Università degli Studi di Pavia

Dipartimento di Fisica

Corso di Laurea Magistrale in Fisica

Comparison of dose distributions in BNCT treatment of brain tumors using human or AI-segmented medical images

MSc thesis by:
Cristina Pezzi

Supervisor:
Prof. Ian Postuma
Co-supervisors:
Dr. Setareh Fatemi
Dr. Barbara Marcaccio

A.A. 2022-2023

"La creatività è soprattutto la capacità
di porsi continuamente delle domande."
Piero Angela

Abstract

Questa tesi valuta l'esito del trattamento con Terapia di Cattura Neutronica del Boro (BNCT) utilizzando algoritmi di Intelligenza Artificiale (AI) per delineare i volumi tumorali. Fornire una accurata e precisa segmentazione del tumore è cruciale per un trattamento mirato ed efficace. Tuttavia, questa procedura può presentare problemi e difficoltà che possono influenzare la buona riuscita del trattamento. L'utilizzo di algoritmi avanzati di AI, come le reti neurali convoluzionali U-Net (CNNs), potrebbe compensare tali ostacoli, consentendo il riconoscimento automatico delle regioni tumorali dalle immagini di risonanza magnetica (MRI) e tomografia computerizzata (CT). Questo approccio non solo ridurrebbe il carico di lavoro dei radiologi, ma potrebbe anche facilitare il processo di segmentazione, consentendo una valutazione più rapida delle regioni di interesse e un'ottimizzazione del piano di trattamento. Inoltre, l'applicazione dell'AI potrebbe essere particolarmente utile anche per la ricerca, fornendo un elevato numero di immagini mediche segmentate in breve tempo. Questo sarebbe particolarmente prezioso nello studio di nuove terapie, compresa la BNCT, fornendo ai ricercatori un elevato numero di immagini mediche segmentate in relativamente breve tempo.

Per questo lavoro di tesi ho ottenuto la segmentazione del tumore attraverso l'algoritmo di intelligenza artificiale addestrato nel progetto AI_MIGHT. La tesi è incentrata sull' applicazione di questi algoritmi su immagini mediche di pazienti affetti da Glioblastoma Multiforme (GBM), un tumore cerebrale caratterizzato da elevata aggressività e tendenza alla recidiva. A causa della resistenza alla radioterapia convenzionale, la BNCT è considerata una terapia promettente per il trattamento del GBM. Le immagini mediche dei pazienti sono state ottenute da un database pubblico, al quale sono stati aggiunti due ulteriori casi provenienti da pazienti che effettivamente sono stati sottoposti a un trattamento di BNCT.

Lo scopo del lavoro è la valutazione dell'efficacia di questi algoritmi di segmentazione automatica attraverso un'analisi dosimetrica, mettendo a confronto i valori di dose assorbita nei pazienti, calcolata nel caso in cui il tumore fosse segmentato manualmente o tramite l'ausilio di rete neurale. Per simulare il trattamento ho impiegato IT_STARTS, un software innovativo per i piani di trattamento BNCT sviluppato a Pavia. Nel corso della tesi, sono mostrati i risultati di tale confronto, i quali hanno evidenziato una correlazione tra il rapporto delle dosi assorbite nelle due diverse segmentazioni e il coefficiente di Dice, che verrà approfondito nel Capitolo 4. Vengono, inoltre, tratte conclusioni riguardo l'applicabilità di questo approccio nei diversi casi analizzati. Infine, si delinea il lavoro

futuro necessario per approfondire questa analisi ed estendere il metodo descritto al trattamento BNCT di altri tipi di tumore. Questo lavoro fa parte del progetto AI_MIGHT dedicato all'implementazione di software per la segmentazione automatica di immagini mediche per l'ottimizzazione del trattamento BNCT.

Abstract

This thesis evaluates the outcome of Boron Neutron Capture Therapy (BNCT) treatment using Artificial Intelligence (AI) algorithms to contour tumor volumes. Providing accurate and precise tumor segmentation is crucial for targeted and effective treatment. However, this procedure can pose challenges that may affect the success of the treatment. The use of advanced AI algorithms, such as U-Net Convolutional Neural Networks (CNNs), could address these obstacles by enabling automatic recognition of tumor regions from Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) images. This approach would not only reduce the workload of radiologists but also facilitate the segmentation process, allowing for a quicker assessment of regions of interest and optimization of the treatment plan. Additionally, the application of AI could be particularly beneficial for research, providing a large number of segmented medical images in a short period, which would be valuable in studying new therapies, including BNCT.

For this work I obtained the tumor segmentation through the AI algorithm trained in the AI_MIGHT project. The dissertation focused on medical images of patients suffering from Glioblastoma Multiforme (GBM), a brain tumor characterized by high aggressiveness and recurrence propensity. Due to its resistance to conventional radiotherapy, BNCT is considered a promising therapy for GBM treatment. Medical images of patients were obtained from a public database, and images of two additional clinical cases who had undergone BNCT were also analyzed.

The purpose of the work was to assess the effectiveness of these automatic segmentation algorithms through a dosimetric analysis, comparing the dose absorbed in patients where the tumor was segmented manually versus using neural network assistance. For this purpose, I used IT_STARTS, an innovative Treatment Planning System for BNCT developed in Pavia. Throughout the thesis, the results of this comparison are presented, which highlighted a correlation between the ratio of absorbed doses in the two different segmentations and the Dice coefficient, to be further explored in Chapter 4. Furthermore, conclusions on the applicability of this approach in various analyzed cases are drawn. Moreover, future work to deepen the analysis and to enlarge the methods to other tumors in BNCT treatment is outlined. This work is part of the AI_MIGHT project dedicated to implementing software for medical image segmentation to optimize BNCT treatment planning.

