term of the Poisson distribution and if lambda denotes the average number of clonogens per tumour after irradiation it follows that $TCP = e^{-\lambda}$.
Munro and Gilbert assumed also that the number of surviving cells per tumour was a negative exponential function of dose. On the basis of this assumption, we obtain for the tumour the characteristic sigmoid dose-response curve as in Figure 1.2. Later, the simple exponential dose-survival curve was replaced by the linear quadratic (LQ) model and thus we arrive at what could be called

the standard model of local tumour control:

$$TCP = e^{-N_0 e^{(-\alpha D - \beta d D)}}$$

In the equation, N_0 is the number of clonogens per tumour before irradiation and the second exponential is the survival fraction after a total dose D given with a dose per fraction d according to the LQ model. Thus when we multiply these two quantities we obtain the (average) number of surviving clonogens per tumour and this is inserted into the previous Poisson expression. Finally, N_0 can be approximated as a function of tumour volume and the clonogenic cell density (i.e. $clonogens/cm^3$ of tumour tissue).

A number of factors can contribute to the probability of local tumour control after fractionated radiotherapy. These factors have been summarized by Withers [22] as "the four Rs of Radiotherapy": recovery from sublethal damage, cell-cycle redistribution, cellular repopulation and tumour reoxygenation. Steel has suggested intrinsic cellular radiosensitivity as a fifth "R" to account for the different tolerance of tissues to fractionated irradiation [23].
It is worth pointing out that these effects were studied on tumour cell line but they can be extended (with some differences) to healthy tissue cells, for this reason some of the pictures presented in the next pages are the expression of normal cell lines behaviour.

Recovery from sub-lethal damage
As discussed in the introduction, the main target of ionizing radiation is the DNA chain. Evidence from a wide variety of recovery experiments, both on in-vitro cell lines and in-vivo normal and tumour tissues, showed that the majority of strand breaks in DNA are satisfactorily repaired. The repair of cellular damage between radiation doses is the major mechanism underlying the clinical observation that a larger total dose can be tolerated when the radiation dose is fractionated. The shoulder of the survival curve reflects accumulation of sublethal damage that can be repaired as shown in Figure 1.3 (a) [24]. The repair capacity of the cells of many tissues in-vivo has been demonstrated using cell survival and functional assays in-vivo [16]. An increase in total dose is required to give the same level of biological damage when a single dose (D_1) is split into 2 doses (total dose D_2) with a time interval between them. The capacity of different cell populations to undergo a successful repair process is reflected by the width of the shoulder on their survival curve. The difference in dose ($D_2 - D_1$) is a measure of the repair by the cells in the tissue. However, we need to take into account that the repair process restores the functionality of the macromolecules. The rejoining of the DNA fragments does not necessarily guarantee the gene function. In that case we have a so called genetic mutation.

Redistribution of Cells through the Division Cycle
Studies on a large variety of cell lines highlighted that radiosensitivity varies during the cell cycle. Indeed, using the cell synchronization technique has

made it possible to investigate the survival curve of the cell population in every phase. In the study about the irradiation of HeLa and chinese hamster cell lines published by Sinclair in 1965 [25], this was summarized as follows (see also Figure 1.3 (c)):

- cells in mitotic phase are generally the most radiosensitive ones,
- if G1 has an appreciable time length, there is usually a resistant period which declines toward the S phase,
- in most cell lines, resistance increases during S-phase with a maximum increase in the last part of the phase,
- in most cell lines, the G2 phase is almost as sensitive as mitosis.

It is important to highlight that the first three of these conclusions have been corroborated in many studies, but the last one was not supported by later experimental findings. Radioresistance in S phase could be explained by the large amount of undamaged sister templates which can be used during the homologous recombination. Sensitivity in G2 and mitosis is possibly due to the checkpoints in G2 phase. Indeed, there are two checkpoints in irradiated G2 cells. The *G2/M checkpoint* occurs early after radiation. It is transient, ATM (Ataxia telangiectasia mutated)-dependent and dose-independent (between 1 and 10 Gy). This checkpoint controls the entry into mitosis of cells that were in G2-phase at the time of irradiation. The *G2 accumulation* checkpoint is independent of ATM, but dependent on dose and Ataxia telangiectasia Rad3 related. It ensures that cells that pass through earlier cell-cycle phases with DNA damage do not enter mitosis [26]. The G2 accumulation following exposure to ionizing radiation probably allows damaged DNA to be repaired prior to mitosis, since DNA repair activity has been detected during the radiation-induced G2 delay and has been related to cellular radiosensitivity [27] [28]. On the other hand, cells that are in G2 phase during irradiation did not have sufficient time to repair the DNA damage and are stopped in the G2/M checkpoint.

Reoxygenation
The response of cells to ionizing radiation is strongly influenced by oxygen [29]. It has an impact on the survival curves for cultured mammalian cells, as shown in Figure 1.4 (a). The reason of this dependency is explained by the oxygen fixation hypothesis. When a tissue is irradiated, there is a direct biological action on the DNA molecules and indirect damage mediated by other cellular molecules, mainly by free radicals produced in water. These radicals are reactive molecules, they produce chemical changes and can lead to biological damages. Free radicals are also unstable and will react rapidly with oxygen, if present, producing new molecules that permanently fix the DNA damage produced by free radicals, leading to irreparable damage. Frequently, the abnormal neo-vasculature within the tumour is not able to meet the increasing

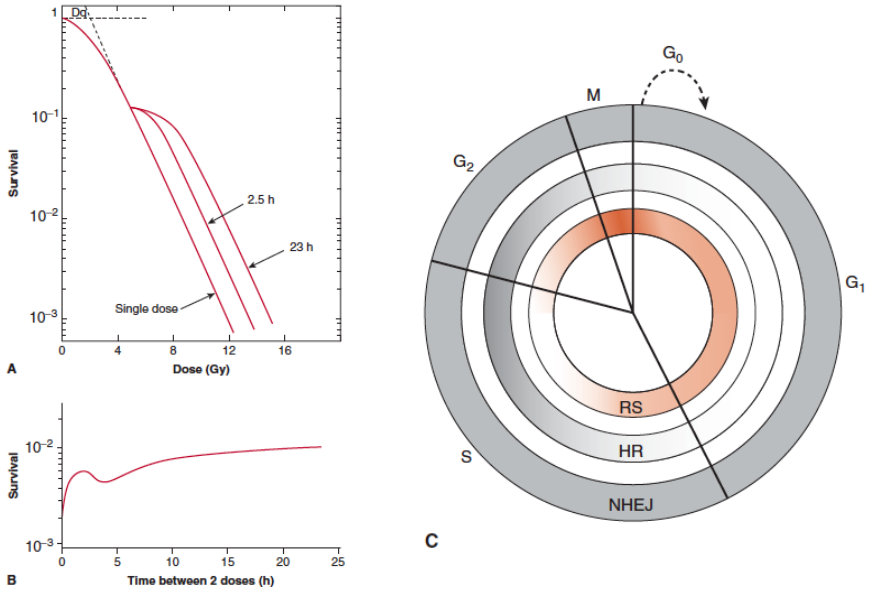


Figure 1.3: *Illustration of the repair of sublethal damage that occurs between 2 radiation treatments. A) Survival curves for a single-dose treatment or for treatments involving a fixed first dose followed, after 2.5 or 23 hours of incubation (at 37° C), by a range of second doses. B) Pattern of survival observed when 2 fixed doses of irradiation are given with a varying time interval of incubation (at 37° C) between them. C) Diagram indicating the active repair mechanisms during the various cell cycle phases and relative radiosensitivity. HR, Homologous recombination which occurs during the S and G2 phases of the cell cycle; NHEJ, nonhomologous endjoining recombination that occurs in all phases of the cell cycle; RS, relative radiosensitivity. Dark shading indicates activity of the particular repair pathway. Dark red shading indicates most radiosensitive portions of the cell cycle. Images from reference [16]*

nutrient demands of the expanding tumour mass. As a consequence, there are some microareas that are nutrient deprived, acidic and oxygen deficient, but still viable for a limited time. Thus, after a single dose of radiation, most of the aerobic cancer cells will be killed. However, there are different mechanisms with different time-scales which act in the changes of the fraction of hypoxic cells in the tumour, reducing the general radioresistance.
Indeed, cells at a certain distance from blood vessels ($100 - 150\mu m$) are beyond the maximal diffusion distance for oxygen and, thus, oxygen fixation can not contribute properly to the efficacy of the radiation damage. The so-called *Oxygen Enhancement Ratio* (OER) indicates that normoxic cells are 2-3 times more radiosensitive than the hypoxic cells.

Repopulation
Each fraction during radiotherapy reduces the total population of the clonogenic tumour cells. Surviving cells can repopulate the tumour by proliferation. As a consequence, any prolongation of the overall treatment time results in a higher number of clonogenic cells that needs to be inactivated and thereby requires a higher radiation dose to achieve local tumour control (see Figure 1.4 (b) for a graphical example). Direct consequence of this mechanism is the need to include into the modelling of TCP the so-called time factor of fractionated radiotherapy [30].This approach was also exported in NTCP studies [31, 32].

The 5th "R": Individual Radiosensitivity
The term radiosensitivity is a compound word consisting of two stems: "radiation" and "sensitivity", as it refers to a "measure" of individual variability which comes from the effect of a certain radiation on a specific subject (the target) that is made by its own genetic characteristics [33]. The first "measure" or dimostration of an "intrinsic" radiosensitivity was proposed by Fertil and Malaise [34] by considering the clonogenic survival fraction at 2 Gy as a parameter of radiosensitivity. Steel and colleagues, some years later, published a paper [23] where they defined this concept as essential for understanding the tratment efficacy. For this reason, they proposed *Radiosensitivity* as the fifth "R" of radiobiology. Now, radiosensitivity has a different connotation, one of the most spread is associated to the variation of the steepness of the acute radiation survival curve among different tumours. Moving from clinical to cellular radiosensitivity it had been possible to identify the syndromes and genetic mutations which strongly impact the individual radiosensitivity (ATM mutation, Fanconi Anemia, Nijmegen breakage syndrome and so on). In the past, patients with these conditions succumbed from whole body irradiation suggesting the necessity of predicting individual radiosensitivity. It is essential for the class of high-risk patients, as the ones with syndromes and mutations, but it could be important even in the case of intermediate class risk patients who are more common in the clinical practice. From this point of view, there are companies all over the world that are performing tools that may help with

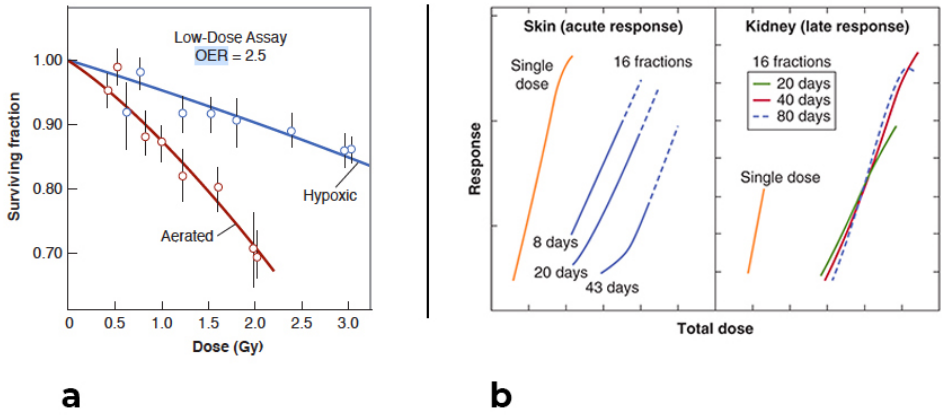


Figure 1.4: *a) The ratio of doses under hypoxic to aerated conditions necessary to produce the same level of cell killing is called the oxygen enhancement ratio (OER). It has a value of about 2.5 at X-ray doses less than about 2 to 3 Gy; b) Illustration of the effect of repopulation during fractionated treatment of skin or kidney. Treatment was a single dose or 16 equal fractions given in different overall times as indicated. Acute skin response was assessed using a numerical scoring technique and kidney response was determined by reduction in ethylenediaminetetraacetic acid (EDTA) clearance. For both tissues the fractionated treatment results in the curves moving to the right (higher doses) due to repair. For the acute skin reactions, extending the time over which a course of 16 fractions is given results in a further increase in the total dose required for a given level of response (isoeffective dose). In contrast, for late response of kidney there is no change in the isoeffective dose for 16 fractions regardless of whether the treatment is given over 20 or 80 days. Images from reference [16]*

the radiosensitivity classification of the patients.

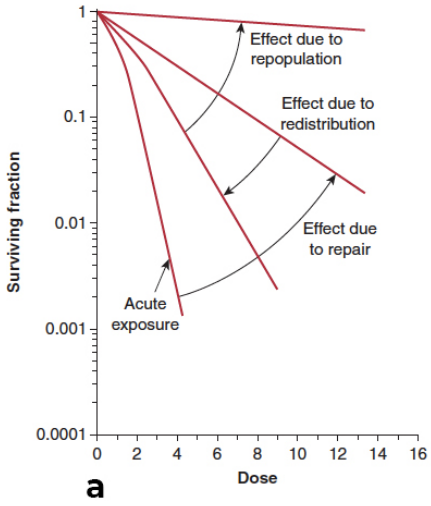


Figure 1.5: *Schematic to illustrate the influence on the survival curve following continuous low-dose rate irradiation, of the processes of cellular repair, redistribution, and repopulation. Image from reference [16].*

Tumour Volume
It is noteworthy to discuss also the impact of the tumour volume. Indeed, large tumours are more difficult to cure than small tumours. This observation is related to several reasons. First, the geometrical dimension is directly proportional to the number of clonogenic cells. Second, the greater the volume the greater the hypoxic region into the tumour. Finally, the dose distribution of large tumours does not reach the curative value due to the dose constraints of the adjacent healthy tissues. As a consequence, both experimental data and clinical studies support the tumour volume as an important factor which influences the tumour control probability after radiotherapy.

1.2 Radiation Effects on Healty Tissues

The fraction of surviving cell determines the success or failure of the treatment regimen as far as the tumour is concerned because a single surviving cell may be the focus for the regrowth of the tumor. For normal tissue this is not the whole story. In fact, the tolerance of healthy tissue to radiation depends on the ability of clonogenic cells to maintain the sufficient number of mature cells which guarantees the organ function. Despite optimum conformation of the treatment fields to the tumour and precise treatment planning and application, the target volume in curative radiotherapy necessarily includes a substantial amount of normal tissues, due to four main reasons. First, the tumour infiltrates microscopically into the normal organs, which necessarily must be included into the high-dose region. Second, some normal tissues such as

soft tissue and blood vessels are naturally irradiated to the prescription dose. Third, there is a limit due to the physical interaction of ionizing radiation with matter, which implies that structures in the entrance and exit channels of the radiation beam may be exposed to clinical relevant doses. Finally, the concept of *Planning Target Volume* (PTV) enlarges the geometrical definition of the *Clinical Target Volume* (CTV) in order to take into account the patients movements (inter and intra-fraction), the set-up and beam uncertainties, forcing a small volume of healthy tissue to the prescription dose. Therefore, as already discussed in the definition of the therapeutic window, effective curative radiotherapy is necessarily associated with an accepted risk of developing side-effects, in order to achieve the best tumour control.
The radiobiological target/stem cell concept postulates that the radiation tolerance of any organ or tissue is defined by the number and intrinsic sensitivity of the tissue specific target cells [35].
Early (acute) symptoms are manifested during or shortly after radiotherapy (usually set to 90 days after the onset of the treatment). Late effects (chronic) have a longer latent time, usually years. However, in some structures, interactions between acute and chronic effects can occur, resulting in the so called "consequential" late effects. This happens when direct impairment of a protective structure in early phase, such as epithelial cell depletion, favors indirect mechanical and chemical exposure in the late phase aggravating the radiation response. Consequential late effects are observed in intestine, urinary tract, oral mucosa and particularly stressed skin localizations [35, 36, 37, 38, 39].

Early radiation effects (H-Type tissues)
Understandably, the early sequelae after radiation exposure were the first ones to be investigated in the first decades of the 20th century. Their manifestation focused immediately the attention of physicians and radiobiologists toward the concept of tolerable prescription dose to the tumour. Indeed, severe acute treatment sequelae can be dose-limiting, and hence may affect tumor control. Particularly, the severe toxicity of the skin and the epidermis were frequent in the orthovoltage era, when a peak occured in the first millimeters of the entrance site of the beam. However, early radiation effects are still relevant even after the technological progress in the dose delivery. Few cases of severe symptoms are still present in the treatment of critical districts and new efforts in this research topic should be done to avoid the interruption of the radiotherapy during the weeks and successfully complete the treatment.
The first component in the early toxicity is the inflammatory response, which alters the function of the target cells. It is characterized by local vasculoconnective (e.g. vasodilation, vascular leakage, erythema), immunological reactions and an increased protein expression, typically pro-inflammatory molecules. It precedes the phase of hypoplasia (see for example the studies on the epidermal, mucosa or leukopenia and cell depletion [36, 37]). *Hierarchical* tissues are the structures that are more affected by early damage. Cell populations into

these tissues are made of stem cells, functional cells and maturing partially differentiated cells (in between the first two types of cells), with stem cells constantly giving rise to maturing cells which eventually fully differentiate and become functional cells. Hierarchical tissues include most epithelial layers and the bone marrow. From this point of view, the cellular turnover becomes very important, mediated by the differentiation of the tissue stem cells. They are indifferentiated cells that can correctly restore the integrity of an organ after an insult. Hence, the radiation tolerance of a tissue is characterized and defined by the number of stem cells and by their intrinsic radiosensitivity. The equilibrium between cell production and cell loss is based on the division pattern of stem cells. On average, each stem cell results in one cell that remains in the stem cell pool and one cell which eventually differentiates, the so-called asymmetrical division [40]. In most turnover tissues, transit cells by far dominate the proliferative cell population; the relationship between the number of transit cells and stem cells depends on the number of transit divisions. It varies from tissue to tissue, as well as the turnover time. It is also associated with the time course of the early radiation response. Moreover, the radiosensitivity of the cells decreases during the differentiation process. Different grades of severity of early toxicity are based on different levels of cell depletion as illustrated in Figure 1.6. From the picture we can also see as the cell loss rate depends on the turnover time of the tissue, and is independent of the treatment; the latent time until a clinical response is reached is tissue-dependent but independent of dose. It is worth noting that the latent time, i.e. the time from the onset of a stimulus to that of the response, is usually longer than the turnover time. This happens because of the residual proliferative capacity of sterilized cells even after medium-high doses. For example, ulceration in human mucosa after a fractionated dose of 20 Gy appears in 10 days, even if the turnover time for that tissue is in the range of 5 days [41]. Finally, we introduced the concept of asymetrical division, but it has to be considered that when the radiation treatment is in a hypofractionation scheme or in accelerated fractionation, the number of surviving stem cells (more senstitive to radiation compared with differentiated cells) is compromised and, thus, the tissue environment does not provide the needed signals to activate the differentiation, and in order to restore the stem cell pool the process goes on in a symmetrical division.

Manuscripts discussing the side effects in different cancer districts demonstrated an association between acute and late toxicities as a possible result of *consequential component* [42]. In the example of gastro-intestinal (GI) toxicity after prostate cancer radiotherapy, it is often described the fact that acute toxicity leads to inflammation, leading to leakage of intestinal contents or damage of GI mucosa and eventual fibrosis or bleeding manifesting as late toxicity. Other authors believe that the association is only a result of shared dosimetric risk factors, or also of the inherent properties in individuals, such as genetics or comorbidities that lead to greater tissue sensitivity to radiation both acutely and chronically. However, the general evidence from the literature is

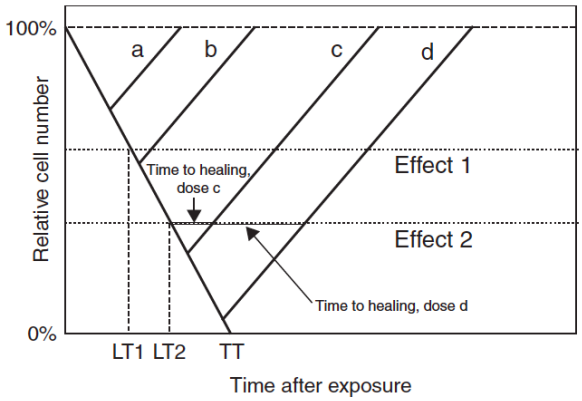


Figure 1.6: *Radiation-induced cell depletion and clinical manifestation of early radiation effects. Radiation exposure of turnover tissues results in an impairment of cell production, while cell loss continues independent of the treatment. The rate at which cells are lost is determined by the tissue turnover time (TT). If the residual proliferation of sterilized cells (abortive divisions) is not taken into consideration, then a complete loss of cells would be observed after one turnover time. A defined clinical effect 1, which is associated with a specific reduction of the cell number, can occur dependent on the dose (dose level b, c or d), and is not observed at lower radiation doses (dose level a). The latent time to clinical manifestation, however, is independent of dose. A more severe effect level 2 is based on a higher reduction in cell numbers, and hence is observed only at higher doses (c and d). Compared with effect 1, the latent time is longer, but is also independent of dose. In contrast, the time to clinical healing is longer with higher doses (d versus c). Image from reference [36]*

that consequential late effects develop in situations where the early radiation responses are associated with breakdown and loss of physiological protective barrier against mechanical or physical stress [43, 44]. This is, in particular, found for surface epithelia of the upper and lower digestive tract (oral cavity, oesophagus, small and large intestine, rectum) with a combination of mechanical and chemical influences, for the epithelium of the urinary bladder (chemical stress), for the epidermis at localisations with major mechanical wear and tear, and also in the lung.

Chronic Radiation Effects (F-Type Tissues)
In the study of the pathogenesis of late radiation effects it has been identified that the tissues affected by late radiation symptoms are more complex and made up of parenchymal, fibroblasts and endothelium cells. Thus, the response is orchestrated by all of these cells and the tissue is defined as a *flexible* tissue (F-Type). Late tissue responses occur in organs because parenchymal cells normally divide infrequently and hence do not express mitosis-linked death until later times when called upon to divide. They also occur in tissues that manifest early reactions, such as skin/subcutaneous tissue and intestine, but these reactions (subcutaneous fibrosis, intestinal stenosis) are quite different from early reactions in these tissues. Late responses (usually regarded as those which occur more than 3 months after treatment) often limit the dose of radiation that can be delivered to a patient during radiotherapy[2]. Damage can be expressed as diminished or even complete loss of organ function, and is usually progressive over time. The nature and timing of late reactions depend on the tissue involved. Damage to stromal and vascular elements of the tissue and the influx of inflammatory cells may cause secondary parenchymal cell death, resulting in increased cell proliferation and further death of functional cells as they attempt mitosis. The latent period to manifestation of organ dysfunction depends on the dose received, because the higher the initial dose the smaller the fraction of surviving parenchymal cells that can repopulate the tissue. One common late reaction is the slow development of tissue fibrosis that occurs in many tissues (eg, subcutaneous tissue, muscle, lung, gastrointestinal tract), often several years after radiation treatment. Radiation-induced fibrosis is associated with a chronic inflammatory response following irradiation, the aberrant and prolonged expression of the growth factor TGF-β and radiation-induced differentiation of fibroblasts into fibrocytes that produce collagen [45]. Transforming growth factor-β also plays a major role in wound healing and the development of radiation fibrosis has similarities to the healing of chronic

[2]Acute symptoms are usually considered of "minor" importance in determining dose constraints, as in many cases acute symptoms/early reactions are solved within some months after the RT. Two main exceptions to this are: (i) acute symptoms that can cause radiotherapy treatment interruptions (as this can compromise tumor control), for instance, preventing severe mucositis is one of the main goals in several cancer treatments; and the presence of consequential late effects, where minimizing acute symptoms reflects in minimization of late chronic toxicity

wounds [46]. Another common late reaction is progressive vascular damage, including telangiectasia that can be observed in skin and mucosa, and loss of microvasculature leading to atrophy (and fibrosis) that is manifest in skin and other tissues.

Studies on lung toxicity, an important site for late symptoms, showed us that radiation reaction can increase the tissue density on CT scans but also that inflammatory cells and inflammatory cytokines play an important role in lung response to irradiation injury (the relationship between this inflammatory response and the later development of functional symptoms is still unclear). Moreover, the dose required to cause functional impairment in lung depends on the volume irradiated, with small volumes being able to tolerate quite large doses [47].

Concerning the tolerated dose to a volume in order not to develop side effects, large differences were found among the structures. It depends on the architecture of the organs. More details on this concept are presented in the next section.

The volume effect in healthy tissues: serial and parallel organs

In radiotherapy, major advances in radiation technology have resulted in a progressive conformation of high-dose volume to the macro- and micro-scopic tumor tissue. As a consequence, the volumes of normal tissue exposed to mid-high dose were significantly reduced, and dose distribution within these volumes became inhomogeneous (differently from conventional techniques). This implies that the effect of exposure of the fractional volume of a normal tissue to certain doses needs to be considered, rather than the dose to the entire tissue or organ. Consequently, the irradiated volume of an organ must be taken into account as an important parameter that determines the clinical consequences of tissue exposure. Finally, the heterogeneity in the dose distribution possibly allows to identify sub-structures into the organ that could be directly involved in a specific toxicity.

A theoretical framework introduced by Withers et al [48] suggests that late responding tissues can be considered as arrays of *Functional Subunits* (FSU) containing groups of cells that are critical for a function (eg, bronchioli in lung, nephrons in the kidney). These FSU were postulated to be able to be regenerated from a single surviving tissue stem cell. Furthermore, tissues were considered to have these FSU operating in parallel to achieve overall tissue function (such as it occurs in lung, kidney, liver) or in series (such as in spinal cord or intestine) in analogy with electrical circuits. Tissues with a parallel structure of FSU have substantial reserve capacity and, although damage to a small volume may completely inactivate this volume, the remaining regions can maintain function and/or may undergo hypertrophy to replace any loss of function (eg, kidney and liver). Tissues with a serial structure of FSU may cease to function even if a small region of the tissue is irreparably damaged, such as may occur in the spinal cord where localized injury can cause complete

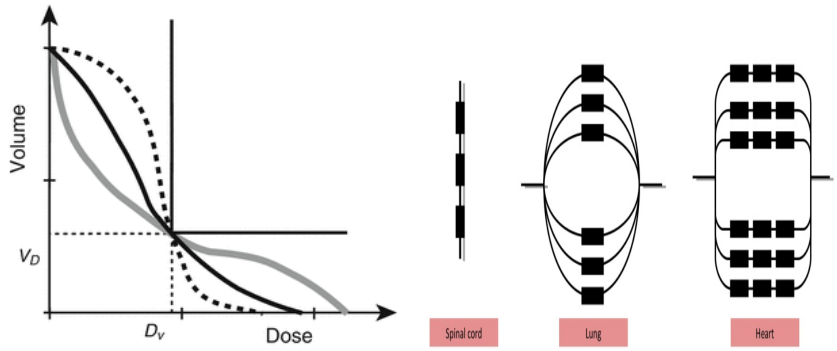


Figure 1.7: *On the left: Representation of limitations of the Dose Volume Histogram (DVH) for the structure with a large volume effect such as parallel-like organs. Three DVH curves producing the different effects will fit for one dose-volume-based constraint. Application of multiple constraints may well control the shape of the DVH as desired. On the right: serial, parallel and serial-parallel organs architecture. Images from reference [49].*

tissue dysfunction and myelopathy, or in the intestine if severe stenosis causes obstruction. In practice, tissues do not fall neatly into these two categories for various reasons, including the common role of the vasculature, the development of inflammatory responses that may extend beyond the treatment field, because FSU may require more than one type of undamaged stem cells for repair and these stem cells may migrate into areas of damage either locally or via circulation. However, the concept that the volume irradiated to high dose is critical to tissue response and that this varies for different organs is well established and used in mathematical models designed to predict normal tissue complication probabilities (NTCP) [50, 51]. The modelling of volume effects on the basis of their serial or parallel organization is useful to explain the apparent paradox that relatively radiosensitive organs, such as kidney and lung, can sustain the loss of more than half their total mass without significant loss of function, whereas relatively radioresistant tissues such as spinal cord can be functionally inactivated by the irradiation of only a small volume.

The first method developed by outcome modellers to catch and describe the importance of the dose-volume distribution was to include the dose-points of the Dose Volume Histogram (DVH) into a mathematical model in order to identify the regions of the DVH that were more responsible for the development of sequelae. Radiosensitivity of the organ was then mathematically included into the NTCP models when it has been decided to move from punctual dose parameters to equivalent uniform dose (EUD), which includes a volume parameter, n, that gives us a description of the the organ behaviour. More details about this dose parameter are given in Chapter 3.

Table 1.1: *Comparing differences in the overall treatment time and local control for the three cohorts included in the DAHANCA head and neck study of treatment time prolongation. Redrawn from [16]*

Total Dose (Gy)	Dose (Gy)	Comment	Overall Time (wk)	3 years Local Control
66-68	2	Split Course	9.5	32%
66-68	2	5 fr/wk	6.5	52%
66-68	2	6 fr/wk	5.5	62%

1.3 The importance of fractionation

Radiobiological experiments performed in France in the first decades of the last century estabilished the need of multifraction regimens in conventional radiation therapy. The studies on the sterilization of the rams performed in France in 1920s proved that if radiation was delivered in daily fractions over a period of time, sterilization was possible without skin damage. The general rule was suddenly confirmed also in humans. The efficacy of fractionation is based on the difference in cell kinetics between tumour and normal cells, more generally on the concepts behind the 4Rs of radiotherapy. In brief, dividing a dose into several fractions spares normal tissues (and not the tumor) because of repair of sublethal damage between dose fractions and repopulation of cells, if the overall time is adequate.

Overall treatment time and its influence on early and late-responding tissues

The effect of the overall treatment time is strictly related to the so called accelerated repopulation phenomenon. It is the consequence of a cytotoxic agent, like radiation or chemo-therapy, which triggers the surviving cells in the tumour to divide faster than before. Evidence of this process has been found in animal and human tumours [52, 53]. The analysis of the study of Withers et al [54] on head and neck cancer suggested that cell repopulation accelerates at 28 days and a dose increment of 0.6 Gy per day is required to compensate this process. Particularly, for fast growing tumours the treatment should be completed soon after it began. As a consequence, it may be better to delay the initiation of treatment than to introduce delays during the treatment (see in Table 1.1 the consequence on local tumour control of the 1-week stop in the DAHANCA head and neck study).

In general, if the overall treatment time is too long, the effectiveness of dose fractions released after beginning of repopulation acceleration is compromised. Overgaard and colleagues [55] demonstrated the importance of the overall treatment time by a retrospective comparison of three consecutive trials of the Danish cooperative group. A summary of the study is shown in Table 1.1.

An estimate of the loss in tumour local control for each day of treatment time prolongation has been determined for head and neck and cervix cancer; resulting values are 1.4% (0.4% to 2.5%) and 0.5%(0.3% to 1.1%). Such a rapid

proliferation does not occur for carcinoma of the breast or prostate where the overall treatment time is not so critical.

Currently, with the term conventional fractionation we define a treatment with daily doses of 1.8-2 Gy, given in five fractions per week and with a prescription dose ranging from 40 Gy to 80 Gy (according to the volume of the tumour). The general rules were developed on empirical basis by Fletcher and colleagues [56], and have been the gold standard for curative radiotherapy over the last decades in most of the American and European hospitals. Dose escalation and different daily doses are currently on study, compatibly with the uncomplicated local tumour control. Particularly, the incoming conformal techniques have increased the interest toward dose escalation.

A description of reasons most influencing the suitable number of fractions is given in this section (hyperfractionation vs hypofractionation). As a starting axiom we quote the Withers definition [48]: "Fraction size is the dominant factor in determining late effects; overall treatment time has little influence. By contrast, fraction size and overall treatment time both determine the response of acutely responding tissues".
In clinical practice, hyperfractionation is applied to increase the total dose compared to the values of prescription dose in conventional fractionation. Accelerated treatment is based on the idea that giving the same dose in half the time by delivering two fractions each day should increase tumor control without affecting the incidence of late effects, because they are not dependent on overall time. Two large multicentric randomized clinical trials on head and neck squamous cell carcinoma were organized by the European Organisation for Cancer Research (EORTC) and Radiation Therapy Oncology Group (RTOG). EORTC compared hyperfractionated treatment of 1.15 Gy for a total dose of 80.5 Gy given in two fractions per day with an interval time of 4-6 hours with a conventional treatment of 70 Gy at 2 Gy/fr [57]. Figure 1.8

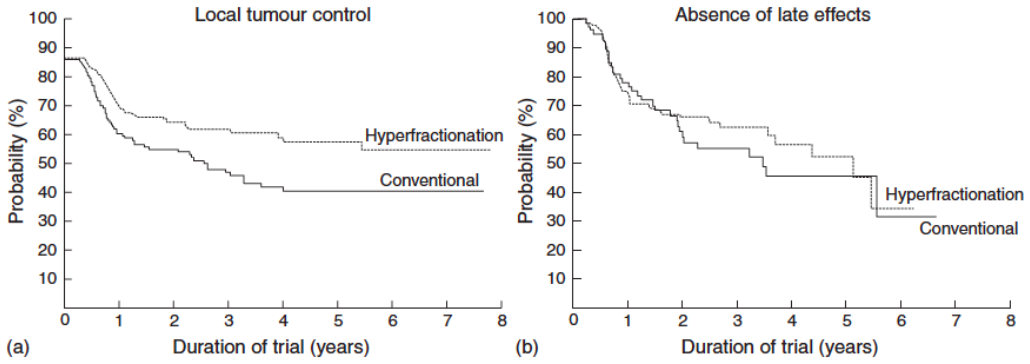


Figure 1.8: *Results of the EORTC (22791) trial of dose-escalated hyperfractionation. (a) Loco-regional tumour control (log-rank p=0.02); (b) patients free of late radiation effects, grade 2 or worse (log-rank p=0.72).. Image from reference [58]*

shows the survival analysis for local tumour control and late tissue damage. For the latter no difference was found between the two regimens, while an increase of about 19% was found in long term local tumour control. Thus, survival appeared higher than in conventional treatment but it did not reach a statistical significance. In the RTOG trial (68 fractions of 1.2 Gy, two fractions per day, 6 hours apart, total dose 81.6 Gy), local tumour control was increased by 8% in the hyperfractionation arm compared with conventional fractionation using 2 Gy fractions to 70 Gy in the same overall time of 7 weeks [59]. No significant difference was found in the overall survival, while the prevalence of grade 3 late effects was significantly increased after hyperfractionation. Both trials confirmed the radiobiological expectation of a better local tumour control when the dose is escalated with hyperfractionation (suggesting a high α/β ratio for squamous cell carcinoma of the head and neck, see next chapter for further details). A possible reason of the toxicity occurrence could be the range of time of multiple fractions, which should be given as far apart as possible and not closer than 6 hours [60]. Indeed, a sufficient time is needed to allow complete repair between fractions after each dose (see Figure 1.9). Another more recent study that reaches the same conclusions was the Continuous, Hyperfractionated, Accelerated RadioTherapy (CHART), where locally advanced non-small cell lung cancer patients enrolled in the study underwent three fractions per day (Figure 1.10) [61]. The efforts of these studies provided much new information, and some of the protocols had an impressive gain in tumour control. Nevertheless, multiple fractions per day have never become mainstream in radiotherapy practice because of the logistics involved.

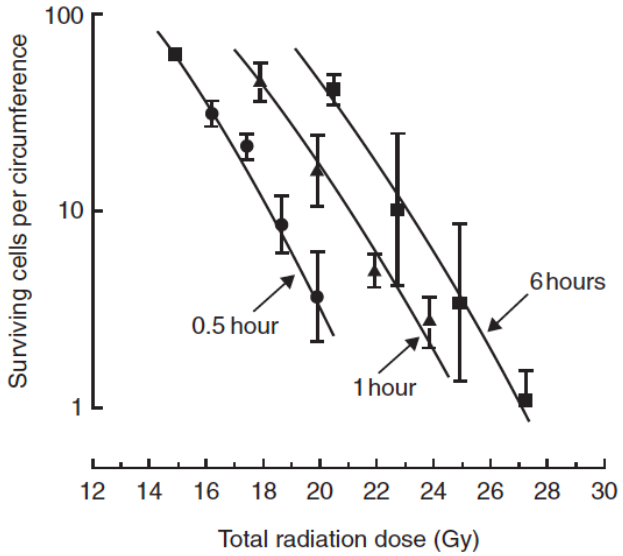


Figure 1.9: *Effect of interfraction interval on intestinal radiation damage in mice. The total dose required in five fractions for a given level of effect is less for short intervals, illustrating incomplete repair between fractions. Image from reference [62]*

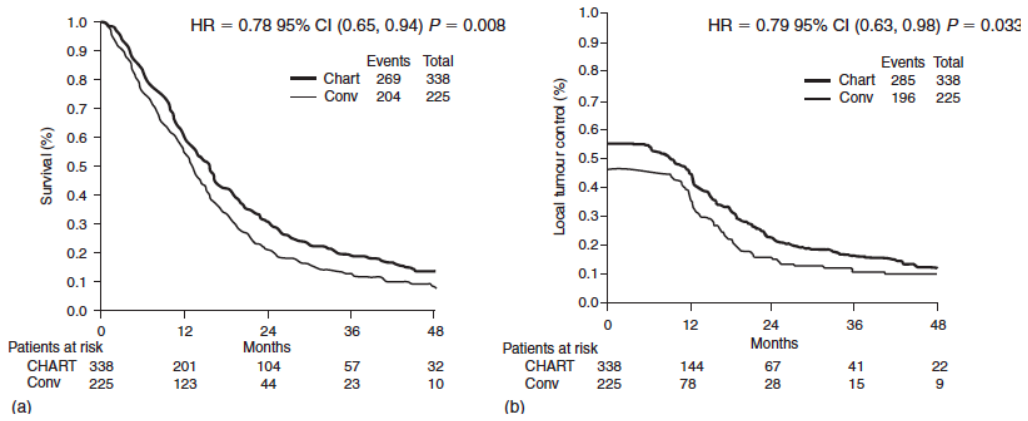


Figure 1.10: *Results of a phase III randomized trial of CHART (continuous hyperfractionated accelerated radiotherapy) in non-small cell lung cancer. (a) Overall survival; (b) local tumour control of patients treated by CHART or by conventional radiotherapy (CHART results are indicated by the heavier line). HR, hazard ratio. Images from reference [58]*

On the other hand, we speak about hypofractionation when a dose higher than 2 Gy/fr is given to the patient. The total number of fractions is reduced, hence the prefix 'hypo'. The radiobiological expectation is that, compared with conventional regimens given in the same amount of time, hypofractionation could improve the therapeutic ratio between tumours and late-responding normal tissues when the $\alpha/\beta_{tumor} < \alpha/\beta_{normal}$. A randomised clinical trial was set up for breast cancer. The study included 1410 women who had local tumor excision. Treatment protocols were 50 Gy in 25 fractions, 39 Gy in 13 fractions and 42.9 Gy in 13 fractions, all of them given in 5 weeks. A significant difference (p=0.027) was found in ipsilateral tumor relapse at 10 years between conventional treatment and hypofractionation at 39 Gy in 13 fractions [63].

Further application of hypofractionated regimen can be found also in the treatment of prostate cancer. Indeed, tumor control and acceptable toxicity outcomes from several hypofractionation or brachytherapy analyses support an α/β ratio for prostate cancer that is low (Brenner et al [64] $\alpha/\beta = 1.5Gy$ and Fowler et al [65] $\alpha/\beta = 2Gy$), perhaps even lower than that for the normal organs that ordinarily constrain the delivery of radiation therapy. A plausible hypothesis is that the α/β ratio of a tissue is determined by the proportion of cycling cells compared with cells that are not dividing. Because most prostate tumors are slow growing there is not a major differential in terms of cell division rates between a typical prostate tumor and the surrounding late responding normal tissue. This finding removes the advantage conferred by multiple dose fractions, and led to the suggestion that "if the fractionation sensitivity is the same for the tumor and the surrounding late-responding normal tissue, much smaller numbers of fractions (with an appropriately reduced dose) would be expected to be at least as efficacious, but logistically

and financially advantageous."[64]. Thus, a low α/β ratio for prostate cancer remains an attractive hypothesis supported by several lines of evidence. While clinical data supporting a low α/β ratio is becoming more plentiful, derived estimates are still characterized by wide confidence intervals. These uncertainties will ultimately not be resolved until biochemical control data from large hypofractionation studies with 5 or more years of follow-up become available.

Among the larger randomized trials we have the PROFIT trial (3 Gy in 20 fractions for a total of 60 Gy) [66] and the CHHiP trial [67] (a total dose of either 57 or 60 Gy again in 3 Gy/fraction). The results of these studies, made with a median follow up of 5 or 6 years, showed that hypofractionation was non-inferior in regard to tumor control, and showed no significant differences in terms of late sequelae. In comparison, the RTOG-0415 study [68] used a lower dose per fraction (2.5 Gy), but a much larger total dose (70 Gy). It had a similar median follow up as with the PROFIT and CHHiP trials, and demonstrated non-inferiority for hypofractionation in terms of tumour control, but there was an increase in grade 2+ late *gastrointestinal* (GI) and *genitourinary* (GU) toxicity. Finally the HYPRO trial [69] used a significantly higher dose per fraction (3.4 Gy) and total dose (64.6 Gy) as compared with the PROFIT and CHHiP trials; the HYPRO trial showed non superiority of the hypofractionation arm in term of tumour control, but also showed higher grade 2+ GI (but not GU) acute toxicity and higher grade 3+ GU (but not GI) toxicity; these toxicity increases can be understood in terms of the high dose per fraction and dose that were used in the HYPRO file.

Brenner and Hall in their review, which considered a total of 8 trials of hypofractionated treatments (including the 4 previously described), concluded with two observations:

- of the four studies, the only one to show significant increased normal tissue toxicity used the largest dose per fraction;

- more in general, the only study that showed significant superiority (as opposed to non-inferiority) for hypofractionation in terms of tumor control was the trial by Yeoh et al [70]. It was the first trial in literature enrolling patients already in 1996 and tumour volume was irradiated with 2D radiotherapy. However there are several caveats; first, it was a comparatively small study with only 217 patients, and second the comparability in terms of risk levels between the hypofractionated and conventional arms may have been questionable.

Overall, while non-inferiority of hypofractionation in terms of tumor control now seems well established (other more recent trails getting similar results can be found in the review of Ritter et al [71]), it is evident that the "superiority trials", based on the hypothesis that moderate hypofractionation would increase tumor control efficacy compared with conventional fractionation, have produced negative results.

Palliative treatments are also of interest for developing hypofractionation schemes. In this circumstance, life expectancy is limited and late tissue side-effects are not a major concern. These schemes are convenient for the patient but also for the sparing of resources. Moreover, stereotactic body radiotherapy (SBRT) is useful to explore and enlarge the knowledge of the dose/fraction domain. In lung cancer, large fractions of dose on small volumes are frequently applied in clinical practice. Indeed, SBRT is now used for many tumors, not only lung cancer [72]. Moderate hypofractionation is routinely used for curative treatment in many countries. The assumption is that the negative effect of the reduced total dose is in some way mediated by the shortness of the overall treatment time. However, these advantages have to be carefully weighted to avoid an increased risk of late normal tissue injury. Again, clinical trials are extremely needed to assess the therapeutic gain compared with the standard protocol.

1.4 Properties of radiation source to improve the Therapeutic Ratio: hadron-therapy

The actual knowledge of dose-response in patients is still mainly based on *3D Conformal Radiotherapy* (3DCRT) studies. Improvements in new technologies should help the oncologist to conform more precisely the dose to the tumour volume sparing the normal tissues. As a consequence, treatment modality is one of the options to optimize the therapeutic index. This could be achieved through a better dose distribution (*Intensity Modulated Radiotherapy* (IMRT), brachytherapy, heavy ion-therapy) and/or radiobiological efficacy due to the properties of alternative sources (heavy ion-therapy). To better understand these properties a brief introduction to the concepts of *Linear Energy Transfer* (LET), *Relative Biological Effectiveness* (RBE) and *Oxygen Enhancement Ratio* (OER) will follow.

LET

Linear energy transfer is the term used to describe the density of ionization in particle tracks. The special unit usually used for this quantity is kiloelectronvolt per micrometer ($keV/\mu m$) of unit density material. In 1962, the International Commission on Radiological Units defined this quantity as follows: the LET of charged particles in a medium is the quotient of dE/dl, where dE is the average energy locally imparted to the medium by a charged particle of specified energy in traversing a distance dl. It can be defined as *restricted* LET

$$L_\Delta = \left(\frac{dE}{dl}\right)_\Delta \tag{1.2}$$

which takes only into account the energy loss due to collisions with energy transfer smaller than threshold Δ,or as L_∞, meaning that every energy transfer

is included and it is equivalent to the physical concept of stopping power. According to the Bethe-Bloch equation for the loss of energy we can summarize the LET dependency from the treatment particle (the target is always the human) as LET $\propto Z^2/\beta^2$. LET depends on particle charge Z and on its velocity, and so energy (see Table 1.2)[3].

Table 1.2: *Linear Energy Transfer for different type of radiation and different particle energy. Redrawn from [16]*

Type of Radiation	**LET** ($keV/\mu m$)
$^{60}Co\gamma$**-radiation**	**0.2**
250 kVp X-radiation	**2**
10 MeV protons	**4.7**
150 MeV protons	**0.5**
recoil protons from fission neutrons (1 Mev)	**45**
130 MeV/u Carbon Ions	**112**
14 MeV neutrons	**12**
2.5 MeV α -particles	**166**
2 GeV Fe nuclei	**1000**

It is common practice to distinguish between the so-called *sparsely* and *densely* ionizing radiations. The first one have an almost uniform spatial distribution of ionizations, the latters are able to produce a condensed number of ionizations in a smaller region. X-rays and γ-rays (low-LET) are sparsely ionizing radiations, while heavy ions, α particles and protons (high LET) are densely ionizing radiation. In Figure 1.11 (a) we can see how the radiation produces more cell killing per Gray when its LET increases. The figure reports the survival of human T1g kidney cells plotted against dose for eight different radiations, with LET varying from 2 $keV/\mu m$ (250 kVp X-rays) to 165 $keV/\mu m$ (2.5 MeV α-particles). The higher the LET the steeper the survival curve, with a less pronounced shoulder. In the linear quadratic (LQ) description, these almost linear cell survival curves have a higher α/β ratio, thus higher LET radiations usually give responses with higher α/β. However, notice that 2.5 MeV α-particles are less efficient than 4.0 MeV α-particles even though they have a higher LET; this is because of the phenomenon of overkill shown in Figure 1.11 (b).

RBE

The RBE has been introduced to compare the efficacy of a radiation under test (e.g. a high-LET radiation) with respect to a reference radiation. This

[3]It could be helpful to provide also information about the range of such particles even if this was not done in the original reference of the table. We derived ranges for protons by the use of the PSTAR software (nist.gov), 10 MeV protons have a range of 0.1 g/cm^3, while 150 MeV protons have a range of 15.8 g/cm^3. A range of 0.001 g/cm^3 was computed for α-particles by the use of ASTAR software (nist.gov). A range for Carbon Ions was retrieved from the study of Habermhel et al[73]

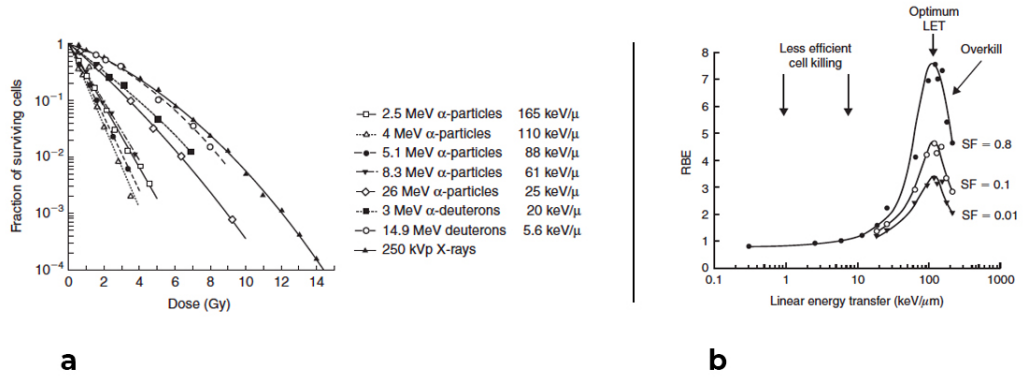


Figure 1.11: *a) Linear Energy Transfer for different type of radiation and different particle energy;b) Dependence of relative biological effectiveness (RBE) on linear energy transfer (LET) and the phenomenon of overkill by very high LET radiations. The RBE has been calculated from Figure 1.11 (a) at cell surviving fraction (SF) levels of 0.8, 0.1 and 0.01. Images from reference [74]*

parameter is defined as:

$$RBE = \frac{dose\ of\ reference\ radiation}{dose\ of\ test\ radiation} \tag{1.3}$$

to give the same biological effect. The reference low-LET radiation is usually $250\,kVp$ X-rays or $60\,Co\,\gamma$-rays since these radiations are readily available. Figure 1.11 (b) shows RBE values for the T1g cells featured in Fig. 1.11 (a), RBE has been calculated at cell survival levels of 0.8, 0.1 and 0.01, illustrating the fact that RBE is not constant but depends on the level of biological damage and hence on the dose level. The RBE also depends on LET, and rises to a maximum at an LET of about $100\,keV/\mu m$, then falls for higher values of LET because of overkill. For a cell to be killed, enough energy must be deposited in the DNA to produce a sufficient number of double-strand breaks. Low-LET radiation is inefficient because more than one particle may have to pass through the cell to produce enough DNA doublestrand breaks. Very high-LET radiation is also inefficient because it deposits more energy per cell, and hence produces more DNA double-strand breaks than are actually needed to kill the cell. These cells are 'overkilled', thus "wasting" a part of the deposited energy, and resulting in a lower likelihood per unit gray. This leads to a reduced biological effect. Radiation of optimal LET deposits the right amount of energy per cell, which produces just enough DNA double-strand breaks to kill the cell. This optimum LET is usually around $100\,keV/\mu m$ but does vary between different cell types and depends on the spectrum of LET values in the radiation beam as well as the mean LET.
In summary, RBE depends on the following: radiation quality (LET), radiation dose, number of dose fractions, dose rate, biologic system or endpoint.

Radiation quality includes the type of radiation and its energy, whether electromagnetic or particulate, and whether charged or uncharged.

OER

OER has been briefly introduced describing the effect of reoxygenation on cells sensitivity. The degree of sensitization is described by OER which is defined (see Figure 1.4 (a)) as the ratio of doses required to give the same biological effect in the absence or the presence of oxygen. For doses of X- or γ-radiation greater than approximately 3 Gy, the OER for a wide range of cell lines in-vitro, and for most tissues in-vivo, is in the range of 2.5 to 3.3. For X- or γ-ray doses less than approximately 3 Gy (ie, in the shoulder region of the survival curve), the OER is reduced in a dose-dependent manner. A reduction of the OER at low doses is clinically important because the individual treatments of a fractionated course of treatment are usually 2 Gy or less. The OER is also dependent on the type of radiation, declining to a value of 1 for radiation with LET values greater than approximately 200 $keV/\mu m$. The oxygen level (pO_2) in most normal tissues ranges between approximately 20 and 80 mm Hg, whereas tumors often contain regions where the pO_2 is less than 5 mmHg. These conditions in solid tumors are primarily caused by the abnormal vasculature that develops during tumor angiogenesis. The blood vessels in solid tumors have highly irregular architecture, and are more widely separated than in normal tissues. A proportion of tumor cells may lie in chronically hypoxic regions beyond the diffusion distance of oxygen. Tumor cells may also be exposed to shorter (often fluctuating) periods (minutes to a few hours) of acute hypoxia as a result of intermittent flow in individual blood vessels. Tumor hypoxia has been observed in a majority of tumors both human and experimental, but has been found to be very heterogeneous both within and among tumors, even in those of similar histopathological type, and usually it does not correlate simply with standard prognostic factors such as tumor size, stage, and grade [75].

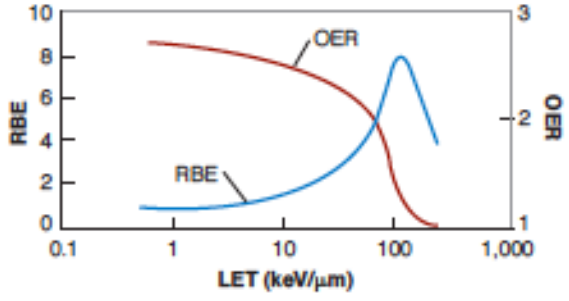


Figure 1.12: *Variation of the oxygen enhancement ratio (OER) and the relative biologic effectiveness (RBE) as a function of the linear energy transfer (LET) of the radiation involved. Note that the rapid increase in RBE and the rapid fall of the OER occur at about the same LET, 100 keV/μm. Image from reference [16]*

Heavy Charged particles

The use of protons for radiotherapy was proposed by R. Wilson in 1946 [76]. Heavy charged particle therapy, in general, is a form of external-beam radiotherapy using beams of energetic protons or positive ions for cancer treatment. The most common type of particle therapy is proton therapy. Particle therapies may have high LET and might contribute to improvements in the therapeutic ratio in several ways. First, because much of the energy of particle beams is deposited in tissue at the end of particle tracks (i.e., in the region of the Bragg peak), they can give improved depth-dose distributions for deep-seated tumors. For example, heavy ion therapy using carbon ions is used in a number of centers across the world. These beams have higher LET than protons and increased RBE. The therapeutic ratio may be improved by particle therapies also because the oxygen enhancement OER is reduced at high LET (see Figure 1.12), so that hypoxic cells in tumors are protected to a lesser degree. The variation in radiosensitivity with position in the cell cycle is also reduced for high-LET radiation and, in general, there is reduced variability in response between different cells. This is partly because cells exhibit reduced capacity for repair following high-LET radiation relative to that following low-LET radiation, leading to an increase in RBE (see Figure 1.12). Compared to protons, carbon ions have the disadvantage that beyond the Bragg peak, the dose does not decrease to zero, because nuclear reactions between the carbon ions and the atoms of the tissue lead to the production of lighter ions that have longer range than the primary iones.

One potential difficulty in using high-LET radiation is that, because late-responding tissues demonstrate greater repair capacity than early responding tissues, the reduction in repair capacity following high-LET irradiation will result in relatively higher RBE values for late responding normal tissues. However, the ability to deliver dose in a finely focused manner using protons or heavy ions combined with IMRT planning techniques reduces the volume of normal tissue exposed to high doses, limiting this concern. Results with protons demonstrate an advantage for treatment of some tumors, such as choroidal melanomas and skull-base tumors, that require precise treatment of a highly localized lesion, and in pediatric tumors, where the dose to normal structures should be decreased as much as possible to avoid side-effects during development [77, 78]. Moreover, the dose distribution of heavy charged particles assumes another important aspect in the treatment of pediatric cancer: it reduces the integral dose to the body, limiting the effect of large volumes exposed to low doses (typical of IMRT with photons) which can lead to radio-induced secondary cancers.

Chapter 2

The Response of Normal Tissue to Radiation

After radiation exposure of tissues, a whole orchestra of events, which can be summarized as "damage processing", is seen well before any clinical change becomes manifest. These cascades are initiated by the induction of free radicals and acute oxidative stress. This results in changes in the activity of transcription factors, and thus in the modification of various intracellular and extracellular signalling pathways. Such changes can be observed in all tissue components (parenchyma, fibroblasts, vasculature), and also in activated macrophages and immune cells. The combination of all these events induces unspecific as well as tissue-specific changes at the cellular/histological level (e.g. cell death, differentiation or proliferation, DNA damage response, chronic oxidative stress, and many others). This integrated response eventually results in pathological changes. Thus, the effects of the radiation dose could be basically seen at three levels of observation:

- *cellular scale*, which includes studies of signaling of radiation damage to DNA and repair mechanism of the double-strand breaks, cell survival analysis to investigate the cell death in cultures, apoptosis assays, blotting techniques to characterize DNA, mRNA, and proteins after irradiation, microarrays analysis and Single-Nucleotide Polymorphisms;

- *tissue domain*, where radio-induced changes of the tissue can be investigated through modern *in-vitro* bioengineering techniques, *ex-vivo* analysis of samples coming from animal excision or patient biopsy and in-vivo quantitative measurements extrapolated and computed from medical imaging such as *Computer Tomography* (CT), *Magnetic Resonance Imaging* (MRI), *Positorn Emission Tomography* (PET), *Cone-Beam CT* (CBCT);

- *patient level*, where the presence of symptoms and their severity can be assessed through questionnaires related to organ dysfunctions or patient

QoL, necessity of medical intervention to solve or alleviate the patient conditions (drugs, physiotherapy, blood transfusion, surgical intervention and many others), medical tools or tests that are able to measure functional parameters (e.g. blood test, saliva collection chamber).

In principle, the larger the observation scale, the more complex the structure of interest and, as a consequence, the longer is the latent time to observe the radiation-induced effects (radioresponsivity). However, since decades of research studies and technological improvement were needed before investigating the tissue and cellular level, the first efforts in understanding the radiation response on human beings were based directly on the patient and his/her symptoms. The studies included in the thesis project followed this chronological evolution and are presented in the same order, i.e., from the macroscale to the microscale.

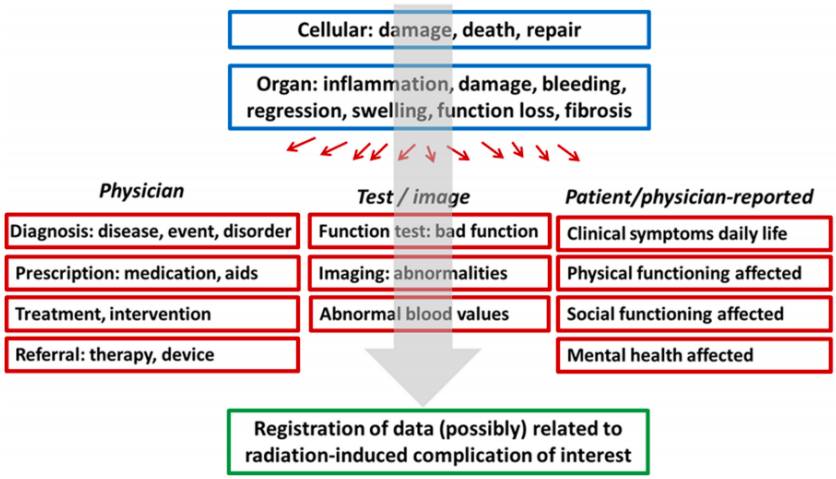


Figure 2.1: *Radiation damage (top two rows) can be indirectly measured or observed in clinical practice (middle three rows). In the database (bottom row) we have to capture all relevant data related to the complications of interest by defining relevant parameters. Image from reference [79]*

2.1 Patient Response

Several systems for documenting normal tissue responses (side effects) to irradiation in patients have been developed to facilitate cross-comparisons between investigators and institutions. These include the *Radiation Therapy Oncology Group* (RTOG)/*European Organization for Research and Treatment of Cancer* (EORTC) classification, the *Common Terminology Criteria for Adverse Events* (CTCAE v4) scale devised by the *National Cancer Institute* (NIH/NCI, 2009) and the *Late Effects Normal Tissue Task Force Subjective, Objective, Management, and Analytic* (LENT/SOMA) system, specifically designed to score late

reactions [80].
The last system had a widespread use in clinical trials in the last 20 years and it includes four main points to assess the patient condition that were defined as follows:

1. Subjective - in which the injury, if any, is recorded from the subject's point of view, that is, as perceived by the patient. This information can be elicited during interviews or derived by asking the patient to complete a carefully designed questionnaire

2. Objective - in which the morbidity is assessed as objectively as possible by the clinician during a clinical examination. In this approach, the clinician may be able to detect signs of tissue dysfunction that are still below the threshold that will give the patient symptoms but are an indication of more serious problems that are developing and will be expressed later

3. Management - which indicates the active intervention that may be taken in an attempt to ameliorate the symptoms

4. Analytic - involving tools by which the tissue function can be assessed even more objectively or with more biological insight than by simple clinician examination.

Moreover, the invasiveness and cost of any tool used to quantify the late effects must be reasonable and proportional to the severity of the symptoms and the possible therapeutic consequences.

In principle, all of these systems are comparable, and the scores from one system may be translated into the scores for another protocol, but with some exceptions. In general, the complications have a grade from 0 (no response) to 5 (lethal). In addition to the scoring system, a crucial aspect for a good evaluation of the outcome is the frequency of assessment (timing of measurements) and the total length of the follow up. Indeed, some symptoms may need five or more years to be manifested (e.g. heart failure), and some other side effects could be expressed for a short range of time and then be solved. From this point of view, most of the studies consider a long window of time for the patient follow up (at least five years), where evaluations are usually made every 6 months (see Figure 2.2). Acute toxicity follows a different cryterion. It is weekly assessed during the whole treatment and, in the most comprehensive trials an extra time point at 3 months is also required. Finally, it is obvious that in order to evaluate the effects of radiation on humans it is strictly important to collect the baseline status of the patient, which will be the standard to compare the follow up with. The baseline could be scored at the beginning of the radiotherapy treatment or at the start of the first oncology treatment, such as chemotherapy, surgery, or hormon-therapy. In this last scenario, a new score immediately before radiotherapy could help reducing the noise due to the presence of a neoadjuvant treatment that may alter the baseline values. A

well-known example is the effect of prostatectomy on the urynary dysfunction or the effect of hormon-therapy on the erectile dysfunction. The dose-response

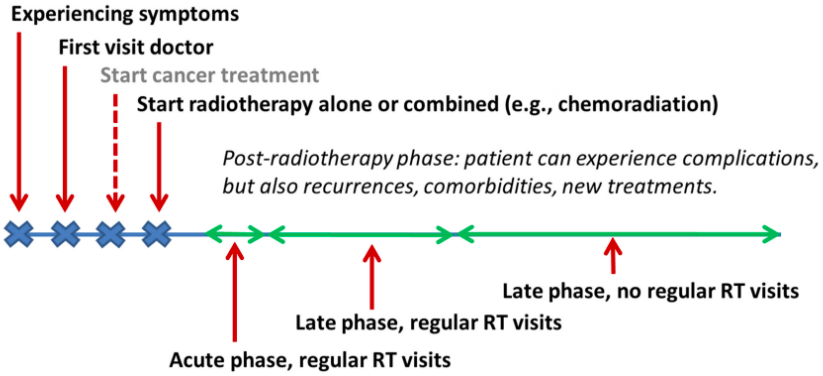


Figure 2.2: *Schematic representation of the time line for a patient receiving radiotherapy; data collection during regular RT visits may cover only partly the time points of interest of the study. Image from reference [79].*

relationship could be investigated by the use of sigmoid shaped functions, as also described in Chapter 3.1. To work with this mathematical function, the considered health outcome is usually dichotomized and converted in a binary value representative of the toxicity of the patient. The symptoms are sometimes described by more than one question and criteria to summarize the values are usually provided with the questionnaire. Then, according to the grade of severity that the study wants to investigate, a threshold is used to convert the number into 0, representative of patients who did not manifest the outcome, or 1, representative of patients who developed the symptom. Furthermore, since the aim is to evaluate the impact of radiation treatment on the patients organs at risk, the first step is to define the organ (presumably) responsible for the investigated clinical outcome. Successively, the dosimetric information for that organ has to be retrieved from the treatment planning system. The basic dosimetric parameters are provided by the software: the points of the *dose-volume histogram* (DVH), the minimum dose, the maximum dose, the mean dose and the integral dose. New features can be derived and computed to perform a more complex analysis (*Equivalent Uniform Dose* (EUD), dose-surface histograms, *dose-surface maps* (DSM) or weekly dose). These features are able to include some radiobiological aspects that go beyond the classic relation between dose and relative or absolute volume. EUD and DSM will be discussed more in details in Section 3.1 and 3.3. Once the best dosimetric parameter has been identified, the fit to the data will give basically an idea of the increasing rate of toxicity for every Gy of increase in the dosemetric parameter. The steeper the dose, the more deterministic is the phenomenon (cause and effect relationship that turns on above a certain threshold dose). Some examples of non-stochastic effect are: hair loss (2-5 Gy), sterility (2-3 Gy), lethality (whole body irradiation of 3-5 Gy). In a less deterministic scenario, there is

more chance for non-dosimetric factors to be responsible for the diversity of the patient response. The average dose-response relationship is described by the dosimetric NTCP model, but a sub-group of patients who are more or less sensitive to radiation could be hidden into the population. This could be explained through non-dosimetric factors: genetic and biological parameters such as gene expression, *Single Nucleotide Polimorphisms* (SNPs) and *micro-RNA* (miRNA); imaging features extracted from the scans included into the clinical protocol CT, MRI, CBCT during the weeks of treatment, PET-CT or even ecography; clinical factors scored during the baseline visit; these factors include patient, treatment and tumor characteristics, but also patient comorbidities. For each complication of interest, at least the known factors from the literature have to be established and included in the data collection procedures.

The approach by Peeters et al [81] is one of the most widespread to enrich the predictive model of toxicity towards a multifactorial NTCP model [1]. Non-dosimetric parameters can be risk or protective factors according to their capability to increase or reduce the probability of sequalae. Their effect is transferred into the NTCP model by the addition of a new costant, the so called *dose modifying factor*, that expresses the weight, and so the importance, of the parameter itself. From a graphical point of view, the inclusion of a new factor is translated into a shift to the right/left of the NTCP curve according to the protective/risk behaviour of the factor (see again Figure 1.2) and the larger is the constant the wider is the horizontal shift. More mathematical details are given in Chapter 3.1, as well as other possibilities to include the extra-dosimetric parameter into a predictive model.

2.2 Tissue Response

2.2.1 In-vivo

The need to detect and characterize cancer in an individual has resulted in a dramatic increase in the use of imaging over the last 20 years. Clinical imaging is now a routine part of diagnosis, staging, guiding localized therapy, and assessing the response to treatment. Cancers occur anatomically among surrounding normal tissues, including critical structures, such as major vessels and nerves, and delineation of the boundary of malignant and nonmalignant tissues is essential for planning surgery and radiation therapy. Cancers also have morphological, physiological, and biochemical heterogeneities which are important for understanding their biology and response to treatment. The ability to explore and define this heterogeneity with modern imaging methods,

[1] Historically the first mathematical method was presented by Gagliardi et al in 2000 [82]. In this study on pneumonitis after breast cancer treatment the team included the age in addition to the dosimetric parameter. Differently from the Dutch team, they fit the three parameters of the NTCP model in the subgroups of patients with age above and below a threshold.

as well as serum and tissue-derived metrics, will enable personalized cancer medicine. Imaging is diverse in that it offers an "anatomical image" of a mass on a CT or a MRI, a functional image of the disease status in PET images of glucose metabolism, and a "microscopic image" used during classification of histological type and grade. Imaging is applied at these multiple levels to help characterizing, understanding, and treating cancer and there is general acceptance that advances in imaging are central in the fight against cancer. Functional and non-functional parameters can be used as non-invasive imaging biomarkers or more easily as descriptors of inflammation, pathological state of fibrosis, swelling or even occlusion of the structure.

However, the utilization of imaging to quantify the RT response has recently emerged both for the tumor and for the tissue surrounding the radiotherapy target. In parallel the use of *radiomics*, i.e., high-throughput mining of quantitative features from standard-of-care medical imaging, has begun to be investigated in many institutes worldwide [83]. Established quantitative imaging approaches to assess the RT response have been explored for both tumors and normal tissues. Concerning the imaging analysis of healthy tissues, possible interests could be the differences in organ morphology and volume dimensions, the changes in density values or in functional parameters but also the variation of more complex imaging features extrapolated by radiomics studies [84].

A reasonable question is whether the quantification of normal tissue toxicity via imaging could improve, or in some cases even replace, traditional toxicity scoring methods that are often more subjective. This could be probably one of the most ambitious aims of the study of the signals coming from medical scans, the union of a quantitative and qualitative information. Proceeding with order is of primary importance to first assess and validate the capability of imaging, whatever the imaging technique, to identify a tissue change due to the radiotherapy treatment. To achieve that, the researcher needs to compare information from the same scan machine (possibly also the same type of machine) at two different time points. Typically, as it happens for the questionnaires, the first time is before the start of radiotherapy (unaltered photo used as a reference). The second point is strictly related to the aim of the investigation, and could be after a certain number of fractions or after the complete treatment (sometimes a third scan could be also included in the study to evaluate the recovery of the radio-induced changes in tissue microstructure). This is a limiting factor, because acquisitions repeated in time are not so spread in tumor treatment protocols. Indeed, it is quite common that patients undergo CT scans just for planning reasons, as well as for MRI which many times is required at the baseline for helping the delineation of structures that otherwise could not be contoured with the CT scan alone. Currently, daily verification of the patient postioning through CBCT are the only guaranteed source of repeated scans in time for most of the tumor and in most of the institute. However, the quality of the image, pixel resolution and sensitivty to the artefacts, make this technique difficult to use for radiomics

studies, even if Aerts and colleagues demonstrated the feasibility of this approach [85]. A dedicated study has to be defined to require and cover the costs of extra scans. Moreover, this practice could lead to an overcharge and congestion of the time slot of the diagnostic machines. Another critical aspect to be mentioned is that some diagnoses are performed in a hospital that is different from the one where the radiotherapy will be then executed. Also, before approaching the quantitative analysis, pre-processing of data should be done to ensure understanding of the image signal quality. For instance, the presence of patient-related artifacts (metal implants) as well as acquisition-related artifacts (e.g. bias field, differing voxel resolutions), could affect the image study and increase the probability of calculating erroneous features [85, 86]. In other words, harmonization and standardization for imaging acquisition may lead to more consistent findings in radiomic studies across institutions. After that, the next step is the image registration of the region of interest between the two or more acquisitions. Both rigid and deformable registrations are commonly used in the course of radiotherapy treatment. The concept of delta-radiomics demands high accuracy in image registration when comparing pre-treatment images with post-treatment images and images during treatment. The registration process can be achieved by manual, automatic, or the combination of both methods. At present, the commonly used automatic registration algorithms include image intensity-based methods, and structure-based methods. In conclusion, this type of research will help to bridge the molecular-clinical gap of translational cancer medicine, with imaging being a powerful research tool with a subsequent role in clinical applications of image-based personalized cancer medicine.

Imaging for small animals

With the growing use of small animal models of disease in cancer research, there has been a corresponding growth in the use of small animal imaging systems to characterize the disease with the aim of understanding cancer in these animal models and its response to novel therapies. Preclinical models of cancer typically include human tumor xenografts and genetically engineered murine or rat models. The use of imaging technologies (micro-MRI, -PET, -CT, -SPECT, and optical imaging) can interrogate local tumor growth, metastatic progression, tumor metabolism, and treatment response to allow for the design and testing of novel clinical regimens [87]. More rarely, these machines are used for investigating the effect of radiation on normal tissues. The use of multiple preclinical, noninvasive imaging techniques also allows the choice of the best way to address a particular oncology question, something that cannot be easily done in the clinic because of cost and patient inconvenience. The scaling of CT systems from man to mouse was made possible by the creation of high-resolution digital detectors (0.1 mm resolution) for radiography and fluoroscopy.

2.2.2 In-vitro

Significant advances in materials, microscale technology, and stem cell biology have enabled the construction of 3D tissues and organs, which will ultimately lead to more effective diagnostics and therapy. *Organoids* and *organs-on-a-chip* evolved from development of biology and bioengineering principles, have emerged as major technological breakthrough and distinct model systems to revolutionize biomedical research by recapitulating the key structural and functional complexity of human organs in-vitro. They are an artificially grown mass of self-organizes cells or tissue that resembles an organ or a micro-anatomy. Due to their properties, these in-vitro techniques could bridge the gap between animal models and monolayer cultures (described in the Section 2.3). Moreover, they can be scaled-up for high throughput testing at a lower cost with fewer ethical concerns [88].

Organoids and organ on a chip allow to overcome some of the limitations of the traditional in-vitro culture models as it is remarked in Figure 2.3. In

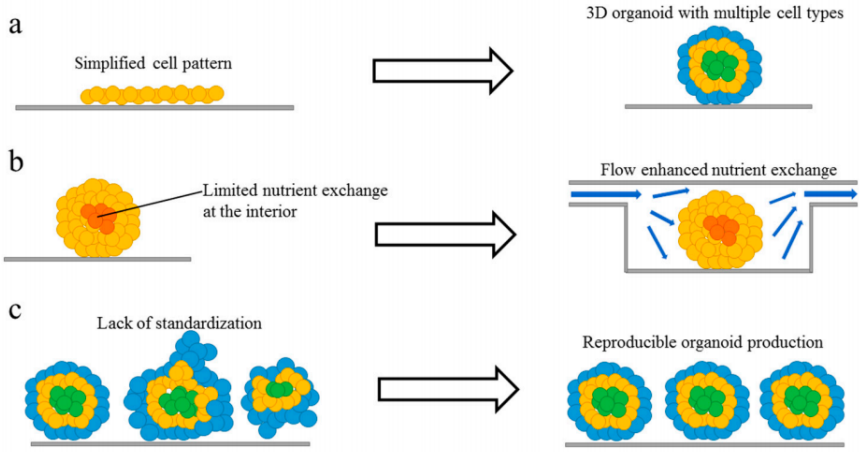


Figure 2.3: *Limitations and goal of current organoid models. (a) Traditional in-vitro models are too simplified; complex organoid models with multiple cell types and 3D architecture can be developed to better recapitulate in-vivo organs. (b) There is a lack of nutrient exchange at the interior of the organoid, introducing flow and improving nutrient and gas exchange will help to create larger and more mature organoids. (c) Current organoid technology has limited uniformity and reproducibility, with better geometrical confinement and environmental control; future organoids production will be more reproducible. Image from reference [89].*

fact, tissue micro-array[2] and organotypic tissue slice culture[3] are too simplified

[2]Tissue microarrays are paraffin blocks produced by extracting cylindrical tissue cores from different paraffin donor blocks and re-embedding these ones into a single recipient (microarray) block at defined array coordinates.

[3]Organotypic tissue culture is defined as the culture of an organ collected from an organism. It is a method allowing the culture of complex tissues or organs. It allows the preservation of the architecture of the cultured organ and most of its cellular interactions.

to represent complex 3D tissues with multiple cell types. This is due to the limited number of cell types and simplified environmental signals in the 2D models. However, in animal tissues and 3D cell culture models, cells and their surrounding microenvironment interact in all three dimensions and enable us to achieve a better understanding of the cellular behavior, both in-vitro and in-vivo. When cultured as 3D models, cells exhibit features that are highly similar to the complex in-vivo conditions and show significant improvements in terms of cell count, cell morphology, cell proliferation, and cell differentiation. *Spheroids*[4] are one of the most commonly used in-vitro 3D tissue culture models. They are formed by cell aggregation and are often used in long-term culture. Upon aggregating into spheroids, cells can establish contacts and create a microenvironment that allows the expression of tissue-like phenotypes. However, most spheroid culture models contain only one cell type and do not completely capture the complex intercellular interaction between different cell types.

Compared with spheroids, organoids develop from stem cells or organ progenitors and self-organizes in a manner similar to in-vivo, following the hierarchical structure of that tissue. Organoids have been developed not only for a large number of cancer types. Indeed, normal tissue stem cell-derived organoids are already in use for gut, salivary glands, mammary glands and liver [90]. However, many current models also do not take into account blood cells, or shear stress by blood flow, stroma, and immune cells. One of the biggest obstacles in growing mature organoids is the restricted nutrient supply, gas exchange and waste removal at the interior of the organoids (Fig. 2.3.b). Since the average diameter of organoids achieved in most studies is usually up to 3 mm, heterogeneity in biological condition is found within the culture model. Finally, the current organoid technology has limited uniformity and reproducibility (Figure 2.3c), making it difficult to be used for toxicity screening or high-throughput testing. This is due to inadequate engineering of the cellular microenvironment and the *extra-cellular matrix* (ECM). To use organoids in high-throughput, the organoids must be generated in a reproducible way.

In 3D organoid models, as organoids increase in size and volume, the core becomes distant from the surface that is in contact with the fresh medium. A simple diffusion process provides insufficient oxygen and nutrient to the growing cells and limits the amount of waste being removed from the cells in the core. Consequently, only cells in contact with fresh medium survive. With microfluidic technology, tissue culture can be carried out in a controlled environment that optimizes temperature, pH, nutrient and oxygen supply and waste removal. Advances in microfluidic technology allow us to engineer the organoids with essential structural and physiological features in a controlled manner and provide microscale structures and parameters that mimic the con-

[4]A multicellular tumor spheroid model was first described in the early 70s and obtained by culture of cancer cell lines under non-adherent conditions. Tumorospheres, is a model of cancer stem cell expansion; tissue-derived tumor spheres and organotypic multicellular spheroids are typically obtained by tumor tissue mechanical dissociation and cutting

Table 2.1: *Comparison of advantages and disadvantages of microfluidic chips and other in-vitro culture models.*

In Vitro Culture Models	Advantages	Disadvantages
2D cell culture (culture dish, transwell membrane and culture flask)	Well estabilished protocol Easy to handle and quantify	Static condition Lack of physical and biochemical cues Large media volume Large variation in nutrients and waste concentration
3D cell culture (engineered culture scaffold, spheroid, microcarrier, tissue biopsy, organoid)	Include cell-cell and cell-ECM interaction Capture the 3d architecture of tissue culture Sensitive to drug treatment	Static condition Inefficient nutrient and waste transport
Microfluidic chip (Organ-on-a-chip)	Fine control over microenvironment Good mass transport provided by fluid flow Ability to integrate with various sensors and actuators	Diffcult to standardize and scale up Require external pumps, tubing, connectors, and valve to operate

ditions in-vivo. Moreover, sensors and actuators can be integrated with the microfluidic devices to enable precise monitoring and control. Optimization of the key parameters, such as the cell-cell and cell-ECM contact, cell type composition, tissue architecture, nutrient exchange, and various physical and electrical stimulation, may greatly minimize batch-to-batch variations and increase fidelity. When accompanied by microfluidic technology, 3D cell cultures can be enhanced to become more complex organ-on-a-chip and organoid on a chip models. Microfluidic organs-on-a-chip platforms have been recently developed to create a variety of biomimetic organ models, such as lung, liver, kidney, heart and neural networks. Microfluidic platforms are often combined with flexible cell culture scaffolds such as hydrogels. They are commonly used to create an in-vivo-like microenvironment for cell culture. They provide functional support for the cells functioning like ECM while promoting survival, proliferation, and differentiation [89]. This is achieved by the natural properties of hydrogels, which have interconnected pores with high water retention that allows nutrients to be transported to the cells efficiently. ECMs provide structural and biochemical support to the cells and promote cell-ECM interaction and growth. Cell-ECM and cell-cell interaction can be tuned by changing the mechanical properties (for example stiffness) and compositions of each material in the scaffold. Microfluidic approaches have advantages over other in-vitro culture models since they offer a better control over the physical and chemical parameters, the design of complex structures and the use of multiple materials to better mimic the in-vivo organs. The advantages and disadvantages of each model are shown in Table 2.1. Finally, once the improvement in technologies and methodologies to recreate similar biological chips will be completed, the initiation of accessible organoid biobanks for research purposes could further advance our understanding of the role of cancer stem cells in therapeutic responses.

2.2.3 Ex-Vivo

Basically, the ex-vivo tissue analysis is a derivation of the already described methods. In fact, when organoids are made up of patient samples, usually coming from a tumor biopsy, we are essentially performing an ex-vivo analysis. The same is happening when, after animal sacrifice, a tissue is analysed by optical microscopy. Both cases are predominantly applied for tumor control studies; since the work presented in chapter 3,4,5 regards normal tissues, no further details will be provided on these ex-vivo methods.

2.3 Cellular Response

The successful use of ionizing radiation to treat cancer results primarily from its ability to cause the death of individual tumour cells. As discussed in Chapter 1, the biological consequences of irradiation, including cell death, are highly influenced by pathways within the *DNA damage response* (DDR) system. The DDR determines not only the sensitivity of cells to die following irradiation, but also the type of cell death that occurs, and the timing of cell death. Because the DDR differs among different types of normal and tumour cells (and probably even within different populations of tumour cells), the manifestation of cell death can also differ widely among different cell types.

Cell survival curve

The cell survival curve is the gold standard to describe the relationship between the radiation dose and the fraction of cells that survive. Reproductive death is generally the endpoint measured with cells cultured in-vitro. A cell may still be physically present and apparently intact, may be able to make proteins or synthesize DNA, and may even be able to struggle through one or two mitoses, but if it has lost the capacity to divide and produce a large number of daughter cells, it is by definition clonogenically dead: it has not survived. A surviving cell that has retained its reproductive integrity and is able to proliferate and produce a large colony (a group of 50 or more cells that have developed from a single cell) is said to be *clonogenic* [16].

The capability of a single cell to produce a large colony that can be seen easily with the naked eye is a convenient proof that it has retained its reproductive integrity. This ability as a function of radiation dose is described by the survival curve. Cells from an actively growing stock culture are prepared into a suspension by the use of trypsin, which causes the cells to round up and detach from the surface of the culture vessel. The number of cells per unit volume of this suspension is counted with an electronic counter. In this way cells may be seeded into a dish; if this dish is incubated for 1 to 2 weeks, each single cell divides many times and forms a colony that can be visible with the naked eye, especially if it is fixed and stained. All the cells in a single colony

are the progeny of a single ancestor. Ideally, the number of colonies should be the same of the number of seeded cells (100%); in reality, the counted colonies may be expected to be in the range of 50% to 90%. Possible explanations for this are: use of sub-optimal growth medium, errors and uncertainties in counting the cell suspension, and the "trauma" of trypsinization and handling. The term plating efficiency indicates the percentage of seeded cells that grow into colonies. The plating efficiency is given by the formula:

$$PE = \frac{Number\,of\,colonies\,counted}{Number\,of\,colonies\,seeded} \times 100 \tag{2.1}$$

In the case of a dish exposed to a certain radiation dose the fraction of surviving cells (SF) is:

$$SF = \frac{Colonies\,counted}{Cells\,seeded \times (PE/100)} \tag{2.2}$$

This process is repeated so that estimates of survival are obtained for a range of doses. The number of cells seeded per dish is adjusted with doses so that a countable number of colonies results: too few cells increase the statistical error; too many cells cannot be counted accurately because they tend to merge one into the other. It is worthful noting that this technique, and the survival curve that results, does not distinguish the modality of cell death, that is, whether the cells died by mitotic or apoptotic death or any other form of interphase death.

Survival curves for mammalian cells are usually presented in the form shown in Figure 2.5, with the dose plotted on a linear scale and the cell surviving fraction on a logarithmic scale. Qualitatively, the shape of the survival curve can be described in relatively simple terms. For sparsely ionizing (low-LET) radiation, such as X-rays, at low doses the survival curve starts out straight on the log-linear plot, with a finite initial slope; that is, the surviving fraction is an exponential function of dose. At higher doses, the curve bends. This bending or curving region extends over a dose range of a few grays. By contrast, for densely ionizing (high-LET) radiations, such as α particles or low-energy neutrons, the cell survival curve is a straight line from the origin; that is, survival approximates to an exponential function of dose. Many biophysical models and theories have been proposed to account for the shape of the mammalian cell survival curve. The linear-quadratic model has taken over as the model of choice to describe survival curves. Although we can regard this as based on pure mathematics (i.e. the simplest formula which describes a curve), it has also been possible to attach radiobiological mechanisms to this model. Indeed, the model can be related to the relation used to describe exchange-type chromosome aberrations that are clearly the result of an interaction between two separate breaks. According to the linear-quadratic model (illustrated in the small square of Figure 2.5) there are two components of cell killing by radiation, one that is proportional to dose and one that is proportional to the

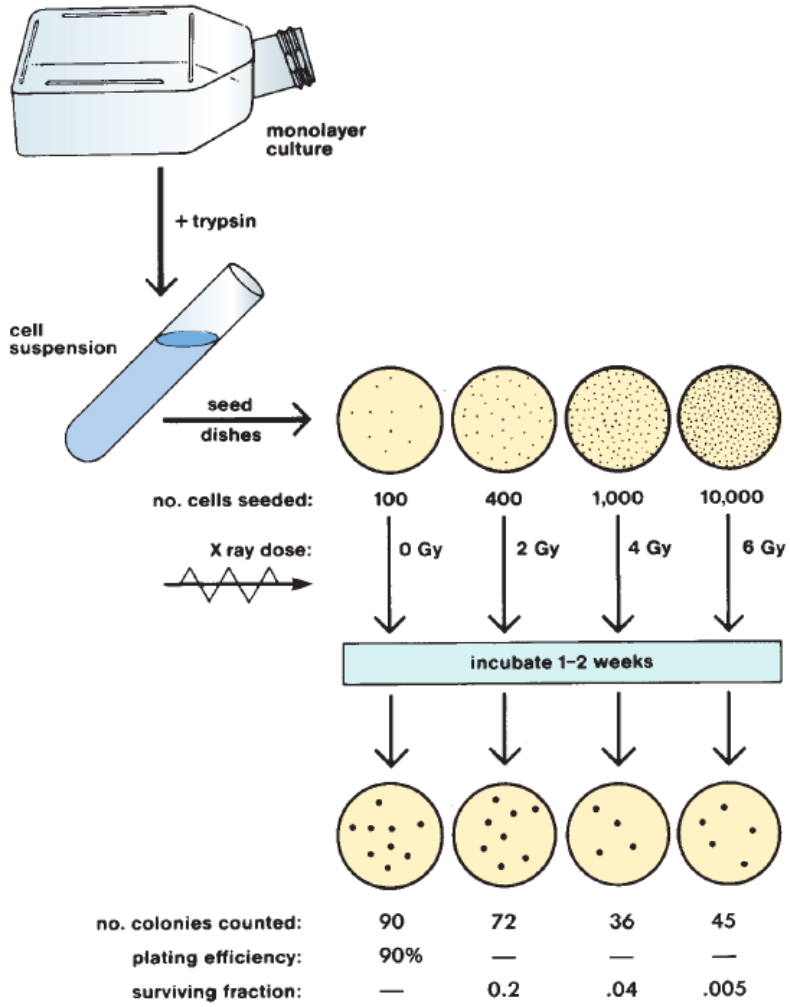


Figure 2.4: *The cell culture technique used to generate a cell survival curve. Cells from a stock culture are prepared into a single-cell suspension by trypsinization, and the cell concentration is counted. Known numbers of cells are inoculated into petri dishes and irradiated. They are then allowed to grow until the surviving cells produce macroscopic colonies that can be counted readily. The number of cells per dish initially inoculated varies with the dose so that the number of colonies surviving is in the range that can be counted conveniently. Surviving fraction is the ratio of colonies produced to cells plated, with a correction necessary for plating efficiency (i.e., for the fact that not all cells plated grow into colonies, even in the absence of radiation). Image from reference [16]*

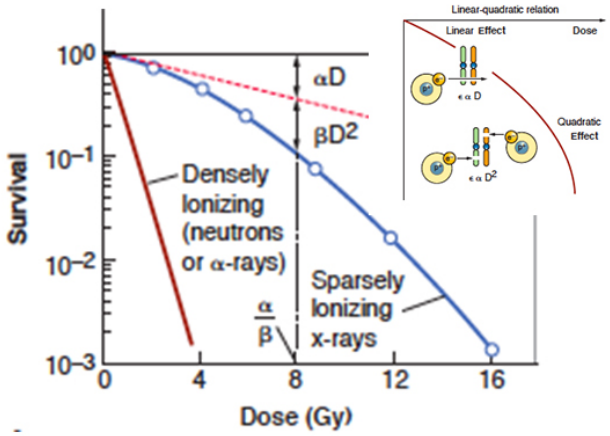


Figure 2.5: *Shape of survival curve for mammalian cells exposed to radiation. The fraction of cells surviving is plotted on a logarithmic scale against dose on a linear scale. For densely ionizing particles, the dose-response curve is a straight line from the origin. The survival curve can be described by just one parameter, the slope. For sparsely ionizing radiation, the dose-response curve has an initial linear slope, followed by a shoulder (at higher doses, the curve tends to become straight again). The experimental data are fitted to a linear-quadratic function. On the small square: relationship between chromosome aberrations and cell survival. Images from reference [16]*

square of dose[5]. By this model, the expression for the cell survival curve is:

$$S(D) = exp(-\alpha D - \beta D^2)$$

in which S is the fraction of cells surviving a dose D, and α and β are parameters depending on radiation quality and cell type. The shape (or bendiness) is determined by the ratio α/β. Since the dimensions of the parameters for α are $Gy^{-}1$ and for β are $Gy^{-}2$, the dimensions of α/β are Gy; this is the dose at which the linear contribution to damage (αD on the logarithmic scale) equals the quadratic contribution (βD^2). The mechanistic idea behind this formula is that the linear component might result from single-track events while the quadratic component might arise from two-track events (see again Figure 2.5). This interpretation is supported by studies of the dose-rate effect, which shows that as dose rate is reduced cell survival curves become straight and tend to extrapolate the initial slope of the high dose-rate curve: the quadratic component of cell killing disappears, leaving only the linear component [23]. This would be expected, since at low dose rate single-track events will occur far apart in time and the probability of interaction between them will be low.

Clonogenic survival is a useful endpoint for measuring the response to ionising radiation for three reasons. Firstly, the regenerative capacity of cells is thought to be the most important factor in determining the response both of renewing normal tissue and of tumour cells. Secondly, clonogenic assays are usually highly reproducible, suitable for many in-vitro cell lines and have been related to gross tumour response in some experimental systems. Finally, the analysis and interpretation of the shape and composition of clonogenic survival curves have provided suggestions as to the underlying mechanisms of cellular response to ionising irradiation. In fact, from the comparison of cell survival curves under different conditions (radiation sources, energy, hypoxia, cell lines, fractionations of the dose, donors, cell phase), it has been possible to understand the role of LET, RBE and OER (including their relationship), the sensitivity of the cell lines (above all for tumor lines), the benefit for the OAR of separating the dose in different fractions, the critical topic of genetic mutations that can affect the patient radiosensitivity (Ataxia-telangiectasia mutation or Fanconi's anemia) and also the different radiosensitivity of the cell phases.

Finally, it is worth mentioning also the presence of non-clonogenic assays. They provide a result in a shorter time and they can be used for cells which do not grow as colonies. Some of these tests are growth assays, which determine the number of viable cells at various times following irradiation in untreated or treated cell cultures. One growth assay that has been optimized for determining the radiosensitivity of human tumor cells is the cell-adhesive matrix

[5]The notion of a component that varies with the square of the dose introduces the concept of dual radiation action [91]. This idea was later interpreted based on the early work with chromosomes in which many chromosome aberrations are clearly the result of two separate breaks.

assay in which primary human tumor cells are plated directly onto culture dishes coated with a combination of cell-adhesive proteins. Overall, the results usually do not correlate easily with results of the clonogenic assay, which remains the gold standard for determining the radiosensitivity of cells in-vitro. The reason for this incongruence is due to the different endpoint: with non-clonogenic assays we are not considering anymore the reproductive integrity of the cells, but only their growth. Nevertheless, quantitative information on the survival fraction can be derived by the cell number as a function of time after irradiation [92].

Apoptosis as a death endpoint

Apoptosis is a highly regulated form of cell death that can be initiated either as a result of conditions occurring within the cell itself (such as those after DNA damage) or from signals generated externally such as those from a surrounding tissue or immune cell [93]. Apoptosis is an essential and normal part of many physiological processes including embryonic development, the immune system, and maintenance of tissue homeostasis. Consequently, alterations in the control of apoptosis contribute to several human diseases, including cancer. Apoptosis is both morphologically and molecularly distinct from other forms of cell death (mitotic catastrophe, necrosis, autophagy, senescense). Morphologically, it is characterized by membrane blebbing, condensation, and digestion of the DNA into small fragments. During this process, cellular contents are also fragmented into many membrane-enclosed apoptotic bodies, which, in-vivo, are taken up by phagocytes. This prevents leakage of potentially damaging cellular proteins and destruction of tissue architecture that is a familiar feature of necrosis. The molecular participants in the apoptotic pathway can be divided into two groups: sensors and effectors. The sensor molecules are involved in making the decision to initiate apoptosis whereas the effectors are responsible for carrying out that decision. Apoptotic cell death is characterized by the sequential activation of several different enzymes known as *caspases,vander3,radxrad*. Apoptosis begins following the activation of a "sensor" caspase such as caspase 8 or 9, which generates the initial signal to induce apoptosis. These caspases subsequently activate a common set of other "effector" caspases (e.g. caspase 3), which then cleave a large set of cellular proteins leading to the ultimate destruction of the cell. Apoptosis that initiates from caspase 8 activation is termed the "extrinsic" pathway because it is normally activated upon the binding of an extracellular ligand and subsequent activation of a death receptor present in the cellular membrane. This extrinsic pathway of apoptosis is not induced by radiation to any significant degree, but is a candidate target for combining novel drugs with radiation. Apoptosis that initiates from caspase 9 is termed the "intrinsic" pathway because it is activated within the cell in response to various forms of cell damage. The activation of caspase 9 is controlled in large part by the balance of pro- and anti-apoptotic proteins that reside in or near the mitochondria.Under normal conditions this

balance is in favour of the anti-apoptotic factors, and activation of caspase 9 is prevented. Conditions that alter this balance lead to activation of caspase 9. After irradiation, this balance can be tipped in favour of apoptosis owing [94]. Apoptotic assays can be performed by measuring the caspase activation in-vitro. However, apoptosis is highly dependent on the balance of the pro and antiapoptotic proteins and this balance varies widely among different cell types and tumours. This explains why irradiation causes apoptosis only in certain normal tissues, while in tumours, an additional mechanism for variation in apoptosis sensitivity arises from the fact that many of the genes that regulate apoptosis are frequently altered in cancer.

DNA damage analysis

As more is learnt about the sub-cellular processes that can lead to cell death following irradiation, new end-points can be used to assess radiosensitivity. Perhaps the "best" correlations between cell death and DNA damage are at the level of chromosomal damage. Detailed analysis of radiation-induced chromosomal aberrations can relate very accurately to cell death in many cells. As a possible predictive clinical test of radiosensitivity this has the advantage that it requires limited cell proliferation and relatively few cells. Micronucleus and premature chromosome condensation assays detect chromosomal fragments and, at least in normal cells, this correlates well with cell death. The use of DNA damage assays, specifically those which measure DNA double strand breaks (DSB), has also been explored for their potential as measures of cellular radiosensitivity. DSBs are detected by specialized proteins which signal to the cell that damage has occurred, thereby initiating the DNA damage response. This response effectively focuses the cell's attention on the damage, stopping other processes such as transcription and cell-cycle progression, and, importantly, initiating repair. *DNA damage-induced nuclear foci* (radiation-induced foci assay) in response to ionizing radiation represents an important assays for detecting DSBs. There are several advantages of assaying for foci formation over other techniques, which include the ease of the protocol and that it can be carried out on both tissue sections and individual cell preparations [16]. Technically, cells/tissues are incubated with a specific antibody raised to the signaling/repair protein of interest, and binding of the antibody is then detected with a secondary antibody, which also carries a fluorescent tag. Fluorescence microscopy detects the location and intensity of the tag, which can then be quantified. The most commonly assayed proteins for foci formation are $\gamma-$H2AX. H2AX is a histone protein, which is rapidly phosphorylated in response to damage to form $\gamma-$H2AX. Staining for the unmodified histone (H2AX) gives a pan nuclear stain or unchanging band on a western blot while $\gamma-$H2AX is rapidly induced on a western blot in response to stress and can be seen to form discreet nuclear foci in damaged cells (see Figure 2.6 (b)). DNA damage-induced increases in $\gamma-$H2AX can also be quantified by flow cytometry. If this value is measured over time, then it also reflects the

kinetics of repair, the number of foci decreases when the DSB are repaired (see Figure 2.6 (b)).