Contents

Introduction

Glioblastoma Multiforme (GBM) is the most frequent and aggressive primary central nervous system tumor, categorized within the group of diffuse astrocytic and oligodendroglial tumors by the 2016 WHO classification of CNS tumors [1]. Research indicates that GBM affects individuals at an estimated incidence rate of around 3-4 cases per 100.000 population per year worldwide. The average onset age of glioblastoma is approximately 65 years, with an incidence at this age of about 10-12 cases per 100.000 population annually. Median survival is around 12 months, with a survival rate of less than 5% at 5 years post-diagnosis. Standard therapy for the treatment of this pathology typically involves surgical resection followed by radiotherapy and chemotherapy. Due to its infiltrating nature and resistance to conventional therapies, the prognosis are often extremely poor, and patients may be enrolled in clinical trials exploring novel treatments.

Boron Neutron Capture Therapy (BNCT) is emerging as a promising approach for the treatment of various cancers, including GBM. Its effectiveness lays on selectively targeting tumor cells while minimizing damage to healthy tissue as a result of a nuclear reaction within the cancer cells. Recent clinical studies investigating BNCT for GBM treatment have shown encouraging outcomes [2]. These studies have indeed demonstrated not only tumor regression but also higher survival rates in patients. However, there are still challenges in optimizing BNCT protocols, enhancing boron delivery to tumor sites, and improving treatment planning and delivery techniques. Furthermore, also the availability of suitable neutron sources and facilities for the treatment remains limited, constraining its widespread clinical implementation.

Albeit less crucial than in photon and hadron therapy, accurate delineation of target volumes is essential for the BNCT success, and it directly impacts on the treatment effectiveness. This process is known as *tumor contouring*. In addition to ensuring optimal targeting of the tumor while minimizing radiation exposure to surrounding healthy tissues and critical structures, tumor contouring facilitates the correlation between the volumetric distribution of the dose in critical targets and the clinical effects observed in patients. Clinicians can thus evaluate the treatment effectiveness, understand the possible unwanted side effects and implement more effective dose limitations. Traditional methods of manual segmentation by physicians have long been considered as the golden standard for tumor contouring. However, the emergence of Artificial Intelligence (AI) offers a promising strategy to improve the state of the art. The current development of

Artificial Neural Network (ANN) has provided new methods of automatic tumor contouring, potentially capable to enhance and speed up the processing of an adequate treatment plan.

This thesis is framed in a project (AI_MIGHT) dedicated to the implementation of AI contouring for the treatment planning in BNCT. In particular, the aim of this work was to simulate treatment planning of BNCT in several clinical cases of GBM, comparing the dosimetry obtained when the target regions were contoured manually with the one obtained automatically. The suitability of such strategy is discussed in light of different figures of merit describing the dose distribution and the clinical outcome. The dissertation is structured as follows.

The opening chapter provides a comprehensive overview of BNCT. The first section focuses on the fundamental physical principles underlying BNCT, including the essential characteristics of boron carriers and neutron beams for optimizing therapy. The following section deals with the dosimetry of BNCT, describing the models currently in use to translate BNCT dose into photon-equivalent units.

In the second chapter the focus shifts to Treatment Planning Systems (TPS) employed in BNCT and their operational mechanisms. After analyzing the differences with the TPS adopetd in conventional radiotherapy in the first section, the second section describes the main characteristics of TPS used in BNCT, providing few examples. Lastly, IT_STARTS, an innovative treatment plan developed in Pavia and adopted for the calculations of this work, is introduced.

The third chapter introduces the application of AI in medicine, with a particular emphasis on cancer imaging and contouring. In order to better understand the functioning of the algorithms used, notions of DL and ANN will be provided in the second section. At the conclusion of this chapter, the concept of U-Net and its architecture are described, explaining how it is applied to medical image contouring. Additionally, the application of U-Net in the AI_MIGHT project is showed.

The fourth chapter reports the clinical cases, the simulations, the results and the analysis conducted. The use of ANN for patient segmentation is evaluated through the result of the in-patient dosimetry. Simulations were performed considered manual and automatic contouring and the differences were critically analyzed. Correlation was studied between dose values and parameters describing the performance of the ANN. Finally, a radiobiological figure of merit based on the tumor control probability of a given dose distribution was introduced to better describe the impact of the contouring on the clinical outcome of the treatment.

Chapter 1

BNCT: Boron Neutron Capture Therapy

1.1 Fundamentals of BNCT

Boron Neutron Capture Therapy, also known as BNCT, is a multidisciplinary and innovative technique for malignant tumor treatment. It is a binary approach as both a ^{10}B-labeled compound and a low-energy neutron beam are needed to deliver therapeutic doses to the target tumor. A BNCT treatment involves basically two steps. First of all, it is necessary to deliver efficiently and selectively a suitable amount of ^{10}B atoms inside the target, which consists of the tumor tissue. Then the tumor must be irradiated with a neutron beam having proper characteristics in terms of spectrum and intensity in order to trigger the wanted nuclear reaction and deliver sufficient dose. The products of the neutron capture reaction on ^{10}B have a high Linear Energy Transfer (LET) and deposit all their energy in the cancer cells. While, due to their short range in tissue, surrounding healthy cells are preserved [3].