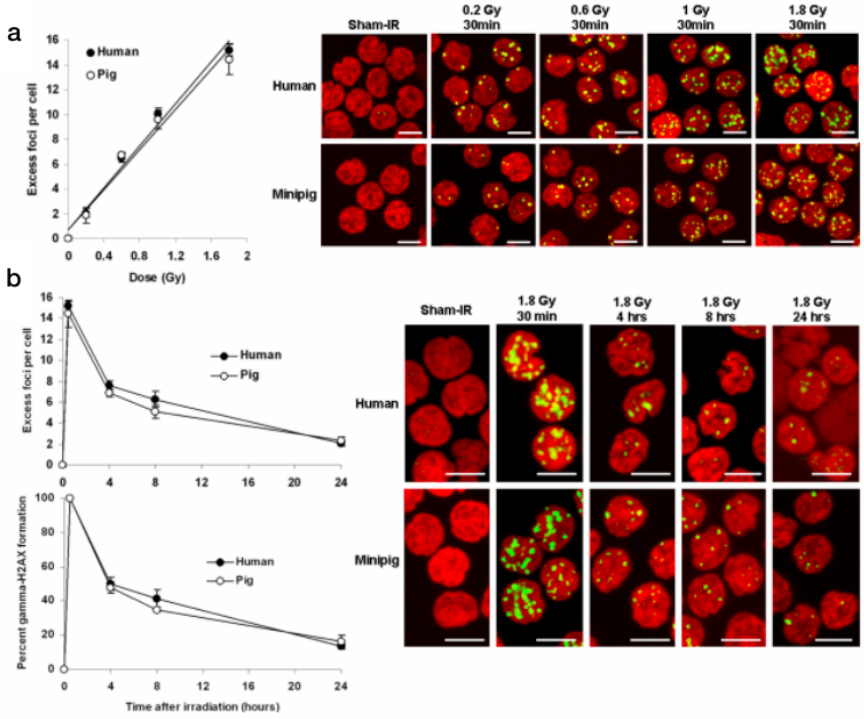


Figure 2.6: *Comparison of $\gamma - H2AX$ foci kinetics in human and minipig lymphocytes after ex-vivo irradiation of blood samples. (A) Human and minipig blood was irradiated with 0.2, 0.6, 1 and 1.8 Gy and excess of $\gamma - H2AX$ was measured at 30 minutes. (B) Human and minipig blood samples were irradiated with 1.8 Gy and incubated at $37°C$ for various lengths of time (kinetics measurement). Image from reference [95]*

Monte Carlo simulations of DNA damage

Cellular responses to radiation damage have been studied for many decades, showing the dependency of DNA damage on the delivered dose, the delivery timeframe and the radiation particle type and energy. Numerous models have been developed to explain these responses across a range of endpoints, including DNA damage, mutations, micronuclei formation, chromosome aberrations and cell survival. Many of these are phenomenological macroscopic models, and simply relate cellular end points to the delivered dose and empirical parameters expressing cell sensitivity, which can depend on the cell line, irradiation conditions and radiation quality. Such phenomenological approaches can capture the overall population-based trends in cell survival that are necessary to describe the effects of radiation therapy, or to estimate effects of exposure to environmental or space radiation.

Efforts to model cell response have focused on damages to the nuclear DNA, which has long been established as the primary radiation target determining

cell viability. The response of cells to radiation has been shown to correlate with the pattern of energy depositions within the nucleus; this correlation is attributed to the resulting differences in patterns and types of DNA damage. Several decades ago, the first studies using Monte Carlo simulations were performed to link the track structure of different radiation modalities with DNA geometries and the probability of damage induction [96]. These studies represent the first attempts to apply track-structure Monte Carlo simulations, to mechanistically understand how radiation energy depositions lead to DNA damage. In an ideal scenario, one would use track-structure simulations of the incident radiation to simulate the physical interactions as a means of obtaining nanometer-scale energy depositions and ionizations in accurate geometric models of the cells and their sub-components (nucleus and DNA). After the physical interactions, the resulting radiolysis products and other ionized molecules react in a physicochemical stage, which is followed by migration of the chemical species. At this stage chemical species can react with each other, be scavenged inside the cells or react with components of the cell, such as the DNA. The simulation finishes by determining the direct (caused by physical interactions) and indirect (caused by chemical reactions) damages to DNA.

Finally, the DNA damage patterns can be used in mechanistic models of DNA repair kinetics to calculate cell viability, accounting for the damage complexity, along with properties of the cell and the surrounding environment, such as repair deficiencies, cell cycle and oxygenation. In recent years, several major developments have led to a surge in attempts to mechanistically describe DNA damage and repair kinetics. An increase in the computational power of standard computers has enabled the simulation of particle tracks in DNA fragments and even whole nuclei. This has been accompanied by improvements in imaging techniques for studying the responses of cells to ionizing radiation, providing an abundance of data showing the importance of repair pathways and their effect on cell viability. Currently, several Monte Carlo simulation codes exist that can provide the nanoscale track structure of particles passing through a medium, which is typically simulated as water but more recently also includes DNA nucleotide material. Tens of codes are used by various groups to simulate the track structure of different types of radiation and then score the resulting initial damages to a cell nucleus.

Thus, these Monte Carlo codes can provide estimates of DNA damages induced both directly (from the initial particle track) and indirectly (from chemical reactions). To fully elucidate the effect of DNA damage induction and repair on cell survival, chromosome aberrations, mutations or other endpoints of interest, the simulated patterns of damage along the DNA strands, as well as their complexity, must then be combined with models that describe the mechanisms of DNA repair [97].

Chapter 3

Results at the Patient Level: Analysis of Different Normal Tissue Complication Probability Models for Prostate Cancer

The previous chapters of the thesis presented an overview of selected radiobiological aspects as a brief introduction of the topics needed to understand the research issues explored during the experiments and studies performed at the National Cancer Institute of Milan (INT) and at the Department of Physics of the University of Pavia. The analysis performed during the three years of PhD basically followed the order they are presented in the thesis, which is the same as the historical path of radiobiological progress about the effects of radiation on normal tissues.

As a general rule, the clinical characteristics of the presented studies are reported in the Appendix; only the information needed to interpret the results will be provided in the main body of the thesis. Moreover, the studies that will be mentioned dealt with the development of late side effects after radiotherapy in *prostate cancer patients* (PCP). Particularly, the toxicity domains for PCP treated with radical or salvage radiotherapy are GI and GU toxicity, and *sexual dysfunctions* (SD). Currently, this last domain is dense of confounding factors which make quantitative studies still complex to perform; for this reason it will not be discussed in the following chapters [1]. Indeed, all the works are on the pelvic district and its OARs.TCP models were not investigated in the thesis project, however, a section with results of some important TCP studies was included in the Appendix.

For a general understanding, it could be helpful to show a representative treatment plan of a patient who undergoes radical RT to the prostate (with

[1] Debates on which are the involved organs are still ongoing. Moreover, hormon-therapy and aging would strongly affect a possible analysis.

the inclusion of the seminal vescicles but without the lymph-nodes irradiation). Studies presented in the continuation of the thesis will involve the structures that are contoured in Figure 3.1. A 3D reconstruction of the same structures is shown in the Appendix of the thesis.

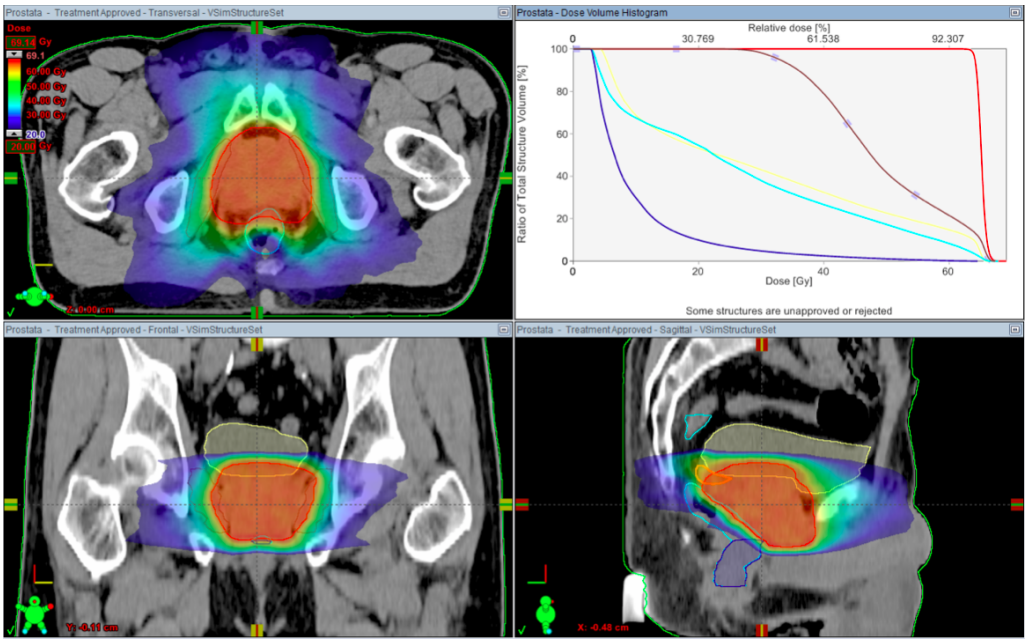


Figure 3.1: *Contours for the bladder are in yellow, the anorectum in cyan, the anal canal in purple and the obturator muscles in bordeaux. In the upper right square a representation of the DVHs is shown. Patients was treated to the prostate and the seminal vescicles with Volumetric Modulated Arc Therapy (VMAT). Total dose was 65 Gy in 25 fraction from 2.6 Gy. Dose distribution represents doses higher than 20 Gy.*

3.1 Mathematics in NTCP Modelling

In this section a mathematical description of the methods that have been used in the PhD work is presented. From in-vitro cell survival experiments exhibiting exponential dependence of survival with dose, a sigmoid-shaped relationship between dose and NTCP can be directly derived [98]. The Lyman model [99] is the most commonly used NTCP model, where the cumulative distribution function of a Gaussian distribution (a Probit function) is chosen to represent the empirical sigmoid dependence of NTCP on dose. In this model, two parameters, D_{50} and m, can be adjusted to change the position and slope of the NTCP curve, respectively.

$$NTCP(D, D_{50}, m) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} exp\left(\frac{-u^2}{2}\right) du \tag{3.1}$$

where

$$t = \frac{D - D_{50}}{mD_{50}}, \tag{3.2}$$

D_{50} is the dose related to 50% toxicity probability, if the whole organ is irradiated uniformly, and m is a parameter controlling the slope of the curve. D_{50} can also be expressed as a function of uniform partial organ irradiation, using a third parameter n which describes the magnitude of volume effect:

$$D_{50}(V) = D_{50}(1)/V^n, \tag{3.3}$$

$D_{50}(1)$ is the D_{50} for the whole irradiation volume. Small values of n correspond to serial volume effects, while large ones correspond to parallel volume effects. As we have already seen in the previous chapter, but also in Figure 3.1, the organ is irradiated with a inhomogeneous dose distribution represented by the use of DVH. To manage all the information included into this statistic of the dose distribution, and also to solve the issue of Fig 1.7 a, the Kutcher-Burman DVH reduction to an effective partial volume (V_{eff}) irradiated at a reference dose (often the maximum) was introduced [100][2]. The model that couples DVH reduction to Lyman model is the Lyman-Kutcher-Burman (LKB) model:

$$V_{eff} = \sum_i v_i \left(\frac{D_i}{D_{max}} \right)^{1/n} \tag{3.4}$$

where: (v_i, D_i), i from 0 Gy to maximum dose are the points of the differential DVH, D_{max} is the maximum dose to the organ, n is the previously cited volume effect parameter. As another approach to reduce a non-uniform DVH to a single parameter, we already mentioned the Equivalent Uniform Dose. EUD is described by the following equation:

$$EUD = \sum_i \left(v_i D_i^n \right)^{1/n} \tag{3.5}$$

When $n \to 0$ EUD tends to the maximum dose, while for $n = 1$ the EUD is equal to the mean dose. Here the DVH is reduced to an effective homogeneous dose (EUD) to the total volume ($DVH \to (D = EUD, v = 1)$). Of note EUD and V_eff are related by $EUD = (V_{eff})^n \cdot D_{max}$. For this reason the Lyman-EUD model (LEUD) can be used as alternative to the LKB model. The LEUD formulation had also a large approval, since it involved doses instead of volumes. This is an advantage because in some analyses the best dosimetric predictor is the EUD with n=1 or 0, which are the mean dose and the maximum dose, and no further calculations are needed to perform a multivariate model

[2]This formulation helps to discriminate which is the less damaging DVH according to the organ sensitivity and organization. However, being a "computational manipulation" of the DVH, it is still affected by the absence of any spatial information (which was already lost within DVH calculation) which could have a great importance in the case of sub-structures with different radiosensitivity are present within the organ.

or to validate an already published model. Even more important is that dose values are much more of immediate understanding for clinicians, and they are also more helpful in treatment plan optimization.
The Logit+EUD model (LOGEUD) [101] was then proposed to fit the toxicity data, using again the EUD as DVH reduction. The logit formula, with log-transformation of the dose variable, describes the dose-response relationship through D_{50} and k[3] (the slope of the curve at D_{50}, which is a surrogate of the m parameter in the LKB model):

$$NTCP(EUD, EUD_{50}, k) = \frac{1}{1 + \left(\frac{EUD_{50}}{EUD}\right)^{k}} \tag{3.6}$$

The advantage of this formulation is that, on the basis of a developed toxicity model, physicians have the possibility to easily compute with a calculator the risk probability for a patient after treatment planning.
Dose modifying factors (DMF), introduced by Peeters et al [81] and used in many works [101, 102, 103, 104] incorporate variables other than dose into LKB models. In the modified LKB model, the D50s for patients with and without the clincial condition are fitted. The ratio of these D50s gives us the DMF, which is a measure for the horizontal shift of the dose-response curve when comparing patients with and without the predisposing feature. According to this definition, the LKB (but similarly the logit-EUD) formulation is modified in the following way:

$$NTCP\,(EUD, DMFs, D_{50}, m) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} exp\left(\frac{-u^2}{2}\right) du \tag{3.7}$$

and

$$t = \frac{EUD - D_{50}}{m(DMF_1 ... DMF_k \times D_{50})} \tag{3.8}$$

where the DMFs reflect the impact of covariates other than dose (e.g., SNPs genotype, copy number variations, smoking status, etc.) on D50 (which increses or decreases accordingly).

In principle, analytical models can appear to be scientifically sound. However, considering the complexity involved in radiotherapy response, mathematically formulating such a process may not be as accurate or complete as intended. As radiotherapy outcomes are determined by complex interactions among patient-specific anatomic and biological and treatment conditions, analytical models which consider limited numbers of variables and sometimes depend on tuning parameters by hand, may not be able to provide a complete or accurate prediction of NTCP. Other methods can be applied to analyze the clinical endpoints. Whatever method is used for toxicity modeling, the inputs

[3]To compare the parameters of the steepness of the sigmoidal function a simple association between m and k can be dervied, $k = 1.6/m$

are variables related to toxicity while the output is the predicted toxicity status. Mathematically, this can be formulated as follows: $f(x, w) : x \rightarrow Y$, where $x \epsilon dn$, is an input variable vector of N-dimension, composed of the input metrics such as dosimetric, clinical and biological variables. Label Y is a scalar representing toxicity status and w denotes the parameters to be optimized in the model. The optimal parameters w^* of model $f(w, x)$ are obtained by optimizing a certain objective function given observed training data. For instance, one can estimate optimal w^* by minimizing the least-squared differences between the model predictions and the observed outcomes or maximizing likelihood function that gave rise to the observed data.
Among these methods it is worthful to mention the regression models, which are largely diffusing also into the clinical research. Logistic regression is possibly the most spread regression technique to fit binomial data. It is a specific type of *generalized linear models* (GLMs) [105], with binomial random component and logit link function. Compared to other GLMs, logistic regression is a more appropriate model for NTCP, as radiation outcomes have been observed to follow an S-shaped (sigmoidal) curve. In logistic regression:

$$f(x_i) = sigmoid(g(x_i)) = \frac{e^{g(x_i)}}{1 + e^{g(x_i)}} \tag{3.9}$$

$g(\cdot)$ is a weighted sum of entries in the input vector, which can be written as:

$$g(x_i) = \beta_0 + \sum_{j=1}^{s} \beta_j x_{ij}, \;\; i = 1, ..., n \;\; j = 1, ..., s \tag{3.10}$$

where n is the number of samples (patients), x_i is the input vector for i_{th} patient, s is the dimension of the input variables vector. $\beta = (\beta_0, \beta_i ....\beta_s)$ are parameters to be optimized by minimizing a cross-entropy loss function, which is defined by taking the negative logarithm of the likelihood:

$$E(w) = \ln p(f|\beta) = -\ln \prod_{i=1}^{n} f_i^{y_i}(1 - f_i)^{1-y_i} = -\sum_{i=1}^{n} [y_i \ln f_i + (1 - y_i) \ln(1 - f_i)] \tag{3.11}$$

Sometimes, to account for the interaction effects between variables (cross-talks)[106], extra terms can be added to the expression of $g(\cdot)$:

$$g(x_i) = \beta_0 \; + \; \sum_{j=1}^{s} \beta_j x_{ij} + \sum_{m=1} \sum_{n=1}^{s} \gamma_{mn} x_{im} x_{in} \tag{3.12}$$

Recently, data-driven models using statistical and machine learning methods, which allow the incorporation of more information into the model building process, have gained popularity in NTCP modeling. In the context of machine learning, the prediction of NTCP can be viewed as a supervised learning problem. The goal of supervised learning is to infer a function that maps inputs to outputs from a labeled training dataset. In supervised learning, each example

is a pair consisting of input object (feature) and output object (label). The learned function can be applied to predict the label of some unseen (out-of-sample) data in the future.
Here, we mention artificial neural networks (ANNs) since this approach has been used in one of the following studies. It is a network which is feed-forward and fully-connected. This method has witnessed renovated interest in recent years with the advent of deep learning methods and their popularity, particularly in computer vision applications. An ANN consists of several layers and neurons, where every neuron in the following layer is connected to all the neurons in its former layer. The connection is unidirectional and no circles exist in the network architecture (see Figure 3.2 (a)). Input nodes, or neurons, are a selection of features from the dataset of the investigated cohort. In a study of dose-response of tissues we have at least one continuous variable which is the dosimetric parameter. Other nodes can be continuous or dichotomous features coming from usual areas of interest such as genetic information, texture features from imaging, comorbidities or biomarkers. Differently from the previously described apporaches, a neural network applies a series of functions to the data. The exact functions will depend on the neural network you are using and on the problem you are trying to solve (regression, classification). Thus, the value of a neuron in the hidden layer and output layer is calculated by taking a weighted sum of all the neurons in its former layer followed by (in the majority of the cases) a non-linear activation function as shown in Figure 3.2 (b) and is given by:

$$a = g\left(\sum_i x_i w_i + b\right) \tag{3.13}$$

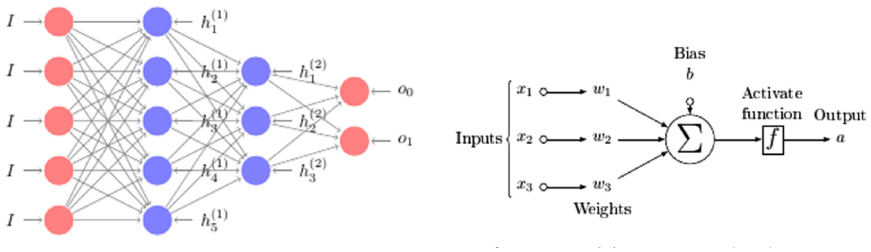


Figure 3.2: *Diagram of multi-layer neural network and of its activation function. Images from [107]*

Some common activation functions are sigmoid $g(t) = \frac{1}{1+e^{-t}}$, ReLU $g(t) = max(0,t)$ and softmax $g_i(t) = \frac{e_i^t}{\sum_j = e_j^t}$. The softmax activation function scales the value of a neuron by a sum of all the neurons in the same layer. It guarantees the sum to be 1, which is appropriate to be implemented in the output layer for solving classification problems. The units in hidden layers can be viewed as learning complex features from data that allow the output layer to be able to better discern one class from another, to generate more accurate decision boundaries. A famous example is the face recognition; units in the first

layers learn edge like features (detect edges at given orientations and positions) and higher layer learn to combine those to become detectors for facial features like the nose, mouth or eyes. The weights of each hidden unit represent those features, and its output (assuming it is a sigmoid) represents the probability that that feature is present in your sample. There are two phases in the training of ANN. One is forward-propagation, which, starting from inputs, obtains the values of all the neurons in hidden layers and in the output layer. The other one is backward propagation which propagates the error signal, computed at the output, all the way back to the inputs, updating weights value. The basic training algorithm of ANN is *stochastic gradients descent* (SGD). However, as there are numerous local minima, advanced SGD algorithms (e.g. Adam [108], which considers momentum and adaptive steps to avoid such pitfalls) yield better prediction performances. Sometimes, to avoid overfitting, the dropout technique [109] is applied in ANN. Dropout is a technique that randomly turns off a certain ratio of neurons during forward propagation such that they do not contribute to the activation of downstream neurons and their weights get temporarily disconnected in subsequent updates. This process has been shown to make ANN more generalizable and mitigates overfitting issues.

Mathematical tools for validation assessment

As prediction models are developed to be applied in new individuals, their value depends on their performance outside the development sample. External validation studies may range from temporal, geographical or clinical settings.
To assess the developed model's performance in the validation sample, *calibration* and *discrimination* are quantified. Calibration reflects the agreement between outcome predictions from the model and the observed outcomes. It is preferably reported graphically with predicted outcome probabilities (on the x-axis) plotted against observed outcome frequencies (on the y-axis). This plot displays the direction and magnitude of model miscalibration across the probability range, which can be combined with estimates of the calibration slope and intercept. A well calibrated model shows predictions lying on the 45° line of the calibration plot with calibration that shows a slope of 1 and intercept of 0. More specifically, the intercept a in the equation $logit(y) = a + b \cdot logit(\hat{y})$ assumes the statistical meaning of *calibration in the large*. It quantifies whether or not the average of predictions corresponds with the average outcome frequency. Values below (or above) this value indicate that the model overestimates (or underestimates, respectively) the outcome. The calibration-in-the-large is often optimal in the development sample of the predictive model. Consequently, it is a useful statistics for identifying whether unexplained differences exist in the outcome frequency of the validation sample. On the other hand, *calibration slope* , denoted as b reflects whether predicted risks are appropriately scaled with respect to each other over the entire range of predicted probabilities ($b = 1$). Typically, $b > 1$ occurs when predicted probabilities do not vary enough, and $0 < b < 1$ occurs when they vary too much. Finally,

it is common to apply statistical tests for the agreement between predicted and observed probabilities using tests such as the Hosmer-Lemeshow test. It has to be underlined that all of these tests have a limited statistical power to evaluate poor calibration and are sensitive to the grouping and sample size, they are often non significant for small size and always significant for large numbers. Furthermore, they do not give indication of magnitude or direction of any miscalibration, hence the preference for calibration plots.
Discrimination refers to the ability of a prediction model to differentiate between those that do or do not experience the outcome event. It is commonly estimated by the so called *concordance index* (c-index). It reflects the probability that for any randomly selected pair of individuals, one with and one without the outcome, the model assigns a higher probability to the individual with the outcome. The c-index is identical to the area under the receiver-operating characteristic curve (AUC) for models with binary endpoints.
Overall performance measures, e.g. R^2 or Brier-Score (a score function that measures the accuracy of probabilistic predictions), are sometimes reported in addition to the traditional measures of discrimination and calibration, although they are less intuitive. For some studies, classification measures such as predictive values, sensitivity and specificity, are computed as performance measures after introducing a probability threshold to define positive/negative predicted outcomes.

However, the general purpose of the studies presented in Chapter 3 was to validate/define dose-response models, i.e. mathematical relationships between dose and rates of observed toxicities. Selection of cutoffs for possible clinical trials was not the aim of this thesis. For this reason, we chose to stick to calibration plot within this chapter.

3.2 Development and Validation of Models for Gastrointestinal Toxicity

In this section three studies performed within this PhD project are presented. All of them are focused on GI toxicity after prostate cancer irradiation.
The first work regards the development and validation of an NTCP model to describe rectal stool frequency and rectal pain, two endpoints that are not so common and need a large population to be investigated (Airopros 0102 + TROG03.04 RADAR).
The work showed in the second part is based on the validation of the NTCP model for late rectal bleeding (LRB) developed in European and American institutes and published in the literature during the last 15 years; clinical factors were also considered in some of these predictive models. Performance was tested on a population of more than 1700 patients (Airopros 0102 + TROG03.04 RADAR + DUE-01).
The third study was still related to the validation process, but on a different

endpoint, which was the fecal incontinence. It is the most investigated rectal side effect after late rectal bleeding, and it was very interesting for us to explore this endpoint on a population treated with new tecnologies (e.g. VMAT, IGRT and localization systems) that scored a higher toxicity occurence compared with the development cohort.

3.2.1 Modelling late stool frequency and rectal pain after radical radiotherapy in prostate cancer patients: results from a large pooled population

The majority of studies devoted to gastrointestinal side effects after radiotherapy for prostate cancer are focused on rectal bleeding, faecal incontinence and overall acute gastrointestinal toxicity. Nevertheless, chronic radio-induced rectal syndrome includes other symptoms, such as urgency, increased stool frequency and rectal pain. There is insufficient knowledge on the incidence of these morbidities and on their relationship with the dose distribution in the rectum and in the anal canal. This is mainly due to their being relatively rare effects; there is difficulty in identifying radiation as the cause of these impairments in an ageing population together with the lack of controlled questionnaire-based prospective scoring describing the pre-radiotherapy baseline situation. However, these symptoms may clearly have a non-negligible impact on the QoL of long-surviving patients and consequently deserve attention. The two considered cohorts were treated at different hospitals, with different dose levels, with different radiotherapy techniques, in different countries and in different time frames. As a consequence, the pooled population presented a wide variety of dosimetric and clinical parameters, with the potential to reach sufficient statistical power to assess the main associations between the selected rare side effects and clinical/dosimetric features.
Development population was created from two high quality multicentre prospective trials (Airopros 0102 and TROG03.04 RADAR) on radiotherapy for prostate cancer. It included 1336 patients, 3-year minimum follow-up, treated with 66-80 Gy (2 Gy/fr, i.e., conventional regimen) in Italian and Oceanian hospitals. Further details about the cohorts can be found in the Appendix. Toxicity was scored with LENT-SOMA-scale. Two toxicity endpoints were considered: grade$\geqslant$ 2 rectal pain (peak toxicity) and mean grade (average score during 3 years of follow-up) in stool frequency>1. This last endpoint was arbitrarly considered as the clinically relevant endpoint as it selects those patients with persitent symptoms (i.e. patients who on average evacuated $\geqslant$ 3 times/day for a 3 year period) more likely to be those whose symptoms are actually due to the radiotherapy. Longitudinal definitions of toxicity were already considered for faecal incontinence [16,17]. Finally, for both endpoints, the baseline questionnaire was used to exclude patients with symptoms that were already present before radiotherapy, while the end of RT questionnaire was used to assess the acute gastrointestinal toxicity following the RTOG/EORTC defini-

tion.
DVHs of anorectum were reduced to EUD. The best-value of the volume parameter n was determined through numerical optimization. A set of EUDs values were computed ranging from 0 to 1 in 0.05 steps. These EUDs were inserted in univariable logistic models and n-value maximizing log likelihood (LLH) was chosen as the most suitable. The MVL analysis was performed to include all covariates that were associated with the endpoint in the univariable analysis (covariates with $p < 0.15$). The odds ratio (OR) was used to express the strength of association of a parameter with the considered symptom. The goodness-of-fit was determined through the Hosmer-Lemeshow (HL) test and the calibration plot (slope coefficient and R^2). The gap between predicted probabilities and observed toxicity rates was evaluated through Brier score. Internal validation was assessed through 10,000 bootstrap resamplings from the original population, while external independent validation was carried out on a more recent Italian population treated with IMRT in 2010-2014. Prescribed dose was between 68 and 80 Gy. Follow-up was performed every 6 months after the end of radiotherapy and toxicity was scored using the Airopros 0102 questionnaire.

The merged dataset consisted of 1337 patients. A total of 1122 (445 patients from Airopros 0102 and 677 from TROG 03.04 RADAR) with complete clinical and dosimetric information and a minimum follow up of 3 years were considered for analysis. 1122 patients were included in the analysis of stool frequency (i.e., patients with no baseline symptoms and with at least 3 out of 6 follow-up points in 3 years), while 677 were considered for rectal pain.
Toxicity rates were as follows:

- Mean stool frequency >1: 4% (45/1122 patients, 5.4% in the TROG 03.04 RADAR population and 2.2% in Airopros 0102)
- Grade2+ rectal pain: 2.2% (21/677 patients, 3.8% in the TROG 03.04 RADAR population and 0% in the Airopros 0102). Due to absence of rectal pain events in the Airopros 0102 population, the analysis of this endpoint was limited to the TROG 03.04 RADAR population.

Figure 3.3 (a) and (b) reports the mean DVHs (together with standard deviation and p-values from the t-test for the different DVH cutoff points) for patients with and without toxicity for both endpoints. EUD was calculated for n-values between 0 and 1 in 0.05 steps.
Figure 3.4 (a) and (b) depicts the EUD values as a function of the volume parameter n for patients with and without the selected toxicity endpoints.
Figure 3.4 (c and d) shows LLH for the univariable logistic model as a function of n. LLH was maximized for n = 1 (i.e., EUD is the mean rectal dose) for mean stool frequency >1 and for n = 0.35 for late rectal pain.

Multivariable model for mean stool frequency >1
Details on univariable analysis including clinical risk factors (cardiovascular

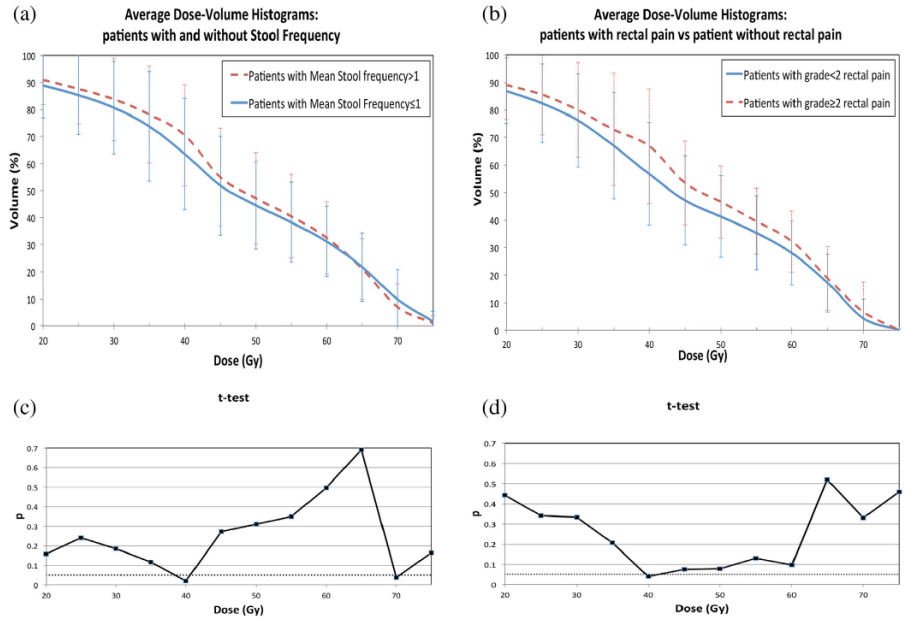


Figure 3.3: *(a) Relative mean rectal dose-volume histogram (DVH) for patients with/without mean stool frequency >1. (b) Relative mean rectal DVH for patients with/without late grade⩾ 2 late rectal pain. (c) p-values for t-tests for differences in DVHs for patients with/without mean stool frequency >1. (d) p-values for t-tests for differences in DVHs for patients with/without late grade ⩾ 2 late rectal pain. Images from reference [110]*

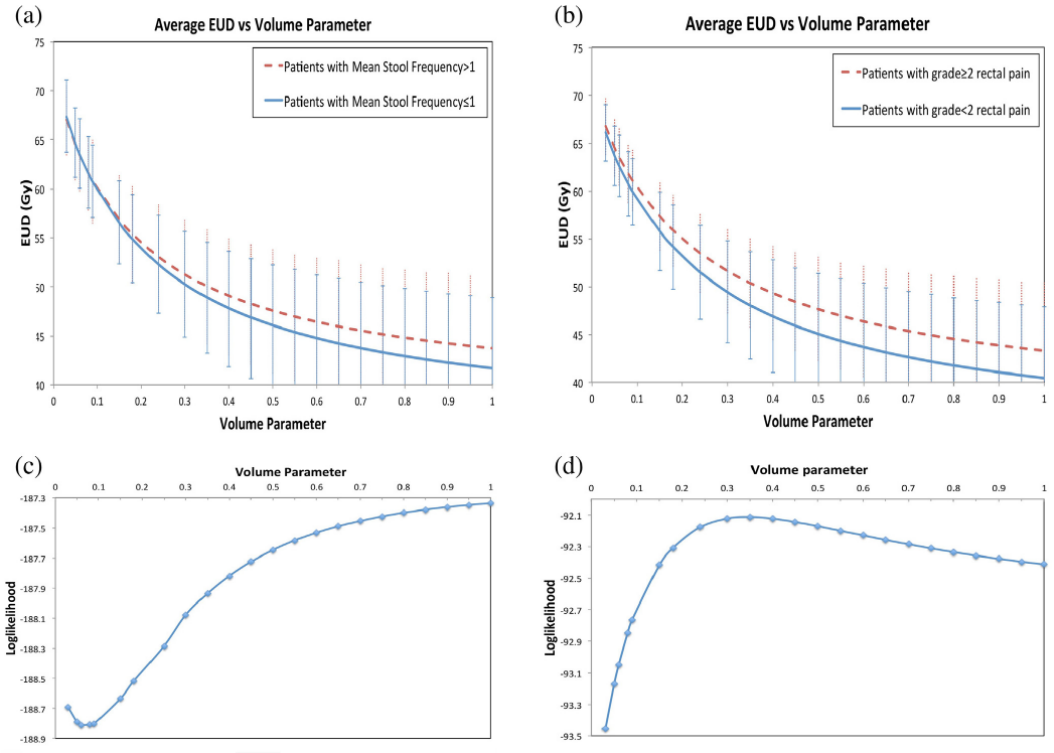


Figure 3.4: *(a) Mean equivalent uniform dose (EUD) as a function of volume parameter n for patients with mean stool frequency >1 and with mean stool frequency 61. (b) EUD as a function of volume parameter n for patients with grade ⩾ 2 late rectal pain and with grade <2 late rectal pain. (c) Log-likelihood (LLH) as a function of n for univariable logistic models associating EUD with mean stool frequency >1. (d) LLH as a function of n for univariable logistic models associating EUD with grade ⩾ 2 late rectal pain. Images from reference [110].*

disease, previous abdominal surgery, use of anticoagulants, diabetes, hormonal therapy, irradiation of pelvic nodes, irradiation of seminal vesicles, presence of acute GI toxicity) are reported in Cicchetti et al [110]. Mean stool frequency was associated with the presence of cardiovascular diseases and the development of grade $\geqslant 2$ acute gastrointestinal toxicity. A MVL resulted with the inclusion of cardiovascular diseases (OR = 1.78) and EUD calculated with n = 1 (OR = 1.04). The calibration plot (presented in the paper) had a slope = 1.03 (intercept = 0, R2 = 0.94), whereas the p-value for the HL test was 0.88; the LLH was 198.6 and Brier score 0.04. Adding grade2+ acute gastrointestinal toxicity to the two variable model did not improve fitting (LLH = 199.1, calibration slope = 1.01, $R^2 = 0.94$, p-value for HL test = 0.82, Brier score 0.04). Details for the models are reported in Table 3.1. Figure 3.5 (a) shows the probability of mean stool frequency >1 as a function of EUD and the presence of cardiovascular diseases. Observed toxicity rates are reported together with model curves. Internal validation confirmed performance measures, calibration slope = 0.93, intercept =-0.02, Brier score = 0.04.
In a previously published analysis on the TROG 03.04 RADAR population, Ebert et al. [12] found significant association between stool frequency, defined as peak toxicity, and the mid-to-low dose range (8-58 Gy) in the anal canal and with mid-to-high doses in the anorectum. Defraene et al. [19] presented a model for incidence of severe frequency (>6 bowel movements/day) including mean dose, while Schaake et al. [20] highlighted the association between anorectal side effects and different anatomical substructures within and around the rectum. Toxicity scoring was in this case based on the CTCAE scale (defined as incidence and not longitudinally). They found that stool frequency was significantly associated with the volume of the levator ani receiving >40 Gy and with the volume of coccygeal muscle irradiated at >45 Gy. The significant volume parameters for EUD calculation were in the range 0.5-1. When considering clinical factors acting as dose-response modifiers, cardiovascular diseases were already found to be a risk factor for enhanced toxicity for a number of other late gastrointestinal endpoints by previously published trials [19,21]. Patients harbouring these types of comorbidities should be granted special attention, optimization and/or prophylactic treatment.

Multivariable model for late rectal pain grade$\geqslant$ 2
For the modelling of late rectal pain, only the TROG 03.04 RADAR population was considered (677 patients, 21/677 toxicity events). The presence of hormone therapy and irradiation of seminal vesicles and lymph nodes could not be considered as covariates in this population due to the previously explained homogeneity of the patients with respect to these factors. Details on the univariable analysis including clinical risk factors are reported in Cicchetti et al [110]. Presence of grade$\geqslant$ 2 acute toxicity was the only clinical factor found to be significantly associated with late rectal pain at the univariable level. Even if a model including acute toxicity cannot be considered as a pre-radiotherapy

Table 3.1: *Multivariable logistic regression details for stool frequency average grade >1 (with 2 and 3 variables) and grade ⩾ 2 rectal pain (with 1 variable and 2 variables). Modified from [110]*

Endpoint Mean Stool Frequency >1 (2 variables)	**β Coefficient**	**OR**	**95% CI for OR**
Constant	-4.81		
EUD (n = 1)	0.04	1.04	1.01-1.08
Cardiovascular disease	0.58	1.79	0.90-3.52
Endpoint: Mean Stool Frequency >1 (3 variables)	**β Coefficient**	**OR**	**95% CI for OR**
Constant	-4.61		
EUD (n = 1)	0.03	1.03	0.99-1.07
Cardiovascular disease	0.62	1.86	1.19-4.84
Grade 2-3 Acute Gastro Intestinal toxicity	0.86	2.40	0.94-3.68
Endpoint Rectal Pain ⩾ 2 (1 variable)	**β Coefficient**	**OR**	**95% CI for OR**
Constant	-7.24		
EUD (n = 0.35)	0.08	1.08	1.00-1.17
Endpoint Rectal Pain ⩾ 2 (2 variables)	**β Coefficient**	**OR**	**95% CI for OR**
Constant	-6.75		
EUD (n = 0.35)	0.06	1.06	0.98-1.15
Grade 2-3 Acute Gastro Intestinal toxicity	1.43	4.18	1.52-11.70

EUD = Equivalent Uniform Dose; OR = Odds Ratio; CI = Confidence Interval

predictive model, we chose to insert this highly explicative variable in a MVL regression. This could provide guidance during early follow up of patients by indicating those at higher risk of having moderate/severe late rectal pain. The final two variable models included grade ⩾ 2 acute gastrointestinal toxicity (OR = 4.2) and EUD calculated with n = 0.35 (OR = 1.06). The calibration plot (pictures in Cicchetti et al [110]) had a slope = 0.98 (intercept = 0, R^2 = 0.63), whereas the p-value for HL test was 0.47. LLH was 89, Brier score 0.03. Inclusion of acute toxicity slightly improved the model in goodness-of-fit with respect to the one only including EUD (calibration slope = 0.96, R^2 = 0.88 and p-value for HL test = 0.65, LLH = 88.9, Brier score = 0.03). Details for the models are reported in Table 3.1. Figure 3.5 (c) shows the probability of late rectal pain grade⩾ 2 as a function of EUD for patients with and without presence of grade⩾ 2 acute gastrointestinal toxicity. Observed toxicity rates are reported together with model curves. Internal validation confirmed performance measures, calibration slope = 1.01, intercept = 0.06, Brier score = 0.03. Late rectal pain was found to be associated with rectal volumes receiving 30-50 Gy and with an EUD calculated with n = 0.35. Patients exhibiting grade ⩾ 2 acute gastrointestinal toxicity were at a higher risk of late moderate/severe pain (OR = 4.2). This could be related to a consequential effect between acute injury and late pain and to a possibly enhanced radiosensitivity of some patients who report worse acute injury and more severe late toxicity. Patients showing intense acute reactions might require a stricter follow up for sudden treatment and mitigation of this type of morbidity. Two previously published works, Thor et al [111] and Schaake et al [112], considered the rectal pain endpoint. Schaake et al were not able to find any associations between pain and dosimetric/clinical factors, whereas Thor et al found that rectal pain

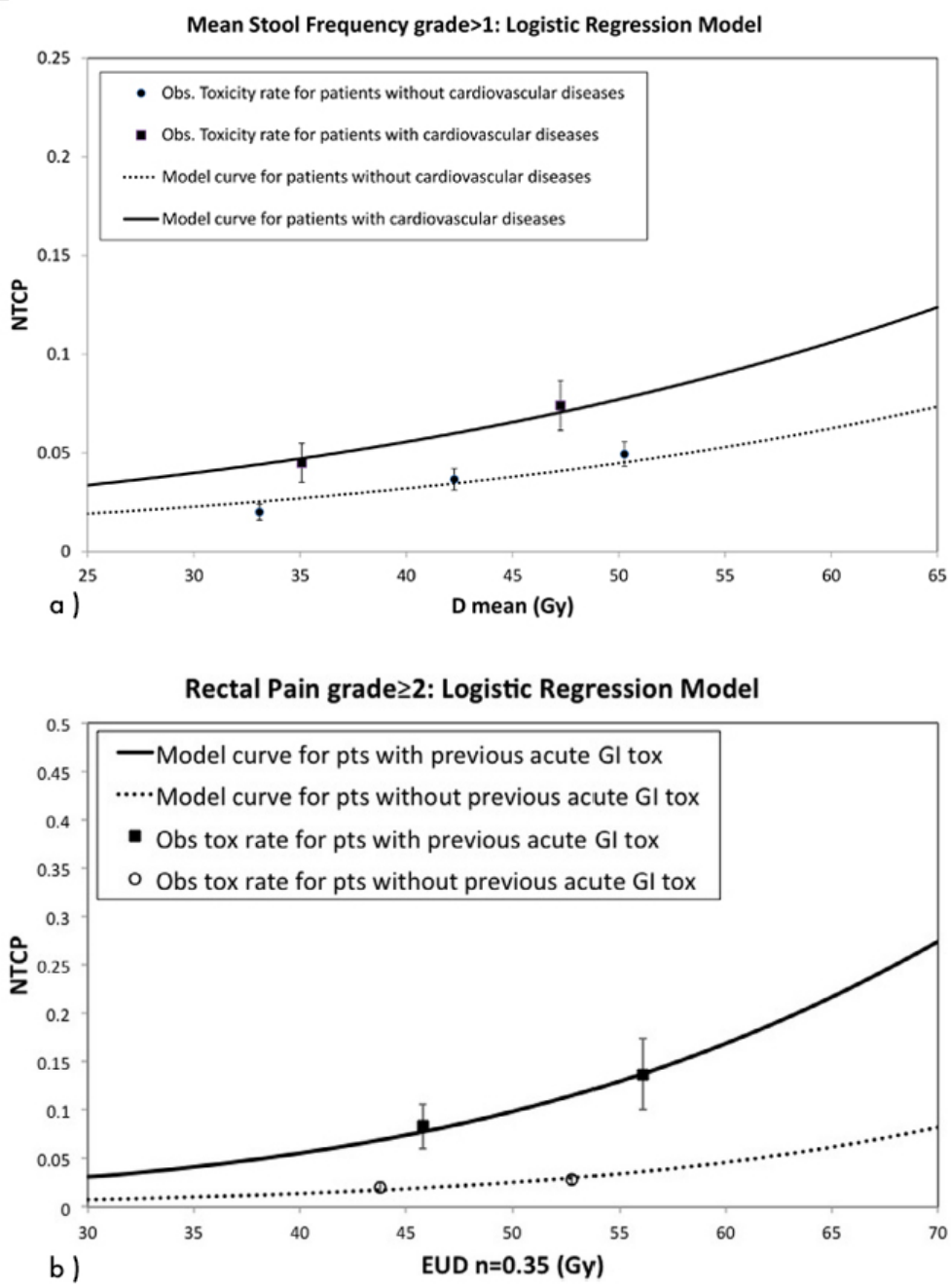


Figure 3.5: *a) Probability of mean stool frequency >1 as a function of equivalent uniform dose (EUD), calculated with n = 1, and of presence of cardiovascular diseases (cardio), b) Probability of grade⩾ 2 late rectal pain as a function of equivalent uniform dose (EUD), calculated with n = 0.35, and of presence of grade⩾ 2 acute gastrointestinal (GI) toxicity. For both plots, continuous lines represent model curves, whereas symbols correspond to toxicity rates observed in the population (error bars represent standard deviations). NTCP = Normal tissue complication probability. Images from reference [110]*

was predicted by intermediate rectal doses combined with low doses to the sphincter anal.

External independent validation
Toxicity incidence in the external population was 6.6% (10/152 pts) for grade$\geqslant$ 2 rectal pain and 17.8% (22/124 pts) for stool frequency average grade >1. Application of the model for rectal pain resulted in calibration slope = 0.95, intercept = 0, R^2 = 0.99 (plot presented in Cicchetti et al [110]) and Brier score = 0.03. For the stool frequency endpoint we found: calibration slope = 0.67, intercept = 0.01, R^2 = 0.89 and Brier score = 0.05.

Findings from analysis of stool frequency and rectal pain are coherent with what was already found in a number of studies when modelling late faecal incontinence [16,17,19]. For all these symptoms there is a relevant indication on the role of mid-range doses to large volumes of the anorectum, which is markedly different from the dose-volume relationship for rectal bleeding, which is related to low volumes receiving high doses. This parallel-like behavior of the anorectum could be biologically related to radioinduced fibrosis and consequent rectal stiffness due to large volumes receiving doses around 30-50 Gy. An interesting study on investigation of the pathophysiology of anorectal radioinduced toxicity considered the relationship between anal/rectum pressures, rectal capacity and sensory functions and side effects [23,24]. They found that radioinduced rectal stiffness seems associated with rectal wall dysfunction and toxicity. A limitation of the present study is related to the lack of information on doses to (sub)structures that are plausibly involved in ano-rectal toxicity, such as dose to the anal canal, to pelvic muscles or to nerves [113]. Validation of the models on an independent external population can be considered as an indication that, as a first step, mean dose to the rectum can be considered as an acceptable surrogate for the description of occurrence of this kind of toxicities.

3.2.2 Validation of prediction models for late rectal bleeding: evidence from a large pooled population of prostate cancer patients

A large amount of quantitative information on dose-volume relationships for rectal bleeding after irradiation for prostate cancer has been collected and analyzed over the last fifteen years. Many large-scale prospective trials have analyzed the association between patient-related/treatment-related parameters and acute or late toxicity to optimize patient selection and treatment planning [114]. NTCP models have been proposed for mild/moderate/severe late rectal bleeding[115, 31, 101, 116, 117, 103, 104]. Some of these models also couple dosimetric and clinical information, with the latter acting as dose-response modifier[31, 101, 104]. The careful application of dose-volume

constraints and the possibility of tuning these constraints according to the characteristics of the individual patient are currently considered to be an effective way to reduce radio-induced morbidity. In addition, increasing demand of reliable quantitative prediction models is prompted by the widespread use of IMRT, with its need for quantitative assessment of dose-volume histograms and of biological-based cost functions to guide inverse planning. Nevertheless, although the above mentioned published NTCP models are already useful in clinical practice, none of them was externally validated to establish its applicability and generalizability in populations other than those used for model development. External validation provides a measure of "generalizability" and "transportability" of the prediction model to populations that are "plausibly related". "Plausibly related" populations can be defined as cohorts that could be slightly different from the one used for model development, e.g. treated at different hospitals, at different dose levels, with different RT techniques, in different countries or in different time frames. Generalizability and transportability are desired properties from both a scientific and practical perspective. Quantifying the confidence and predictive accuracy of model calculations provides the decision-maker with the information necessary for making high-consequence decisions. The more often a model is externally validated and the more diverse these settings are, the more confidence we can gain in use of the model for prospective decision-making and its possible use in interventional trials.
In this study, we aimed at multiple fully independent validations (other investigators), including geographic validations (other places), spectrum transportability (wide range of prescription doses) and treatment technique validation (models developed on three dimensional conformal RT, 3DCRT, while validation includes IMRT). The validation dataset consisted of a pooled population from three large prospective trials[118, 119, 106] representing almost eighteen hundred patients with 3 year minimum follow-up. Particularly, the total cohort was the same of the previous study (TROG03.04 RADAR + Airopros 0102) with the addition of the DUE01 trial: a prospective multicenter study aimed at developing predictive models of urinary toxicity and erectile dysfunction after radical high-dose RT for prostate cancer. However, the rectal toxicity questionnaires were also administrated to patients. All DUE-01 patients were treated with radical intent by IMRT, with/without IGRT. The prescribed doses ranged between 65 and 80 Gy, including conventional (2 Gy/fr) and moderate hypofractionated (2.35-2.75 Gy/fr) schedules. Doses were corrected to 2 Gy-equivalent using the linear-quadratic model and applying an α/β ratio of 3 Gy.

Toxicity endpoint definition and evaluation
Patients were examined at the start of treatment, at the end of radiotherapy and at least every 6 months in the first 3 years of follow-up. Intestinal symptoms were classified according to the LENT/SOMA scoring systems for late radiation morbidity. For the current analysis, we focused on mild, moderate,

and severe LRB and grouped the symptoms as follows:

1. grade 1 (G1): bleeding up to twice a week (and the baseline questionnaire indicating no bleeding)

2. grade 2 (G2): bleeding >2 times/week (and the baseline questionnaire indicating no bleeding)

3. grade 3 (G3): daily bleeding (and the baseline questionnaire indicating no bleeding or bleeding $\leqslant$ 2 times/week) OR need of any number of blood transfusions and/or laser coagulation procedures

We defined bleeders as patients experiencing this event at any time greater than 5 months after RT completion until the 3 year follow-up, even if they recovered. This definition was adopted to better compare the results with the "actuarial" definition of bleeding reported in most studies.

Validation of previously published models
We considered 15 NTCP models published in the literature for the prediction of grade$\geqslant$ 1 (G1+), grade$\geqslant$ 2 (G2+) and grade=3 (G3) LRB with and without the inclusion of dose-modifying factors. Specifically, the considered models included presence of previous abdominal surgery and presence of cardiovascular disease as dose-modifying factors. Details of the selected NTCP models are reported in Table 3.2 (NTCP models without dose-modifying factors) and Table 3.3 (NTCP models with dose-modifying factors).
To compare conventional and hypofractionated schemes, calculations were always performed by applying fraction-size correction to the rectum DVHs ($\alpha/\beta = 3Gy$).

Patient characteristics
The merged dataset consisted of 1633 patients (654 from Airopros 0102, 707 from TROG 03.04 RADAR and 272 from DUE-01) with complete clinical and dosimetric information and a minimum follow-up period of 3 years. G1+ LRB was scored in 465 (28.5%) patients. G2+ and G3 LRB were reported by 255 (15.6%) and 112 (6.8%) patients, respectively. LRB rates stratified for each single trial are shown in Table 3.4. Information about main clinical/treatment characteristics of the patients included in the analysis are reported in Appendix. Figure 3.6 shows the mean rectum DVH for patients without toxicity and patients with G1+/G2+/G3 LRB. The average DVHs were significantly different in the three populations, with the Airopros 0102 patients receiving higher doses. In each population, the DVHs of patients with G2 OR G3 late rectal bleeding were significantly worse compared to the DVHs of patients without toxicity in the range of the percent rectal volume receiving more than 35 Gy (V35Gy) and up to 70 Gy (V70Gy). Notably, there was no significant difference between the mean DVH of patients with G2 bleeding in the TROG

Table 3.2: *Details of normal tissue complication probability models (not including dose-modifying factors) considered in the present work for validation in the Airopros 0102-TROG 03.04 RADAR-DUE01 pooled population.*

Reference	N pts (endpoint rate,%)	Prescribed dose (Gy) RT technique	NTCP Model	D50 (Gy) best fit (68% CI)	m or k best fit (68%CI)	n best fit (68%CI)
Endpoint: grade 1-2-3 late rectal bleeding						
D'Avino et al [115]	84 (25)	76 3DCRT	LKB	87.3 (-11.4,+14.9)	0.37 (-0.11,+0.27)	0.10 (-0.08,+0.16)
Gulliford et al [31]	361 (44)	64-74 3DCRT	LKB	59.2 (-9.3-,+8.8)	0.29 (-0.29,+0.30)	0.17 (-0.17,+0.30)
Endpoint: grade 2-3 late rectal bleeding						
Rancati et al [101]	321 (7)	64-70 3DCRT	LKB	75.7 (-1.5,+1.5)	0.14 (-0.01,+0.01)	0.24 (-0.05,+0.05)
Rancati et al [116]	669 (8)	70-78 3DCRT	Logit EUD	88.9 (-1.3,+1.4)	10.1 (-0.6,+0.6)	0.03 (-0.03,+0.11)
Michalsky et al [117]	1503 (13.5)	60-79.2 3DCRT	LKB	76.9 (-1.6,+1.6)	0.13 (-0.02,+0.02)	0.09 (-0.03,+0.03)
Tucker et al [103]	1010 (15)	68.4-79.2 3DCRT	LKB	79.1 (-1.9,+2.6)	0.15 (-0.03,+0.04)	0.08 (-0.02,+0.04)
Gulliford et al [31]	361 (15)	64-74 3DCRT	LKB	68.9 (-2.1,+2.1)	0.16 (-0.03,+0.04)	0.18 (-0.07,+0.07)
Endpoint: grade 3 late rectal bleeding						
Rancati et al [101]	547 (2)	64-70 3DCRT	LKB	78.6 (-3.7,+3.7)	0.06 (-0.005,+0.005)	0.06 (-0.01,+0.01)
Rancati et al [116]	669 (5)	70-78 3DCRT	Logit EUD	93.1 (-2.0,+2.0)	9.4 (-0.8,+1.4)	0.05 (-.04,+0.05)
Defraene et al [104]	512 (6)	68-78 3DCRT	LKB	79.0 (-5.0,+7.5)	0.15 (-0.03,+0.05)	0.18 (-0.09,0.15)

Table 3.3: *Details of normal tissue complication probability models (including dose-modifying factors) considered in the present work for validation in the Airopros 0102-TROG 03.04 RADAR-DUE01 pooled population.*

Reference	Prescribed dose (Gy) RT technique	NTCP Model	D50 (Gy) best fit (68% CI)	DMF	m or k best fit (68%CI)	n best fit (68%CI)
Endpoint: grade 2-3 late rectal bleeding						
Rancati et al [116]	70-78 3DCRT	Logit EUD	88.4 (-1.3,+1.5)	0.93 Abdominal Surgery	10.7 (-0.7,+1.0)	0.03 (-0.01,+0.02)
Tucker et al [103]	68.4-79.2 3DCRT	LKB	79.1 (-1.9,+2.6)	0.95 Cardiovascular disease	0.15 (-0.03,+0.04)	0.08 (-0.02,+0.04)
Endpoint: grade 3 late rectal bleeding						
Rancati et al [116]	70-78 3DCRT	Logit EUD	91.7 (-2.3,+2.5)	0.90 Abdominal Surgery	10.3 (-0.8,+1.2)	0.05 (-0.02,+0.03)
Defraene et al [104]	68-78 3DCRT	LKB	82.4 (-5.9,+10.0)	0.91 Abdominal Surgery	0.15 (-0.03,+0.05)	0.18 (-0.11,0.14)
Defraene et al [104]	68-78 3DCRT	LKB	82.9 (-5.9,+10.0)	0.91 Abdominal Surgery 0.92 Cardiovascular disease	0.15 (-0.03,+0.05)	0.18 (-0.09,0.15)

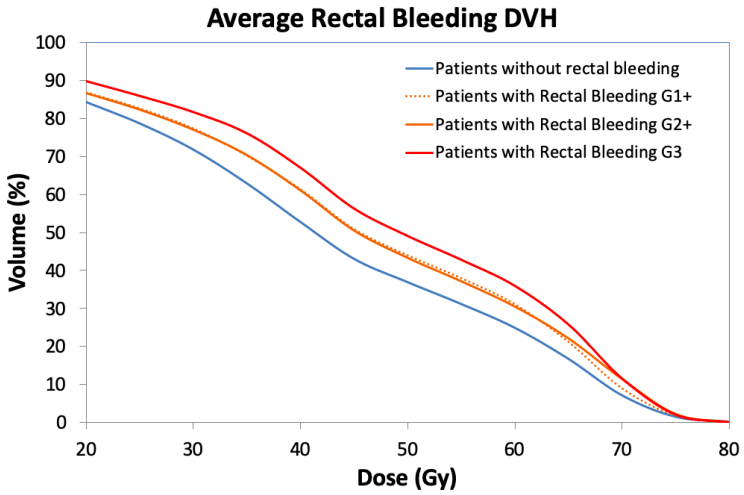


Figure 3.6: *Relative mean rectal dose-volume histogram for patients without late rectal bleeding (blue line), patients with grade 1-2-3 (orange dotted curve), grade 2-3 (orange line) and grade 3 rectal bleeding. The curves are reported for the whole population.*

Table 3.4: *Rates of rectal bleeding in each trial and in the pooled population.*

Population	G1-2-3 (%)	G2-3(%)	G3(%)
TROG 03.04 RADAR	38.2	24.5	8.8
Airopros 0102	18.5	7.8	5.0
DUE-01	27.2	11.4	6.2
Pooled Population	28.5	15.6	6.8
*p-vaue of two-proportion	z-test between population was	<0.05, except for Airopros	0102 VS DUE-01 in G3 LRB

03.04 RADAR population compared to patients with grade 3 toxicity. In contrast, patients with G3 bleeding in the Airopros 0102 population exhibited the worst mean DVH.

Validation of published NTCP models in the merged population

Table 3.5 presents a summary of the evaluation of the performance of the published NTCP models on the merged population Airopros 0102, TROG 03.04 RADAR and DUE-01. Figure 3.7 depicts the calibration plot and the fit calibration line for each model of G1+ (Figure 3.7 (a and b)) and for the G2+ (Figure 3.7 (c to i)). Calibration plot of models concerning severe rectal bleeding are shown in Figure 3.8. Particularly, for what concern the calibration plot, we decided to group patients in 4 isoprobability class (low, low-mid, mid-high, high risk class) according to the range of predicted probability that we derived from the validation models. For instance, if we had a range of predicted probability between 0.05 and 0.45, we divided patients in 4 groups: 0.05-0.14, 0.15-0.24, 0.25-0.34, 0.35-0.45. P-values for HL test were <0.001 in all cases, indicating that the agreement between data and prediction is poor. Most of the models were far from a correct calibration (calibration slope=1 and offset=0). However, increase of toxicity with increasing of the dose can be

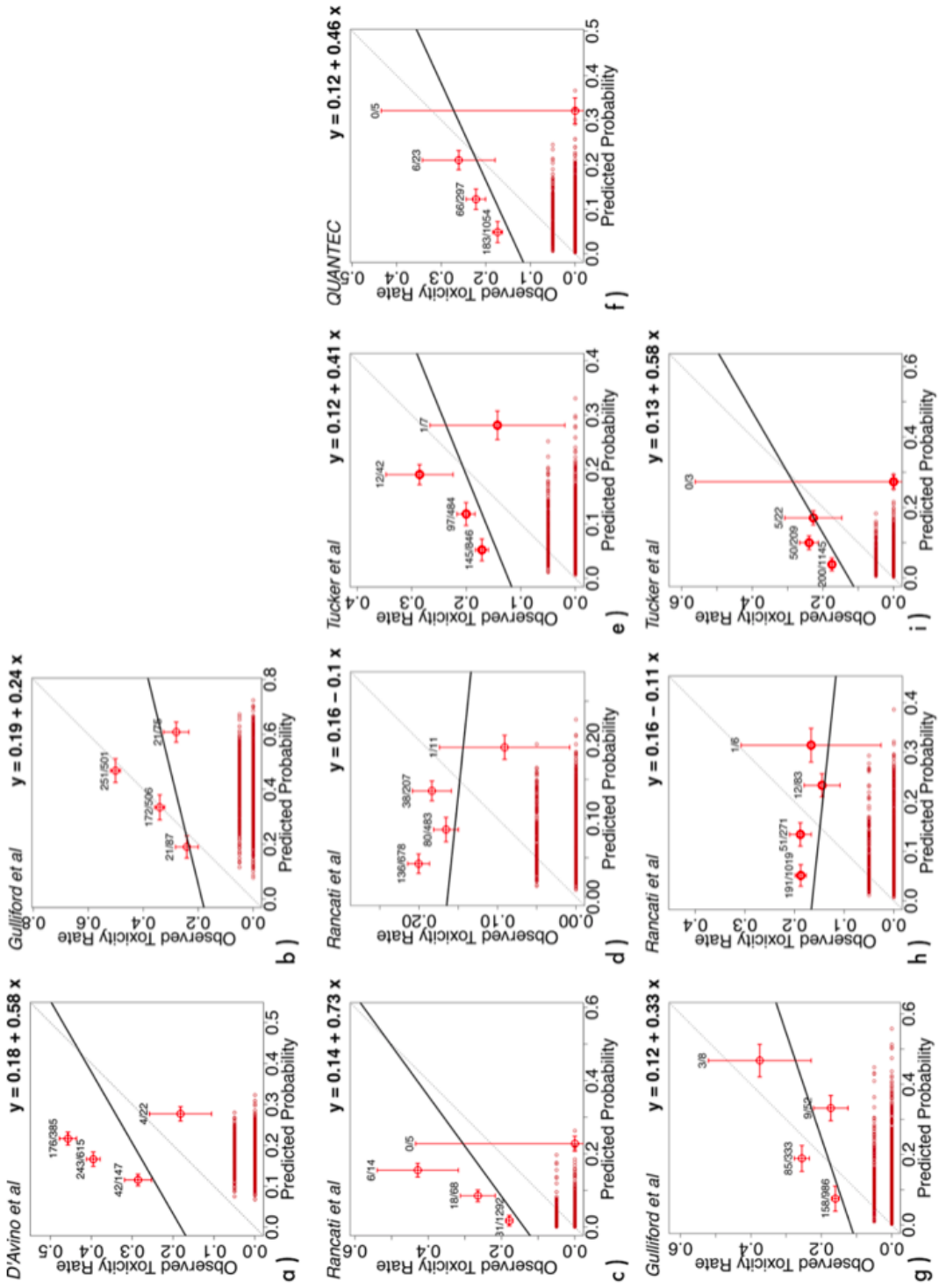


Figure 3.7: *calibration plot and fit equation for Grade 1+ (a,b) and Grade 2+ (c-i) models. The four points represent the isoprobability ranges defined in the text.*

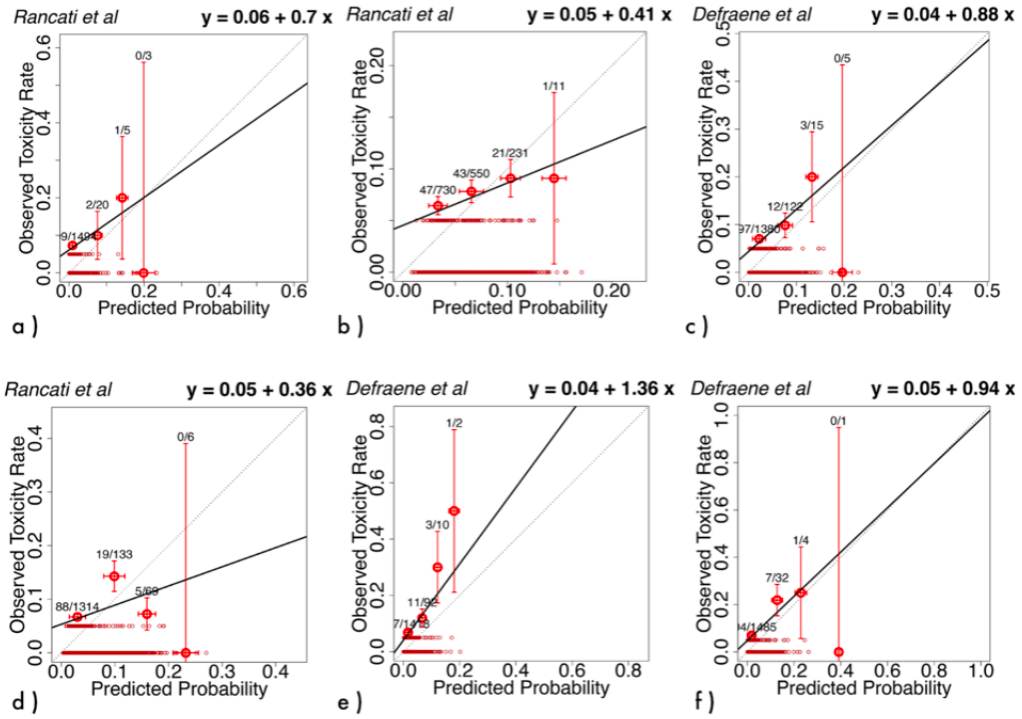


Figure 3.8: *Calibration plot and fit equation for Grade 3. The four points in the plot represent the isoprobability ranges defined in the text.*

Table 3.5: *Calibration plot statistics.*

Reference	Calibration Slope	Offset	Chi square	Brier Score
Grade 1-2-3 late rectal bleeding				
D'avino et al [115]	**0.58**	0.18	135.71	0.215
Gulliford et al [31]	0.24	0.19	119.60	0.223
Grade 2-3 late rectal bleeding				
Rancati et al [101]	**0.73**	0.14	2567.3	0.143
Rancati et al [116]	-0.10	0.11	353.19	0.141
Michalski et al [117]	0.46	0.12	289.50	0.139
Tucker et al [103]	0.41	0.12	198.53	0.128
Gulliford et al [31]	0.33	0.12	104.82	0.125
Rancati et al [116] (DVH+abdominal surgery)	-0.11	0.11	426.10	0.141
Tucker et al [103] (DVH+cardiovascular disease)	**0.58**	0.13	535.48	0.141
Grade 3 late rectal bleeding				
Rancati et al [101]	**0.70**	0.06	Inf	0.064
Rancati et al [116]	0.41	0.05	21.42	0.064
Defraene et al [104]	**0.88**	0.04	143.14	0.064
Rancati et al [116] (DVH+abdominal surgery)	0.36	0.05	74.91	0.055
Defraene et al [104] (DVH+abdominal surgery)	**1.36**	0.04	2662.74	0.064
Defraene et al [104] (DVH+abdominal surgery +cardiovascular disease)	**0.94**	0.05	340.92	0.064

observed for some predictive models in the three grades of toxicity. Seven acceptable models (D'Avino DVH only for G1+; Rancati 2004 DVH only, Tucker DVH + cardiovascular disease for G2+, Defraene (univariate and multivariate models) and Rancati 2011 DVH only for G3) can be selected considering calibration slope in the range of values between 0.55 and 1.45. It is worth noting that for some validations the calibration is made worse by the deteriorating effect of the last point (high risk patients), which is poor in terms of statistical robustness (Figure 3.7 (a, b, c, i) and Figure 3.8 (a,c and f)). To give an idea, Gulliford for G1+ would have an almost perfect calibration in describing the experimental rates of the 94% of the population. Models with inclusion of use of cardiovascular drugs and previous abdominal surgery gave good results in severe bleeding. The study, thus, confirmed that cardiovascular diseases and presence of abdominal surgery are clinical factors that should be considered in estimating the single patient complication probability. Models for G3 bleeding

showed most interesting results, also when clinical factors were included. Possibly, it could be a consequence of the fact that for severe bleeding we had a more homegeneous toxicity rate among the cohorts and also a better discrimination between DVHs of patients with and without the toxicity. However, univariate analysis for previous abdominal surgery and presence of cardiovascular disease did not give significant results (p-value between 0.1 and 0.4, except for cardiovascular disease and G2+ where p-value was 0.04) on this pooled population. On the contrary, dosimetric factors highlighted from the five selected models, i.e., EUD with volume parameter n=0.10, n=0.24/n=0.18 and again n=0.18 for G1-2-3, G2-3 and G3, respectively, are significant risk factors also on the merged cohort (odds ratios performed with univariate analysis were inserted in Appendix).
The consistency of these dosimetric results might allow us to infer that symptoms, such as late rectal bleeding, are described by parameters which point to the right area of the DVH. Furthermore, even if the serial behavior of the rectum is largely acclaimed in radiation oncology, it is important to point out that this massive validation study slightly moves the equivalent uniform dose to the medium-high dose region (compared to the QUANTEC statements), where the n value of the best models were 0.17, 0.18 and 0.24.. This shift could be explained considering that the highest rates of toxicity are in RADAR population, which has a good separation in the average rectal DVHs of bleeders and not bleeders in the region between 55 and 68 Gy.

3.2.3 Predicting Late Fecal Incontinence Risk After Radiation Therapy for Prostate Cancer: New Insights From External Independent Validation

As we discussed in the previous section, in the conformal radiation therapy era, late rectal bleeding was the most frequent and investigated rectal toxicity symptom (results on a PubMed search on the number of papers about specific rectal symptoms in the last 5 and 10 years). With the coming of IMRT and IGRT, and the simultaneous application of appropriate dose-volume constraints to the rectum an important decrease in the insurgence of bleeding was observed (3-year incidence $<15\%$, $<10\%$, and $<5\%$ for mild, moderate, and severe events, respectively [114]). Only recently have researchers begun to consider other symptoms, such as late fecal incontinence (FI) (3). Even if FI occurs less frequently ($\sim 5\%$ of patients), it has an even stronger negative impact on QoL. Krol et al [120] used the expanded prostate index composite bowel function to assess the impact of late anorectal dysfunction on QoL, and FI was the symptom with the largest impact (primarily related to embarrassment), while Bacon et al [121] showed that the level of worry caused by bowel morbidity was greater than that of sexual and urinary dysfunctions. Trying to minimize late fecal incontinence is thus of high importance, in order to guarantee good durable QoL to the long-term prostate cancer survivors. The use of predictive

models is an effective method to realize personalized treatment optimization. Some normal tissue complication probability (NTCP) models for late FI can be found in the literature [104, 116, 111, 122, 32], sometimes also including clinical modifying factors together with dosimetric features [104, 116]. Their use in clinical practice is limited by the lack of external validation, which can establish their applicability in populations other than the one used for model development, with particular interest in application in the IMRT domain. The development of validated models is essential to establish interventional studies aimed at changing patient management to reduce side-effects in cancer survivors. In this study, we aimed at externally validating the NTCP model published by Rancati et al [116], which considers a longitudinal definition of late FI. This definition is of clinical relevance, and is also important for the social wellbeing of patients, as it can take both persistence and severity of incontinence symptoms into account. It was also shown that it can better discriminate events clearly related to the radiation therapy with respect to events not directly due to radiation therapy[32]. The validation population consisted of DUE-01 patients (see details in the Appendix). All doses were corrected to 2 Gy-equivalent dose using the linear-quadratic model, but differently from the previous approach, here, applying an α/β ratio of 4.8 Gy[123]. The linear-quadratic model in its formulation including treatment time correction was also considered (with dose recovered per day $\gamma = 0.7 Gy/day$ and conventional treatment time of 80 Gy considered as reference [124, 125]) to assess if it could better describe the rate of toxicity scored in conventional and hypofractionated treatment involving the same 2 Gy-equivalent doses.

Fecal incontinence scoring and endpoint definition
Toxicity was prospectively assessed before radiation therapy, at the end of treatment and every 6 months thereafter, until 2-year-minimum follow up. For this purpose, a selfreported questionnaire based on LENT/SOMA system was used. The patient-reported questionnaire and the timing schedule were the same as used in the population considered for the NTCP model development [116]. The questionnaire scores incontinence as follows: grade 1, unintentional stool discharge "sometimes" experienced; grade 2, unintentional stool discharge "often" experienced or sporadic use of sanitary pads; and grade 3, daily unintentional stool discharge or use of sanitary pads >2 times/week. Endpoint definition followed the longitudinal characterization of late FI as presented by Fiorino et al [32]. Mean FI during follow up was calculated as the average FI grade during the first 2 years after radiation therapy completion. Patients with at least 3 out of 4 follow up points were included in the analysis (the 2-year follow-up point was mandatory). The resulting synthetic score for the persistence and severity of incontinence symptoms is continuous, and can range from 0 to 3 (0 for patients registered with grade 0 FI at each follow up and 3 for patients with grade 3 FI at each follow-up). Following the choice made for the original NTCP model development, a mean FI grade >1 was considered

as the toxicity endpoint. The NTCP model for late FI proposed by Rancati et al [116] is a logit model:

$$NTCP\,(Dmean) = \frac{1}{1 + \left(\frac{Dmean_{50}}{Dmean}\right)^{k}} \tag{3.14}$$

where Dmean is the mean rectal dose, k is a parameter that determines the slope of the sigmoid dose-response curve and $Dmean_{50}$ is the mean rectal dose that results in 50% probability of experiencing late FI.
In the model including only the mean dose, $Dmean_{50}$ has the same value for all patients, whereas in models including 1 or more clinical factors acting as dose-response modifiers, $Dmean_{50}$ takes on different values for patients with/without the clinical features. The dose modifying factor is defined as the ratio of $Dmean_{50}$ for patients with/without the selected clinical feature. Best fit parameters for the model including only mean rectal dose were: k = 2 (68% confidence interval [CI]: 0.7- 2.8) and $Dmean_{50}$ = 223.6 Gy (68% CI: 201.6-249.8 Gy). Parameters for the model with the inclusion of previous abdominal surgery (SURG) were: k = 2 (68% CI: 1.8-2.2), $Dmean_{50}$ = 223.6 Gy (68% CI: 201.6-249.8 Gy) and DMF = 0.73 (68% CI: 0.56-0.98). In the model with the inclusion of presence of colon disease k = 2 (68% CI: 1.9-2.2), $Dmean_{50}$ is 223.6 Gy (68% CI: 211.2-266.4 Gy) and DMF = 0.64 (68% CI: 0.49-0.89).

Comparison between development and validation population A summary of the characteristics of the considered validation population and of the population (Airopros 0102 trial [118]) used to fit the published NTCP models is shown in Table 3.6. The models were originally trained on a population of patients treated with 3DCRT, whereas the external validation population consisted of patients irradiated with IMRT. Other differences were related to prescription doses (higher in the validation population) and to the fractionation scheme (30% patients in the validation population received moderate hypofractionation). Moreover, differences in geographical and temporal aspects were also encased in this validation study: patients were treated in different centers around Italy and in different decades. Figure 3.9 shows the placement of this kind of study into the wide scenario of all possible validation analyses.

Validation of the effect size for the clinical and dosimetric features in the external validation population was performed by comparing the odds ratios (ORs, in the frame of univariate logistic analysis) for the development (Airopros 0102) and validation populations. Replication of the NTCP models in the independent population was subsequently evaluated by calibration plot (calibration slope and R-squared), Brier score, and Receiver Operating Characteristics Curve analyses.

In total, 229 patients with 2-year follow up and dosimetric/ clinical characteristics were available for the current analysis (details are reported in Ap-

Table 3.6: *Main characteristics of the development and validation populations. Table modified from [126].*

	Development population (Airopros 0102)	Validation population (multicenter setting)
Number of patients	506	228
Years of treatment	2002-2004	2010-2014
Location of institutes	Italy	Italy
Average patient age (68% confidence interval)	73 (68-78)	71 (65-77)
Radiation therapy technique	3DCRT	IMRT
Prescription dose range (Gy)	70-78	65-80
Prescription dose (corrected with $\alpha/\beta = 3$ Gy) range (Gy)	70-78	72.8-83.9
Prescription dose (corrected with $\alpha/\beta = 5$ Gy) range (Gy)	70-78	70.5-81.1
Dose/fraction (Gy/fr)	2.0	2-0-2.75
Rate of longitudinal mean fecal incontinence grade >1 (absolute number of patients %)	21 (4.3%)	25 (11%)
Average mean rectal dose corrected for linear-quadratic model without treatment time correction (Gy) (68% confidence interval)	44.0 (35-53)	37.7 (31.3-43.7)
Previous abdominal surgery rate (%)	8.4	38.9
Previous bowel disease rate (%)	6.1	12.7

pendix). The observed rate of mean FI grade >1 in the first 2 years after radiation therapy completion was 10.9% (25/229 patients): 8.1% and 17.4% in the conventional fractionated and hypofractionated subpopulations, respectively (z-test for proportions, P = 0.04). Details on the distribution of mean rectal dose (which is the relevant dosimetric feature in the considered model) are reported in Figure 3.10. The mean rectal dose and presence of previous abdominal surgery were confirmed as independent risk factors in the validation population, with ORs comparable to those previously published [116, 122, 111, 32]. The OR for mean rectal dose was 1.06 (range: 1.01-1.09) in the validation population versus 1.04 (range: 1.01-1.07) in the developing set, while the OR for SURG was 1.6 (range: 1-2.2) versus 1.9 (range: 1.2-3.6). The independent role of previous colonic disease was not confirmed in the validation population OR=0.3 (range: 0.15- 0.5) versus 2.7 (range: 1.4-5.2): for this reason, the model including colon diseases was not considered for validation. Figure 3.11 presents the calibration plots together with R^2 values for the considered models. The model exclusively including mean dose and that including abdominal surgery as dose-response modifying factor showed both a clear trend (i.e., increased observed toxicity rates with increasing predicted risk), but the absolute toxicity rates were underestimated (i.e., absolute predicted rates were always lower than corresponding absolute observed rates). Brier score values are shown together with the classical decomposition: reliability, resolution and uncertainty (available in Cicchetti et al[126]. The latter was the dominant coefficient in this validation study. Differences among models were exclusively affected by changes in the composition of the sample (conventional population only). ROC curve analysis gave the same area under the curve (AUC) of 0.64 in the development and validation population for the model including mean dose and previous abdominal surgery. It may be hypothesized that this result could be related to a hidden effect of hypofractionation beyond the standard linear quadratic correction. As a matter of fact, hypofractionation results to

be a risk factor in the validation population with OR = 2.4 (range: 1.6-3.7, FI rates 8.1% vs 17.4%, in conventionally treated group vs hypofractionated patients). Even the introduction of treatment time correction (see Cicchetti et al [126] for the details) did not completely explain the discrepancy between observed absolute toxicity rates and estimated rates (Figure 3.11 (c) and 3.11 (d)). To try to understand the origin of this discrepancy, application of the NTCP models to the subset of conventionally treated patients (160 patients, FI rate 8.1%) was also tested. Calibration plots for this subgroup are reported in Figure 3.11 (e) and 3.11 (f). Even in this case underestimation of observed toxicity rates is present. In this work, we validated models for predicting late fecal incontinence by concentrating on a longitudinal definition that considers both the severity and duration of symptoms. Regardless the heterogeneity and differences between the 2 settings (training and validation population, see again Figure 3.10 and Table 3.6), several clinical implications can be drawn from the present analysis. A first important clinical observation is that the rate of late fecal incontinence did not decrease with the use of IMRT. In the Airopros 0102 training population, this rate was 4.3%, while in the recent IMRT population it was 8.1% in the conventionally fractionated subgroup and 17.4% in the hypofractionated patients. With respect to other rectal symptoms, we found a decrease in severe bleeding rate, an increase in pain and stool frequency and similar rates for mild/moderate bleeding and for acute toxicity. Detailed comparison of rectal symptom rates in the training cohort [118] and in the validation population [110] together with average values retrieved from the literature [114, 81] is presented in Cicchetti et al [126]. This finding agrees with the studies of Wortel et al [128] and Al-Mamgani et al [129], who reported reduced rates for urinary toxicity and rectal bleeding using IMRT but no differences for fecal incontinence both as acute and late toxicity. This increase in FI rate is not explained by an increase in mean rectal dose (which is lower in the IMRT population; see Figure 3.10) or by known clinical risk factors, as demonstrated by the model calibration plots, with models correctly predicting increased toxicity rates with addition of risk factors, but still failing in the estimation of absolute toxicity rates. The published NTCP models proved to be robust with respect to confirmation of known risk factors such as rectal mean dose and presence of previous abdominal surgery, with concordance of the odds ratios in the 2 populations. These findings are also in agreement with several works published in the timeframe between the collection of data for model development and the validation study [122, 31, 112, 130, 81, 128]. On the other hand, the presence of diseases of the colon was not confirmed as a risk factor. The full NTCP models had a satisfactory calibration slope, indicating that it is a good tool for the selection of patients at higher risk of developing late FI, even if the absolute predictions were underestimated [4]. As

[4]In Cicchetti al[126] was presented a table with the possible thresholds in mean rectal doses which could be suggested to discriminate patients at higher risk of late FI and to guide treatment optimization

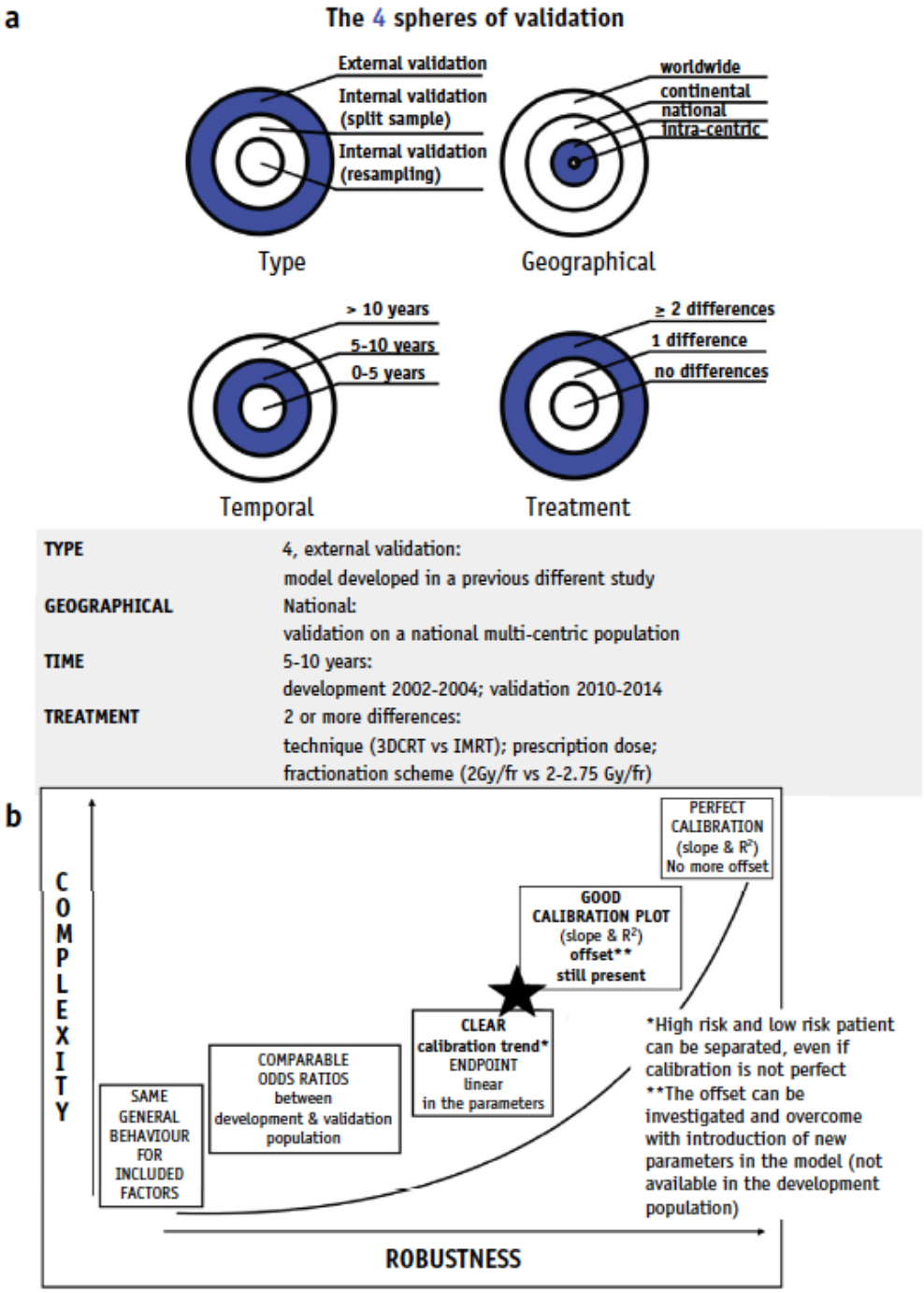


Figure 3.9: *(a) Schematic description of possible validation trials and positioning of the present study into the wide panorama of the validation analyses (solid circular annuli), following TRIPOD Guidelines [127]. (b) Graphical representation of possible course of validation studies as a function of complexity and robustness of the statistical approach. The present work could be placed in between the third and the fourth steps of this validation path (black star in the figure), exhibiting comparable odds ratios and clear calibration trend, but a calibration slope >1 and failing calibration-in-the-large. Image from reference [126]*

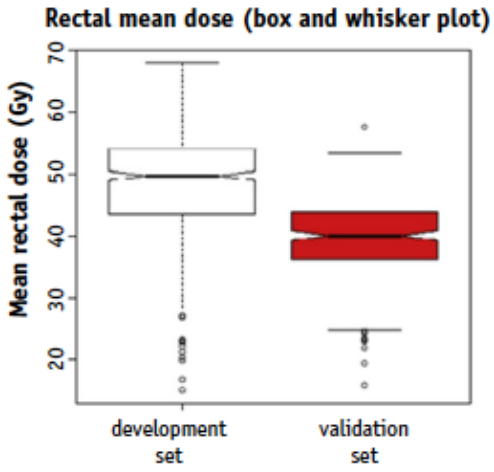


Figure 3.10: *Distribution (box and whisker plot) of mean doses in the development (white) and validation (red) populations. All doses were corrected to 2 Gy-equivalent using the linear-quadratic model and applying an alpha/beta ratio of 4.8 Gy [123] and including treatment time correction with $\gamma = 0.7$ Gy/day [124, 125].Image from reference [126]*

already pointed out, calibration plots clearly showed a systematic underestimation of absolute observed FI rates. An initial hypothesis was that the presence of patients treated with hypofractionated regimens could partly explain this increased rate of late FI (hypofractionation was a risk factor in the validation population with OR = 2.4), suggesting a role of larger daily doses beyond the one established by the linear quadratic model. Indeed, explicit introduction of hypofractionation as a risk feature did not solve the discrepancy between absolute predicted and observed toxicity rates; moreover, this disagreement (even if smaller) was also found in the conventional subgroup. Another possible explanation of underestimation could be the presence of a previously neglected risk factor. Irradiation of pelvic lymph node could play this role, due to the substantial difference in the fraction of patients with pelvic radiation therapy in the 2 populations (5% vs 22%, development vs validation population, respectively). It is reasonable to think that dose to the bowel could play a role in increasing fecal control-like symptoms. Indeed, irradiation of lymph nodes resulted to be a risk factor in the IMRT group (OR = 2.2, range: 1.4-3.4, FI rate 8.9% vs 17.6%, patient without and with pelvic irradiation, respectively), but a model including this further risk factor (together with mean rectal dose and abdominal surgery) did not solve underestimation of toxicity rates (see Cicchetti et al. for further details [126]). A final check on the application of the model to the conventional population without irradiation of lymph nodes also definitely revealed that the same underestimation was still present (detailed results in Cicchetti et al [126]). A further investigated hypothesis was that the higher FI rates could be related to an effect of rectal volumes receiving high doses (above 78 Gy), which could not be revealed by the older populations

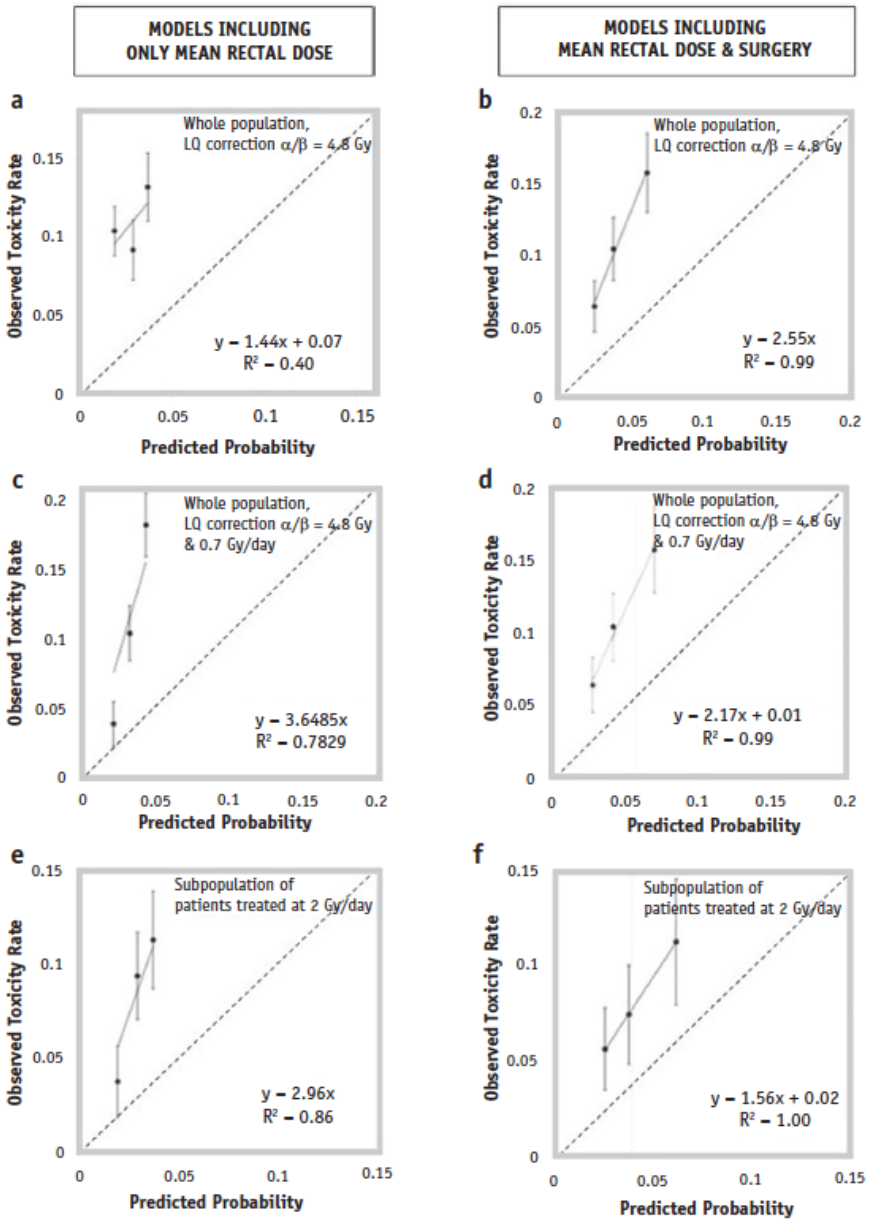


Figure 3.11: *Calibration plots present rate of observed events in a group of patients versus mean predicted probability for the same group. Error bars represent the confidence interval in observed frequencies. The left column of the figure reports calibration plots for models including only the mean rectal dose (a, c, and e), while the right column presents calibration plots for models including mean rectal dose and presence of previous abdominal surgery as additional risk factors (b, d, and f). The first row in the figure presents results for 2 models applied to the whole validation population, with correction of mean rectal doses using the linear-quadratic model and $\alpha/\beta = 4.8Gy$, but no treatment time correction (panels a and b); the second row in the figure corresponds to models applied to the whole validation population, with correction of mean rectal doses using the linear-quadratic model $\alpha/\beta = 4.8Gy$, also including treatment time correction with $\gamma = 0.7Gy/day$ (c and d). The third row presents results for the model applied to the subpopulation of conventionally treated patients (e and f). Image from reference [126]*

treated at lower prescription doses. However, these DVH regions were not significant risk factors (OR = 1 for V75Gy and OR = 0.96 for V80Gy). Otherwise, in a complementary line of reasoning, the responsibility of the increased FI rates could be attributed to decreasing volumes of the spared structures into the pelvic region. Consistent with this hypothesis, V5Gy was a risk factor with OR = 1.05, but with P-value = 0.15. Another possibility could be an enhanced effect of IMRT spreading of low/medium doses (below 20-30 Gy) into the pelvic floor, thus irradiating tissues/organs involved in the development of incontinence symptoms, such as small bowel, pelvic muscles, pudendal nerve or perirectal fat space [112, 131, 129, 113]. Detailed analysis of doses in the perirectal space is needed to appreciate the validity of this hypothesis. A further work facing this aspect and the detailed analysis of dose-maps in the ano-rectal region is in progress. In order to better understand if miscalibration was due to differences between simulated rectal DVHs and accumulated rectal doses, sub-analysis was performed, stratifying patients with respect to the presence of IGRT or not. The IGRT population consisted of 128 pts (toxicity rate 8.5%), 63 were conventionally fractionated (toxicity rate 6.3%) and 65 had HF schemes (toxicity rate 17%). The no-IGRT population consisted mainly of conventionally fractionated patients (97/101, toxicity rate 9.3%). There was no difference in model calibration for the stratified populations, indicating no relevant effect of possible differences between planned and accumulated rectal doses. Details are reported in Cicchetti et al [126]. As a final consideration, Figure 3.9 (b) depicts the possible course of validation studies as a function of complexity and robustness of the statistical approach. The present work could be placed in between the third and the fourth steps of this validation path, exhibiting comparable odds ratios and clear calibration trend, but a calibration slope >1 and failing calibration in the large [127]. Additional investigation of other important factors is needed to regain calibration in the large, together with further efforts to model the possible effect of hypofractionation beyond the linear-quadratic model, and the possible role of organs/structures, other than the rectum and the anal canal, in the insurgence of fecal incontinence symptoms. Weaknesses of this study include the size of the validation population, which is only about one-half of the population used for the development; this is increasing uncertainties in the model performance estimates. Another possible limitation is related to differences in co-morbidity rates in the 2 populations; this heterogeneity could enhance the mis-calibration in absolute toxicity rates. In conclusion:

- IMRT did not result in a decrease of incidence and severity of late fecal incontinence
- The application of the models obtained by the 3DCRT era in modern RT practice shed light on possible effects of hypofractionation on radioinduced fecal incontinence, with fractionation correction following the linear-quadratic model apparently being insufficient to consider the effect of larger daily doses

- New scenarios were also opened, with possible need to consider doses outside the anorectal region in order to prevent late fecal incontinence.

Further investigation on larger prospectively followed populations is needed to confirm these results and to understand why late FI was not decreased in the recent IMRT population.

3.3 Beyond the standard approach

In this section three works are presented in summary focusing more on the innovation part of each study rather than on the clinical results.

3.3.1 Artificial Neural Network in toxicity modelling

The first study is about the development of a *Ready-to-Use Graphical Tool Based on Artificial Neural Network* (ANN) Classification [132]. It was applied for the Prediction of Late Fecal Incontinence on the Airopros 0102 population.

ANNs have a great potential for modeling complex nonlinear relationships between independent and dependent variables, but their effectiveness and superiority over classical classification methods is a matter of debate [133, 134, 135]. A possible disadvantage of ANN-based models is that they could have an increased risk of overfitting that would reduce their generalizability [136]. Yet, the major argument against ANN-based predictive models is that they are commonly considered impenetrable so-called "black boxes" because of their complexity and low accessibility, and are therefore seldom used in clinical practice. Clinicians are reluctant to use models that act like black boxes (i.e., for which there are not clearly available factors such as the direction of association between the endpoint and the model features [are features risk factors or protective factors?], the size of the effect, and the possible interaction among different features). If the final aim is to have models somehow included into decision support systems to help clinicians, the use of black boxes would be a problem because they hide the reasons that led to decision (even from the point of view of some national laws). However, the term ANN includes many different approaches, some of which can address this last relevant issue. A basic approach is the feed-forward ANN with backpropagation training, which performs realistic nonlinear multiple regressions in a reliable manner if correctly applied. The term realistic suggests that the regression rule that can be derived is reasonable and is not affected by problems that frequently arise from nonlinear systems, such as overfitting and poor generalization of the data, whereas reliable means that the results can be considered robust and are not affected by variability in the network structure. Thus, the aim of the present study was to develop a feed-forward ANN for predicting late fecal incontinence (LFI) after high-dose prostate cancer RT. The ANN would reduce the risk of overfitting and establish a method to develop ready-to-use ANN-based graphical calculating devices (hereafter called graphical classification tools [GCTs])

that might be easily used in clinical practice. To our knowledge, this study is the first one to propose the development of such tools.