BNCT acts completely differently from conventional radiotherapy and hadrontherapy, where an extremely focused beam of photons (or hadrons) is needed to hit and destroy tumor cells. In particle therapy, the beam energy is set in such a way that the maximum dose is absorbed within the tumor, located at a specific depth, by adjusting the position of the Bragg peak. On the other hand, BNCT relies on the localized accumulation of boron in the different tissues, which is the factor that ensures the selectivity of the treatment.

The idea to use neutrons to treat cancers dates back about almost a century ago, after the neutron discovery from James Chadwick in 1932. Several difficulties have prevented a regular and broad clinical use in the past [4]. However, BNCT is recently experiencing a new era of expansion thanks to technological advances and the latest clinical results are particularly promising [2, 5, 6].

1.1.1 Nuclear reaction

The physical principles of BNCT were studied in 1936 by Gordon Locher [7], and it was applied since the early 1950s using nuclear reactors as neutron source.

BNCT is based on the following nuclear reactions:

$$\begin{aligned} {}^{10}B + n \longrightarrow {}^{11}B \longrightarrow {}^{7}Li^{*} + {}^{4}He + 2.31 \text{ MeV} & \\ {}^{7}Li^{*} \longrightarrow {}^{7}Li + \gamma + 0.478 \text{ MeV} \qquad [BR = 93.9\%] & \end{aligned} \tag{1.1}$$

$${}^{10}B + n \longrightarrow {}^{11}B \longrightarrow {}^{7}Li + {}^{4}He + 2.79 \text{ MeV} \qquad [BR = 6.1\%] \tag{1.2}$$

which occur with a cross section of $\sigma = 3837$ barns for thermal neutrons having energy of $E_n = 0.025$ eV.

The selectivity of BNCT is thus based on the preferential uptake of 10-boron in the tumor. Neutrons, however, interact also with other elements in the biological tissues, producing a mixed radiation field that deposits dose in a non-selective way. This inevitable dose component can be limited not to exceed the tolerance of normal tissues, while the tumor absorbs a higher dose when the boron concentration is sufficiently high.

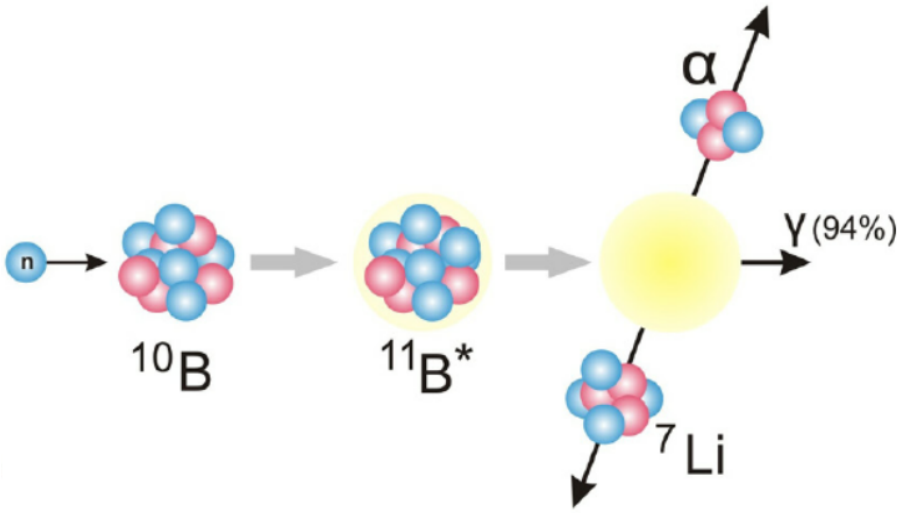


Figure 1.1: *Nuclear reaction between* ^{10}B *and thermal neutrons. Illustration from [8].*

As shown in Figure 1.1, the ^{10}B nucleus captures the neutron, and the unstable ^{11}B breaks into a recoil atom of ^{7}Li and an alpha particle. With a probability of 93.9%, the lithium nucleus is produced at an excited state, and so it immediately decays, emitting a 478 keV photon. The lithium ion and the alpha particle, emitted back-to-back, have a range in biological tissues of approximately 5.2 and 7.5 μm respectively. The summed range of those two particles is comparable to the average size of a mammalian cell ($\sim$10 μm).

This means that, if the neutron capture reactions take place in a cell, most of the energy deposition occurs inside that specific cell. The damage is thus limited to cells containing ^{10}B, as shown in Figure 1.2.