Within this PhD project, the tool was applied to the Airopros 0102 population. In this study, 598 men recruited in 2 national multicenter trials were analyzed. Information was recorded on comorbidity, previous abdominal surgery, use of drugs, and dose distribution. Fecal incontinence was prospectively evaluated through self-reported questionnaires. To develop the ANN, the study population was randomly split into training (n = 300), validation (n = 149), and test (n = 149) sets. Mean grade of longitudinal LFI>1 (expressed as in the previous study) was considered as the endpoint. A suitable subset of variables able to better predict LFI was selected by simulating 100,000 ANN configurations. The search for the definitive ANN was then performed by varying the number of inputs and hidden neurons from 4 to 5 and from 1 to 9, respectively. A final classification model was established as the average of the best 5 among 500 ANNs with the same architecture. An ANN-based graphical method to compute LFI prediction was developed to include one continuous and n dichotomous variables. For each of the selected 5 ANNs, the classification threshold *th* was defined as the lowest value corresponding to a sensitivity $< 80\%$. This high sensitivity value was chosen to be confident that almost all late chronic LFI are correctly identified as positive events by our model. Because the investigated toxicity endpoint highly impairs quality of life, high sensitivity was favored with respect to specificity. The resulting further classification capabilities were also recorded for positive and negative predictive values. A method to graphically display the output of ANNs was developed with one continuous and n dichotomous variables through a ready-to-use graphical classification tool (GCT). According to this method, dose-response curves were generated following 2 steps:

1. establish a virtual set of inputs covering all possible combinations of input variables. To do this, the continuous variable (usually related to the dose) is sampled to a finite set of discretized values;

2. calculate and record the numerical output of the ANN for every combination of the virtual set of inputs.

The resulting values, grouped together for each possible combination of the n dichotomous inputs, may be represented as dose-response curves (i.e., dose = continuous variable; response = ANN output). Once the dose-response curves were generated and a classification threshold *th* was defined (i.e., a specific ANN output value that separates between LFI yes/no), the dichotomous classification output of the ANN can be established for the whole virtual set of inputs. The dichotomous classification can also be graphically represented as a single ready-to-use GCT after the subdivision of the whole spectrum of possible inputs. To distinguish between the n dichotomous inputs, a score of 1, $2, 4...2^{n-1}$ is assigned to their possible positive occurrence; otherwise, a value of

0 is assigned. The sum of these scores, the dichotomous variable score (DVS), is a value ranging between 0 and 2^{n-1}, representing all possible combinations of dichotomous inputs. Figure 3.12 illustrates how the DVS is generated with an example in the case of 4 dichotomous inputs (i.e., n = 4).

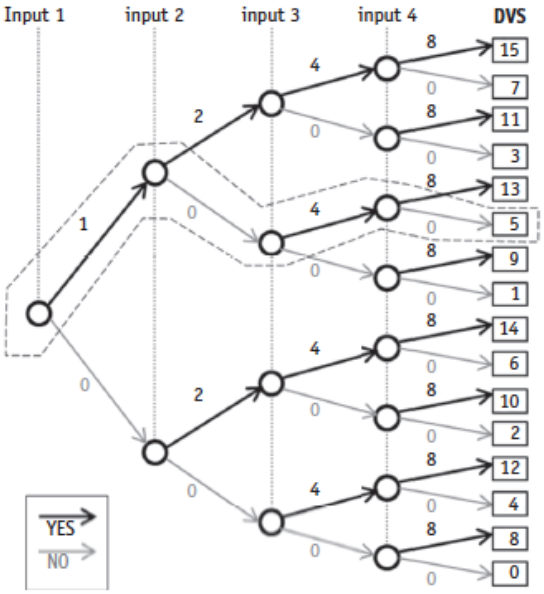


Figure 3.12: *Example of dichotomous variable score (DVS) generation with 4 dichotomous inputs (i.e., n = 4). A score of 1, 2, 4, and 8 (i.e., $2^{n-1} = 2^3$) is assigned to the positive occurrence (black arrow) of input 1, 2, 3, and 4, respectively. In case of negative occurrence (gray arrow), a score of 0 is assigned. The sum of the scores, the DVS is a value ranging between 0 and 15 (i.e., $2^{n-1} = 2^{4-1}$). To better clarify how DVS is calculated, the dotted contour shows a specific example with inputs 1 and 3 being positive and inputs 2 and 4 being negative. Scores for inputs 1 and 3 are 1 and 4, respectively, and scores for inputs 2 and 4 are 0. The resulting DVS, obtained as the sum of the single scores, is 5. No other combination of input variables provides the same DVS. Image from reference [132].*

The assignment of a combination of input variables to a specific value of DVS does not represent any ranking or classification of them, nor does it have any meaning with respect to their predictive value or importance. No correlation is expected between LFI and DVS. DVS simply is beneficial in assigning a value between 0 and 2^{n-1} to any possible combination of dichotomous inputs, which then is important for the generation of the ready-to-use GCT. In principle, the original choice of which of the input variables is scored with $1,2,4...2^{n-1}$ in case of positive occurrence is random. Figure 3.13 (a) shows the variation in the performance of the ANN model depending on the number of hidden neurons, having set the number of input variables to 5. The choice of more than 4 hidden units would not increase the corresponding ANN performance. Then, setting the number of hidden neurons to 4, ANN performance was investigated when varying the number of inputs. Figure 3.13 (b) shows

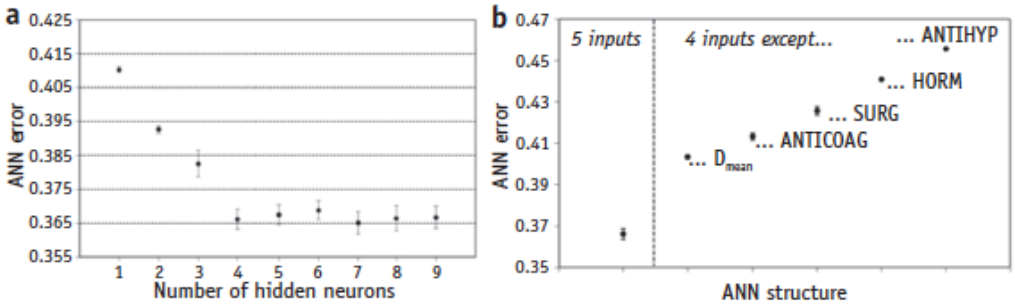


Figure 3.13: *Variation of the artificial neural network (ANN) normalized classification error, varying the number of (a) hidden neurons and (b) inputs. Abbreviations: ANTICOAG = use of anticoagulants; ANTIHYP = use of antihypertensive drugs; Dmean = mean dose to the rectum; HORM = use of neoadjuvant and adjuvant hormone therapy; SURG = previous abdominal surgery. Image from reference [132].*

that each supplementary input from 4 to 5 increases the classification accuracy. Based on these results, the ANN architecture with 5 inputs and 4 neurons in the hidden layer was finally selected the model that best balances generalization capabilities with optimal classification performance. The 5 resulting input variables were the Dmean, previous abdominal surgery, use of anticoagulants, use of antihypertensive drugs, and use of neoadjuvant/adjuvant hormone therapy (see Figure 3.14). Notably, the proposed method to generate GCTs cannot be applied if more than one continuous variable is selected as an input to the ANN. Although the method is therefore not applicable to every ANN, this approach may find many applications in RT by taking one dosimetry surrogate to assess a dose-response curve and considering the remaining clinical predictors as dichotomous variables.

3.3.2 Development and validation of an meta-model for late rectal bleeding

The second study was based on the same population of the LRB validation. We have already highlighted in that section that a common characteristic of the published predictive models for LRB was the global consistency of radiobiological parameters (n, k or m and D_{50}). The same is also true for the reccurring of specific clinical factors. Indeed, considering also manuscripts with a more clinical approach, where non-dosimetric factors were investigated alone in univariate analysis without taking into account the dosimetric variables, it was possible to identify a group of risk factors consistently identified by different papers.

This common aspects suggested us the idea to perform a *meta-analysis*; we considered a model procedure that went beyond the maximum likelihood

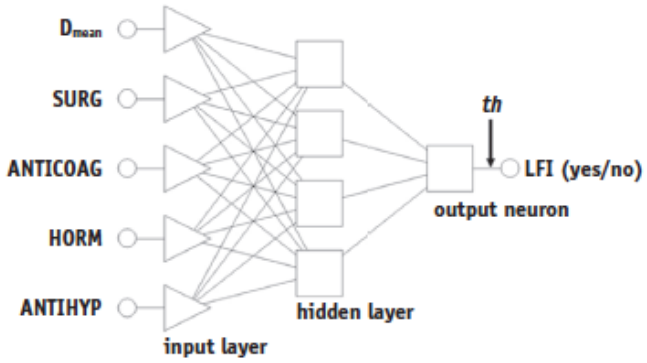


Figure 3.14: *Schematic representation of the feed-forward ANN structure for prediction of LFI after high-dose prostate radiation therapy. Th is the decisional threshold on the output neuron to dichotomize to LFI yes/LFI no for the numerical ANN output. Abbreviations: ANN = artificial neural network; ANTICOAG = use of anticoagulants; ANTIHYP = use of antihypertensive drugs; Dmean = mean dose to the rectum; HORM = use of neoadjuvant and adjuvant hormone therapy; LFI = late fecal incontinence; SURG = previous abdominal surgery. Image from reference [132].*

model fitting of the data. In other words we developed a predictive model for each grade of rectal bleeding where dosimetric and non-dosimetric prameters, were retrieved from a literature search. We followed two different approaches to determine "meta-values" for model parameters.

We collected the dosimetric informations from the most recent EUD models (only dosimetric models) considered in the study presented in Section 3.2.2. Particularly, Gulliford et al [31]and D'avino et al [115] for LRB G1+, Rancati et al [116], Michalsky et al (QUANTEC) [117] and Gulliford et al [31] for LRB G2+ and finally, Defraene et al [104] and Rancati et al for LRB G3 [116]. Details about the model parameters were already presented in Table 3.2 and 3.3. Since we decided to work with a EUD-logit model, parameters of the steeepness expressed by the m value (LKB approach) were converted into k through the relationship $k = 1.6/m$. Dosimetric coefficients (n, k, D_{50}) were resolved by weighted mean of published values, using their standard deviation as weight. The procedure was applied to each grade of severity. Results of the computation are shown in Table 3.7.

Concerning the clinical factors, we used the evidence found in Defraene et al [104] and Rancati et al [116] about the effect of previous abdominal surgery (SURG) and presence of cardiovascular diseases (CARDIO). These features, expressed by DMF, were differently weighted taking into account the prevalence of the features and the size of the study since DMF were usually reported in the literature without confidence intervals (limiting the possibility of using the previous approach). We also identified diabetes (DIAB) and hormon-therapy (HT) as risk factors coming from pubblications [137, 138, 139, 140, 141]. The information that we retrieved for these features were the ORs,

Table 3.7: *Dosimetric and non-dosimetric parameters with their coefficient computed without fitting the data*

Meta-model dosimetric and non-dosimetric parameters					
Endpoint	**D_{50}**	**k**	**n**	**Factor**	**DMF**
LRB G1+	70.2 (58.7-81.7)	5.4 (5.2-5.7)	0.12 (0.10-0.14)	SURG	0.92
LRB G2+	79.9 (79.0-80.8)	10.6 (10.1-11.1)	0.1 (0.07-0.13)	CARDIO	0.94
LRB G3	89.7 (87.0-90.5)	9.7 (9.3-10.1)	0.09 (0.05-0.14)	HT	0.96
				DIAB	0.96

the incidence in the population of patients with that factor, the toxiciy rate of occurrence. We summed up the populations related to each factor into one large cohort (two in total: one for diabetes and one for hormon-therapy). For that, the confusion matrix[5] was computed together with a new "avereged" OR for these two features. Finally, OR was converted into a DMF using the Appelt and Voegelius approach [142]. DMF for the whole set of clinical parameters can be found in Table 3.7.

In summary:

- A shift of around 10 Gy in D_{50} is necessary to describe decreasing probability of increasing severity of LRB
- volume parameter for LRB G2+ is almost in agreement with QUANTEC findings and high doses to small volumes are more important where increasing the endpoint severity
- steepness between LRB G2+ and LRB G3 is almost comparable, while steepness in LRB G1 is more shallow
- presence of previous abdominal surgery is the most impacting factor, while hormon-therapy and diabetes have a lower impact on the NTCP model.

A validation of the models was performed on the same population of Section 3.2.2 (Airopros 0102 + TROG03.04 RADAR + DUE-01). Calibration plots for G1+, G2+ and G3 are shown in Figure 3.15. We decided to validate the "meta-model" with the only dosimetric information (column on the left in Figure 3.15) and to compare it with the 5-variable "meta-model" (EUD + SURG + CARDIO + HT + DIAB, column on the right in Figure 3.15),

[5]In predictive analytics, a table of confusion (sometimes also called a confusion matrix), is a table with two rows and two columns that reports the number of false positives, false negatives, true positives, and true negatives. This allows more detailed analysis than mere proportion of correct classifications (accuracy).

where the whole set of variable parameters was derived from literature without fitting the data. Dosimetric models were able to describe the general trend of the cohort; however, relevant offsets have been recorded in LRB G2+ and LRB 3 ($\sim$ 10% and $\sim$ 3%, respectively). Inclusion of clinical factors improves slightly the fit for LRB G1+. Indeed, the model predicted very well toxicity rate below 35% (which involved 87% of the population) while partially failing at higher probabilities. It helps more in describing the toxicity of the most severe score (LRB G3). In this description, the meta-model is able to predict toxicity in 85% of the population with slope of 0.83 and offset of 0.01.

Last point (predicted high risk of toxicity) seems to be affected by the necessity of attenuating the DMF values for patients in very high risk class, in which 3 or more risk factors are present. Indeed, DMF or ORs retrieved from literature, were computed taking into account the effect of one single factor on the population. Considering the impact of two or more factors as we are doing in the meta-model, we should also know the relationship between these factors (in the development set) before to proceed with a linear multiplication of the DMF_i (in the validation set). Indeed, for a patient with 4 risk factors the standard equation should be:

$$NTCP\left(EUD_{n=0.09}\right) = \frac{1}{1 + \left(\frac{DMF_{SURG} \times DMF_{DIAB} \times DMF_{HT} \times DMF_{CARDIO} \times EUD_{50}}{EUD_{n=0.09}}\right)^{k}} \tag{3.15}$$

According to this approach, a patient with an average value of EUD (55 Gy) should have an increased probability of 12% (17% - 5%I) for LRB G3 (where 5% is the value from the dosimetric NTCP model and 17% is the value from the multivariate NTCP model) . There are no reasons for a possible sinergistic effect in between this set of parameters. As a consequence, it is more realistic to believe that in a multivariable model the role of each parameter is overrated compared to its weight found in univariate analysis.
However, it is worth to note that meta-modelling is a flexible process that can be improved and enriched during time, including new studies and factors till the achievement of satisfactory calibration coefficient (slope and R^2 as close as possible to 1).

3.3.3 Analysis of gastro-intestinal toxicity by mean of rectal Dose-Surface-Maps

In this chapter we have seen how many investigators have studied the risk of rectal complication from radiotherapy in terms of DVH data. However, the main restriction of DVHs is that the spatial distribution of dose is disregarded. Indeed, several limitations arise when only scalar values or DVHs are used within predictive models:

- A DVH is limited to a previously defined organ,

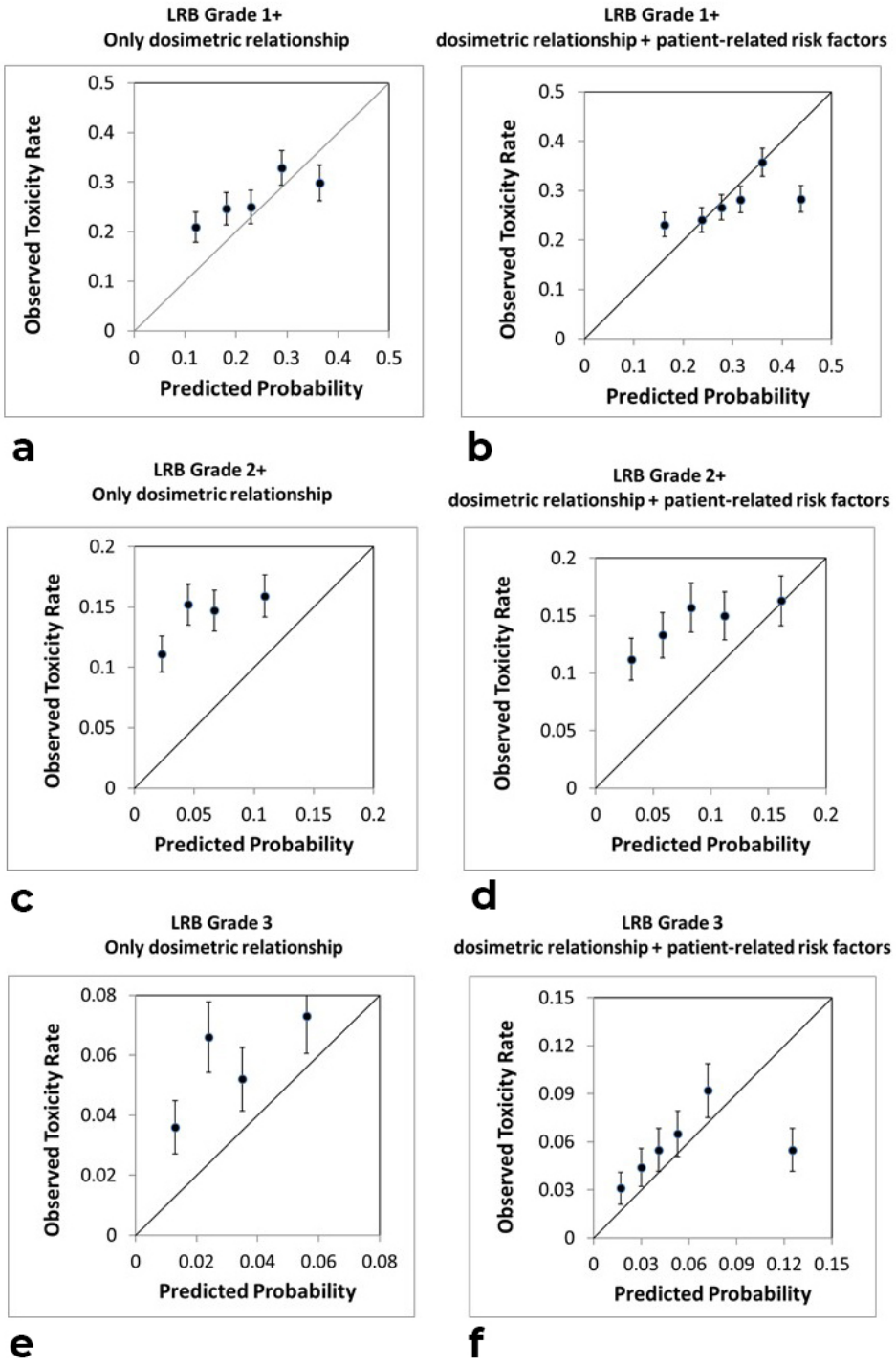


Figure 3.15: *Calibration plots for LRB G1+, G2+ and G3. Column on the left depicts the dosimetric model, while on the right column the multivariable model is presented"*

- By construction, different 3D dose distributions may lead to the same DVH,
- The information on the spatial distribution of dose is lost by merely considering the organ volume, thus ignoring the local 3D variations and thus the likely heterogeneous intra-organ radiosensitivity,
- Correlation exists between adjacent DVH bins,
- More broadly, toxicity is a complex multiparametric phenomenon that may involve structures at different scales, from sub-organ parcels to large structures or regions whose response may additionally depend on individual radiosensitivity.

These limitations may explain, partially, the limited prediction capability of DVH-based models. To address some of the limitations of organ DVHs, recent approaches aim at investigating more localized dose-toxicity relationships by analyzing the dose at lower spatial scales. Therefore several studies have recently considered spatial features of the 3D dose distribution in order to improve the performance of predictive models [143]. Our team presented an

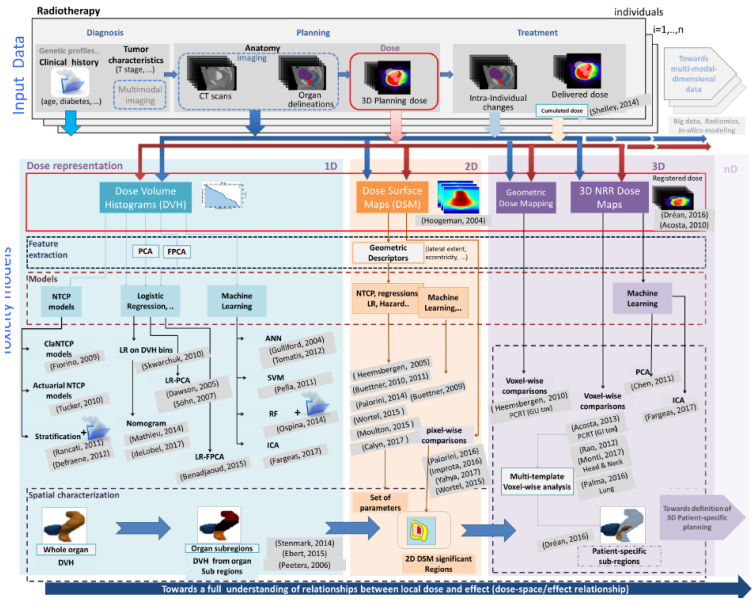


Figure 3.16: *Overview of the state of the art of predictive models of toxicity. They are all based on population data and appear divided in three different categories depending on how the 3-dimensional (3D) dose and patients' characteristics are considered. Either through the dose-volume histogram (DVH), dose-surface maps, or 3D Non-Rigidly registered dose maps.*

NTCP model based on the DSM of the urinary bladder [144, 145] and worked also in the validation of that model on the TROG 03.04 RADAR cohort[146].

Due to the good results found in Palorini et al and Yaha et al, we decided to apply the map analysis also to find spatial association for the rectal toxicity.

For this investigation we used again the endpoint of longitudinal faecal incontinence in the DUE-01 cohort.

The tubular structure of the rectum allows the delivered dose to be easily represented by DSM, by virtually unfolding the rectum and visualizing the resulting dose distribution in 2D. Each pixel in the 2D image corresponds to a portion of the organ wall with the local dose computed, for instance, by interpolation at that 3D point. By construction, these DSMs reflect the dose of the organ wall. DSMs can be exploited via the extraction of geometric features from iso-dose curves [122, 147] or through direct pixel-wise comparisons [128, 144, 145, 148]. This last approach has been implemented also in the presented study to investigate the maps of the rectum and anal canal.

The anal canal was defined as the structure extending 3 cm cranially from the anal sphincter. The rectum was defined from the anal canal to the point where the structure starts turning anteriorly. The dose map of the tubolar surface was unfolded along the posterior axis and converted into a 2-D dose map. The lateral (left-right direction) extent of the map varies with the circumfernce of the outline of each image slice, and corresponds to a snapshot of the anatomy at the time of acquisition of the treatment-planning CT. Therefore, the dose-map was represented both with this anatomical axis and with a relative lateral axis of 200 pixels, using linear one-dimensional interpolation (see Figure 3.21). This latter representation will be referred to as laterally normalized DSM (referring to the geometry, not dose). As in the validation study, even in this analysis a conversion to 2 Gy-equivalent dose has been done (using $\alpha/\beta = 3Gy$) to compare the conventional scheme with the hypofractionated regimen. Obviously, the conversion has been performed to pixel level. For all patients with and without toxicity, respectively,the average value of the dose in each pixel was calculated after aligning the dose maps for each of the OARs. This resulted in two 'mean dose maps' (for each OAR), where any difference in the dose given to patients with and without toxicity can be detected. The anorectal DSMs were aligned anatomically at the caudal end of the anal canal, the rectal DSMs at the caudal end of the rectum and the anal DSMs at the cranial end of the anal canal. Laterally all DSMs were aligned along the central axis of the organ. An alternative alignment, based on the dose distribution, was also tested. The null-hypothesis that there is no difference between the dose distributions of the patients with and without toxicity was tested. As suggested in [122], the test statistic used was derived from the difference in mean dose for each group per pixel. Moreover, p-value adjustment for multiple correlations [6] via permutation test was performed on the p-value map. Further, the pixels in the test-statistic map for the observed data with values corresponding to

[6]The adjusted p-value is the smallest familywise significance level at which a particular comparison will be declared statistically significant as part of the multiple comparison testing.

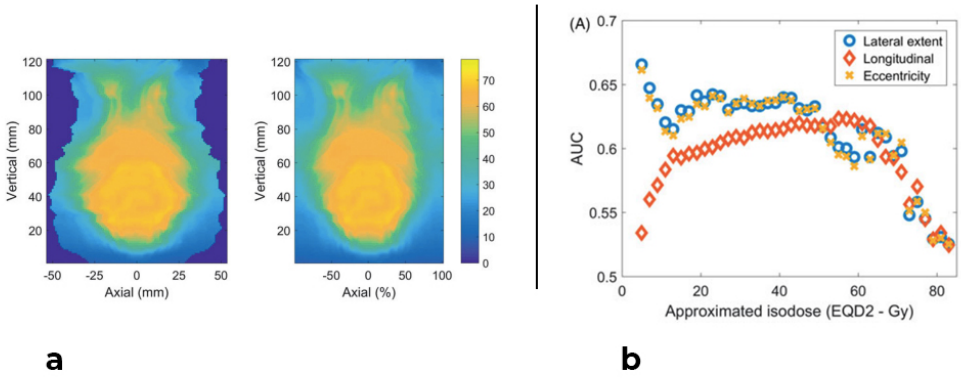


Figure 3.17: *(a)Example of DSM of the anal canal and rectum represented with an anatomical lateral axis (left) and a normalized lateral axis (right), respectively. The colorwash indicates the EQD2 in Gy. (b) AUC of models based on the lateral extent, longitudinal extent and eccentricity of ellipses approximating different isodoses in the OAR. Image from reference [143]*

adjusted p-values < 0.05 were identified. As argued by Palorini et al [144], the use of this adjusted p-value may be overly conservative for this type of application. Therefore, also the nonadjusted p-value map is reported, based on a pixel-wise Student's t-test. The anatomical alignment of DSMs resulted in lower p-values compared to the dose-based alignment and it was used to perform the analysis of the discussed results.

The most significant difference in DSMs was found for the anal canal, with an anatomical lateral axis, reaching an adjusted p value below 0.01. Figure 3.17 (b) shows that the patients with incontinence generally had a larger portion of the anal canal irradiated, specifically the caudal part of the organ. Thus, the irradiated volume of the anal canal, rather than the maximum dose, seems to determine the risk of fecal incontinence. The rectal DSMs did not display a significant difference between patients with and without fecal incontinence. In addition, a number of geometric features were extracted from the DSMs, considering a range of dose thresholds (following the method proposed by Buettner et al [147]). The shape of the dose was approximated by fitting an ellipse to each of 40 isodoses from 5 Gy to 83 Gy in steps of 2 Gy (in EQD2). The longitudinal (i.e., cranio-caudal) extent, the lateral extent and the eccentricity were recorded for all isodoses for each patient, and a univariate NTCP model based on each candidate feature was fitted to the outcome data in order to find the best dose predictor. The model used was the cumulative normal function, as in the well-known Lyman model [99]. Using a Matlab script, internal validation was performed with bootstrapping (500 cycles) and the mean AUC value over the bootstrap cycles was used as a measure of model performance.

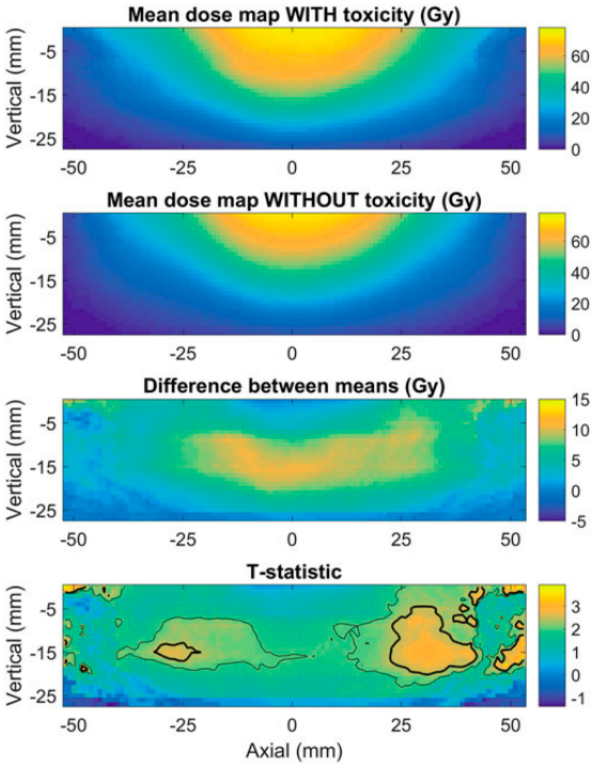


Figure 3.18: *Mean-dose maps of the anal canal for patients with and without fecal incontinence, the pixel-wise difference between the mean-dose maps, and the test-statistic map. The outer- (thin) and inner (thick) lines indicate where the test statistic corresponds to p-values and adjusted p-values, respectively, below .05. Image from reference [143]*

The AUC values of the univariate NTCP models, averaged over the bootstrap cycles, are presented in Figure 3.17 (b). One model was fitted for each of the spatial parameters characterizing the shape of the isodoses . According to these results, fecal incontinence is best discriminated by the lateral extent of the low- to medium isodoses (5–49 Gy), but also the longitudinal extent of isodoses around 60 Gy scored highly. In conclusion, the risk of causing late fecal incontinence strongly depends on the portion of irradiated anal canal, which is well quantified by the lateral extent of the isodoses, especially at low doses.

Chapter 4

Results at the Tissue Level: in-vivo, ex-vivo and in-vitro analysis of Radiation Damage in the Pelvic Region

The results presented in this chapter are derived from three very different investigations about the direct effect of radiation on tissues. They are representative of three methodological approaches for studying the radiation-induced damage at the organ level:

- in-vivo: MRI images pre and post RT for a small cohort of prostate cancer patients were analysed to find possible imaging biomarkers able to describe the tissue changes induced by radiation;
- ex-vivo: in a mouse study, we irradiated the urinary bladder and analyzed the bladder wall through the microscope after mouse sacrifice;
- in-vitro: a micro-fluidic chip of endothelial cells was created to explore the impact of radiation on microvasculature.

The results showed in Section 4.1 were also published [149], while Section 4.2 and 4.3 report results of a preliminary study and preliminary result of a 5-year project funded by the Italian Association for Cancer Research (AIRC), respectively. The studies were focused on radiation damage, but at the same time their continuation will also investigate the possible role of this damage on the development of adverse events.

4.1 Analysis of MRI in prostate cancer patients

Multiparametric MRI (Mp-MRI) has been introduced as an imaging technique able to monitor the tumor response to oncological therapies during the follow-up. Additionally, it is also possible to exploit this source of information to characterize the normal tissues surrounding the tumor and to assess their response to radiation. Indeed, Mp-MRI could have a compelling role in this context, due to its potentiality to fully characterize the soft tissues from different points of view, such as their anatomical, structural, and functional properties described by the information carried out by different acquisition techniques. In the case of prostate cancer anatomical structures such as the pelvic floor muscles, usually disregarded in NTCP studies, can be investigated by MRI since they could play an important role especially in the continence function. Moreover, the combination of mp-MRI and radiomic techniques allows the quantification of the tissue's spatial organization by the use of texture analysis. It attempts to quantify intuitive qualities described by spatial variation in pixel intensities (e.g. gray levels). This can improve the knowledge of the muscle response to radiation, by an assessment of the structural changes, which may be dependent on the received dose.

In this study, we aimed to exploit mp-MRI information and radiomic techniques for a quantitative assessment of the local radiation effect on the muscular tissue of internal obturators, as OAR potentially involved in urinary toxicity. For this purpose, texture analysis was applied on T2-w and postcontrast T1-w images acquired before RT and 12 months after its completion and structural changes were quantified in subregions related to different planned doses.
We considered thirteen patients treated for prostate cancer with exclusive radical external beam RT at the National Cancer Institute in Milan. MRI imaging was performed before RT and 12 months after RT completion using a 1.5 T scanner. Analysis was performed in collaboration with the *National Research Council* (CNR).

Image processing and regions of interest (ROIs) identification
T2-w and T1-w images were first corrected for magnetic field inhomogeneities by using the nonparametric nonuniform intensity normalization (N3) algorithm. It is a non-uniformity correction technique that finds a multiplicative field that maximizes the frequency content of the intensity distribution of the corrected scan[150]. Second, a normalization step was performed between T2-w-pre- and T2-w-post images and between T1-w-pre and T1-w-post images for each patient, in order to reduce potential errors due to the nonquantitative value of signal intensity, using the histogram matching method[1].

[1]In image processing, histogram matching is the transformation of an image so that its histogram matches a specified histogram.

Contours of internal obturator muscles were manually delineated on T2-w-prescans. To preserve the original quantitative information of the images for the next textural analysis, a contour propagation step was adopted to automatically identify ROIs in each acquired image (see Figure 4.1 (a) for a schematic overview of the image registration process). For this purpose, a deformable transformation was estimated between T2-w-pre and T2-w-postimages using the nonrigid registration method implemented in the open-source software Elastix[151]. The chosen registration method consisted of a rigid realignment followed by the classical Free-Form Deformation based on B-splines, with the parameters optimization available in the Elastix implementation, as reported in Broggi et al[152] The same image registration approach was adopted to estimate the deformation between T2-w and T1-w post-CA images, acquired in the same session. This registration was required to recover spatial mismatches occurred during the acquisition session, caused by patient movements. The estimated deformation field was then used to automatically deforme contours of internal obturator muscles previously delineated on T2-w-pre to obtain the same structures on T2-w-post and T1-w post-CA, using an optimized home-made version of the Transformix function implemented in Elastix, based on mesh deformation. The mesh, once deformed, was cut on the corresponding planes of the moving image to obtain the propagated binary masks on each slice. The accuracy of this method was estimated in this same district by comparing automatic and manual contours of different organs, showing good agreement between them, with a global error limited to the voxel size.

Deformation was also estimated between both T2-w images and planning CT data, using the same image registration method previously described, but optimized for a multimodal registration. In this case, the deformation field was then applied on the three-dimensional dose distribution associated to the planning CT. It was thus possible to have a dose distribution on obturator muscles directly on MRI images. The quality of the estimated deformation field was evaluated in terms of jacobian index[2]. Jacobian values in this region were always positive and in the range 0.7–1.3, indicating little deformations and a regular and physically reliable field.

Using the deformed dose map superimposed on the muscle volume, two different subregions were identified (see Figure 4.1 (b)): a region receiving low-doses (below 55 Gy, L_ROI) and a region receiving medium-high doses (above 55 Gy, H_ROI). The 55 Gy threshold was chosen following previously published results indicating doses to muscles between 50 and 65 Gy as associated to increased toxicity risk [149]. Furthermore, 55 Gy is the currently used dose-constraint to muscles involved in head-and-neck radiotherapy, and it was very close to the mean simulated dose to the obturators over the considered patient population (54.6 Gy). The 50–55 Gy value can be also considered as a sort of natural threshold between low doses and medium-high doses, in a radiotherapy treatment prescribing 78 Gy to the target volume. Finally, in each

[2]It is the determinant of the matrix that defines the deformation field

muscle volume, from the deformed dose maps, isodose regions were identified at 5 Gy steps, from 30 to 70 Gy (see Figure 4.1 (c)).

Textural features computation
Eighteen textural features were extracted (a) in the whole muscle region, manually delineated or automatically propagated, (b) in the L_ROI and H_ROI, and (c) in each isodose region, both on T2-w and T1-w post-CA images, before and after RT. In particular, the following parameters were calculated:

1. First-order statistical features calculated from the gray-level histogram of the whole ROI using 256 bins (mean value, variance, 95th percentile, entropy, skewness, and kurtosis).

2. Second-order statistical features, based on co-occurrence matrices (GLCM) (energy, correlation, homogeneity, entropy, contrast, and dissimilarity). The co-occurrence matrices were calculated for each section in a 2D fashion (considering only ROIs with an area of at least 20 pixels) using a quantization value of 64 levels in the four principal directions at 1 pixel distance and then averaged. Quantization was performed using the formula:
$$I(x,y) = 64 \times \frac{I(x,y)_{original} - I_{min}}{I_{max} - I_{min}} \tag{4.1}$$
where I_{max} and I_{min} denote maximum and minimum intensity of the original image.

3. Higher order statistical features calculated from the *Neighborhood Gray-Tone Difference Matrix* (NGTDM) (coarseness, busyness, complexity, contrast, and strength). Accordingly, NGTDM was calculated for each slice using 64 levels.

4. Fractal dimension (FD), estimated using the semi-variogram method[153]

The choice of 2D computation of textural matrices and FD was made for the very different voxel size in the third dimension with respect to the in-plane resolution (3 mm vs. 0.456 mm). The final feature value was calculated as the median value of each parameter among all slices, as proposed by Assefa et al[154] in order to limit the effect of the strong voxel size anisotropy in MRI images. Texture analysis was performed using the "in-house" software previously developed in the Python programming language. To verify the robustness of textural features with respect to ROI size,a Spearman correlation analysis was performed between histogram-based parameters and ROI volumes and between the other features and the median ROI area of each structure. Moreover, differences in ROI volumes were calculated between T2-w and T1-w, pre- and post-RT and H_{ROI} and L_{ROI}, for a verification of volumes conservation in the features comparison.

The effects of radiotherapy on muscular tissue were assessed as changes in textural features between pre- and post-RT images, considering the whole ROI

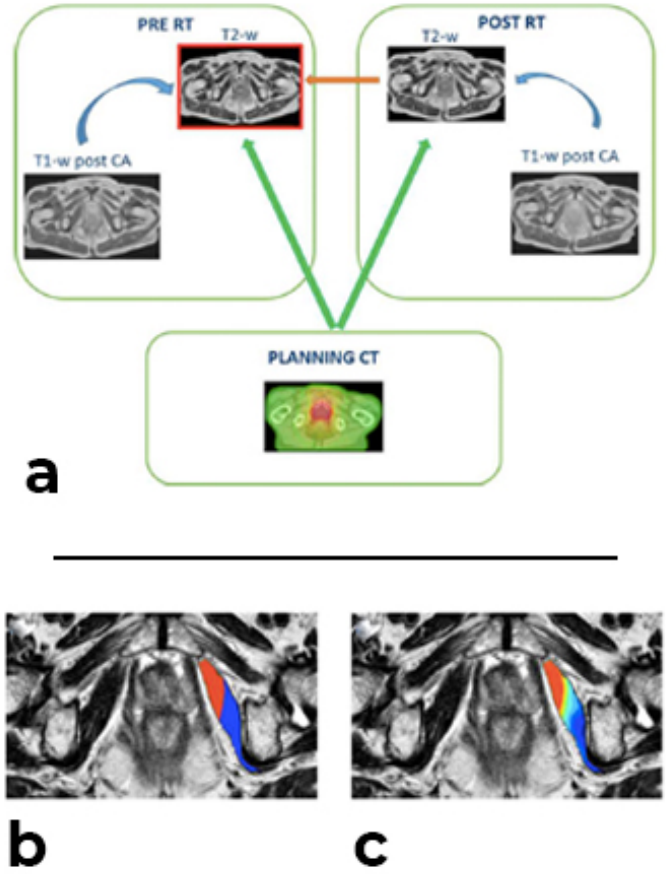


Figure 4.1: *(a) Schematic overview of the image registration and contour propagation processes. T2-w images acquired before radiotherapy (in red) were considered as the reference image. Blue arrows indicate intrasession deformable image registration, used for contour propagation; orange arrow indicates deformable image registration between T2-w-pre-RT and T2-w-post-RT, used for contour propagation; green arrows indicate deformable image registration between planning CT and T2-w (pre- and post-RT), used for dose deformation. (b) Upper portion of the image: medium-high dose region (above 55 Gy); Lower portion of the image: low-dose region (below 55 Gy). (c) isodose regions, identified from 30 to 70 Gy, with 5 Gy steps, from the lower to the upper portions of the image (from blue to red in the online version). Image from reference [149]*

and the H_{ROI} and L_{ROI}. Moreover, in order to verify the influence of the dose and possible anatomical variations within the volume, differences between the two subregions were calculated before and after RT. Differences were calculated as:

$$\Delta\, feature = \frac{feature_2 - feature_1}{feature_1} \times 100$$

where $feature_1$ and $feature_2$ corresponded to pre- and post- RT values respectively, when the effect of irradiation was investigated; $feature_1$ and $feature_2$ corresponded to values calculated in the H_{ROI} and L_{ROI} respectively, when differences in the two subregions were considered. Differences were evaluated through the Wilcoxon signed rank test; P-values were adjusted with the Bonferroni correction for multiple comparison, and statistical significance was set at adjusted $P < 0.05$. Finally, dose–response curves were built by considering the features calculated in the different isodose regions. Features correlating with ROI size were removed from the analysis. ROI volumes were generally conserved after contour propagation, since the percentage volume variation between T1-w and T2-w scans was $0.1 \pm 0.3\%$ for the pretreatment acquisitions, and $0.2 \pm 0.3\%$ for the post-treatment acquisitions; the decrease in volume after RT was $4.9 \pm 8.1\%$ for T2-w images and $4.7 \pm 8.5\%$ for T1-w images. When the whole ROI was considered, a general signal increase was detected after the treatment (the T2-w signal increased by 59% and the T1-w post-CA enhancement increased by 19% 12 months after RT). Moreover, a change in the spatial organization was identified, as highlighted by an increase of the indices related to the homogeneity and a decrease in contrast, complexity, and fractal dimension as shown in Table 4.1. In general, the differences in T2-w signal were more significant and stronger than the differences in T1-w post-CA enhancement.

Subregion analysis

When the two subregions were separately analyzed, the ROI size was significantly different between H_{ROI} and L_{ROI}. In particular, H_{ROI} presented a volume about 50% higher than L_{ROI}. It was shown that more significant or, at least, stronger differences were present in the region receiving medium-high doses[149]. The H_{ROI} resulted more hyperintense, more homogeneous and less complex; this enhancement was also visually perceptible looking at the MRI scans, as qualitatively shown in Figure 4.2 (red circles). Also considering the histograms shape, it is possible to appreciate the differences between pre and post-treatment, as numerically quantified by the firstorder statistical features. Nonetheless, also the low-dose region presented the same changes after irradiation, although less evident. In fact, considering the variations on T2-w images, the mean intensity increased by 77% (corrected $P < 0.001$) on the H_{ROI} whilst the increase in L_{ROI} was only of 29.7% (corrected $P < 0.001$). This condition is repeated for $correlation_{GLCM}$ (18.3% vs. 8.94%), $contrast_{GLCM}$ (-50.1% vs. -39.5%), $coarseness_{NGTDM}$ (89.7% vs. 47.9%), $busyness_{NGTDM}$ (-46.3%

Table 4.1: *Differences in textural features calculated on T1-w and T2-w images considering the whole ROI. Table modified from reference [149]*

		Obturator muscles	
	δ post-pre RT	**T1-w(%)**	**T2-w(%)**
Histogram	Mean	18.9 ± 12.8^a	58.6 ± 20.9^a
	Variance	189 ± 212^a	212 ± 84.3^a
	95th percentile	29.8 ± 22.1^a	77.7 ± 23.5^a
	Entropy	2.20 ± 5.54	5.60 ± 5.68^a
	Skewness	-4.51 ± 25.8	-8.68 ± 23.6
	Kurtosis	13.5 ± 103	40.8 ± 319
GLCM	Energy	7.40 ± 12.5	15.4 ± 8.96^a
	Correlation	10.2 ± 10.4^a	17.8 ± 6.70^a
	Homogenity	13.2 ± 11.4^a	30.5 ± 10.6^a
	Entropy	-1.41 ± 4.20	-3.52 ± 2.18^a
	Contrast	0.02 ± 31.6	-35.8 ± 13.9^a
	Dissimilarity	-71.9 ± 12.7^a	-23.4 ± 7.91^a
NGTDM	Contrast	2.60 ± 46.2	-23.5 ± 21.3^a
	Busyness	-16.5 ± 22.5	-27.0 ± 22.8^a
	Complexity	207 ± 269^a	-29.3 ± 9.23^a
	Strength	53.0 ± 44.8^a	82.8 ± 33.2^a
	FD	-1.59 ± 1.68^a	-2.60 ± 1.10^a
a = Bonferroni corrected P<0.05			

vs. -20.2%), $strength_{NGTDM}$ (124% vs. 64.6%), and FD (-3.49% vs. -1.43%) (corrected $P < 0.05$ in all cases).Again, as for the whole ROI analysis, features calculated on T2w-MRI exhibited more enhanced differences in textural features with respect to T1-w post-CA.

In order to verify the spatial effects of irradiation, the differences between the two subregions before and after RT were also analysed. At the baseline, the obturator muscles presented some significant differences between the anterior and posterior part, especially in the mean intensity and in certain GLCM features thus highlighting a noncompletely homogeneous anatomical pattern. These differences became more evident after irradiation, when the increase in T2-w signal raised from 10% to 53% (corrected $P = 0.0034$), and also the large majority of the other indices increased and differences were associated to lower and significant P-values.

This condition can be better assessed by the dose-feature profiles reported in Figure 4.3 (c,d,e), where differences between pre- and post-RT of three selected parameters in each isodose region are shown. In particular, the T2-w mean signal intensity showed an exponential relationship with dose [$\Delta T2 - w_{signal} = 0.4 \cdot exp(0.08 \cdot dose)$], the 95th percentile of the T2-w signal exhibited a linear increase in the range 35–65 Gy ($\Delta 95thperc = -55 \pm 1.7 \cdot dose$), and the GLCM T2-w contrast showed a quasi-linear decrease. These behaviors, especially the exponential trend of the mean value, highlighted the influence of irradiation

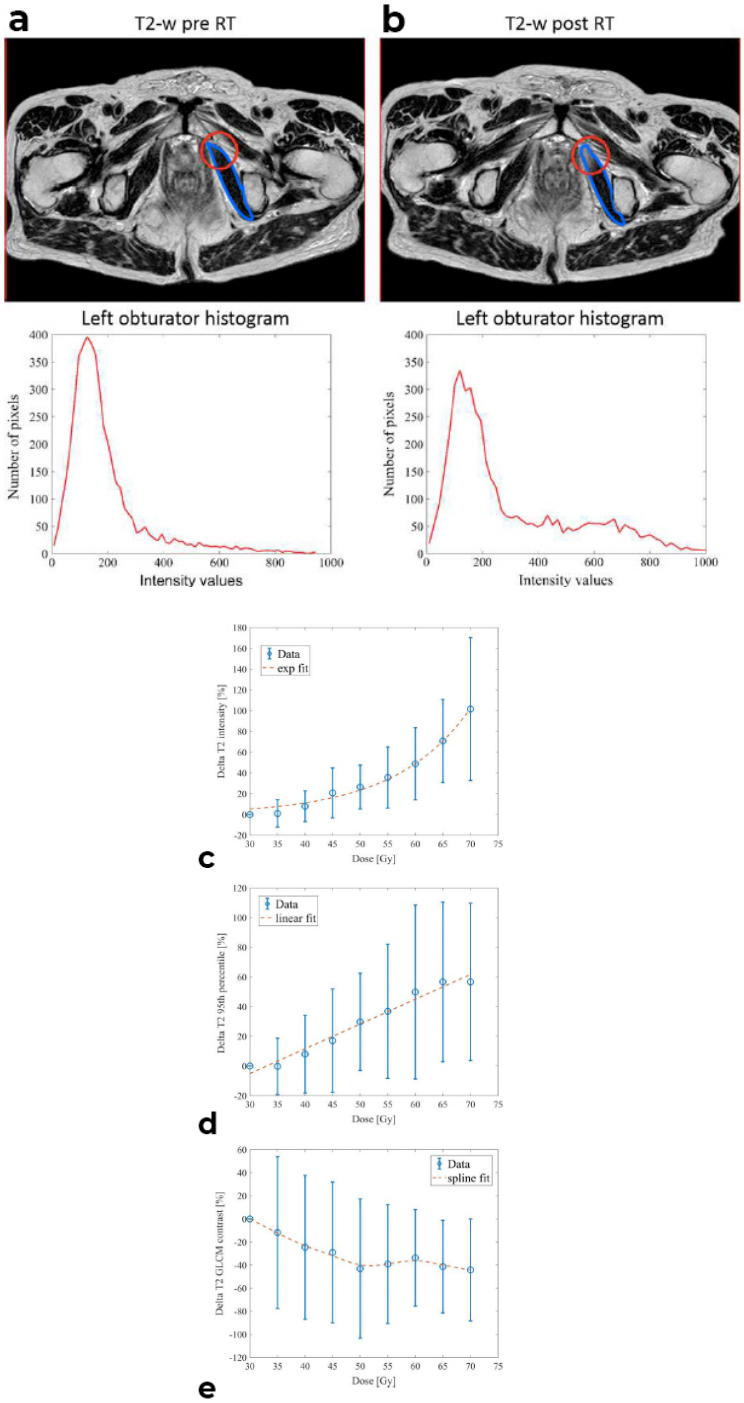


Figure 4.2: *(a) An example of T2-w images before and (b) 12 months after RT and the histogram of the corresponding left obturator muscle (contoured in blue). In the circle, an increase of signal intensity after treatment is highlighted, as also visible in the increased pixel number in the right part of the histogram. Dose-feature profiles (between 30 and 70 Gy) for the variation (as a percentage) of: the T2-w mean signal intensity (c), the 95th percentile of the T2-w histogram (d) and the GLCM contrast (e) on the entire set of paients.*

on the MRI signal.