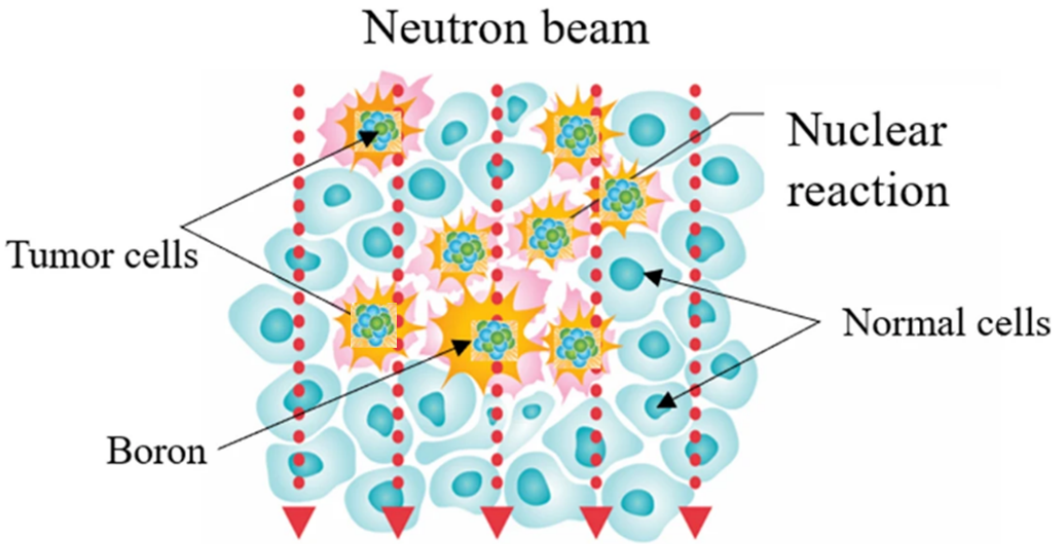


Figure 1.2: *BNCT principle of selectivity whereby the $^{10}B(n,\alpha)^{7}Li$ nuclear reaction only takes place inside tumor cells. Illustration from [9].*

The high LET of the alpha and ^{7}Li ions ensures an effective cell killing. As a matter of fact, experiments have shown that 2 to 6 alpha particles are sufficient to induce a lethal lesion in a cell, depending on the different cell lines [10]. This is due to the high energy transfer in a short path, with a dense dose deposition pattern, which irreversibly damages the DNA and leads to cell death.

Summarizing, BNCT effectiveness is based on the possibility to obtain a sufficient tumor-to-normal tissue boron concentration ratio and on the interaction of thermal neutrons inside the tumor cells.

1.1.2 Boron carriers

As both ^{7}Li and ^{4}He produced in neutron boron capture are characterized by very high biological effectiveness and a short range, the micro-distribution of boron inside the tissues is fundamental to provide an **effective** and **selective** therapy. It is necessary for boron atoms to accumulate preferentially in the target, and for this reason, injecting any boron formulation is not enough [11].

Two ^{10}B delivery agents, the amino acid *p-boronophenylalanine*, known as BPA, and the *sulfhydryl borane* $Na_2B_{12}H_{11}SH$, BSH, have been used for the BNCT clinical treatments.

These boron compounds, shown in Figure 1.3, were selected because they are characterized by negligible toxicity, a certain degree of tumor specificity, and a suitable accumulation of boron inside the target (the tumor concentrations of ^{10}B should be $\sim$ 20 ppm), long retention inside the tumor and, on the contrary, a fast and complete clearance from healthy tissue [12]. These properties allow for enhancement of the tumor-to-normal tissue concentration ratio and for a higher *therapeutic ratio*, which is the ratio between the minimum dose in the tumor and the maximum dose in the healthy tissues.

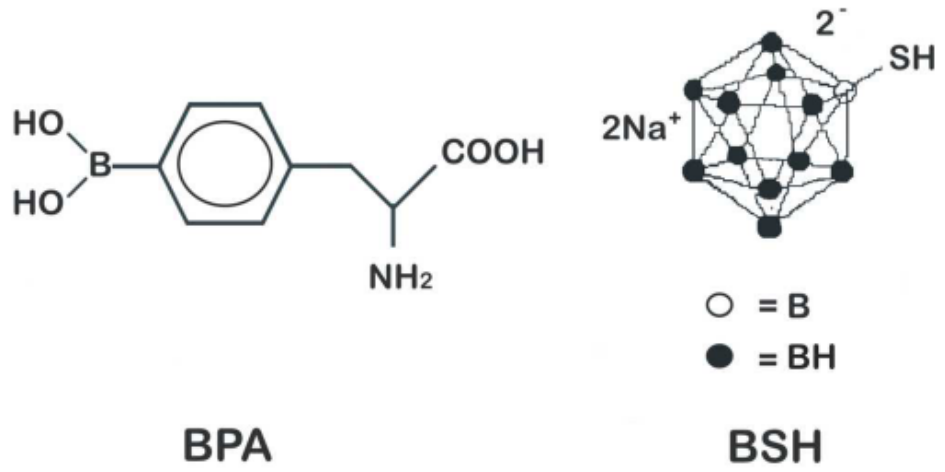


Figure 1.3: *Chemical structure of BPA and BSH. Illustration from [13].*

BPA is the drug currently used in clinical trials. It was used for the first time by *Mishima et al.* in 1987 as a boron delivery agent for the treatment of cutaneous melanomas [14], as it is a precursor of melanin. It can be taken up by cancer cells as a result of enhanced cellular amino acid transport and remains in melanoma cells more than in other cells as it interacts with the melanin polymerization process of melanoma cells [15]. The effectiveness of BPA for the treatment of brain tumors is based on its capability to be actively transported across the blood-brain barrier (BBB) into the normal brain, so it can reach even the most infiltrating gliomas. Preclinical studies in the 9L rat gliosarcoma brain tumor model at BNL showed for the first time that the F-BPA, a more soluble BPA fructose complex, produced a tumor-to-normal brain boron concentration ratio of 5:1 [16, 17].

On the other hand, BSH works in a totally different way to reach the target volume. Unlike BPA, it is not able to actively cross the BBB, but it reaches the tumor cells by a passive diffusion process through the tumor vasculature, which is not tight enough to stop it. Also in this case, reports have shown a tumor:blood boron concentration ratio higher than 1, showing a certain selectivity. Studies have proved the effectiveness of BSH for the treatment of intracranial tumors and its extremely low toxicity [18, 12], while for highly infiltrated tumors, such as glioblastomas, the capability for the boron to cross the BBB is essential, therefore BPA is more suitable.

Both BSH and BPA are approved for clinical use in protocols worldwide, and currently

researchers are studying new promising delivery agents exploiting different strategies to improve the accumulation of boron in the tumor, such as the labeling of the borated formulation with tumor-specific antibodies [12]. Efforts are being devoted all over the world to develop new borated formulations able to increase the BNCT effectiveness.

1.1.3 Neutron beam

In addition to boron delivery agents, an adequate neutron field has to be created in the boron-loaded tumor cells. Unlike radiotherapy and hadrontherapy, BNCT neutron beams do not need to be precisely conformed to the tumor volume, but they should be wide enough to cover and irradiate the whole tumor and possible invisible spreads. In general, circular apertures of 12 to 14 cm in diameter are used for clinical treatments, although they vary depending on the organ to irradiate. However, for brain tumors, neutron beams having a diameter of up to 17 cm have been proposed [19].

Neutron energy has a fundamental role in the treatment. First of all, as said before, the ^{10}B(n,α)^{7}Li reactions (1.1) and (1.2) occur with a higher cross section, as shown in Figure 1.4 , when *thermal neutrons* interact with boron nuclei, so the neutron beam must have low energy. However, the beam energy also depends on the depth of the tumor. As a matter of fact, thermal beams cannot reach deep-seated tumors due to the attenuation in the tissues. In this case, an *epithermal neutron beam* is required.

As shown in Figure 1.5, the neutron flux radically changes with the depth inside a certain material. For thermal beams, the maximum thermal neutron flux is at the beginning and then it exponentially decreases. The epithermal beam, instead, thermalises while interacting. This generates a thermal neutron flux whose maximum is at a depth of 2-3 cm. This peak can be shifted further by increasing the energy of the beam. This means that for deep-seated target volumes, epithermal beams will generally be the best, while for shallow targets, such as melanomas, thermal or mixed thermal and epithermal (hyperthermal) beams will be used [19].

The main features that characterize thermal and epithermal beams are **intensity** and **quality**. Considering the eventual presence of fast neutrons and photons inside the epithermal neutron beam, the recommended values of the beam quality have been published by IAEA in the TecDoc in 2023 [21]:

(i) maximum contamination of thermal neutrons inside an epithermal beam

$$\frac{\phi_{th}}{\phi_{epi}} \leq 0.05$$

(ii) minimum degree of collimation as neutron current over epithermal neutron flux

$$\frac{J}{\phi_{epi}} \geq 0.7$$

(iii) maximum ratio between gamma dose rate and epithermal neutron flux

$$\frac{\gamma}{\phi_{epi}} \leq 7 \cdot 10^{-13}\ cm^2\ Gy$$

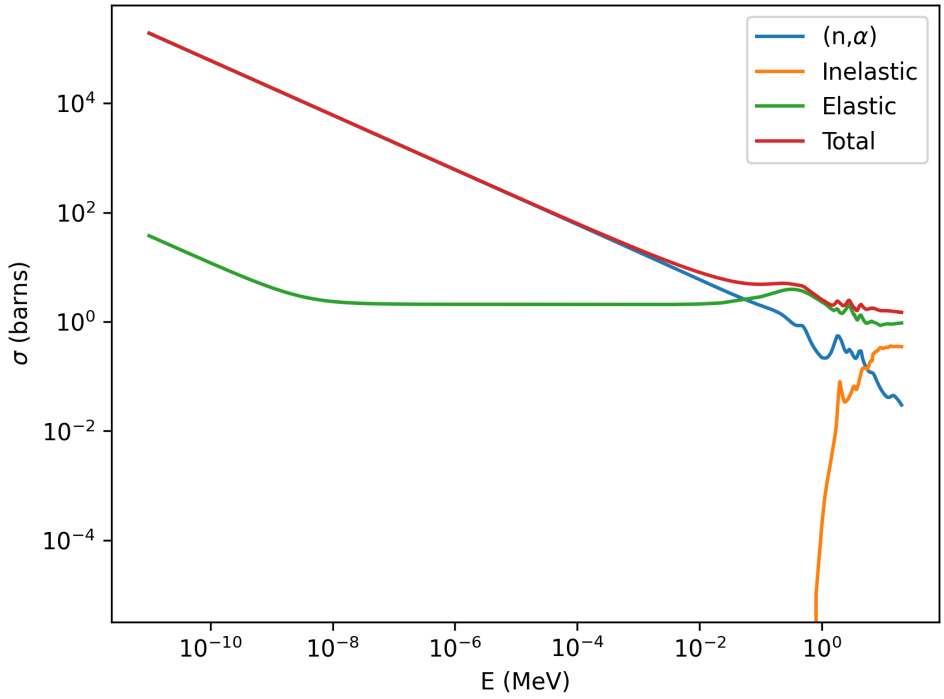


Figure 1.4: *Neutron cross sections with boron reactions. For low energies the main reaction that takes place is the neutron capture. Illustration made using data from ENDF/B-VIII [20].*