In this study, accordingly to other works, we found an increase in the MRI signal intensity after irradiation and, through the use of texture analysis, a more homogeneous spatial organization, especially in the region receiving a high dose. Before treatment, the muscle volume presented a non-completely homogeneous pattern, as highlighted by some significantly different features between the two considered subregions. However, stronger changes were found after irradiation and the clear exponential relationship between mean signal intensity and radiation dose resulting from the present study may confirm the hypothesis that high signal intensity can help to detect the muscle insult by radiotherapy. The main innovation of our work consisted of the assessment of textural features, in addition to the simple mean intensity. This approach allowed to quantify the tissue spatial organization within the whole structure before and after irradiation. Although the mean intensity value remains the most robust and strongest feature to highlight the radiation- induced changes, the other parameters contributed to a deeper understanding of the inflammatory reaction and the consequent changes in the T2-w signal. A strength of our work is related to the assessment of the spatial relationship between local dose and MRI changes using the deformed dose maps, in order to better understand the dose–response of muscular tissue at a local level. We found stronger changes in the region receiving >55 Gy. However, looking at the dose–response curves shown in Figure 4.2, it seems that dose has not a threshold effect, but it plays differently in the various features. In fact, it exhibited an exponential effect for the mean signal intensity, while it presented a quasi-linear trend for the 95th percentile and GLCM contrast.
The limitations of this work are mainly in the limited number of enrolled patients, although the reported changes after RT were highly significant. This study was intended as a pilot analysis, which should be confirmed in a currently ongoing larger and external dataset. Another limitation is the lack of correlations with toxicity data acquired 12 months after the end of RT. Future works will be directed to test the relationship between texture analysis findings and the presence of urinary toxicity.

4.2 Immunohistochemical and pathological analysis of irradiated urinary bladder in mice

The introduction of the last generations of accelerators, that provide VMAT and precise IGRT, along with a deeper knowledge of healthy tissue response to radiation have greatly enhanced the treatment outcome with a concomitant reduction of side effects, especially for rectal and gastrointestinal toxicities. We have already discussed about the use of fractionation as a key tool for increasing the therapeutic ratio. However, the better sparing of normal tissues

achieved with the most recent RT modalities has reawakened the interest of the radiation oncology community towards hypo-fractionated schemes, with the aim of reducing both patient discomfort and treatment costs. Recently, various protocols, from moderate to extreme hypo-fractionated schemes, have been suggested. However, variations from conventional schedules should be considered with caution, since extreme reduction of the number of fractions, without the support of robust data on tissue radiobiological behaviour, might potentially lead to unacceptable doses to the healthy tissues. When considering irradiation for prostate cancer, this is particularly important for urinary toxicities. Recent results from clinical trials with intermediate or long follow-up show that, differently from gastro-intestinal toxicities, urinary toxicity was generally not reduced even after the introduction of VMAT and IGRT techniques, with a significant impact on the patient QoL. This is mainly due to the difficulties in sparing the bladder, which is adjacent to the prostate and necessarily included in the planning target volume, and to the enhanced bladder sensitivity to the highest doses; but also to the lack of predictive models of urinary toxicity, mainly due to the unavailability of sufficiently powered prospective studies.

How the bladder responds to variation in fraction sizes is still an open question. Our result in DUE-01 trial [144] is in agreement with the large majority of clinical trials, that suggest that hypo-fractionated and accelerated techniques appear to significantly increase the risk of normal tissue injury. The same findings were highlighted by the study "REQUITE: Validating Predictive Models and Biomarkers of Radiotherapy Toxicity to Reduce Side-Effects and Improve Quality of Life in Cancer Survivors" [155]. It is a multicentric european project which enrolled 1760 prostate cancer patients from 7 countries. Different protocols were adopted in the several institutes including dose per fractions in between 2-3.3 Gy and total prescription dose varying between 59-80 Gy. As a first approximation, grouping the patients from this large cohort into conventially and non-conventinally treated, the rates of moderate-severe urinary toxicity are 22.0% and 37.8%. Nonetheless, this is in contrast with the only α/β values reported in literature that were derived from a few old animal experiments. Bentzen et al found an α/β ratio equal to 5.8 (3.6 – 8.8) Gy and suggested that late radiation injury in the mouse urinary bladder is one of the least sensitive late endpoints with respect to changes in dose per fraction[60]. A comparison between present clinical studies and old animal studies is not straightforward. The clinical endpoint is usually determined with patient reported questionnaires that include a variety of symptoms (such as increased frequency, urgency, bleeding or late incontinence, etc.), while animal endpoint was defined as a reduction of 50% of mouse bladder volume with respect to baseline, measured with transurethral bladder filling at a fixed pressure. Furthermore, at the time of laboratory experiments only traditional irradiators were available, with a fixed radiation source and simple static shielding techniques that were very far from current clinical practice. Last, most importantly,

the confidence intervals of α/β ratios that can be derived from modern practice observational clinical trials are very large, since, in patients, it is not possible to systematically span a wide range of fractionation schemes or prescribed doses. Given this last point, experiments on small animals still represent a necessary step for accurately determining the sensitivity of bladder to fractionation. Modern micro-linac systems open the possibility of reproducing experiments on small animals by overcoming some of the major difficulties encountered in the old experiments.

Within this PhD project we decided to explore the endpoint of acute inflammatory urinary reaction that typically arises immediately after the treatment and usually resolves in few weeks. However, it is well established that in a significant percentage of PC patients, acute symptoms evolve in chronic diseases from 6 months to 10 years. The study further aimed at evaluating current laboratory tests and imaging techniques which might be able to promptly and accurately identify the inflammation patterns through in-vivo non-invasive techniques. Within this objective, we identified three possible analyses.
First, well established physicochemical and bacteriological tests on urine samples, collected through metabolic cages in order to avoid invasive procedures (such as the use of catheters), was used to identify the presence of inflammatory markers and their kinetics after animal irradiation.
Second, an innovative method based on *Bioluminescence and Fluorescence imaging* (BFi) was tested. Bioluminescence CCD cameras detect photons emitted from luciferase enzymes in appropriate animal models and allow in-vivo investigations to detect, localize and quantify specific biologic processes. The NFkB-luc2 reporter mouse is a model recently developed at Università degli Studi di Milano that allows the tracking of the transcription factor (TF) NFkB, the master regulator of the expression of most of the cytokines responsible for the initial phase of acute inflammation. This model was chosen for mouse bladder irradiation with the micro-linac in order to study the activation of inflammatory pathways throughout all the animal body.
Third, ex-vivo pathological analyses on dissected specimens were also performed by an experienced animal pathologist in order to quantify the inflammatory cells in bladder and in the kidneys, but also in other tissues such as testes and seminal vesicles that can be highly affected by radiation.
Finally, a local immunohistochemical (IHC) analysis of the bladder by Ki67 staining was also performed.
Highly conformal image-guided treatment plans for bladder irradiation were delivered with the micro-linac X-RAD 225Cx SmART system. It consists of a cone beam CT imaging system and a therapeutic X-ray source (199kV, 15 mA) into a single platform, housed within a self-shielded cabinet. In particular, a bladder target volume was defined through the CBCT imaging system and an optimal configuration of collimators and therapeutic beams was chosen in order to uniformly deliver the prescribed dose to the bladder volume, while minimizing the dose to the surrounding tissues. Before going through

Table 4.2: *Description of four irradiation schemes to the urinary bladder, including also corrections to Equivalent Dose to 2 Gy/fr (EQD2) according to the Linear-Quadratic model and computed with three values of α/β (1,3,5 $Gy^{-}1$).*

Scheme	N Animals (irradiated+control)	Dose per fraction (Gy)	N of fractions	Equivalent Dose to 2 Gy/fraction if $\alpha/\beta = 1$ (Gy)	Equivalent Dose to 2 Gy/fraction if $\alpha/\beta = 3$ (Gy)	Equivalent Dose to 2 Gy/fraction if $\alpha/\beta = 5$ (Gy)
A	9 (7+2)	2	4	8	8	8
B	9 (8+1)	3	4	17.7	14.4	13
C	9 (7+2)	4	4	36.4	24.0	18.8
D	8 (7+1)	5	3	31.2	22.4	18.7

the tests, the microlinac apparatus followed a preliminary calibration of the imaging panel detector, of the X-ray tube and acquistion of a correction matrix to correct mechanical offset.

The mouse bladder was irradiated with highly conformal beams (see Figure 4.3 for a general overview of the treatment settings)in four different fractionated schedules, at fraction doses that span the range of values employed in clinical practice. The schemes are described in Table 4.2. Details of the treatment machine and planning system are shown in Figure 4.3. Mice underwent

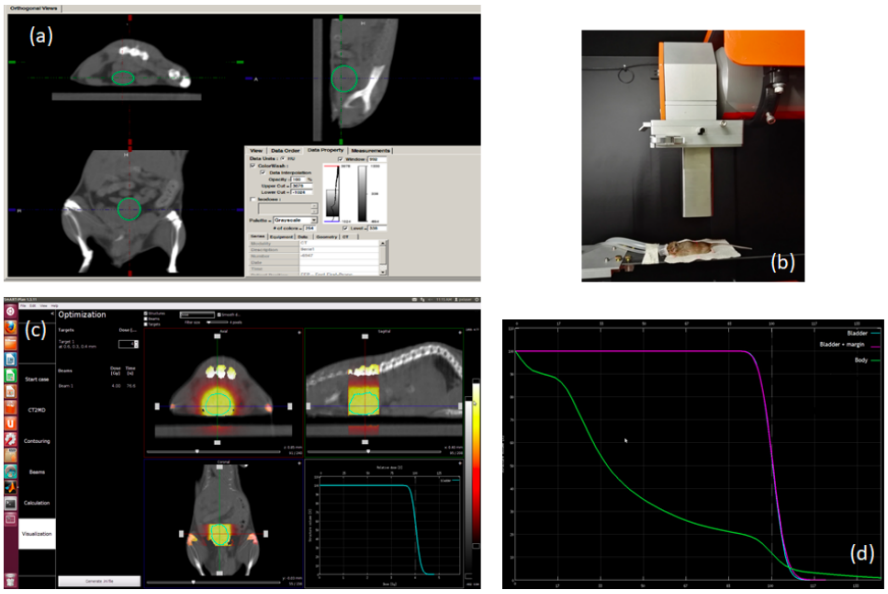


Figure 4.3: *(a) identification of bladder on cone-beam CT; (b) experimental set-up for mouse irradiation; (c) dose distribution using a conformal full arc centered on the mouse bladder; (d) dose-volume histograms for bladder and total body.*

BFi before RT, at the end of RT and 25 days after the treatment. Five ROIs were defined in the mice: total body, thymus, abdominal region, genital area and head. Photons within the ROIs were counted for 5 minutes. The resulting imaging was a 2D picture as the example of Figure 4.4. Particularly, it needs to be specified that the whole body does not take into account the animal paws, and the abdominal region was defined as the squared region minus the

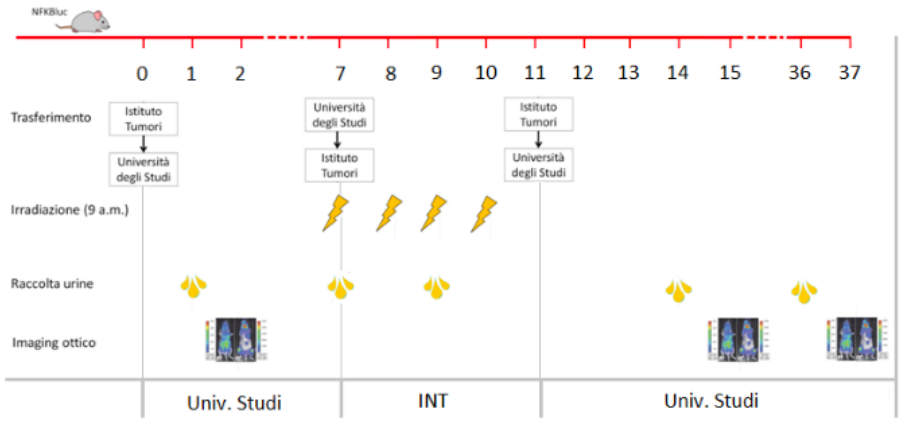


Figure 4.4: *Experimental design of the study. Time line starts when mice exit from the enclosure and enter in the experimental procedure. First day of irradiation is the 7th.*

circle of the genital area.

To analyse the picture output we begun grouping the 7 control mice and getting an average of the signal within the 5 ROIs at different time points. On the basis of this average for the control group we computed the fold change (FC). This is defined as the ratio between the two quantities; for quantities A and B, the FC of B with respect to A is B/A. It is often used when analysing multiple measurements of a biological system taken at different times, as the change described by the ratio between the time points is easier to interpret than the difference. In our case we defined two average fold changes:

$$FC_{4days} = \frac{Avg\ signal_{4days}}{Avg\ signal_{baseline}} \tag{4.2}$$

and

$$FC_{25days} = \frac{Avg\ signal_{25days}}{Avg\ signal_{baseline}} \tag{4.3}$$

For both values we derived the relative standard deviation (σ) and defined the thresholds $th_{1,i} = FC_i + \sigma_i$ and $th_{2,i} = FC_i + 2\sigma_i$ as representative of a mild and moderate increase due to external reasons not attributable to the procedure of the study desing, thus, in our case to the impact of RT. For example, average number of photon in the whole body at the baseline was 4^5, FCs were 1.39 ($\sigma = 0.59$) and 2.22 ($\sigma = 0.85$) at 4 days and 25 days, respectively. Similar number were found for Thymus and Abdomen, while different variations were registered for both genital area and head: around 1.60 ($\sigma = 0.65$) at 4 days and 1.75 ($\sigma = 0.75$) at 25 days. Finally, we counted for each group (A,B,C,D) and each ROI the number of mice with FC values higher than the two thresholds th1 and th2. The results of a moderate effect and severe effect are included in Table 4.3 and Table 4.4, respectively.

For mouse-like animal, with short range of life (2 years) and high cell turnover, a period of 25 days could be, presumably, representative of the tran-

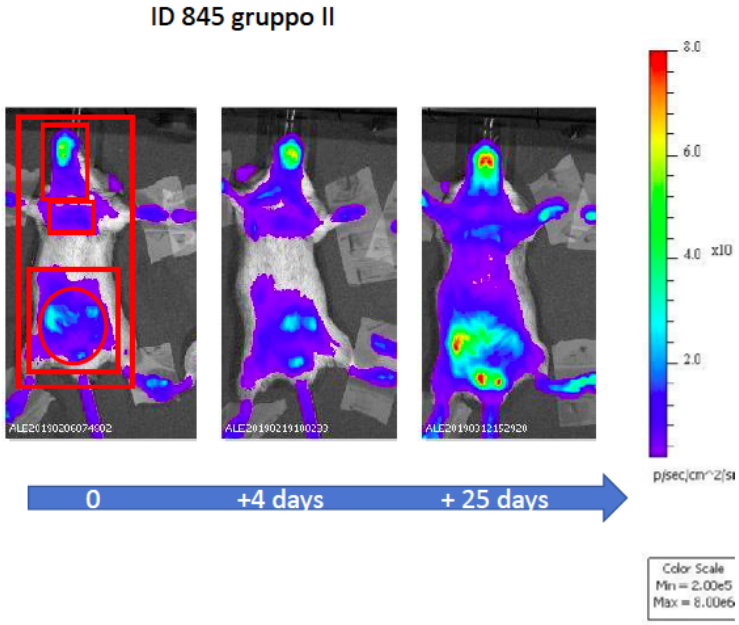


Figure 4.5: *Example of 2D images from in-vivo bioluminescence imaging; output at baseline, 4 days and 25 days. Baseline image includes regions of interest defined in the study.*

Table 4.3: *Number of mice whose fold change exceeded of more than 1σ (moderate effect) the fold change computed for the control group in that specific ROI. First and second value in each column are associated to the time point of 4 days and 25 days, respectively.*

	Whole $(4-25_{days})$ th_1	Thymus $(4-25_{days})$ th_1	Abdominal Area $(4-25_{days})$ th_1	Genital Area $(4-25_{days})$ th_1	Head $(4-25_{days})$ th_1
A	0-4	1-4	0-4	1-5	2-4
B	4-1	1-3	4-0	2-1	2-1
C	0-2	0-4	0-3	0-7	0-3
D	5-1	2-2	4-1	4-2	4-1

Table 4.4: *Number of mice whose fold change exceeded of more than 2σ (severe effect) the fold change computed for the control group in that specific ROI. First and second value in each column are associated to the time point of 4 days and 25 days, respectively.*

	Whole $(4-25_{days})$ th_2	Thymus $(4-25_{days})$ th_2	Abdominal Area $(4-25_{days})$ th_2	Genital Area $(4-25_{days})$ th_2	Head $(4-25_{days})$ th_2
A	0-4	0-4	0-2	1-4	1-4
B	2-1	0-2	1-0	2-0	1-1
C	0-2	0-2	0-2	0-4	0-1
D	4-1	1-1	1-1	2-1	1-0

Table 4.5: *Number of evidences from pathology report for different organs in groups A,B,C, D and in control group.*

Group	Bladder	Testes	Kidney	Other	Total
A	0	7	2	5	14
B	0	7	5	2	14
C	1	2	3	0	6
D	0	1	2	2	5
Control	0	0	2	0	2

sition time between acute and late phase response. Particularly, there is a large consensus in defining the beginning of late events after 3-4 months from the end of the RT. On the other hand, acute events can be scored up to 2-4 weeks after treatment [156, 157] but there are experiments that extend the acute phase at the cutpoint of 3-4 months[158]. Looking at the data with this perspective, there is a trend for schemes A (4x2Gy) and C (4x4Gy) to express an inflammatory response in the tardive time point, while mice in group D (3x5Gy) were more fast in exhibiting the inflammatory condition, with a general recovery at 25 days. Moreover, mice in this group showed a moderate variation at 4 days in all the ROIs, but the severe increase counted in the whole body, and not followed in the other ROIs revealed a spread bioluminescense activation, even outside the contoured regions. This could be explained by the fact that a dose per day of 5 Gy exceeds the limit threshold of most organs within the body. The findings in groups A,C and D suggested an acute-response tissue behaviour of the bladder. However, group B (4x3Gy) showed hybrid variations. It was in phase with A and C in the thymus, while it was more aligned with group D in the other ROIs. Furthermore, in very few cases the variations at 4 days were confirmed at 25 days, rejecting the thesis of consequentiality between acute and late inflammatory response, which is not necessarily correlated with the side effects manifestation. From this point of view, it could be helpful to compare the results of the BFi with the ex-vivo pathological results. In Table 4.5 we summarized the pathology report in 4 columns associated to the bladder, the testes, the kidney and all other organs (preputial gland, lung, skin, stomach). For this analysis we considered only the moderate or severe impairment of the organs.

It is not possible to compare more in detail, with these numbers, the specific tissue analysis with the BFi values in the different ROIs and, obviously, we can not infer the insurgency time from this analysis.

Surprisingly, a severe pathologic condition was found in the mice of group A. This is not something we could expect, but it gives us some indications. A possible bias could be the necessity for the team to reach the operating speed. Indeed, this issue could affect the time under anesthetic gas or also the stress due to the operations with mice. Another option could be the volume of the bladders within the group A, which impacts on the total integral dose. We

retrieved this information from the CT plan but we did not find significant differences with groups B,C and D. However, the agreement with the BFi analysis of group A represents one of the strongest evidence from this preliminary study which aimed at validating the BFi as a possible useful non-invasive technique. The majority of the pathological findings, such as tissue infiltration, mineralization and degeneration or atrophy are tardive effects. This represents a good agreement between the analysis in the whole body at 25 days in BFi and in the histological report. At the same time, the value of Table 4.4 for whole body in group D at 4 days could explain an inflammatory response to the direct effect of RT in the case it is delivered with high dose as for group D which is not followed by a large number of structural changes in the tissue. As a proof for this speculation, we should also take into account that measurements for group D were performed one day after the end of RT, i.e., at 4 days instead of 3, for technical reasons. It means that, possibly, the number of photons at 4 days was also mitigated.

Finally, IHC evaluation of Ki67+ expression on the bladder was also performed. Ki67+ is a protein associated with cellular proliferation. Rationale of the measurement is that the epithelial cells of the bladder have a low proliferative capacity. As a consequence, high signal of Ki67+ should describe an increase in the proliferation velocity of the mucosa caused by a local damage. Samples of the bladder were stained by immunohistochemistry with Ki67+ primary antibody (see Figure 4.6). For this analysis, groups were splitted in low dose (group A and B) and high dose (group C and D). Staining was higher in irradiated mice, and showed an association with dose, though not statistically significant (mean values 3.4% vs 5.6% vs 9.8%, for controls vs low dose vs high dose, p=0.10). More interestingly, bladder Ki67+ staining (expressed as a percentage) was also associated with increased luminescence (expressed by the FC) in the urinary region as measured 25 day after irradiation (r=0.34, p=0.05).

In conclusion, we tested the response of 35 mice to different schemes of treatment for a uniform dose delivered to the urinary bladder. Due to the limitation in number and the not huge differences in the schemes, it was not possible to derive robust radiobiological information. We registered an unexpected number of adverse events in the group which received the lowest physical dose and we also found a different trend in the group with the highest dose per fraction and one less fraction in total. Even with some technical limitations [3], bioluminescence and fluorescence imaging were able to describe the inflammatory response to radiation. A percentage of 64% (18/28) and 57%(16/28) of mice showed values of luminescence that were higher than the control group at 4 days and 25 days, respectively. This variation was partially in agreement with pathological results. Morevoer, IHC assay indicated that in-vivo imaging

[3]If the administered luciferin is not abundantly present, light emission may not be a true representation of luciferase activity. Another issue with luminescence is the limited and wavelength-dependent transmission of light through the animal tissues. As a general rule, there is an approximate 10-fold loss of photon intensity for each centimeter of tissue depth.

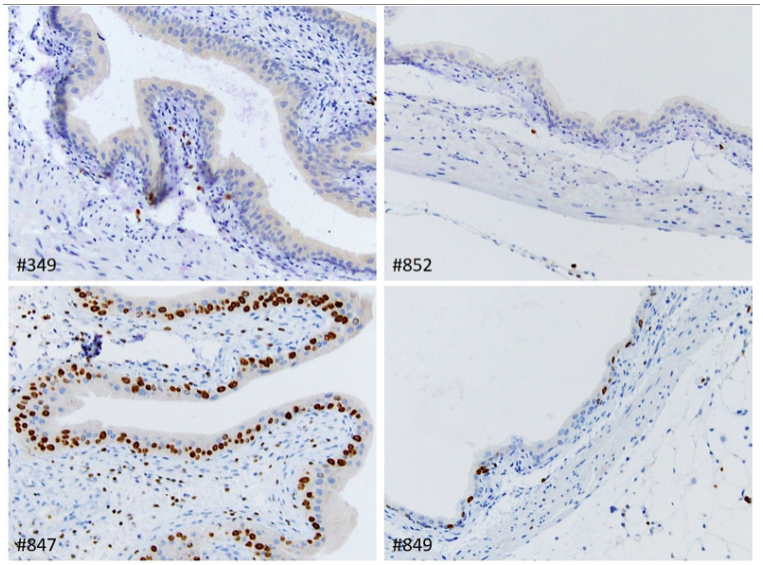


Figure 4.6: *Urinary bladder, immunohistochemical (IHC) for Ki67 in control mice (#349 and #852) and group D treated mice (#847 and #849) showing increased percentage of Ki67-positive epithelial cells. 200x magnification*

could be a good surrogate for ex-vivo findings as described by Ki67+. For what concerns the urine analysis, a direct comparison with BFI could not be possible since the majority of the samples collected in this preliminary study were contaminated by other agents (mainly sediment of food pellet).

4.3 Analysis of Radiation Impact on Microvasculature by mean of Microfluidic Chips

In many studies, changes in the vascular function have been observed before the occurrence of late fibrosis and atrophy tissues[159]. This supports the concept that vascular damage plays an important role in all late radition injury. In addition, several trials found an increased risk of normal tissue toxicity for patients who had medical conditions or specific habits that adversely affect the stability of the vascular system (hypertension, diabetes, smoking, alcohol abuse, obesity, use of cholesterol-lowering drugs, use of drugs for cardiac morbidity)[79, 106, 116, 144]. Understanding how ionizing radiation affects vasculature in the tumor microenvironment and in normal tissues is crucial but it has been limited by the difficulty in visualizing the response of these interdependent components over time and in-vivo. Statistical models (i.e. only based on extensive collection of data and on search for association between covariates and endpoint) are not suitable to describe this complex microenvironment.

On the basis of a recent funding by the *Italian Association for Cancer Research* (AIRC 21479). we proposed to develop a three-dimensional mechanistic

model of the damage induced by RT on the microcirculation and of its effects on the tumour microenvironment and the surrounding normal tissue. The goal is to derive parameters of the model by experimental evaluation of radiation effects on microfluidic devices holding functional microvascular networks and in a clinical setting of prostate, head-and-neck and breast cancer patients. Particularly, in-vivo measurments are also planned to be performed by the use of sublingual microscopy and a spectro-photometer at skin level.
3D computational modelling is achieved in strong collaboration with Modelling and Scientific Computing (MOX), Department of Mathematics, Politecnico di Milano. In the first year of the project, the microvasculature computational modelling, already developed at MOX for other purposes, was extended in order to include the damage to miscrovaculature due to irradiation, and to accurately quantify the radiation induced variations of the parameters that characterize the interaction of the vascular network with the surrounding tissue. The model accounted for: (a) capillaries-to-tissue interactions taking into consideration capillary wall permeability; (b) hydraulic and oncotic pressures; (c) microvasculature morphology and capillary density; (d) blood properties along with hematocrit heterogeneity within the vessels network; (e) extravascular properties; (f) the presence of lymphatic system; (g) the hematocrit dependent flow properties (Fåhræus-Lindqvist effect). The model was implemented using GetFEM++, an open source C++ library, which provides finite element method for solving potentially coupled systems of linear and nonlinear partial differential equations.

Among the entire vascular network, the capillary bed (CB) is the most affected by radiation. In summary, the functional units of the CB go from the meta-arterioles at the inlet to the small venule at the outlet, passing through the smallest entities, the capillaries. The units have different dimensions (r), thickness of the wall (d) and elasticity (Young's modulus, Y), due to the presence or not of smooth muscles. The model was able to describe the different units according to the range of the input parameters r, d and Y.

In this preliminary study we started from the simplest frame to include into the computational model a perturbation due to the RT effect. The model, actually, was already able to describe a more complex network, but here we worked with an y-shape bifurcation, which is set at the beginning of the capillary bed. The response of tissues to radiation is, basically, an inflammatory response. As any other inflammation process, it begins with heat and redness of the area (early effect, range of time: hours-days), then it is followed by the swelling of the structures and the pain (intermediate effect, range of time: days-months) and finally, the loss of the organ functionality (late effect, years). We aimed at describing the very first hours after single fraction.
In the first seconds after irradiation, there is a phase of vasocostriction or hemostasis. Subsequently, it is followed by a prolonged vasodilation, which leads to increased blood flow in the area and to a state of redness and heat in the tumor micro-environment. *Vasodilation* was the first perturbation that

we considered. The second effect taken into account was the *enhanced permeability* of the endothelium wall, so the capability of macro-molecules to pass through the layer where before the RT they were not able to. From this point of view it is worth to mention the in-vitro study by Kabacik and colleagues[160] where a monolayer of endothelial cells (EC) was irradiated with different dose points. Here, macromolecules of different size were perfused in the samples. They came up with a mechanistic model of the effect of ionizing radiation on the endothelial membrane. According to their model, reactive oxygen species (produced by irradiation) are able to increase the number of Calcium Ions, those activate the ADAM10 protein that can destroy the endothelial junctions (acting on the cell adhesion molecule E-cadherin). The higher the dose, the more copious the amount of cargo transferred across the EC monolayer, above all when largest macromolecules (70kDa) are perfused. On the contrary, no effect after a certain dose were evident for the smallest size of molecules (0.45 kDa).
Vasodilation and the enhanced permeability play a crucial role in the accumulation of fluid in the interstitial space. As a consequence, we considered also the *exudation*, i.e., the leakeage of blood components such as platelets, fibrin, leukocytes or proteins into the interstitial space and also in the blood wall. In the course of time, exudation could lead to edema and to the swelling of the membrane (due to the infiltration of components), then to capillary or venules occlusions (less rigid than arterioles) and finally to microcirculation impairment, with the consequent parenchymal damage of the organs due to the lack of an appropriate level of nutrients.
These three effects can be translated in the perturbation of the following parameters:

1. an increase in the arterial pressure, Pv, to mimic vasodilation, from the usual 30 to 25 mmHg to 32 to 27 mmHg;

2. a step variation of the wall hydraulic conductivity, Lp, between 10^{-15} (almost impermeable) and $10^{-13}m^2sKg^-1$ (effect of radiation on the endothelial cells junction);

3. as a consequence of the fluid accumulation in the extravascular space during treatment we increased the interstitial pressure, Pt, from the usual small negative values to +3mmHg (where 5 mmHg is the standard edema value).

The preliminary results of the computational task can be described by Figure 4.7 where we can see the extravascular domain and the bifurcation in the case with $L_p = 10^{-13}m^2sKg^-1$. On the left, the imposed values of the pressure at the border (32 and 27 mmHg) and also the increasing transmural pressure from 3.00 to 3.037 mmHg (almost 1% after one single cicle) were highlighted. On the right you can see the arrows showing the liquid motion from the small arterioles to the extravascular space. Variation in parameters can describe increasing damage related to increased accumulated dose; this could be reached

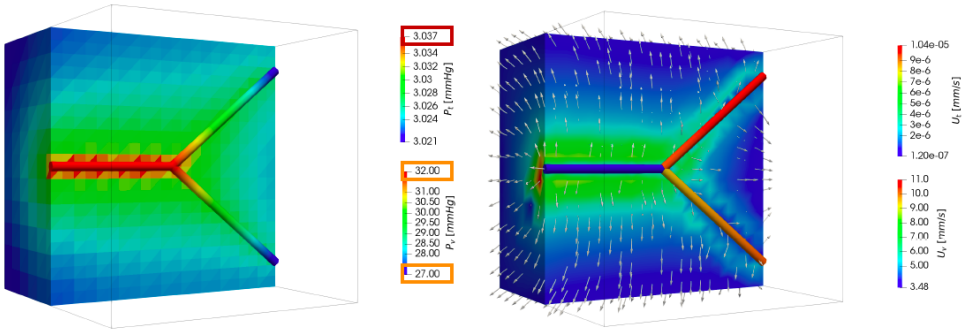


Figure 4.7: *Output of the computational algorithm. Pictures were created in the case with* $L_p = 10^{-13} m^2 s K g^{-}1$ *. In the 3D representation on the left is shown the arterial pressure of the bifurcation together with the interstitial pressure, on the right the two velocity of the system.*

at different times after beginning of RT according to dose per fraction, source of radiation, concomitant treatments and the patient baseline situation.
Concerning the in-vitro setup, in the last decade, significant advancements were made both in the microfabrication and materials science, leading to the design of vascularized tissues. A microfluidic device was considered to describe functional and perfusable microvascular networks. It consists of two lateral media channels and a central gel channel. It is made of PDMS (polydimethyl-siloxane) through soft lithography techniques and micropatterned silicon wafers. The PDMS structure is bonded to a glass covership after oxygen plasma treatment. Moreover, microfluidic channels are treated to promote matrix adhesion. Three main activities have regarded the in-vitro setup.
First, we ran preliminary tests to identify possible radiotherapy-induced effects on the chip and not related to the microvascular network (MVN). Indeed, the full in-vitro setup is composed of cells (EC and supporting cells), media, 3D gel/scaffold, and the chip structure (PDMS and glass). Aim of such preliminary tests was to identify possible alterations of these components when exposed to radiation, so that effects due to the MVN irradiation can be separated in future experiments. For this reason, we considered ad hoc experiments with microfluidic chips not including cells, measuring mechanical properties alterations, gel structure changes (by imaging), solute release within media (e.g. gel debris or reactive oxygen species), and cell viability by using the treated media. Furthermore, to better analyse possible mechanical alteration employing mechanical testing (e.g. compression), a bigger version of the chip has also been realized, so that gel samples reached affordable dimensions for the tests.

A second activity was focused on the experimental design. In particular, we defined ECs for MVN generation since different options can be considered. We chose to test MVNs obtained from two different EC type: human umbilical vein endothelial cells (HUVEC) and human microvascular cells (HMEC). The former are the most commonly used cells for MVN because of their robustness

and their ability to form networks by a vasculogenic-like process (protocols for their handling are available in the literature and widely accepted). The latter are more appropriate for this experimental analysis since they are truly microvascular cells. Anyway, their use is less common than HUVEC and their handling is more difficult. Finally, we identified a portable incubator as a critical component to mantain cell viability during the treatment. This component assumes even greater importance given that we will consider different treatment delivery techniques, comparing a single dose to a clinically significant fractionation scheme.
In the continuation of the work, irradiation will be performed between 0.5 Gy and 4 Gy at 0.5 Gy steps (8 dose points). For each dose point, three different uniform field geometries will be considered: (i) a field encompassing the whole vascular system; (ii) a field involving half vascular system at its periphery and (iii) a field involving half vasculature at its center. For each experimental configuration (dose+geometry) 6 chips will be considered, for a total of 144 chips to be irradiated. Six non-irradiated chips will be considered as control group, for a total of 150 chips.
To fully characterize the problem specific in-vivo measurements will be performed on patients. This characterization will allow us to derive supplementary data but also to correct the computational model by secondary effects that can not be described by a closed system such as the one of the in-vitro experiments. Thus, from the clinical setting two tools will be used to obtain in-vivo measurements from a cohort of head-and-neck and breast cancer patients (126 to be enrolled in 26 months for both the diseases). Skin reflectance measurements will be acquired with a spectrophotometric imaging system able to register multiple images of the skin at different wavelenghts (483-950 nm). This measure will be performed in both districts. H&N patients will have also measures of microvascular density, red cell content and red cell flow velocities through GlycoCheck sublingual microscopy. Weekly dose to sublingual region will be calculated from the planned dose distribution. In addition, the variety in tumour location will allow a large spectrum of dose values coupled to in-vivo measurement of microvascular properties. For both measurements, the time points will be: before RT, every week during RT and at 1-3-6-12-18-24-30-36 months.
To provide some practical example, we will have the chance to measure "blood" velocity in both in-vitro and in-vivo experiments. In the first case we will perfuse a bolus injection of a solution of a specific molecular weight (we will test several kDa), while in the second one we well consider the velocity of red blood cells (cells with larger size in the blood and with intrinsic elastic properties of cell membrane). These two settings will provide discrepancies in the results due to the differences in the viscosity of the system (both for the elasticity of the hematocrit but also for the presence of a realistic melting pot of substances such as leukocytes and platelets). Moreover, the presence of a realistic interstitial space in clinical measurements could contribute altering the angio-

genesis, and consequently, the microvascular density. In general the in-vitro experiments will supply a solid description of the direct effect of the ionizing radiation on the endothelial layer and a range of values for some fluid dynamic and mechanical properties of the network. Clinical setting will furnish a more reliable description of these properties and, indirectly, will remove a certain set of values from the range derived by in-vitro measures giving at the same time a possible comprehension of secondary effects that are not considered in microfluidic chips. In conclusion, measurements collected in the clinical evaluation phase will be compared with the computational model predictions in order to have a quantitative feedback on the model equation parameters, possibly leading to decreased model uncertainty. A quantitative comparison of patient data with model outputs represents a significant challenge because clinical data and model prediction may refer to very different spatial and temporal scales. To tackle this difficulty, we will identify correspondences between model outputs and measured clinical data. Then, a qualitative comparison will be performed, looking at the trends of variation of the model outputs and clinical data with respect to the patient variability.

Chapter 5

Result at Cell Level: in-vitro studies and Monte Carlo Simulations

In Chapter 4.1 we presented a study on the texture analysis of T1 and T2 weighted magnetic resonance images to investigate RT induced changes in the obturator muscles [149]. An evidence from that work were the exponential relationship between mean signal intensity and local RT dose, confirming the hypothesis that imaging can be used to determine in an objective way the muscle insult by RT. Moreover, taking into account that there are more than 600 muscles in the human body (40% of the body mass) and that they have ubiquitous nature, it is difficult to irradiate a tumor target without including skeletal muscles in the medium-high dose region. For this reason, RT has the potential of influencing the biology of human skeletal muscle cells (HSkMCs) and to produce negative effects both locally (muscle specific) and regionally (impairment of physiological functioning, as also speculated in Section 3.2.3 [126]). Thus, it is somehow surprising that the effects of radiation on HSkMCs are poorly studied and understood. On the basis of these considerations, a proposal entitled "Survival and Radiation Damage Analysis of Human Skeletal Muscle Cells after Photon and Heavy Ion irradiation: Experimental Data and Monte Carlo Simulations" was designed. The project was funded by 5x1000 Fondazione IRCCS Istituto Nazionale dei Tumori di Milano for a total of 39.000 €[1]. It was developed in partnership with Università di Pavia, INT and *Centro Nazionale di Adroterapia Oncologica* (CNAO). Aim of the proposal was to fully characterize the survival dose-response of HSkMCs and to include these information into a Monte Carlo simulation (MC) of RT induced cell death. In this feasibility project we aim at investigating the possibility of using HSkMCs for in-vitro experiments on cell survival and cell damage after irradiation. An important aspect to highlight is that it was the first time that this cell line was

[1]Principal Investigator of the study: Alessandro Cicchetti.

used for a cell survival assay. As a consenquence of that, to reduce the uncertainties on the cell line characteristics we decided to purchase immortalised HSkMCs by a commercial vendor. An immortalised cell line is a population of cells from a multicellular organism which would normally not proliferate indefinitely but, due to mutation, have evaded normal cellular senescence and instead can keep undergoing division. The cells can therefore be grown for prolonged periods in-vitro. Particularly, *Human telomerase reverse transcriptase* (hTERT) was the approach used on the HSkMC line. Indeed, telomerase is a ribonucleoprotein that is able to extend the DNA sequence of telomeres, which make the cells to undergo infinite cell divisions through abating the senescence process. This protein is inactive in most somatic cells, but when TERT is exogenously expressed, the cells are able to maintain sufficient telomere lengths to avoid replicative senescence.

The main tasks of the study are described in the following sections:

- Cell survival curves, where the analysis of cell survival was performed for Carbon Ion (samples were irradiated at CNAO) and for photons (samples irradiated at INT);
- Functional analysis of DNA damage and cell death, which included the analysis of γ-H2AX foci and the western blot assay for apoptosis;
- Monte Carlo simulations, where a comparison between the cell survival curves and the results of a biophysical model was performed in order to derive biological parameters.

All these investigations included single dose-point irradiations performed between 0 and 6 Gy. However, a task about the effect of fractionation schemes was also considered in the proposal. The aim was to compare different fractionation schedules with a total physical dose equal to 6 Gy (1x6 Gy, 2x3 Gy and 3x2 Gy). This part of the project is still on going and it is not described in the thesis.

5.1 Cell Survival

One of the main endpoints in the evaluation of radiation treatment is the assessment of cell survival curve after irradiation. HSkMCs cells were never studied under the irradiation effects. As a consenquence, the first step of the project consisted of studying HSkMCs viability and clonogenicity. Immortalized HSkMCs deriving from healthy donor were characterized. First of all, to evaluate the cell proliferation rate, the doubling time of each cell line was assessed, through the measurement of the time it took for the cells to duplicate. We calculated that the cell line duplicated in about 36 h, which is consistent with the indications of the company that provided us the cells. Then,

HSkMC cells were evaluated for their ability to form colonies. To this end, the cells were plated in 6-well poly-lysine coated plates at different cell densities (50-100-200-400-800-1000-2000-4000 cells/well) and let to grow up to 20 days. Unfortunately, still maintaining their proliferative potential, the cells were not able to form colonies at any of the considered densities, rather showing an irregular distribution, likely due to their typical elongated shape. To manage

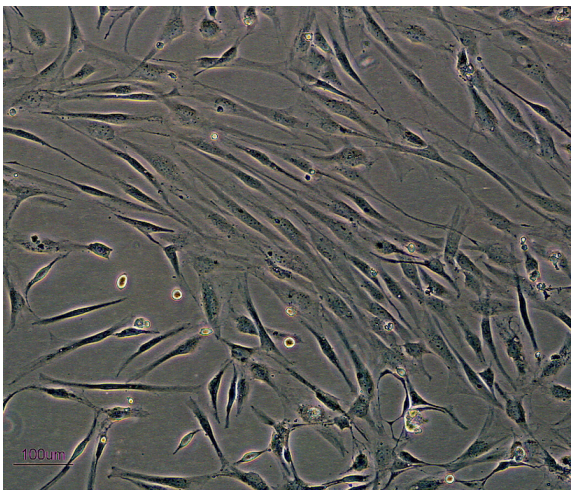


Figure 5.1: *Colony formation assay for immortalised Human Skeletal Muscle Cells. Cells in the microscope image appear as singular individual bodies rather than a cell colony. Image from reference [92].*

this, we performed the clonogenic assay on wells coated with a layer of soft agar, a gel in which cells are partially trapped, thus minimizing/limiting their elongation. Briefly, 2 ml of medium mixed with sterile 0.5% agar (diluted in RPMI medium) was plated in each of the 6-wells. When the agar coating was ready, 1 ml of HSkMC cells mixed up with 0.5% agar (1:1 ratio) were plated in each well and let to solidify for 1 h at room temperature. When ready, the cells were incubated at 37° and let to grow up to 20 days. When observed under the microsope, the cells appeared trapped in the opaque and vacuolated, indicating that they were almost dead, and no colonies were formed. To evaluate the cell response to the late effects of radiation treatment on HSkMCs bypassing the lack of colony formation ability, we were needed to contrive a surrogate assay able to provide us the information on muscle cells sensitivity to the radiation treatment, other than the conventional clonogenic assay. Therefore, we proposed to derive proliferation data by the use of non-clonogenic assays as proposed by Price and McMIllan [92]. Indeed, the two researchers proposed the so-called *growth assays* for measuring the response of cells to ionising radiation in the case of cell lines which do not grow as colonies. Growth assays estimate survival by comparing the total number of viable cells in treated and untreated cultures. Various methods of measuring the number of viable cells in culture were developed and summarised in the paper [92]. Growth curves are constructed from serial measurements of cell numbers (see Figure 5.2). Simply

comparing cell numbers at any fixed time after irradiation will not produce meaningful surviving fractions. This is due both to the dose-dependent lag period after irradiation before regrowth is obtained (see in Figure 5.2 the different time for the minimal value in curve X Gy and Y Gy, where Y>X) and to the time taken by control cultures to reach confluence. In fact, the endpoint in a growth assays is the ability of the total cell population to regain the growth rate of the control population. Surviving fractions are thus obtained when treated cultures attain exponential regrowth at the control growth rate. When the cell number is counted on just one day using one inoculum size, spurious results are obtained. If the timing is too short after treatment, doomed cells will have not been cleared from the culture and an overestimate of survival will be obtained; and if its too long, control cells will have reached confluence again leading to an over-estimate of the surviving cell number. Figure 5.3 shows an example of the survival curves which can be obtained in these situations. To

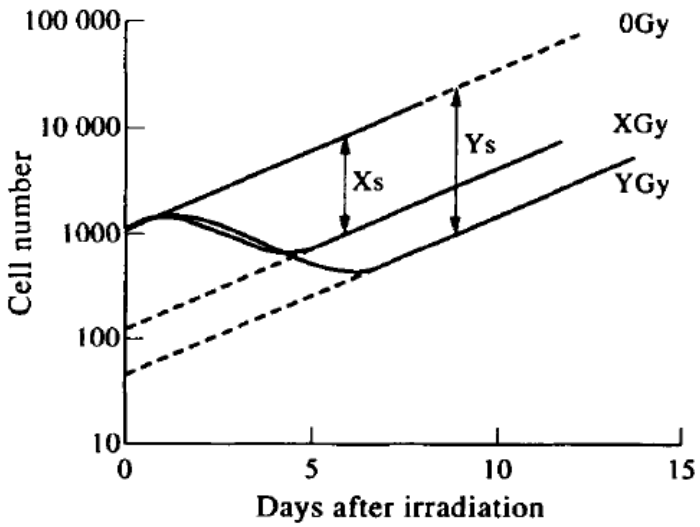


Figure 5.2: *Calculation of radiation-induced cell kill from cell growth curves. Cell numbers are counted at regular intervals and growth curves are constructed for untreated cells (O Gy) and cells treated with increasing doses of irradiation (X Gy and Y Gy). Survival fractions at X Gy and Y Gy are measured as Xs and Ys.*

overcome this, the *vertical displacement of growth curves* approach was used to derive cell survival curves. When treated cultures regrow at the same rate as controls, evaluation of the vertical displacement of the curves can provide an estimate of the level of cell kill (Figure 5.2). Multiple estimations of survival can be made at various time intervals during the exponential growth phase. Problems arise with this method if treated cultures have not achieved exponential growth at the time when the rate of growth of the control cultures is decreasing due to confluency. Forward extrapolation of control curves may be required to overcome this. Also, a single day may not be sufficient to obtain surviving fractions for the full range of doses: the cultures exposed to lower

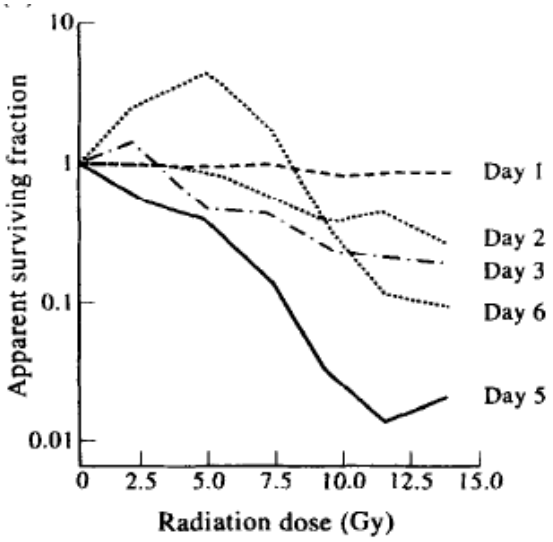


Figure 5.3: *Apparent survival curves for analysis of growth data at different times. An example of the radiation effect on a human bladder carcinoma cell line, MGH-UI, where the ratio of cell numbers has been calculated on several days. Image from reference [92].*

doses may have regained exponential regrowth, but the cultures exposed to higher doses are still in the lag phase, so that the first part of the curve will be an accurate measure of survival but the latter part will not.