(iv) maximum ratio between fast neutron dose rate and epithermal neutron flux

$$\frac{\phi_{fast}}{\phi_{epi}} \leq 2 \cdot 10^{-13}\ cm^2\ Gy$$

(v) minimum epithermal neutron flux to trigger enough nuclear reactions with boron in a reasonable treatment time

$$\phi_{epi} \geq 5 \cdot 10^8 cm^{-2}\ s^{-1}\ .$$

To produce such neutron flux, the beam can be generated from nuclear reactors [22] or particle accelerators. The latter are becoming more and more common worldwide [23, 24, 25] as they present many more advantages compared to reactors: they are smaller and can be installed inside hospitals, requiring simpler licensing, operation and maintenance [26]. They can be certified as clinical devices paving the road towards clinical trials. Reactors are normally installations used for other purposes and not employed for clinical research alone.

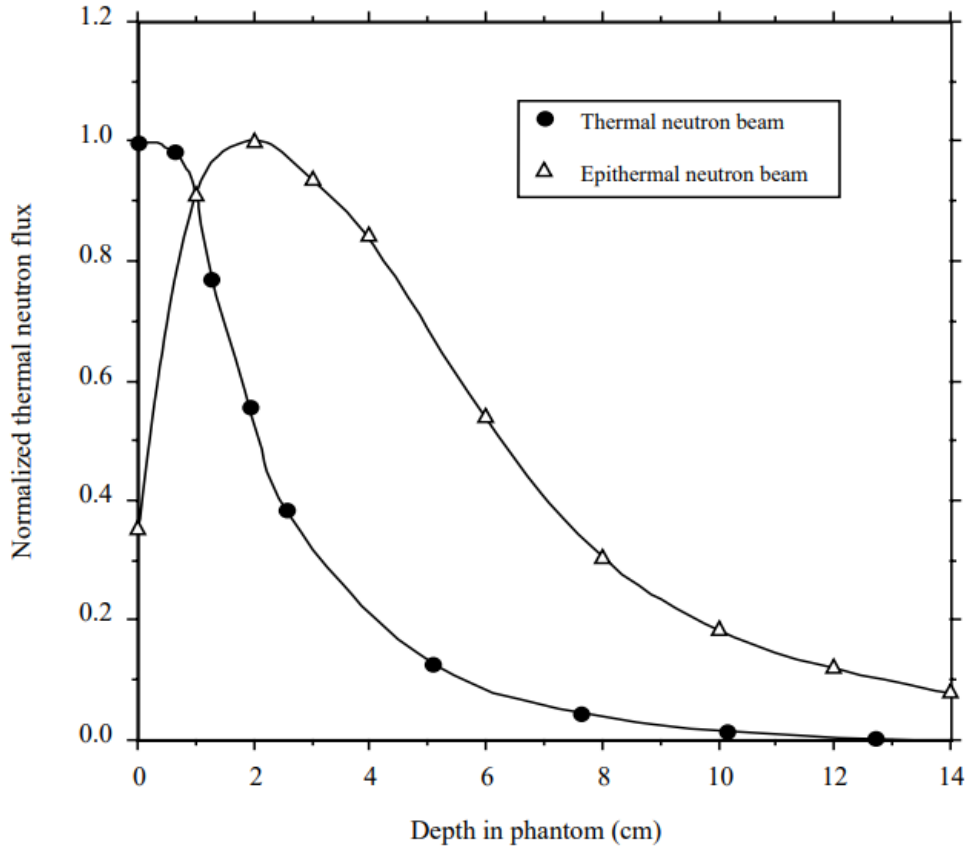


Figure 1.5: *Comparison of flux-depth distributions for thermal and epithermal neutrons in a tissue equivalent phantom. Illustration from [19].*

Accelerator-based BNCT, known as AB-BNCT, is based on the production of neutrons through nuclear reactions between accelerated charged particles and a specific target. The most common accelerators rely on ^{7}Li(p,n)^{7}Be, ^{9}Be(p,n)^{9}B, and ^{9}Be(d,n)^{10}B endothermic and exothermic reactions. In these cases, protons or deuterons are accelerated and directed against a lithium or beryllium target. The produced neutron beam is not suitable for patient treatment as-is because the energy is too high (typically around 200 keV for Li and 1.5 MeV for Be), and the collimation is not adequate. It is thus necessary to optimize the spectrum by "slowing down" the average energy and by removing the unwanted component of the beam with a structure called *Beam Shaping Assembly* (BSA). Neutrons are passively moderated with proper materials and geometries, to tailor the beam for patient irradiation [27].

There are currently about 30 projects related to the installation, commissioning and use of accelerators for clinical BNCT in the world (`https://isnct.net`). The most advanced projects are those based on Sumitomo Heavy industries 30 MeV cyclotron coupled to a Be target in Japan, where the clinical application is already covered by the Health

National System for the treatment of Head & Neck tumors [28]. In China a clinical trial is currently ongoing using a linear accelerator manufactured by TLS and run by Neuboron Medtech. Other accelerators are being commissioned for clinical use in Finland, South Korea, Japan. In Italy, two projects of clinical facilities are being developed: one in CNAO (Pavia) based on a commercial solution by TLS and one at the University of Campania L. Vanvitelli based on the Radiofrequency Quadrupole accelerator and Be target designed and built at the National Laboratory of Legnaro of INFN [29, 30, 31]. The work described in this thesis is framed in this project and it has been carried out using the designed neutron beam described in [27].