A time point of 16 days from cell exposure to radiation was defined. It was chosen assessing cell proliferation up to 16 days from cell exposure to radiation. Briefly, cells were plated in 25 cm^2 flasks and, after 48h, irradiated at different doses (0.5-1-1.5-2-4-6 Gy) with photons. Irradiated cells were then detached, counted with the Coulter cell counter and re-plated at different cell-concentrations (2000-4000 cells/ml/well) in 6-well plates, and let to grow up to 16 days. At different time points from irradiation (1-4-7-12-16 days) cells were counted with the same Coulter in order to assess the effect of the treatment on cell proliferation at different stages. From this preliminary test, we found that after 16 days the cell number was over the seeded one for all dose-points.

To determine the HSkMCs response as a function of the source irradiation, cells were exposed to photons or carbon ions. In the first case, the irradiations were delivered at INT by the use of the Co^{60}-irradiatior, which exposed the cultures to a constant dose rate of 4.8 Gy/min. The longer the permanence into the irradiation box of the machine, the higher the absorbed dose. On the other hand, carbon ion irradiation was performed at CNAO. Cells were inserted in a water phantom and irradiated during the night (out of the scheduled time for clinical treatments) in one of the three treatment rooms. Since muscles could be included as not into the PTV, we decided to investigate the dose response of HSkMCs placed both in the entrance channel of the Bragg curve (4 cm, LET = 14 keV/μm) and in the middle (15 cm, LET = 45 keV/μm)

Table 5.1: *Survival ratios for Human Skeletal Muscle Cell line with different radiation source and different LET. Not applicable (n.a.) was inserted in the table where cell cultures did not reach the confluence at 16 days after irradiation.*

	Survival Ratio (ratio, ± 68 % CI)		
Dose (Gy)	**HSkMCs (γ irradiation)**	**HSkMCs (Carbon ions irradiation (low LET))**	**HSkMCs (Carbon ions irradiation (High LET))**
0	1	1	1
0.5	0.75 ± 0.05	0.79 ± 0.02	n.a
1	0.69 ± 0.05	0.53 ± 0.03	0.61 ± 0.03
1.5	0.51 ± 0.06	0.48 ± 0.03	0.25 ± 0.01
2	0.31 ± 0.06	0.32 ± 0.03	0.11 ± 0.02
4	0.10 ± 0.10	0.03 ± 0.07	0.01 ± 0.02
6	0.02 ± 0.17		

of a spread out Bragg peak (SOBP, 12 to 18 cm). This choice was made taking into account the DVH of obturator muscles as shown in Figure 3.1, where it is clear that a not negligible part of the structure is irradiated at values of dose close to the prescription one. Cells were irradiated at 6 dose points (0, 0.5, 1, 1.5, 2, 4, 6 Gy) for photons and the carbon in in the entrance channel, 5 dose points (0, 0.5, 1, 1.5, 2, 4 Gy) in the SOBP. Physical doses were considered, thus no conversion was applied on the basis of the RBE of carbon ions. Finally, for statistical significance, all experiments were carried out in triplicate. The obtained results are illustrated in Table 5.1. Not all the measurements could be accepted and included in the table since at 16 days there still were cultures which did not reach confluence. Particularly, this was true for the cultures irradiated with 4 Gy or more. In fact, as already said, the lag period after irradiation is dose-dependent, and in our case also LET dependent. As a consequence if a 16 day time was acceptable for photons, it could be too short for Carbon ions. We could not considered either 4 Gy in the irradiation of HSkMCs with carbon ions and with flask in the middle of SOBP, while when the cells were placed in the entrance channel we excluded values for 6 Gy.

5.2 Functional Analysis of DNA Damage and Cell Death

The measurement of γ-H2AX foci levels in cells provides a sensitive and reliable method for quantitation of the radiation-induced DNA damage response. Even for this assay we planned to irradiate HSkMCs samples both with photons and with carbon ions. A dose of 4 Gy was selected. This is a widespread value for testing the DNA repair kinetics by mean of γ-H2AX foci. In fact, it guarantees a good amount of DSBs without reaching saturation levels at

the peak of the signal (usually within 15-60 min). After irradiation, cells were fixed by paraformaldehyde[2] at different time points (0 min - 30 min - 60 min and 48 hours) to study the γ-H2AX induction and loss kinetics. Since no data are reported in the literature about immunoflorescence of HSkMCs after irradiation, we decided to set the last point of the analysis at 48h in order to be sure to reach the plateau, which is representative of the un-repaired damages. After fixation cells were permeabilized with Triton and then blocked in Phosphate Saline Buffers. Cells were then stained with DAPI (4,6-diamidino-2-phenylindole). DAPI is a blue fluorescent probe that fluoresces brightly upon selectively binding to the minor groove of double stranded DNA, where its fluorescence is approximately 20-fold greater than in the nonbound state. Its selectivity for DNA and the high cell permeability allows efficient staining of nuclei with little background from the cytoplasm. It is excited by the violet (405 nm) laser line and it is commonly used as a nuclear segmentation tool in high-content imaging analysis. After that, γ-H2AX foci were stained through the use of an associated green fluorescent antibody (530 nm).
The experiment we designed was with the use of photon irradiation. To perform the analysis we defined with the confocal microscope a set of *Field of View* (FoV) inside the Nunc flask. In each selected FoV we detected the number of phosphorylations (which occurs specifically at sites of DNA DSBs) within the nucleus. The average number for each time point was derived, while the ratio of unrepaired damage was computed dividing every counts by the longitudinal maximum, that for this experiments was obtained at 30 min. Indeed, the ratios at 0 (considered as background), 30 min, 60 min and 48 hours (considered as plateau) were 10.0%, 100.0%, 48.5% and 23.1%, respectively. The kinetics curve together with a set of images of the stained cells is shown in Figure 5.4. The last point of the kinetic plot was of high interest for the project because its radiobiological meaning (i.e. the capability of the cell to complete a certain amount of DSBs repair) might be regarded as an input for the Monte Carlo simulations. More details about this possible implication are described in the next section.

Moreover, apoptotic cell death was investigated by the western blot analysis of Caspase-3. It is one of the most common approach to evaluate the apoptosis since Caspase-3 activation is induced by both the intrinsic pathway (DNA damage from external stress such as irradiation) and the extrinsic pathway (endogenous ligands that bind to their corresponding death receptors) as

[2]The broad objective of fixation is to preserve cells and tissue components in a "life-like state" and to do this in such a way as to allow for the preparation of stained samples. In this way we can produce consistent chemical and physical characteristics in tested samples which allow to perform biological assays and also comparisons. These observations allow us a view of a dynamic ever-changing environment "fixed" at a particular point in time. For practical purposes fixation aims to prevent or arrest the degenerative processes which commence as soon as a tissue is deprived of its blood supply, which must be removed to make an effective study of the cell line of interest.

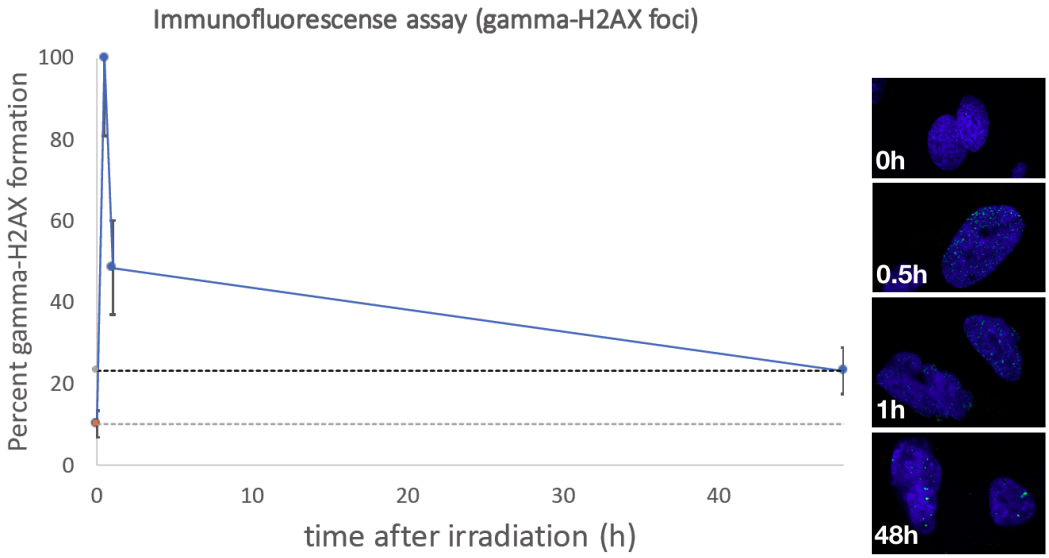


Figure 5.4: *Graphical representation of the kinetics curve and microscope images of γ-H2AX foci of Human SKeletal Muscle Cells after photon irradiation (4 Gy) at 4 different time points: 0 min, 30 min, 1 hour and 48 hours.*

shown in Figure 5.5. We decided to first assess qualitatively if the apoptotic pathways were significant in HSkMCs at a reference dose of 4 Gy. If so, a set of measurements at different dose-points were planned in order to separate from the "raw" cell survival curve the contribution of apoptotic death and, consequently, perform on these new ratios the fit of the radiobiological parameters by Monte Carlo simulation. Caspases are responsible for cleavage of numerous cellular proteins, leading to the biochemical and morphological hallmarks of apoptosis. Here, we tested Caspase-3 cleavage, *Poly (ADP-ribose) polymerase* (PARP) and PARP cleavage. The PARP role consists in detecting and initiating an immediate cellular response to radiation-induced SSB by signaling the enzymatic machinery involved in the SSB repair. Cleaved Caspase-3 is able to inactivate PARP (the so-called cleaved PARP) leading to programmed cell death in systems where DNA damage is extensive. In Figure 5.6 the comparison of PARP and cleaved PARP between control sample and treated cells at 4 Gy is shown. Measure was performed 20 minutes after the irradiation. No differences were found between the two samples in this Western blot analysis. *Glyceraldehyde 3-phosphate dehydrogenase* (GAPDH) was the protein used to normalize the abundance of control and irradiated samples. The contribution of cleaved Caspase-3 is not shown since it was even less evident than the cleaved PARP one.

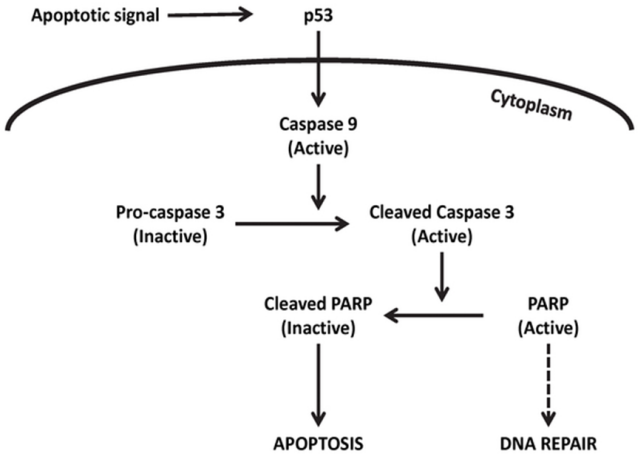


Figure 5.5: *Schematic diagram of the apoptotic pathway tested in the study.*

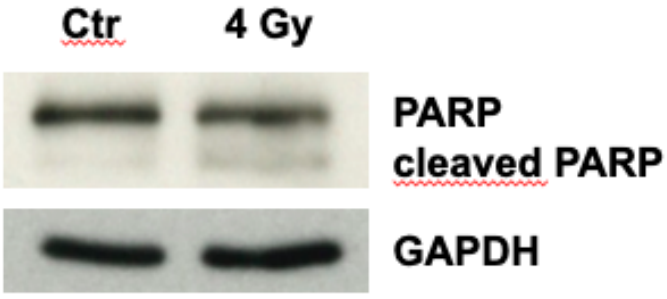


Figure 5.6: *Western blot assay for Human Skeletal Muscle Cell line: comparison between control and cells irradiated with 4 Gy of photons.*

5.3 Monte Carlo Simulations

A biophysical model of radiation damage, implemented as a Monte Carlo simulation code, was used to describe cell death. The code, which is called BIANCA (*BIophysical ANalysis of Cell death and chromosome Aberrations*), has been developed at Università di Pavia and at Istituto Nazionale di Fisica Nucleare - Sezione di Pavia [161]. It is characterized by the following basic assumptions: (i) ionizing radiation can induce DNA *critical lesions* (CLs), where by definition each CL disrupts the continuity of the chromatin fibre producing two main independent chromosome fragments; (ii) distance-dependent mis-rejoining of fragments, or fragment un-rejoining, gives rise to chromosome aberrations; iii) certain aberration types (i.e. dicentrics, rings and large deletions[3]) lead to clonogenic cell death. A schematic representation of these aberration types is reported in Figure 5.7, while a detailed discussion on these assumptions can be found in [161, 162].

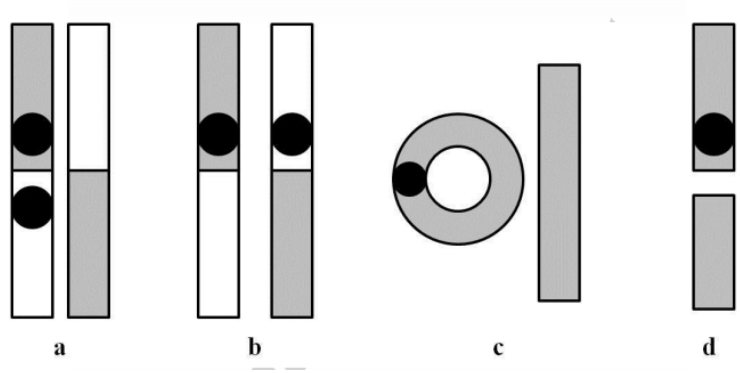


Figure 5.7: *Schematic representation of a dicentric (a), a translocation (b), a centric ring (c) and a terminal deletion (d). Image from reference [161]*

The *CL yield* (expressed as the mean number of CLs per Gy and per cell, which can be converted into CLs per unit length of particle track in case of ion irradiation) is an adjustable parameter of the biophysical model. Its value mainly depends on radiation quality (i.e. particle type and energy), but it is also modulated by the target cell features. Indeed, the value of the CL parameter, rather than relating to the number of initially induced DNA (cluster) damages, also incorporates further biological factors such as the specific cell repair proficiency. The CL yield therefore represents the number of initial damages that, following both the initial damage severity and the subsequent biological factors, break the chromosome in two large independent fragments that participate in the formation of chromosome aberrations, which in turn can lead to cell death. In general, the CL yield tends to increase with the radiation LET (except at very high LET values) and with the cell radiosensitivity.
Concerning the distance-dependence of chromosome fragment end-joining, a

[3]Where "large" means larger than a few Mega base-pairs

step function was defined. Thus, a threshold distance d has been fixed to the mean distance between two adjacent chromosome territories. As a consequence, each chromosome fragment is assumed to have a certain probability, f, to remain un-rejoined even if there are possible "partners" within the threshold distance d. The value of f is assumed to be dependent on the cell-line but independent of radiation quality, and f is the second, and last, adjustable parameter. Finally, the assumed relationship between cell death and chromosome aberrations is derived from experimental observations indicating a one-to-one relationship between the logarithm of the surviving fraction and the mean number of dicentrics, rings and deletions visible in metaphase [163].
The main input data for running a simulation for chromosome aberrations and cell survival consist of the radiation type (photons, light ions or heavy ions), absorbed dose and CL yield, as well as the cell nucleus shape (which can be spherical or cylindrical) and size. For each cell irradiated with a given photon dose, an actual number of CLs is extracted from a Poisson distribution, and these CLs are uniformly distributed in the cell nucleus. On the contrary, for each cell irradiated with a given dose of an ion type of given LET, an actual number of primary ions traversing the cell nucleus is extracted from a Poisson distribution, and for each primary ion an actual number of CLs is also extracted from a Poisson distribution. In case of irradiation with carbon ions each CL has a 50% probability of being induced at a certain radial distance from the primary ion, to take into account the effects of the so-called "delta rays". Afterwards, the chromosomes and chromosome-arms hit by each CL are identified, the process of chromosome fragment end-joining (or un-rejoining) is simulated, and the scoring of different aberration categories is reproduced. Finally, the cell is counted as a surviving cell if it does not contain any "lethal aberration" (i.e. dicentric, ring or large deletion), otherwise it is counted as a dead cell. The process is repeated until the required statistical significance is obtained, and the repetition for different dose values provides simulated dose-response curves for chromosome aberrations and cell survival.
The code has been used in the past to describe the clonogenic cell death of V79 cells and human fibroblasts. To adapt the BIANCA II Monte Carlo simulation code to deal with HSkMCs a geometrical study was set up. By the use of a confocal microscope we acquired a set of 2D images (as the ones showed in Figure 5.4) of stained nuclei. The images were separated by steps of 0.2 μm. We computed through the software ImageJ the average height of the nuclei within a FoV and also the average major and minor axis of the nucleus 2D projection in order to approximate the nucleus as a cylinder with elliptical base. The results of the study were 6.6, 18.8 and 11.6 μm, respectively.
As we saw in Section 5.1, the cells were exposed to ^{60}Co γ-Rays and Carbon ions with two different LET values (14 and 45 keV/μm).

Monte Carlo simulations for photons irradiation

As a preliminary estimate we fited the experimental data by the use of the LQ Model. The derived α/β ratio was 8.5 (α= 0.409 and β= 0.048). Thus, we started imposing the f = 0.18 since this value has been previously used for human normal cells irradiated with photons which expressed a high α/β (α/β=9) (AG01522)[164]. Consequently, the CL parameter was adjusted in order to describe the experimental data. This was achieved by fixing the CL yield as CL = 2.3 $Gy^{-1} \cdot Cell^{-1}$.
As a second approach, we experimentally derived the f value from the immunofluorescence analysis at 4 Gy. The percentage of the unrepaired damage was defined as the difference between the value at 48h and the background. The obtained value (as described in Figure 5.4) was f = 0.13 (0.23 - 0.10). Again, we adjusted the CL parameter in order to describe the survival ratios, which happened for CL = 2.6 $Gy^{-1} \cdot Cell^{-1}$. Experimental data and Monte Carlo simulations[4] for both the cases are shown in Figure 5.8. As we expected, the difference in the f values (f = 0.13 vs f = 0.18) is associated to a different shape of the curve. Particularly, the larger the f the smaller the quadratic component of the curve. This results in the fact that the curves cross each other. In any case, both simulations are consistent with the data, suggesting that the immunofluorescence assay of γ-H2AX foci can be used to estimate the value of f.

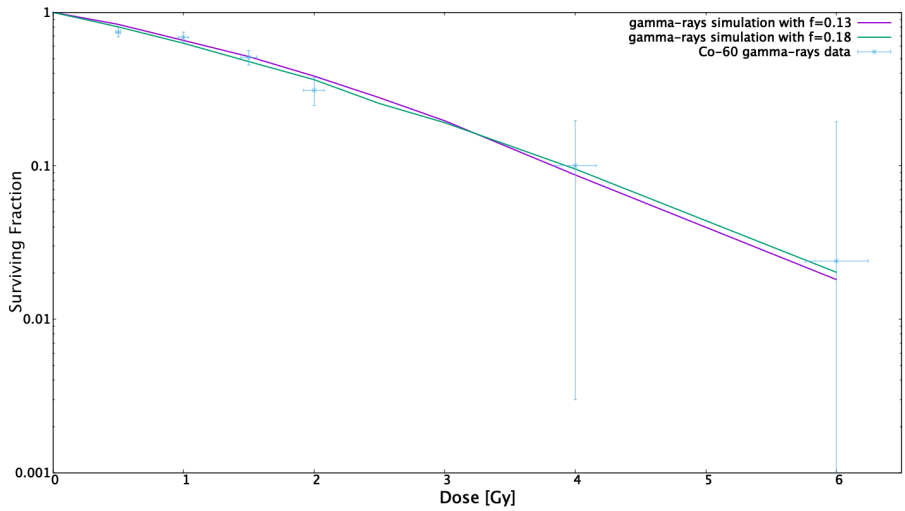


Figure 5.8: *Experimental data and Monte Carlo simulations for Human Skeletal Muscle cells irradiated with ^{60}Co γ-rays. Vertical error bars represent 68% CI of the surviving fractions. Horizontal error bars take into account the uncertainty of the dose (4%). Monte Carlo simulations have an error less than 5%*

[4]The number of cells built in the simulations was chosen so as to ensure an error less than 0.05. This approach was also applied to the C-ions simulations.

5.3.1 Monte Carlo simulations for carbon ions

According to the model the f parameter should be independent on radiation quality, thus, its values were left unchanged in the ions simulations. On the other hand, the CL yield is influenced by the radiation source and has to be computed for each radiation quality. More specifically, we derived the CL parameter by considering the formula developed in Carante et al [164] to fully predict the ion-survival of the cell line of interest based on the ion-survival of a reference cell line, as well as the photon response of both. More specifically, the CL yield to predict the survival of the cell line of interest following exposure to a given radiation quality (i.e., a given ion type and energy) can be derived as follows:

$$\frac{CL}{\mu m} = \left(\frac{CL}{\mu m}\right)_{ref} \cdot \left[\left(\frac{CL}{Gy \cdot Cell}\right) \Big/ \left(\frac{CL}{Gy \cdot Cell}\right)_{ref}\right] \cdot \frac{V_{ref}}{V} \tag{5.1}$$

Here, $(CL/\mu m)_{ref}$ is the CL yield used for the reference cell line exposed to the same radiation quality, $(CL \cdot Gy^{-1}Cell^{-1})$ and $(CL \cdot Gy^{-1}Cell^{-1})_{ref}$ are the CL yields used to simulate photon exposure for the cell line of interest and the reference cell line, respectively, and finally V_{ref} and V is the nucleus volume used for the reference cell line and the cell line of interest, respectively.

In that specific case, Carante and colleagues constructed a radiobiological database describing the survival of V79 cells which were considered as a reference cell line. In this work, applying the formula to the CL values of the simulations with photons we obtained the new values to be used as code inputs for carbon ions irradiation at low and high LET:

- low LET (14 $keV/\mu m$)
    1. for simulation with f = 0.18 –> CL = 0.0057 $CL/\mu m$
    2. for simulation with f = 0.13 –> CL = 0.00645 $CL/\mu m$
- high LET (45 $keV/\mu m$)
    1. for simulation with f = 0.18 –> CL = 0.030 $CL/\mu m$
    2. for simulation with f = 0.13 –> CL = 0.034 $CL/\mu m$

In the experiment at low LET irradiation (flask placed in the entrance channel of the beam) both simulations (f = 0.18 and f = 0.13) are in line with data, especially considering that these are full predictions with no parameters adjustment. The agreement between the model and the data is not fullfilled for the points at 1.5 and 2 Gy. However, these two experimental points are very similar to the corresponding points for γ-rays (see Table 5.1), whereas in general one should expect that low LET C-ions are more effective than ^{60}Co γ-rays.
Concerning high LET (flask placed in the middle of the SOBP), both sim-

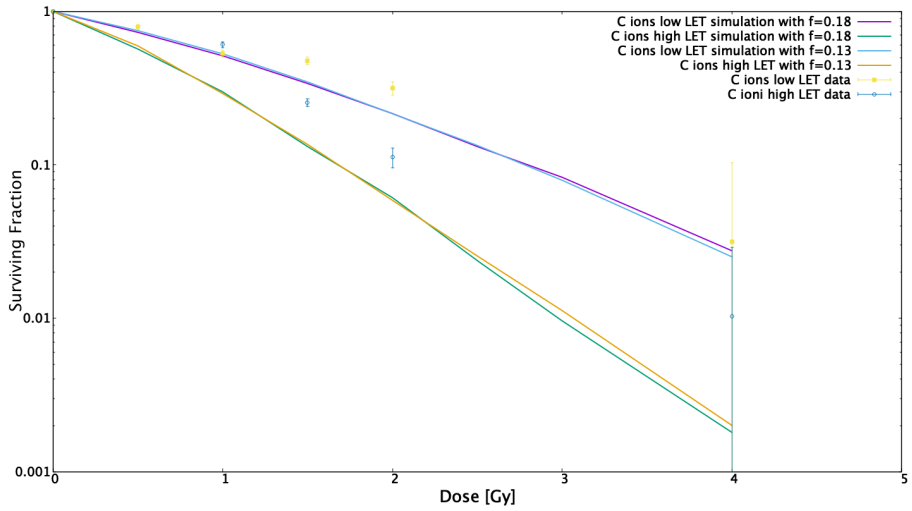


Figure 5.9: *Experimental data and Monte Carlo simulations for Human Skeletal Muscle cells irradiated with carbon ions. Cells were irradiated in two different settings with low LET (14 keV/µm, upper lines and points) and high LET (45 keV/µm, lower lines and points) carbon ions. Vertical error bars represent 68% CI of the surviving fractions. Monte Carlo simulations have an error less than 5%.*

ulations are overestimating the effectiveness. One of the possible reasons is related to the fact that these full predictions are based on a CL database that was specifically constructed for cells with low α/β ratio, which is not the case of the cells considered in this work. Indeed cells with a low α/β ratio, like V79 cells, tend to show a more pronounced difference between the response at low LET and that at high LET. More specifically, while at low LET these cells are rather radio-resistant, for these cells the increase of RBE with LET is generally steeper with respect to cells with a high α/β ratio. Since the Monte Carlo simulations performed in this work are based on a reference cell line (V79) that is characterized by a low α/β ratio, this steeper increase of RBE with LET is unavoidably reflected in the results. This explains why, when the LET increases from 14 to 45 keV/micron, the simulations overestimate the effectiveness (i.e., underestimate the cell surviving fraction) with respect to the experimental data, which were obtained with cells with a high α/β ratio and thus show a slower increase in effectiveness with increasing LET. We expected that the construction of a database for cells with high α/β ratio, which is one of the foreseen future developments, will lead to a better agreement with the data.

Conclusions and future perspectives

In this PhD thesis the biological effects of ionizing radiation on normal tissues involved in radiotherapy for prostate cancer were investigated. Analyses were performed at different levels, from the macroscopic scale of organ symptoms to the microscopic one of cell death. The studies developed in this project aimed at integrating evidence from the literature with new research modalities, including new mathematical approaches, the combination of different technologies and assays, also including computational systems. Concerning the organ level, a validation of different Normal Tissue Complication Probability Models was performed for toxicity domains where robust predictive models have already been developed in the past, such as late rectal bleeding and faecal incontinence. Analyses on patients' toxicity carried out on large international trials (within a parternship with University of Western Australia) gave us the possibility to enlarge the field of research and to investigate also symptoms with a low rate of toxicity, such as late stool frequency and rectal pain (incidence <5%). Moreover, the already mentioned toxicity endpoint of late faecal incontinence was also analysed by the use of novel approaches. In a first study, an Artificial Neural Network was trained to derive a prediction tool including risk factors from dosimetric and clinical variables, whereas in a second approach the same outcome was investigated by inclusion into the model of more sophisticated dose parameters able to overcome the limitations of the classical Dose Volume Histogram. In this work (within a partnership with Karolinska Hospital) an analysis of spatial information for dose distributions was carried out on Dose Surface Maps from the rectum and the anal canal. The conclusions from the majority of these studies highlighted that secondary organ structures or sub-structures, which are usually not taken into account in predictive modelling, could play a role in the development of the organ dysfunction. Following this idea, we studied the obturator muscles (in partnership with the National Research Council), which are possibly involved in the development of bowel and urinary dysfunctions, and the effect of radiation on the microvasculature (in partnership with Politecnico di Milano and EOC in Lugano), which plays

an important role in the manifestation of a variety of late radio-induced effects (scale of years). The first aspect (radiation-induced modifications in obturator muscles) was investigated by the use of Magnetic Resonance images, particularly by the extraction of radiomic features from T1-weighted and T-2 weighted aquisitions, the second one (effect of radiation on the microvasculature) was studied by a computational model and also by the use of microfluidic chip. A preliminary pilot study on mice investigated the possibility of detecting the effect of dose and dose fractionation on bladder functionality/inflammation using in vivo bio-fluorescent imaging . In vivo results were able to capture what was described ex-vivo by pathological analysis and immunohistochemical assay. The central message emerging from these works was that the large panorama of modern non-invasive tools can provide useful information for understanding the radiation-induced changes in the tissues, from conditions like fibrosis or oedema to the organ dysfunction. Finally, on the basis of the results of the investigation on obturator muscles, we decided to downscale the analysis to the cell level, setting up an in vitro study using immortalised Human Skeletal Muscle Cells. A radiobiological characterization of this cell line has never been performed before. The cells were irradiated with photons (at Fondazione IRCCS Istituto Nazionale dei Tumori) and carbon ions (at Centro Nazionale di Adroterapia Oncologica) and three types of analysis were performed: cell survival assay, functional analysis of DNA damage and Monte Carlo simulations (with a model developed at Universita di Pavia and INFN – Sezione di Pavia). The results obtained in the experiments provided useful information to be used as input for the simulations. Overall, these studies highlighted the difficulty of getting robust models for the complex issue of the response of tissues to ionizing radiation, in a field that includes a magnitude of parameters. This complexity even expands when going from a journal paper towards the clinical practice. The motto of precision medicine, "the right treatment for the right patient given at the right time", fails to emphasize the importance of careful validation of predictive models, which is essential for a successful implementation of personalised medicine. To face to this issue, the idea along this PhD project was that the analysis of multiple aspects could bring a reduction in the model noise and provide more robust models. The development of such a process needs (i) a variety of expertise in the team (or active collaborations with other Institutes or Universities), (ii) financial funding to perform extra protocol investigations (e.g. blood biomarkers, extra imaging scans for radiomics, plastic chips, immunological assays), (iii) infrastructures to collect and manage a large amount of data, (iv) data harmonization when information comes from different institutes and (v) the sharing of data. Even after that, a collaborative effort is needed with a common goal of assessing the discovery (ensuring that data are available to the larger scientific community) in order to complete the transfer of scientific knowledge into the clinical space. It is crystal clear that this is a very challenging task, but it remains even more clear that changing our trajectory to more efficient, collaborative, and valid methods is required

to produce a profound shift in how we approach personalised medicine.

Appendix

Tumour Control Probability models for prostate cancer

Two main models were presented in the literature. The first one, named the "Marsden" TCP model [165, 166], is a mechanistic model assuming firstly that a tumour is only "controlled" (i.e., eliminated) when every single clonogen has been "killed" (i.e., rendered incapable of further division), and further that all the clonogens have the same radiosensitivity and that this remains constant from fraction to fraction. Because in general we do not know the radiosensitivity of the tumour clonogens of the single patient, the "Marsden" TCP model uses a mean α and an assumed standard deviation σ_α over the population. For prostate cancer and for the standard 74 Gy treatment a first set of best fit parameters were obtained by Nahum and Uzan [167] from a fit to the clinical outcomes reported by Dearnaley et al [168]: $\alpha = 0.262\,Gy^{-1}$,$\sigma_\alpha = 0.045\,Gy^{-1}$, $\alpha/\beta = 10\,Gy$, $\rho_{clon} = 107\,cm^{-3}$. The second one is a study from the Memorial Sloan Kettering Cancer Center also provided best fit parameters for the Marsden TCP models for different prognostic groups of prostate cancer [169], details are given in the table here below.

Patient subgroup	Webb model
Low risk	$\alpha = 0.105\ \mathrm{Gy}^{-1}$ [0.03, 0.22] $\sigma(\alpha) = 0.0053\ \mathrm{Gy}^{-1}$ [0, 0.05] $\rho_c = 10.55\ \mathrm{cm}^{-3}$ [$2{\cdot}10^{-2}$, $4{\cdot}10^{4}$] M = −17.9210
Intermediate risk	$\alpha = 0.120\ \mathrm{Gy}^{-1}$ [0.08, 0.19] $\sigma(\alpha) = 0.000054\ \mathrm{Gy}^{-1}$ [0, 0.03] $\rho_c = 53.53\ \mathrm{cm}^{-3}$ [2, $5{\cdot}10^{3}$] M = −19.8720
High risk	$\alpha = 0.107\ \mathrm{Gy}^{-1}$ [0.05, 0.18] $\sigma(\alpha) = 0.000006\ \mathrm{Gy}^{-1}$ [0, 0.02] $\rho_c = 54.4140\ \mathrm{cm}^{-3}$ [$3{\cdot}10^{-1}$, $3{\cdot}10^{4}$] M = −17.7626
Low T stage	$\alpha = 0.090\ \mathrm{Gy}^{-1}$ [0.05, 0.14] $\sigma(\alpha) = 0.000054\ \mathrm{Gy}^{-1}$ [0, 0.03] $\rho_c = 3.415\ \mathrm{cm}^{-3}$ [$1{\cdot}10^{-1}$, $2{\cdot}10^{2}$] M = −31.1824
High T stage	$\alpha = 0.157\ \mathrm{Gy}^{-1}$ [0.10, 0.22] $\sigma(\alpha) = 0.0056\ \mathrm{Gy}^{-1}$ [0, 0.03] $\rho_c = 3020.50\ \mathrm{cm}^{-3}$ [$3{\cdot}10^{1}$, $6{\cdot}10^{5}$] M = −25.9879

* 68% confidence intervals are given in brackets.

Figure 10: *Fit from Memorial Sloan Kettering study using Marsden Model*

The same group from the MSKCC also provided a fit for TCP as described through a sigmoid-shaped dose–response curve using a logistic regression model [170] They characterized the fitted dose–response curves by TCD_{50}, the dose corresponding to a TCP of 50%, and by γ_{50}, the normalized slope of the dose–response curve around a TCP of 50%. Results for the different prognostic groups are given in the table here below.

Patient subgroup	TCD_{50} (Gy)	Confidence intervals on TCD_{50}	γ_{50}	Confidence intervals on γ_{50}
All patients (103 patients)	70.5	68%: 68.5–72.1 Gy 95%: 65.0–73.3 Gy	2.9	68%: 2.1–3.8 95%: 1.3–4.8
Low T-stage (57 patients)	68.6	68%: 66.3–70.2 Gy 95%: 60.0–71.3 Gy	4.0	68%: 2.6–6.0 95%: 1.3–8.5
High T-stage (46 patients)	75.6	68%: 73.9–76.9 Gy 95%: 71.8–78.0 Gy	4.9	68%: 3.7–7.0 95%: 2.4–9.8
Low Gleason (50 patients)	68.5	68%: 65.9–70.3 Gy 95%: 60.2–71.7 Gy	3.9	68%: 2.6–5.8 95%: 1.5–8.2
High Gleason (52 patients)	74.1	68%: 72.1–75.7 Gy 95%: 68.0–77.2 Gy	3.6	68%: 2.4–5.3 95%: 1.1–7.1
Low PSA (32 patients)	69.3	68%: 66.3–71.2 Gy 95%: 55.9–72.4 Gy	5.2	68%: 3.1–10.6 95%: 1.0–26.3
High PSA (71 patients)	71.9	68%: 69.2–73.8 Gy 95%: 63.6–75.4 Gy	2.6	68%: 1.7–3.7 95%: 0.8–4.9

low T-stage <T2c high T-stage ≥T2c
low Gleason score ≤6 high Gleason score >6
low pre-treatment PSA level ≤10 ng/ml high PSA level >10 ng/ml

Figure 11: *Fit from Memorial Sloan Kettering study using sigmoid-shape dose-response curve*

The AIROPROS 0102 study

AIROPROS 0102 study was aimed to prospectively assess predictors of late rectal toxicity after RT for prostate cancer. In total, 1132 patients treated in 22 different institutions entered the study between July 2002 and March 2004. The selection criteria were: (a) histologically confirmed prostate adenocarcinoma; (b) participating centres using three-dimensional radical (no post-prostatectomy) radiotherapy with prescription doses $\geqslant$ 70 Gy, at 1.8–2 Gy/fr (median dose 74 Gy, range:70–78). Information regarding co-morbidity (diabetes, hypertension, previous disease of the colon, presence of haemorrhoids), previous abdominal surgery (rectum-sigma resection, kidney resection, cholecystectomy, appendectomy), TURP or TURV, and use of drugs (hormonal therapy, antihypertensives, anticoagulants) was recorded. Treatment planning was performed on CT scans using a slice interval 65 mm. Patients were scanned with an empty rectum to enhance the reliability of the rectal DVH. An anatomy-based definition of the rectum was followed: the rectum (considered as a solid organ) was drawn on CT slices starting just above the anal verge and continuing until it turns into the sigmoid colon. If rectal volume (including filling) was $\geqslant$ 100 cm^3 , repeat scanning was suggested. A dummy run on rectum contouring variability was performed and this definition was found to be sufficiently robust. Rectal DVHs of the entire treatment were recorded for all patients. The current analysis refers to 718 patients with a 36-month complete follow-up. For each patient the prescribed dose for each irradiated volume (pelvis, seminal vesicles, and prostate), maximum and mean rectal doses were considered; rectal DVHs of the whole treatment were recorded for all patients and the percent fractions of rectum receiving more than 20, 30, 40, 50, 60, 70, 75 Gy (named V20 Gy–V75 Gy) were considered. Tables 1 shows the distribution of clinical, dosimetric and technical parameters in the subset of patients considered in the present analysis. Of particular importance is the delineation of the rectum: an anatomically based definition (from the anus to the point where it turns into the sigmoid) was used by the participating centres and validated by a dummy-run investigation that was mandatory before entering the study.

The TROG 03.04 RADAR trial

The RADAR trial determines whether adjuvant androgen suppression, bisphosphonates and radiation dose escalation for localised prostate cancer (PC) may improve oncologic outcomes. This study examines whether these measures increase rectal and urinary dysfunction and are secondary trial endpoints. Men with histologically confirmed adenocarcinoma of the prostate without lymph node or systemic metastases with T stage 2b and above or T stage 2a primary tumours of Gleason score $\geqslant$ 7 histology and baseline PSA levels $\geqslant$ 10 ng/mL immediately prior to randomisation were eligible to participate after providing their written informed consent. After stratification by treatment cen-

Table 1. Patient characteristics and distribution of dosimetric parameters. V30Gy–V75Gy: per cent volume of the rectum receiving more than 30–75Gy.

	Yes (%)	No (%)
Previous abdominal surgery	6.6	93.4
Use of anticoagulants/antiaggregants	17.4	82.6
Use of antihypertensives	46.1	53.9
Cardiovascular disease (other than hypertension)	3.1	96.9
Diabetes	6.2	93.8
Haemorrhoids	19.0	81.0
Hormonal therapy	89.4	10.6
Pelvic node irradiation	5.5	94.5
Seminal vesicles irradiation	79.0	21.0

	Minimum	Maximum	25th percentile	50th percentile	75th percentile
Follow-up time (years)	6.4	8.7	6.8	7.3	7.6
ICRU dose (Gy)	70.0	79.2	72.0	74.0	76.0
Dose to the pelvis (Gy)	41.4	50.4	45.0	45.0	45.0
Dose to seminal vesicles (Gy)	27.0	76.0	60.0	66.6	72.0
Rectal V20Gy (%)	28.0	100.0	88.6	97.7	100.0
Rectal V30Gy (%)	18.8	100.0	79.1	93.4	100.0
Rectal V40Gy (%)	14.0	100.0	60.1	78.1	92.3
Rectal V50Gy (%)	10.1	100.0	39.7	52.1	62.1
Rectal V60Gy (%)	6.7	93.3	27.6	37.6	52.2
Rectal V70Gy (%)	0.0	80.5	10.3	18.2	26.6
Rectal V75Gy (%)	0.0	32.2	0.0	0.2	6.5
Maximum rectal dose (Gy)	65.2	85.2	72.5	74.6	76.4
Mean rectal dose (Gy)	15.4	69.0	45.7	52.8	58.4

tre, baseline PSA level (10/10–20/⩾20), Gleason score (⩽6/⩾), and T stage (T2/T3,4), random allocation using the minimisation technique occurred at the Central Office in Newcastle, Australia to one of four treatment arms. Additional stratification by HDR brachytherapy boost centre is outlined below. All men received 6 months of leuprorelin (22.5 mg i.m. 3 monthly) commencing at randomisation, 5 months before RT to the prostate and seminal vesicles (but not to pelvic lymph nodes). Men in the control arm received no further treatment (i.e. short term AS [STAS]). Men in the other androgen suppression only treatment arm received an additional 12 months of leuprorelin (22.5 mg i.m. 3 monthly) (i.e. intermediate term AS [ITAS]). Men allocated to the two bisphosphonate treatment arms received zoledronic acid 4 mg i.v. every 3 months for 18 months starting at randomisation with 6 months leuprorelin (22.5 mg i.m.) starting 5 months before RT (STAS + Z) or the same AS plus RT followed by 12 months leuprorelin (22.5 mg i.m. 3 monthly) (ITAS + Z). The trial's main endpoint is prostate cancer specific mortality (PCSM) and was reported alongside other oncological endpoints in 2018 [171]. In this report we describe the effects of the study drugs and radiation dose escalation on rectal and urinary dysfunctional symptoms. These are protocol specified secondary trial endpoints scheduled for analysis in 2012 after 5 years minimum follow up.

In the RADAR trial a regulated radiation dose escalation programme was achieved by requiring participating centres to select their preferred dosing options from a pre-determined range of doses and techniques. The dosing options were 66, 70 and 74 Gy to the ICRU point using external beam alone (EBRT only) and 46 Gy to the ICRU point using external beam followed by a high dose rate brachytherapy (HDRB) boost dose of 19.5 Gy using three fractions of 6.5 Gy. Before enrolling subjects or selecting dose and technique each centre had to demonstrate that it could meet the pre-determined accuracy of dose delivery criteria specified in the trial protocol. These criteria were dose dependent in stringency and are described in more detail in the Web appendix. Having selected a specific dose level, each centre continued to use this dose from thereon. Dose escalation to the next dose level was permitted at a treatment centre if it could meet the more stringent treatment accuracy criteria specified for that dose level. Full details of the derivation of radiation target volumes, dose volume histogram constraints and set up accuracy requirements are provided in the relevant portions of the RADAR trial protocol. Briefly the GTV needed to encompass the prostate only in subjects with T stage <3b, Gleason score <8 and iPSA <20 tumours. The seminal vesicles needed to be included in the GTV of subjects with tumours exceeding one or more of these criteria. In centres not equipped to treat conformally one PTV (i.e. without size reduction) required treatment with minimum margins built around the GTV of 10 mm, excepting posteriorly where the margin could be as small as 5 mm. In centres equipped to treat conformally, two PTVs could be treated if the treating investigator considered necessary i.e. using a PTV 1 to between

46 and 60 Gy, followed by a size reduction to PTV 2 where the remaining dose (to 70 or 74 Gy) would be delivered. Margins built around GTV 1 to achieve PTV 1 were set at 10–15 mm, except posteriorly where they were 5–10 mm. Margins around GTV 2 to achieve PTV 2 were set at <10 mm, except posteriorly where they were ⩽5 mm. Rectal dose constraints mandated that no more than 30% of the rectal volume should receive 70 Gy, and no more than 30% of the femoral heads should receive 60 Gy. In subjects receiving a HDRB boost the anterior rectal wall was required to receive no more than 70% of the reference dose and the prostatic urethra no more than 120%. The use of multiple co-planar fields was allowable either with the subject prone or supine. Beam energies needed to be 6 MV or greater.

Table 1
Characteristics of the 1071 subjects' ages, tumours and radiation doses.

	STAS		STAS + Z		ITAS		ITAS + Z		HDRB		66 Gy		70GY		74GY		Total	
Randomised	268		268		268		267		237		125		427		262		1071	
Age																		
Median	69.3		69.1		68.6		68.2		68.0		69.5		69.9		66.4			
Interquartile range	63.9–73.3		63.8–73.3		63.3–73.2		63.6–72.3		66.1–72.8		64.1–73.2		64.6–74.1		60.6–71.4			
T stage																		
T2	170	(63.4)	171	(63.8)	170	(63.4)	169	(63.3)	85	(35.9)	94	(75.2)	311	(72.8)	179	(68.3)	680	(63.5)
T3,4	98	(36.6)	97	(36.2)	98	(36.6)	98	(36.7)	152	(64.1)	31	(24.8)	116	(83)	83	(31.7)	391	(36.5)
Gleason score																		
<7	26	(9.7)	25	(9.3)	25	(9.3)	25	(9.4)	4	(1.7)	18	(14.4)	48	(11.2)	29	(11.1)	101	(9.4)
=7	155	(57.8)	155	(57.8)	138	(51.5)	151	(56.6)	112	(47.3)	73	(58.4)	261	(61.1)	143	(54.6)	599	(55.9)
>7	87	(32.5)	88	(32.8)	105	(39.2)	91	(34.1)	121	(51.1)	34	(27.4)	118	(27.6)	90	(34.4)	371	(34.6)
iPSA group																		
<10	74	(27.6)	74	(27.6)	72	(26.9)	73	(27.3)	60	(25.3)	38	(30.4)	120	(28.1)	68	(26.0)	293	(27.4)
10–<20	107	(39.9)	107	(39.9)	104	(38.8)	107	(40.1)	92	(38.8)	52	(41.6)	179	(41.9)	97	(37.0)	425	(39.7)
⩾20	87	(32.5)	87	(32.5)	92	(34.3)	87	(32.6)	85	(35.9)	35	(28.0)	128	(30.0)	97	(37.0)	353	(33.0)
RT dose																		
HDRB	57	(21.3)	57	(21.3)	61	(22.8)	62	(23.2)	NA		NA		NA		NA		237	(22.1)
66 Gy	30	(11.2)	30	(11.2)	32	(11.9)	33	(12.6)	NA		NA		NA		NA		125	(11.7)
70 Gy	111	(41.4)	108	(40.3)	106	(39.6)	102	(38.2)	NA		NA		NA		NA		427	(39.9)
74 Gy	68	(25.4)	65	(24.3)	64	(23.9)	65	(24.3)	NA		NA		NA		NA		262	(24.5)
Not given/per protocol	2	(0.7)	8	(3.0)	4	(1.5)	6	(2.2)	NA		NA		NA		NA		20	(1.9)

The DUE-01 trial

A prospective trial started in 2010, aiming at developing models for urinary toxicity and erectile dysfunction after radiotherapy for prostate cancer. This analysis is finalised at highlighting correlations between clinical/dosimetric factors and acute urinary specific symptoms, as measured by single questions of the International Prostate Symptom Score (IPSS).Gastrointestinal toxicity was defined as secondary endpoints. DUE01 is a prospective, multi-centric cohort study that enrolled patients in 9 Italian Institutes. Main requirements of the trial are: dose $\geqslant$ 70 Gy with conventional fractionation (CONV) or $\geqslant$ 65 Gy with moderate hypo-fractionation (HYPO, $\leqslant$2.7 Gy/fr); no prostatectomy prior to radiotherapy. The planning CT was performed without contrast liquid with the bladder full or half-full. Before the start of radiotherapy, several clinical data were prospectively collected, including: age, body-mass-index, diabetes, previous pelvic/abdominal surgery, previous transurethral resection of the prostate, use of drugs (e.g. antihypertensives, anticoagulants, anti-aggregants), smoking (yes/no), coffee (cups/day), clinical stage, initial and

Table 2
Dysfunctional rectal symptoms: comparisons of the trial arms, radiation doses and techniques.

Timing	Model[a]	Variable	Endpoint[b]									
			CTC		Bleeding		Urgency		Frequency		Incontinence	
			OR	*p*-Value	OR	*p*-Value	OR	*p*-Value	OR	*p*-Value	OR	*p*-Value
End RT	Model 1	STAS	1	–	1	–	1	–	1	–	1	–
		STAS + Z	0.97	0.870	1.11	0.643	1.31	0.155	0.67	0.030	1.08	0.725
		ITAS	1.03	0.860	0.88	0.583	1.09	0.655	0.74	0.113	0.92	0.691
		ITAS + Z	1.20	0.345	1.00	0.993	1.13	0.528	0.93	0.684	1.00	0.994
		HDRB	1	–	1	–	1	–	1	–	1	–
		66 Gy	**4.91**	**<0.001**	**3.17**	**<0.001**	**2.48**	**<0.001**	1.81	0.011	1.10	0.715
		70 Gy	**4.43**	**<0.001**	**3.70**	**<0.001**	**3.14**	**<0.001**	**2.54**	**<0.001**	1.00	0.981
		74 Gy	**3.74**	**<0.001**	**3.38**	**<0.001**	**2.58**	**<0.001**	**1.79**	**0.002**	1.16	0.498
	Model 2	EBRT	1	–	1	–	1	–	1	–	1	–
		HDRB	0.23	**<0.001**	**0.28**	**<0.001**	**0.35**	**<0.001**	**0.46**	**<0.001**	0.94	0.739
	Model 3	% rectal volume > 60 Gy[c]	1.01	0.041	1.01	0.367	**1.02**	**0.001**	**1.03**	**<0.001**	1.02	0.012
18 months	Model 1	STAS	1	–	1	–	1	–	1	–	1	–
		STAS + Z	1.09	0.720	1.63	0.046	0.79	0.272	0.79	0.272	0.94	0.800
		ITAS	0.96	0.866	1.52	0.090	1.08	0.695	1.08	0.695	1.22	0.374
		ITAS + Z	0.77	0.291	1.25	0.379	0.79	0.265	0.79	0.265	1.39	0.142
		HDRB	1	–	1	–	1	–	1	–	1	–
		66 Gy	1.56	0.121	1.54	0.146	0.92	0.766	1.13	0.540	1.52	0.127
		70 Gy	0.93	0.734	1.14	0.568	1.15	0.456	1.09	0.650	1.42	0.102
		74 Gy	1.25	0.357	1.53	0.080	0.82	0.350	1.17	0.426	1.03	0.902
	Model 2	EBRT	1	–	1	–	1	–	1	–	1	–
		HDRB	0.90	0.592	0.76	0.178	1.00	0.990	1.28	0.115	0.77	0.176
	Model 3	% rectal volume > 60 Gy[c]	1.02	0.026	**1.02**	**0.005**	1.01	0.048	1.00	0.969	1.02	0.019
36 months	Model 1	STAS	1	–	1	–	1	–	1	–	1	–
		STAS + Z	1.25	0.400	1.04	0.875	1.09	0.705	1.02	0.922	1.24	0.416
		ITAS	1.13	0.640	0.84	0.475	0.98	0.935	1.16	0.475	1.39	0.190
		ITAS + Z	1.21	0.455	0.97	0.898	1.10	0.676	0.98	0.931	1.20	0.475
		HDRB	1	–	1	–	1	–	1	–	1	–
		66 Gy	1.68	0.105	1.87	0.041	1.86	0.043	0.76	0.309	1.97	0.024
		70 Gy	1.23	0.417	1.94	0.005	**2.73**	**<0.001**	1.59	0.015	1.34	0.231
		74 Gy	1.85	0.022	1.50	0.126	1.88	0.015	0.88	0.562	1.24	0.420
	Model 2	EBRT	1	–	1	–	1	–	1	–	1	–
		HDRB	0.67	0.085	**0.56**	**0.009**	**0.44**	**<0.001**	0.84	0.334	0.71	0.126
	Model 3	% rectal volume > 60Gy[c]	**1.03**	**0.002**	**1.03**	**<0.001**	1.01	0.093	1.01	0.059	1.01	0.093

Abbreviations: CTC, Common Toxicity Criteria (version 2); OR, odds ratio; IPSS, International Prostate Symptom Score; STAS, short term (6 months) androgen suppression; ITAS, intermediate term (18 months) androgen suppression; Z, zoledronic acid; EBRT, external beam radiotherapy; HDRB, high dose rate brachytherapy; Gy, Gray.
Bolded values are considered statistically significant.

[a] Model 1 co-variables include treatment arm, age and pre-existing rectal symptoms, medical conditions and habits (described under Methods), and EBRT doses controlled by HDRB; Model 2 also includes HDRB controlled by EBRT; Model 3 In subjects receiving EBRT only this model includes% of rectal volume receiving > 60 Gy (continuous variable).

[b] Endpoints: CTC = CTC proctopathy score ⩾ 1; Bleeding ⩾ 1 = grade 1 (occult), grade 2 (>2/week), grade 3 (daily), grade 4 (gross haemorrhaging) [Lent Soma scale]; Urgency & tenesmus ⩾ 1 = grade 1 (occasional), grade 2 (intermittent), grade 3 (persistent), grade 4 (refractory) [Lent Soma scale]; Frequency ⩾ 1 = grade 1 (2–4/day), grade 2 (5–8/day), grade 3 (>8/day), grade 4 (uncontrolled diarrhoea) [Lent Soma scale]; Incontinence ⩾ 1 = score 1 (a little), 2 (quite a bit), 3 (very much) [IPSS Question 11: "Have you had any unintentional release (leakage) of stools?"].

[c] Percentage of rectal volume receiving greater than 60 Gy (continuous variable).

pre-radiotherapy PSA, type and duration of hormonal therapy, intensity modulated RT (IMRT, yes/no). Dosimetric parameters were also collected and full planning data were exported to a dedicated software (VODCA); bladder DVHs/DSHs were calculated both in percent/absolute values (cc/cm2). Moreover, since different fractionation schemes were allowed, weekly histograms were also considered (DVHw/DSHw); this approach derives from the evidence that the weekly dose highly impacts on the acute side effects , and DSHw were chosen as main dosimetric predictors.

Table 1
Patient characteristics.

	All patients ($n = 262$)	CONV patients ($n = 120$)	HYPO patients ($n = 142$)	p-Value (t-test or z-test)
Age (years, median, range)	71 (46–82)	69 (46–79)	71 (56–82)	*0.004*
BMI (Median, range)	26 (19–42)	27 (21–41)	26 (19–42)	0.12
Diabetes (yes,%)	35 (13.4)	20 (16.7)	15 (10.6)	0.21
Use of antihypertensives (yes,%)	124 (47.3)	65 (54.2)	59 (41.5)	*0.05*
Previous abdominal surgery (yes,%)	109 (41.6)	57 (47.5)	52 (36.6)	0.1
Cardiovascular diseases (yes,%)	70 (26.7)	39 (32.5)	31 (21.8)	*0.05*
Gastro-intestinal diseases (yes,%)	12 (4.6)	6 (5.0)	6 (4.2)	0.76
Urological diseases (yes,%)	13 (5.0)	8 (6.7)	5 (3.5)	0.24
Use of anticoagulants (yes,%)	19 (7.3)	8 (6.7)	11 (7.7)	0.94
Use of antiaggregants (yes,%)	75 (28.6)	42 (35.0)	33(23.2)	*0.05*
Use of 5alpha reductase inhibitors (yes,%)	21 (8.0)	11 (9.2)	10 (6.2)	0.53
Use of alpha blockers (yes,%)	41 (15.6)	19 (15.8)	22 (13.6)	0.94
Use of drugs for gastric diseases (yes,%)	21 (8.0)	10 (8.3)	11 (6.8)	0.87
Use of drugs for cardiovascular diseases (yes,%)	87 (33.2)	46 (38.3)	41 (25.3)	0.11
Use of cholesterol lowering drugs (yes,%)	33 (12.6)	18 (15.0)	15 (9.3)	0.28
Use of antidepressant drugs (yes,%)	8 (3.1)	5 (4.2)	3 (1.9)	0.34
Smoke (yes,%)	43 (16.4)	14 (11.7)	29 (20.4)	0.08
Alcohol (yes,%)	113 (43.1)	47 (39.2)	66 (46.5)	0.29
TURP (yes,%)	24 (9.2)	7 (5.8)	17 (12.0)	0.13
Clinical stage = T1 (yes,%)	137 (52.3)	73 (60.8)	64 (45.1)	*0.02*
Clinical stage = T2 (yes,%)	82 (31.3)	30 (25.0)	52 (36.6)	0.06
Clinical stage = T3 (yes,%)	23 (8.8)	11 (9.2)	12 (8.5)	0.98
Prescription dose (Gy, median, range)		78 (74–80)	71.4 (59.8–75.2)	
Dose/fraction (Gy, median, range)		2 (1.8–2)	2.5 (2.35–2.7)	
Irradiation of pelvic nodes (yes,%)	78 (29.8)	26 (21.7)	52 (36.6)	*0.01*
Range of pelvic dose (Gy)		50–54	48–54.4	
Irradiation of seminal vesicles (yes,%)	161 (61.5)	71 (59.2)	90 (63.4)	0.57
Range of seminal vesicle dose (Gy)		60–80	51.8–75.2	
Any neoadjuvant hormone therapy (yes,%)	136 (51.9)	68 (56.7)	68 (47.9)	0.16
Exclusive neoadjuvant antiandrogen (yes,%)	75 (28.6)	43 (35.8)	32 (22.5)	*0.02*
Exclusive Neoadjuvant LH-RH analogue (yes,%)	24 (9.2)	15 (12.5)	9 (20.4)	0.09
Neoadjuvant total androgen blockade (yes,%)	37 (14.1)	10 (8.3)	27 (19.0)	*0.01*
CTV (cc, median, range)	48.8 (15.9–134)	39.6 (18.9–100)	53.4 (15.9–134)	*0.005*
PTV (cc, median, range)	133.8 (43–279)	120.3 (43–279)	141.6 (44.3–270.8)	*0.01*
Bladder volume (cc, median, range)	194.4 (45.9–890)	195.4 (50.2–890)	194.4 (45.9–889)	0.88

CONV = conventionally fractionated; HYPO = hypofractionated; BMI = body-mass-index; TURP = previous transurethral resection of the prostate; CTV = clinical target volume; PTV = clinical target volume; LH-RH = Luteinizing hormone-releasing hormone.
t-test was used to compare continuous variable distribution in conventionally treated patients vs hypofractionated patients. z-test for proportions was used for comparison of rates of factor variables in the two populations.
Numbers in italics indicate significant p-values (< 0.05).

Rectal Pain and Stool Frequency Validation study

Validations were computed fixing the offset at 0. We decided to evaluate the model more by the use of the R^2 rather than the slope of the model itself. The original idea that we pursuited was the fact that in an univariate model with dose as parameter, at the 0 Gy the probability of side effects should be 0%. However, this procedure was not the best choice due to the possible differences between development and validation dataset, differences that can be manifested by the presence of an offset. We updated numbers for stool frequency, while the calibration of the predictive model for rectal pain was almost without offset.

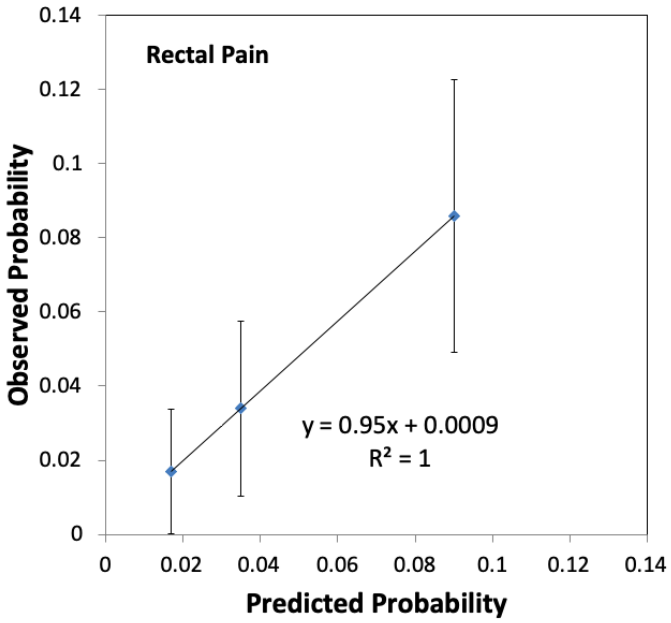


Figure 12: *Calibration plot for late rectal pain in the DUE-01 population. Offset for this calibration plot was almost 0.*

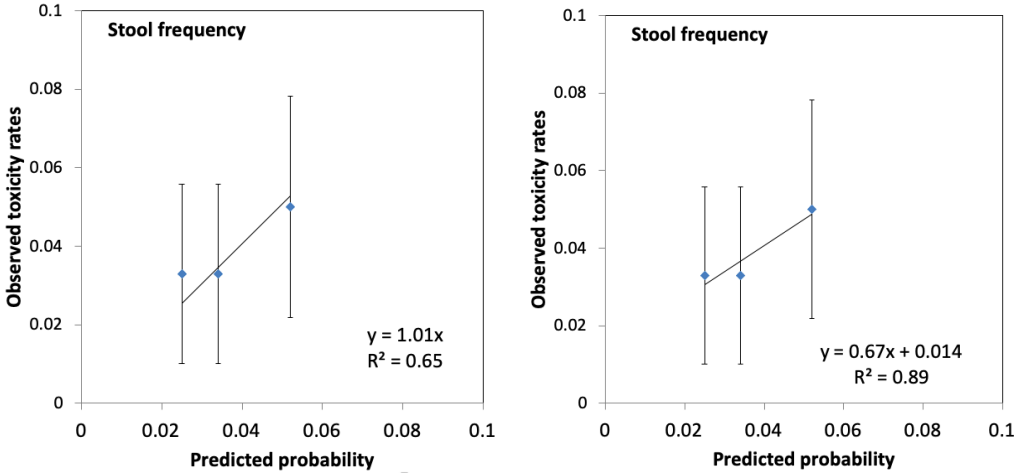
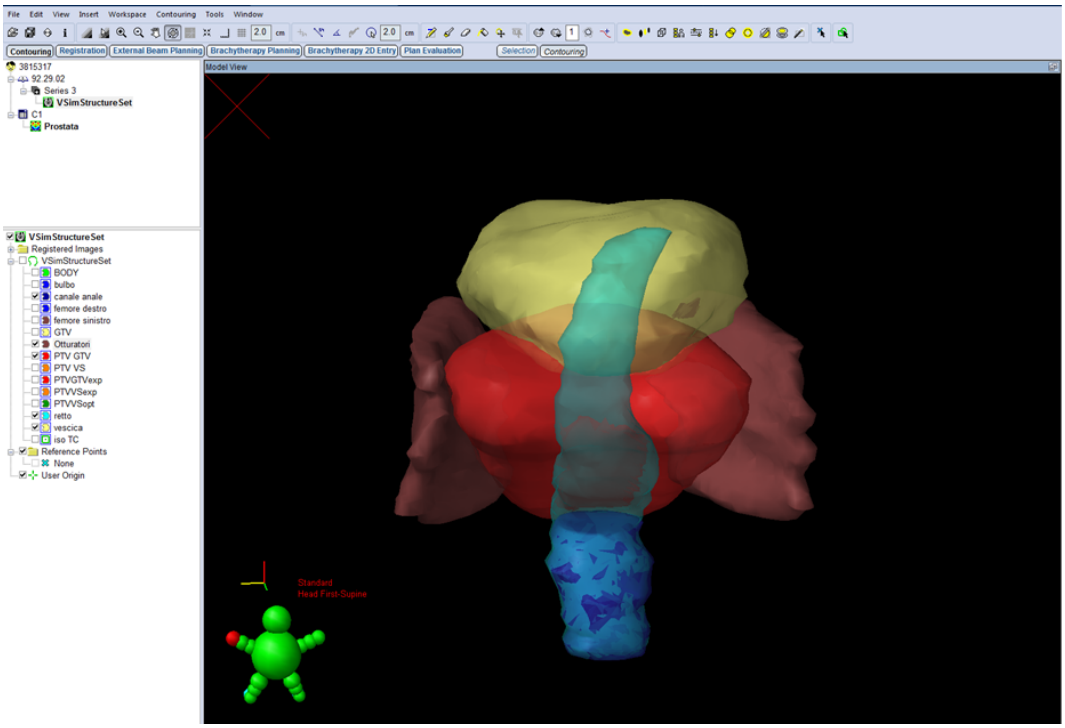


Figure 13: *Calibration plots for late stool frequency in the DUE-01 population. Plots represent fit with and without fixing the offset (left and right respectively).*

Structures and substructures analysed in this work

In the following picture a 3D representation of the prostate and of the organs at risk that were investigated within the PhD project is depicted . The red structure is the Planning Tumour Volume of the prostate. For what concerns the normal tissues: the cyan organ in transparency is the anorectum (anal canal + rectum) (largely studied in Section 3.2 and 3.3), the solid one is the anal canal (specifically analysed in Section 3.3.3), the yellow structure is the urinary bladder (studied in vivo and ex-vivo in Section 4.2) and the two structures embracing the prostate in light brown are the obturator muscles (studied by the use of imaging in Section 4.1 and in vitro and in silico in Chapter 5). Finally, the capillary network (which is part of all of these structures) was explored in Section 4.3.



Bibliography

[1] W K Roentgen. On a new kind of rays. *Science*, 3(59):227–231, 2 1896.

[2] J Clifford Perry. Roengten rays - their application to medicine and surgery. *New Jersey Dentat Society*, 7(30):307–316, 11 1896.

[3] V Despeignes. Observation concernant un cas de cancer de l'estomac traité par l'es rayons rontgen. *Lyon Médical : Gazette Médicale et Journal de Médecine Réunis*, pages 503–506, 12 1896.

[4] N Foray et al. Individual response to ionizing radiation. *Mutation Research/Reviews in Mutation Research*, 770:369–386, 9 2016.

[5] M L Bouchacourt. Sur là diffèrence de sensibilitè aux rayons de roentgen de la peau des diffèrents sujects, et, sur le mème sujet des diffèrents règions du corps. *Proceedings of the 40th Congress of the French Association for the Advancement of Science, Lyon*, pages 942–947, 1911.

[6] N Foray et al. 100 years of individual radiosensitivity: How we have forgotten the evidence. *Radiology*, 264(3):627–631, 9 2012.

[7] G Vogine and N Foray. The law of bergonie and tribondeau: A nice formula for a first approximation. *International Journal of Radiation Biology*, pages 1–7, 8 2012.

[8] R Stockton. The unbelievable true story of america's radium girls. *Wall Street Journal*, 11 2017.

[9] R Winslow. The radium water worked fine until his jaw came off. *Wall Street Journal*, page A1, 8 1932.

[10] J D Watson and F H C Crick. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 4 1953.

[11] T T Puck and P I Marcus. Action of x-rays on mammalian cells. *Journal of experimental medicine*, 103:653–666, 1956.

[12] R L Warters and K G Hofer. Radionuclide toxicity in cultured mammalian cells: Elucidation of the primary site for radiation-induced division delay. *Radiation Research*, 69(2):348–359, 2 1957.

[13] R L Warters et al. Radionuclide toxicity in cultured mammalian cells: Elucidation of the primary site of radiation damage. *Current topics in radiation research quarterly*, 12:389–407, 1977.

[14] E C Halperin et al. *Perez and Brady's Principles and Practice of Radiation Oncology*. Wolters Kluwer Health/Lippincott Williams and Wilkins, 5 edition, 2008.

[15] H Holthusen. Erfahrungen uber die vertraglichkeitsgrenze fur rongenstrahlen und deren nutzanwendung zur verhutung von schaden. *Strahlentherapie*, 57:254–259, 1936.

[16] E Giaccia and A Hall. *Radiobiology for Radiologist*. Wolters Kluwer Health/Lippincott Williams and Wilkins, 7 edition, 2012.

[17] D S Chang et al. *Therapeutic Ratio in Basic Radiotherapy Physics and Biology*. Springer, 2014.

[18] L B Marks et al. Quantitative analyses of normal tissue effects in the clinic. *International Journal of Radiation Oncology, Biology and Physics*, 76(3):S1–S160, 2010.

[19] T R Munro and C W Gilbert. The relation between tumour lethal doses and the radiosensitivity of tumour cells. *Radiation Research*, 34:264–251, 4 1961.

[20] M Baumann et al. Response of human squamous cell carcinoma xenografts of different sizes to irradiation: relationship of clonogenic cells, cellular radiation sensitivity in vivo, and tumor rescuing units. *Radiation Research Society*, 123(3):325–330, 9 1990.

[21] M Krause et al. Preclinical evaluation of moleculartargeted anticancer agents for radiotherapy. *Radiotherapy Oncology*, 80:112–120, 2006.

[22] H Withers. The four r's of radiotherapy. *Advances in Radiation Biology*, 5:241–271, 1975.

[23] G Steel et al. The 5rs of radiobiology. *International Journal of Radiation Biology*, 56(6):1045–1048, 2009.

[24] M M Elkind and H Sutton. Radiation response of mammalian cells grown in culture: Repair of x-ray damage in surviving chinese hamster cells. *Radiation Research*, 13:556–593, 1960.

[25] W K Sinclair and R A Morton. X-ray and ultraviolet sensitivity of synchronized chinese hamster cells at various stages of the cell cycle. *Biophysical Journal*, 5:1–25, 1965.

[26] B Xu et al. Two molecularly distinct g(2)/m checkpoints are induced by ionizing irradiation. *Molecular and Cellular Biology*, 22:1049–1059, 2002.

[27] H Nagasawa et al. Relationship between gamma-ray-induced g2/m delay and cellular radiosensitivity. *International Journal of Radiotion Biology*, 66:373–379, 1994.

[28] GD Kao et al. Detection of repair activity during the dna damage-induced g2 delay in human cancer cells. *Oncogene*, 20:3486–3496, 2001.

[29] R H Thomlinson. L H Gray. The histological structure of some human lung cancers and the possible implications for radiotherapy. *British Journal of Cancer*, 9:539–549, 1955.

[30] I R Vogelius and S M Bentzen. Meta-analysis of the alpha/beta ratio for prostate cancer in the presence of an overall time factor: Bad news, good news, or no news? *International Journal of Radiationtherapy, Oncology, Biology and Physics*, 85:89–94, 2013.

[31] S L Gulliford et al. Parameters for the lyman kutcher burman (lkb) model of normal tissue complication probability (ntcp) for specific rectal complications observed in clinical practise. radiotherapy and oncology. *International Journal of Radiationtherapy, Oncology, Biology and Physics*, 102:347–351, 2012.

[32] C Fiorino et al. Late fecal incontinence after high-dose radiotherapy for prostate cancer: Better prediction using longitudinal definitions. *International Journal of Radiationtherapy, Oncology, Biology and Physics*, 83:38–45, 2012.

[33] A Wojcik et al. Considerations on the use of the terms radiosensitivity and radiosusceptibility. *Society for Radiobiological Protection*, 38:N25–N29, 2018.

[34] B Fertil and E P Malaise. Intrinsic radiosensitivity of human cell lines is correlated with radioresponsiveness of human tumors: analysis of 101 published survival curves. *International Journal of Radiationtherapy, Oncology, Biology and Physics*, 11:1699–1707, 1985.

[35] W Dorr and M Schmidt. *Normal tissue radiobiology. In: A. Brahme, J.H. Hendry (Eds), Comprehensive Biomedical Physics.* Elsevier, Amsterdam, 7 edition, 2014.

[36] W Dorr. *Pathogenesis of normal-tissue side-effects. In: M. Joiner and A. Van der Kogel (Eds), Basic Clincial Radiobiology.* Hodder Arnold, 4 edition, 2009.

[37] W Dorr. Radiobiology of tissue reactions. *Annual of ICRP*, 44:58–68, 2015.

[38] H P Rodemann and M Bamberg. Cellular basis of radiation-induced fibrosis. *Radiother Oncol*, 35:83–90, 1995.

[39] J Yarnold and M C Brotons. Pathogenetic mechanisms in radiation fibrosis. *Radiother Oncol*, 97:149–161, 2010.

[40] W Dorr et al. Effects of stimulated repopulation on oral mucositis during conventional radiotherapy. *Radiotherapy and Oncology*, 37:100–107, 1995.

[41] E Van der Schueren et al. Radiotherapy by multiple fractions per day (mfd) in head and neck cancer: acute reactions of skin and mucosa. *International Journal of Radiation Oncology and Biology*, 19:301–311, 1990.

[42] M S Peach. Systematic review of the relationship between acute and late gastrointestinal toxicity after radiotherapy for prostate cancer. *Prostate Cancer*, 2015.

[43] W Dorr and H Hendry. Consequential late effects in normal tissues. *Radiotherapy and Oncology*, 61(3):223–231, 2001.

[44] W Dorr et al. Radiation induced lung reactions in breast cancer therapy. *Strahlenther Onkology*, 181:567–573, 2005.

[45] L Hakenjos et al. Radiation induced lung reactions in breast cancer therapy. *Strahlenther Onkology*, 76:503–509, 2000.

[46] J W Denham and M Hauer Jansen. The radiotherapeutic injury - a complex 'wound'. *Radiotherapy and Oncology*, 63:129–145, 2002.

[47] L Marks et al. Radiation dose volume effects in the lung. *International Journal of Radiation Oncology, Biology and Physics*, 76:S70–S76, 2012.

[48] H R Withers et al. Treatment volume and tissue tolerance. *International Journal of Radiation Oncology, Biology and Physics*, 14:751–759, 1988.

[49] T Shibata. *Treatment Planning for Head and Neck Malignancies. Y Nishimura, R Komaki (Eds) In: Intensity-Modulated Radiation Therapy: Clinical Evidence and Techniques.* Hodder Arnold, 4 edition, 2009.

[50] S Bentzen et al. Quantitative analyses of normal tissue effects in the clinic (quantec): an introduction to the scientific issues. *International Journal of Radiation Oncology, Biology and Physics*, 76:S3–S9, 2010.

[51] L B Marks et al. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology, Biology and Physics*, 76:S10–S19, 2010.

[52] A C Begg et al. The value of pretreatment cell kinetic parameters as predictors for radiotherapy outcome in head and neck cancer: a multicenter analysis. *Radiotherapy and Oncology*, 50:13–23, 1999.

[53] W Dorr. Three as of repopulation during fractionated irradiation of squamous epithelia: Asymmetry loss, acceleration of stem-cell divisions and abortive divisions. *International Journal of Radiation Oncology, Biology and Physics*, 72:653–43, 1997.

[54] H R Withers et al. The hazard of accelerated tumor clonogen repopulation during radiotherapy. *Acta Oncologica*, 27:131–146, 1988.

[55] J Overgaards et al. Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: Dahanca 6 and 7 randomised controlled trial. *Lancet*, 362:933–940, 2003.

[56] G H Fletcher et al. Hypofractionation: lessons from complications. *RadiationTherapy and Oncology*, 20:10–15, 1991.

[57] J C Horiot et al. Hyperfractionation versus conventional fractionation in oropharyngeal carcinoma: final analysis of a randomized trial of the eortc cooperative group of radiotherapy. *Radiotherapy and Oncology*, 25:231–241, 1992.

[58] M Baumann and G Gregoire. *Modified fractionation. In: M. Joiner and A. Van der Kogel (Eds), Basic Clincial Radiobiology.* Hodder Arnold, 4 edition, 2009.

[59] K Fu et al. A radiation therapy oncology group (rtog) phase iii randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of rtog 9003. *International Journal of Radiation Oncology, Biology and Physics*, 48:7–16, 2000.

[60] S M Bentzen et al. Repair halftimes estimated from observations of treatmentrelated morbidity after chart or conventional radiotherapy in head and neck cancer. *Radiotherapy and Oncology*, 53:219–226, 1999.

[61] S Dische et al. A randomised multicentre trial of chart versus conventional radiotherapy in head and neck cancer. *Radiotherapy and Oncology*, 44:123–126, 1999.

[62] M C Joiner and S M Bentzen. *Fraction: the Linear-Quadratic Approach. In: M. Joiner and A. Van der Kogel (Eds), Basic Clincial Radiobiology.* Hodder Arnold, 4 edition, 2009.

[63] J R Owen et al. Effect of radiotherapy fraction size on tumour control in patients with early-stage breast cancer after local tumour excision: long-term results of a randomised trial. *Lancet Oncology*, 7:467–471, 2006.

[64] DJ Brenner and EJ Hall. Hypofractionation in prostate cancer radiotherapy. *Translational Cancer Research,(Prostate Cancer: Current Understanding and Future Directions)*, 7, 2018.

[65] J Fowler et al. Is α/β for prostate tumors really low? *International Journal of Radiation Oncology,Biology,Physics*, 50:1021–1031, 2001.

[66] CN Catton et al. Randomized trial of a hypofractionated radiation regimen for the treatment of localized prostate cancer. *Journal of Clinical Oncology*, 50:1884–1890, 2017.

[67] D Dearnaley et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 chhip trial. *Lancet Oncology*, 17:1047–10600, 2016.

[68] WR Lee et al. Randomized phase iii noninferiority study comparing two radiotherapy fractionation schedules in patients with low-risk prostate cancer. *Journal of Clinical Oncology*, 34:2325–2332, 2016.

[69] S Aluwini et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (hypro): acute toxicity results from a randomised non-inferiority phase 3 trial. *Lancet Oncology*, 16:274–283, 2015.

[70] EE Yeoh et al. Hypofractionated versus conventionally fractionated radiotherapy for prostate carcinoma: final results of phase iii randomized trial. *International Journal of Radiation Oncology,Biology,Physics*, 16:1271–1278, 2011.

[71] M Ritter et al. Hypofractionation for prostate cancer. *Cancer Journal*, 15:1–6, 2009.

[72] S H Benedict et al. Stereotactic body radiation therapy: the report of aapm task group 101. *Medical Physics*, 37:4078–4101, 2010.

[73] D Habermel et al. The relative biological effectiveness for carbon and oxygen ion beams using the raster-scanning technique in hepatocellular carcinoma cell lines. *PLoS One*, 9, 2014.

[74] M C Joiner. *Linear Energy Transfer and Relative Biological Effectiveness. In: M. Joiner and A. Van der Kogel (Eds), Basic Clincial Radiobiology.* Hodder Arnold, 4 edition, 2009.

[75] P Vaupel et al. Treatment resistance of solid tumors: role of hypoxia and anemia. *Medical Oncology*, 18:243–259, 2001.

[76] R Wilson. Radiological use of fast protons. *Radiology*, 47(5), 1946.

[77] H Suit. The gray lecture 2001: coming technical advances in radiation oncology. *International Journal of Radiation Oncology, Biology and Physics*, 53:798–809, 2002.

[78] T F DeLaney. Proton therapy in the clinic. *Frontiers for Radiation Therapy and Oncology*, 43:465–485, 2011.

[79] W Heemsbergen and M Witte. *The Importance of the Quality of Data. In: C Fiorino and T Rancati (Eds), Modelling Radiotherapy Side Effects.* Taylor and Francis Group, 1 edition, 2019.

[80] U Hoeller et al. Cosmesis from the patient's and the doctor's view. *International Journal of Radiation Oncology, Biology and Physics*, 57:345–354, 2003.

[81] S T Peeters et al. Rectal bleeding, fecal incontinence, and high stool frequency after conformal radiotherapy for prostate cancer: normal tissue complication probability modeling. *International Journal of Radiation Oncology, Biology and Physics*, 66:11–19, 2006.

[82] G Gagliardi et al. Radiation pneumonitis after breast cancer irradiation: analysis of the complication probability using the relative seriality model. *International Journal of Radiation Oncology, Biology and Physics*, 46:373–381, 2000.

[83] P Lambin et al. Radiomics: extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48:441–446, 2012.

[84] J Deasy and M Thor. *Quantitative Imaging for Assessing and Predicting Toxicity. In: C Fiorino and T Rancati (Eds), Modelling Radiotherapy Side Effects.* Taylor and Francis Group, 1 edition, 2019.

[85] S S F Yip and H J W L Aerts. Applications and limitations of radiomics. *Physics in Medicine and Biology*, 61:R150–R166, 2016.

[86] M L Welch et al. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiotherapy and Oncology*, 130:2–9, 2019.

[87] R Weissleder and M J Pittet. Imaging in the era of molecular biology. *Nature*, 452:580–589, 2008.

[88] F Yu et al. Engineering microfluidic organoid-on-a-chip platforms. *Micromachines*, 165(10), 2019.

[89] H Liu et al. Advances in hydrogels in organoids and organs-on-a-chip. *Advanced Materials*, pages 1–28, 2019.

[90] P Nagle et al. Patient-derived tumor organoids for prediction of cancer treatment response. *Seminars in Cancer Biology*, 53:258–264, 2018.

[91] A M Kellerer and H H Rossi. The theory of dual radiation action. *Radiation Research*, 8:85–158, 1972.

[92] P Price and T J McMillan. The use of non-clonogenic assays in measuring the response of cells in vitro to ionising radiation. *European Journal of Cancer*, 30:838–841, 1994.

[93] R C Taylor et al. Apoptosis: controlled demolition at the cellular level. *Nature Reviews Molecular Cell Biology*, 9:231–241, 2008.

[94] B G Wouters. *Cell death after irradiation: how, when and why cells die. In: M. Joiner and A. Van der Kogel (Eds), Basic Clincial Radiobiology.* Hodder Arnold, 4 edition, 2009.

[95] M Moroni et al. Evaluation of the gamma-h2ax assay for radiation biodosimetry in a swine model. *International Journal of Molecular Sciences*, 14:14119–14135, 2013.

[96] J Schuemann et al. A new standard dna damage (sdd) data format. *Radiation Research*, 191:76–92, 2019.

[97] H Nikjoo et al. Perspectives in radiation biophysics: From radiation track structure simulation to mechanistic models of dna damage and repair. *Radiation Physics and Chemistry*, 128:3–10, 2016.

[98] H Willers and KD Held. Introduction to clinical radiation biology. *Hematol Oncol Clin North Am.*, 20:1–24, 2006.

[99] J T Lyman et al. Complication probability as assessed from dose-volume histograms. *Radiation Research*, 8:13–19, 1985.

[100] G J Kutcher and C Burman. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method. *International Journal of Radiation Oncology, Biology and Physics*, 16:1623–1630, 1989.

[101] T Rancati et al. Fitting late rectal bleeding data using different ntcp models: results from an italian multi-centric study (airopros0101). *International Journal of Radiation Oncology, Biology and Physics*, 73:21–32, 2004.

[102] A S Coates et al. Tailoring therapies–improving the management of early breast cancer: St gallen international expert consensus on the primary therapy of early breast cancer 2015. *Annals of Oncology*, 26:1533–1546, 2015.

[103] S Tucker et al. Late rectal toxicity on rtog 94-06: analysis using a mixture lyman model. *International Journal of Radiation Oncology, Biology and Physics*, 78:1253–1260, 2010.

[104] G Defraene et al. The benefits of including clinical factors in rectal normal tissue complication probability modeling after radiotherapy for prostate cancer. *International Journal of Radiation Oncology, Biology and Physics*, 82:1233–1242, 2012.

[105] P McCullagh and J A Nelder. *Generalized Linear Models*. Boca Raton, Fla: Chapman & Hall/CRC, 2000.

[106] F Palorini et al. Multi-variable models of large international prostate symptom score worsening at the end of therapy in prostate cancer radiotherapy. *Radiotherapy and Oncology*, 118:92–98, 2016.

[107] S Cui et al. *Building a Predictive Model of Toxicity. In: C Fiorino and T Rancati (Eds), Modelling Radiotherapy Side Effects*. Taylor and Francis Group, 1 edition, 2019.

[108] D Kingma and J L Ba. Adam: a method for stocastic optimization. *Published as a conference paper at ICLR 2015*, 2015.

[109] N Srivastava et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.

[110] A Cicchetti et al. Modelling late stool frequency and rectal pain after radical radiotherapy in prostate cancer patients: Results from a large pooled population. *Physica Medica*, 32:1690–1697, 2016.

[111] M Thor. Relationships between dose to the gastro-intestinal tract and patient-reported symptom domains after radiotherapy for localized prostate cancer. *Acta Oncologica*, 54:1326–1334, 2015.

[112] W Schaake et al. Normal tissue complication probability (ntcp) models for late rectal bleeding, stool frequency and fecal incontinence after radiotherapy in prostate cancer patients. *Radiotherapy and Oncology*, 119:381–387, 2016.

[113] K R Trott et al. Biological mechanisms of normal tissue damage: importance for the design of ntcp models. *Radiotherapy and Oncology*, 105:79–85, 2012.

[114] V Landoni et al. Predicting toxicity in radiotherapy for prostate cancer. *Physica Medica*, 32:521–532, 2016.

[115] V D'avino et al. Prediction of gastrointestinal toxicity after external beam radiotherapy for localized prostate cancer. *Radiation Oncology*, 10(80), 2015.

[116] T Rancati et al. Inclusion of clinical risk factors into ntcp modelling of late rectal toxicity after high dose radiotherapy for prostate cancer. *Radiotherapy and Oncology*, 100:124–130, 2011.

[117] J M Michalski et al. Radiation dose-volume effects in radiation-induced rectal injury. *International Journal of Radiation Oncology, Biology and Physics*, 76:S123–S129, 2010.

[118] G Fellin et al. Clinical and dosimetric predictors of late rectal toxicity after conformal radiation for localized prostate cancer: Results of a large multicenter observational study. *Radiotherapy and Oncology*, 93:197–202, 2009.

[119] TROG clincial trial summary. Trog 03.04 - randomised trial investigating the effect on survival and psa control of different durations of adjuvant androgen deprivation in association with definitive radiation treatment for localised carcinoma of the prostate (radar). *Lancet*, 30:82049–82063, 2005.

[120] R Krol et al. Anorectal symptoms after prostate radiotherapy. *Nederlands Tijdschrift voor Geneeskunde*, 160:A9635, 2015.

[121] C G Bacon et al. The association of treatment-related symptoms with quality-of-life outcomes for localized prostate carcinoma patients. *Cancer*, 2002.

[122] F Buettner et al. The dose-response of the anal sphincter region-an analysis of data from the mrc rt01 trial. *Radiotherapy and Oncology*, 103:347–352, 2012.

[123] S Tucker et al. Estimation of α/β for late rectal toxicity based on rtog 94-06. *International Journal of Radiation Oncology,Biology,Physics*, 2:600–605, 2011.

[124] I R Vogelius and S Bentzen. Meta-analysis of the alpha/beta ratio for prostate cancer in the presence of an overall time factor: Bad news, good news, or no news? *International Journal of Radiation Oncology Biology and Physics*, 85:89–94, 2013.

[125] J F Fowler. 21 years of biologically effective dose. *British Journal of Radiology*, 83:554–568, 2010.

[126] A Cicchetti et al. Predicting late fecal incontinence risk after radiation therapy for prostate cancer: New insights from external independent validation. *International Journal of Radiation Oncology Biology and Physics*, 102:127–136, 2018.

[127] K G Moons et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): Explanation and elaboration. *Annals of internal medicine*, 162:W1–W73, 2015.

[128] R C Wortel et al. Late side effects after image guided intensity modulated radiation therapy compared to 3dconformal radiation therapy for prostate cancer: Results from 2 prospective cohort. *International Journal of Radiation Oncology Biology and Physics*, 95:680–689, 2016.

[129] A Al-Mamgani et al. Role of intensity-modulated radiotherapy in reducing toxicity in dose escalation for localized prostate cancer. *International Journal of Radiation Oncology Biology and Physics*, 73:685–691, 2009.

[130] R Di Franco et al. Rectal/urinary toxicity after hypofractionated vs conventional radiotherapy in low/intermediate risk localized prostate cancer: Systematic review and meta analysis. *Oncotarget*, 8:17383–17395, 2017.

[131] S Gulliford et al. Radiotherapy dose-distribution to the perirectal fat space (prs)is related to gastrointestinal control-related complications. *Clinical and Translational Radiation Oncology*, 7:62–70, 2017.

[132] M Carrara et al. Development of a ready-to-use graphical tool based on artificial neural network classification. *International Journal of Radiation Oncology and Biology*, 102:1533–1542, 2018.

[133] S Tomatis et al. Late rectal bleeding after 3d-crt for prostate cancer: Development of a neural-network-based predictive model. *Physics in Medicine and Biology*, 57:1399–1412, 2012.

[134] F Cabitza et al. Unintended consequences of machine learning in medicine. *JAMA*, 318:517–518, 2017.

[135] E S Berner et al. Benefits and risks of machine learning decision support systems. *JAMA*, 318:2353–2357, 2017.

[136] A Van der Schaaf et al. Benefits and risks of machine learning decision support systems. *International Journal of Radiation Oncology, Biology and Physics*, 91:468–471, 2015.

[137] S J Feigenberg et al. 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 and Biology*, 62:397–405, 2005.

[138] C Fiorino et al. Relationship between dvhs and late rectal bleeding after radiotherapy for prostate cancer: analysis of a large gropu of patients pooled from three institutions. *Radiotherapy and Oncology*, 64:1–12, 2002.

[139] G Sanguineti et al. Adjuvant androgen deprivation impacts late rectal toxicity after conformal radiotherapy of prostate carcinoma. *British Journal of Cancer*, 17(86), 2002.

[140] M W Swarchuk et al. Late rectal toxicity after conformal radiotherapy of prostate cancer: multivariate analysis and dose-response. *International Journal of Radiation Oncology and Biology*, 47:103–113, 2000.

[141] D M Herold et al. Diabetes mellitus: a predictor for late radiation morbidity. *International Journal of Radiation Oncology and Biology*, 43:475–459, 1999.

[142] A Appelt et I R Vogelius. A method to adjust radiation dose-response relationships for clinical risk factors. *Radiotherapy and Oncology*, 102:352–354, 2012.

[143] E Onjukka et al. Patterns in ano-rectal dose maps and the risk of late toxicity after prostate imrt. *Acta Oncologica*, 2019.

[144] F Palorini et al. First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy. *Radiotherapy and Oncology*, 119:123–128, 2016.

[145] I Improta et al. Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer. *Physica Medica*, 32:1681–1689, 2016.

[146] N Yahya et al. Independent external validation of predictive models for urinary dysfunction following external beam radiotherapy of the prostate: Issues in model development and reporting. *Radiotherapy and Oncology*, 120:339–345, 2016.

[147] F Buettner et al. 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*, 54(21), 2009.

[148] R C Wortel et al. Dose-surface maps identifying local dose-effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer. *Radiotherapy and Oncology*, 117:515–520, 2015.

[149] E Scalco et al. Texture analysis of t1-w and t2-w mr images allows a quantitative evaluation of radiation-induced changes of internal obturator muscles after radiotherapy for prostate cancer. *Medical Physics*, 45:1518–1528, 2018.

[150] N J Tustison et al. N4itk: improved n3 bias correction. *IEEE Trans Med Imaging*, 29:1310–1320, 2010.

[151] S Klein. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*, 29:196–205, 2010.

[152] S Broggi. A comparative evaluation of 3 different free-form deformable image registration and contour propagation methods for head and neck mri. *Technology in Cancer Res Treatment*, 16:373–381, 2017.

[153] P Soille. On the validity of fractal dimension measurements in image analysis. *Journal of Visual Communication Image Representation*, 7:217–229, 1996.

[154] D Assefa. Robust texture features for response monitoring of glioblastoma multiforme on t1-weighted and t2-flair mr images: a preliminary investigation in terms of identification and segmentation. *Medical Physics*, 37:1722–1736, 2010.

[155] P Sebold. Requite: A prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. *Radotherapy and Oncolgoy*, 138:59–67, 2019.

[156] L Yang et al. Pathophysiological responses in rat and mouse models of radiation-induced brain injury. *Mol Neurobiol*, 54:1022–1032, 2017.

[157] II Verginadis et al. A novel mouse model to study image-guided, radiation-induced intestinal injury and preclinical screening of radioprotectors. *Cancer Research*, 77:908–917, 2017.

[158] M Dombrowsky et al. Acute skin damage and late radiation-induced fibrosis and inflammation in murine ears after high-dose irradiation. *Cancers*, 11:727, 2019.

[159] M P Law. *Radiation-induced vascular injury and its relation to late effect in normal tissues. In: J T Lett and H Adler (Eds), Advances in Radiation Biology.* Taylor and Francis Group, 9 edition, 1981.

[160] S Kabacik and K Raj. Ionising radiation increases permeability of endothelium through adam10-mediated cleavage of ve-cadherin. *Oncotarget*, 30:82049–82063, 2017.

[161] F Ballarini and M. Carante. Chromosome aberrations and cell death by ionizing radiation: Evolution of a biophysical model. *Radiation Physics and Chemistry*, 2016.

[162] M Carante and F Ballarini. Calculating variations in biological effectiveness for a 62 mev proton beam. *Front. Oncol.*, 76(6), 2016.

[163] M Conforth and J Bedford. A quantitative comparison of potentially lethal damage repair and the rejoining of interphase chromosome breaks in low passage normal human fibroblasts. *Radiat. Res.*, 111:385–405, 1987.

[164] M Carante et al. First benchmarking of the bianca model for cell survival prediction in a clinical hadron therapy scenario. *Physics in Medicine and Biology*, pages https://doi.org/10.1088/1361–6560/ab490f, 2019.

[165] S Webb and AE Nahum. A model for calculating tumour control probability in radiotherapy including the effects of inhomogeneous distributions of dose and clonogenic cell density. *Physics in Medicine and Biology*, 38:653–656, 1993.

[166] AE Nahum and B Sanchez-Nieto. Tumour control probability modelling: basic principles and applications in treatment planning. *Physica Medica*, 17:13–23, 2001.

[167] AE Nahum and J Uzan. (radio)biological optimization of external-beam radiotherapy. *Comput Math Methods Med*, 2012.

[168] DP Dearnaley et al. Escalated dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the mrc rt01 randomised controlled trial. *Lancet Oncology*, 8:457–487, 2007.

[169] S Levegrun et al. Fitting tumor control probability models to biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer: pitfalls indeducing radiobiological parameters for tumors from clinical data. *International Journal of Radiation Oncology, Biology and Physics*, 51:1064–1080, 2001.

[170] S Levegrun et al. Risk group dependence of dose-response for biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer. *Radiother Oncol.*, 63:11–26, 2002.

[171] DJ Joseph et al. Ten year final results of the trog 03.04 (radar) randomised phase 3 trial evaluating duration of androgen suppression +

zoledronate for locally advanced prostate cancer. *Journal of Clinical Oncology*, 36:1–10, 2018.

List of publications

1. Cicchetti A. et al., Modelling late stool frequency and rectal pain after radical radiotherapy in prostate cancer patients: Results from a large pooled population. *Physica Medica 2016. DOI: 10.1016/j.ejmp.2016.09.018.* I.F.=2.24

2. Cicchetti A. et al., Predicting late faecal incontinence risk after radiotherapy for prostate cancer: new insights from external independent validation. *IJRBP 2018. DOI:10.1016/j.ijrobp.2018.05.013.* I.F:=5.55

3. Scalco E. et al., Texture Analysis of T1-w and T2-w MR Images Allows a Quantitative Evaluation of Radiation-Induced Changes of Internal Obturator Muscles after Radiotherapy for Prostate Cancer. *Med Phys. 2018. DOI: 10.1002/mp.12798.* I.F:=2.62

4. Bedini N., Cicchetti A. et al., Evaluation of Mediators Associated with the Inflammatory Response in Prostate Cancer Patients Undergoing Radiotherapy. *Disease Markers. DOI:10.11155/2017/9128128.* I.F.=2.95

5. Carrara M. et al., Development of a Ready-to-Use Graphical Tool Based on Artificial Neural Network Classification: Application for the Prediction of Late Fecal Incontinence After Prostate Cancer Radiation Therapy. *Int J Radiat Oncol Biol Phys. 2018. doi: 10.1016/j.ijrobp.2018.07.2014.* I.F:=5.55

6. Alessi A. et al., Role of pretreatment 18F-FDG PET/CT parameters in predicting outcome of non-endemic EBV DNA-related nasopharyngeal cancer (NPC) patients treated with IMRT and chemotherapy. *Radiol Med. 2018. DOI:10.1007/s11547-018- 0980-6.* I.F.=1.82

7. Orlandi E. et al., Multivariable model for predicting acute oral mucositis during combined IMRT and chemotherapy for locally advanced nasopharyngeal cancer patients. *Oral Oncol. 2018. DOI: 10.1016/j.oraloncology.* 2018.10.006. I.F.=4.79

8. Cicchetti A. et al., Predicting late faecal incontinence risk after radiotherapy for prostate cancer: new insights from external independent validation. Reply to Loganadane *IJRBP 2019*. I.F:=5.55

9. Seibold P. et al., REQUITE: A prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. *Radiother. Oncol.* I.F:=5.22

10. Onjukka E. et al., Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT. *Acta Oncol.* I.F:=3.47