1.2 BNCT dosimetry

The radiobiology and the dosimetry of BNCT are more complex than those of other radiation therapy modalities. As said in Section 1.1.1, BNCT is based on the nuclear capture reaction between thermal neutrons and ^{10}B atoms, and the resulting ions are responsible for dose deposition. However, ^{10}B(n,α)^{7}Li interaction is not the only one involved during the treatment. Neutrons can also interact with other elements in human tissues producing particles characterized by different physical properties and biological effectiveness [19]. These particles deliver an unavoidable and non-negligible background dose to both tumor and normal tissues. Essentially, BNCT is characterized by a mixed radiation field. Dose contributions of different particles have to be summed by taking into account their biological effectiveness.

1.2.1 BNCT dose components

According to ICRU Report 44 [32], soft tissue is mainly composed of four atoms in the following concentrations: oxygen ^{10}O (76.2%), carbon ^{12}C (11.1%), hydrogen ^{1}H (10.1%) and nitrogen ^{14}N (2.6%). For dose estimation in BNCT, the *hydrogen dose* and *nitrogen dose*, namely the doses due to the interactions of neutrons with hydrogen and nitrogen atoms, have a key role because of their higher cross sections and KERMA factors, while the dose delivered by the nuclear reactions involving oxygen and carbon is negligible [33], as shown in Figure 1.6. Elastic scattering between neutrons and hydrogen nucleus has a lower cross section than capture. However, the dose contribution is high because there are many hydrogen atoms and because the recoil protons, especially above 10 keV, release a non-negligible dose.

While penetrating the tissues, epithermal and fast neutrons are thermalized through elastic scattering with hydrogen atoms. In this ^{1}H(n,n')p reaction, the incoming neutron loses on average half of its energy and a high-LET recoil proton is set off. On the contrary, thermal neutrons are mainly involved in *capture processes*, such as ^{1}H(n,γ)^{2}H and ^{14}N(n,p)^{14}C. In the first case, a low-LET photon is produced at an energy of $E_\gamma = 2.2$ MeV, while the nitrogen capture process releases an intermediate-LET proton having $E_p = 590$ keV [21].

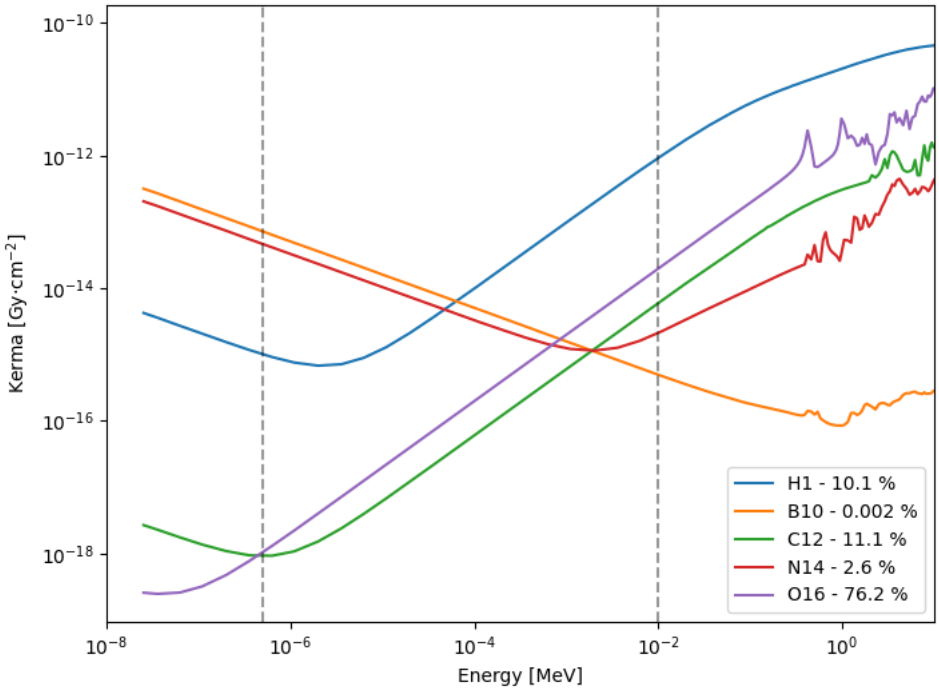


Figure 1.6: *Neutron cross-section of main tissue elements from ENDF/B-VIII.0 [34]. Boron, hydrogen and nitrogen are more likely to react with thermal energy neutrons compared to carbon and oxygen, whose reactions can be negligible in terms of dose calculation. The dotted lines separate the various thermal, epithermal and high energy regimes.*

The BNCT dose components can be summarized in four groups whose depth-dose profiles, for an epithermal neutron beam, are shown in Figure 1.7:

(i) $\mathbf{D}_B$ *boron dose component*, which is the dose delivered due to the high-LET light ions produced in ^{10}B(n,α)^{7}Li reaction. This is the major contribute to the total dose [33].

(ii) $\mathbf{D}_{th}$ *thermal neutron dose component*, which is the dose delivered by thermal neutrons ($E_{th} < 0.5$ eV) while interacting with tissue. As shown in Figure 1.7, the main process depositing dose due to thermal neutrons other than boron capture is ^{14}N(n,p)^{14}C.

(iii) $\mathbf{D}_f$ *fast neutron dose component*, which is the dose contribute due to the thermalization of fast neutrons ($E_f > 10$ keV) in which protons are emitted with an energy that is half that of incoming neutrons on average. As shown in Figure 1.7, it is the highest component in the skin and decreases exponentially with depth as neutrons are slowed down.

(iv) $\mathbf{D}_\gamma$ *gamma dose component*, which is the dose delivered by photons produced in the ^{1}H(n,γ)^{2}H reaction, in the^{10}B(n,α)^{7}Li reaction and by photons present in the incident neutron beam.

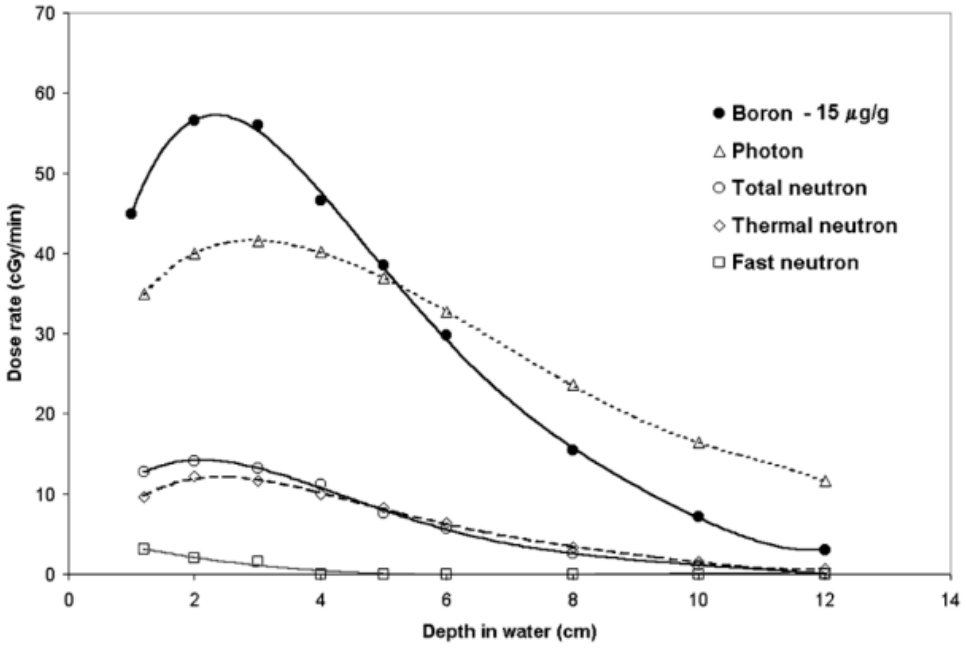


Figure 1.7: *Depth dose profiles representative of an epithermal neutron beam at the MITR-II FCB Research Reactor in United States [35]. Illustration from [3].*

Given this mixed-radiation field, the issue is to identify a radiobiological model able to describe how the dose components interact and cause a certain effect in biological tissues. As a matter of fact, the establishment of a suitable model has a key role in the development of a correct *treatment planning*, ensuring the capacity to prescribe the therapeutic dose to the tumor and, at the same time, a safe dose value to the surrounding healthy tissues.

As mentioned in Section 1.2.1, different nuclear reactions are involved during a BNCT treatment, which delivers a specific dose. For this reason, the total absorbed dose delivered by BNCT is given by the arithmetical sum of these dose components:

$$D = D_B + D_{th} + D_f + D_\gamma \tag{1.3}$$

However, BNCT clinical protocols make use of *photon-equivalent doses* to make predictions of the therapy outcome based on the knowledge gained in photon therapy [36]. For this reason, it is required to convert BNCT dose value in photon-equivalent units. Traditionally, the BNCT dose translated into photon units has been indicated as *photon-equivalent dose* or *biologically-weighted dose* D_w. A large variety of units to indicate the administered dose in photon equivalent units have been established by the BNCT community, such as Gy-Eq, or the more specifical Gy(RBE) and Gy(IsoE), depending on the adopted model [21].

For the conversion, it is necessary to consider that the BNCT mixed radiation field comprises radiations with different Linear Energy Transfer (LET) and Relative Biological Effectiveness (RBE). This parameter, which is LET dependent, is defined as the ratio of the absorbed dose of a reference radiation (often taken to be ^{60}Co gamma rays), D_R, to the absorbed dose of the particular radiation being examined, D_X, that is required to obtain the same level of biological effectiveness [37]:

$$RBE_X = \frac{D_R}{D_X} \tag{1.4}$$

Essentially, it indicates the biologically effectiveness of a radiation under study compared to photons.

1.2.2 The RBE-weighted Dose model

The absorbed BNCT dose D can be converted into weighted dose D_w, by multiplying each dose component by the respective (dose and dose rate independent) RBE factors. This radiobiological model is known as **RBE-weighted dose model** [38]. Regarding the boron dose component D_B, it is, however, fundamental to take into account also the boron micro-distribution inside the tumor cells. As mentioned in Section 1.1.2, the selectivity and the effectiveness of BNCT depend on the way boron atoms are accumulated inside the target volume. For this reason, the boron dose contribution requires to be weighted by another parameter, called *Compound Biological Effectiveness* (CBE), whose value varies with the boron compounds, the target tissue or cell, and the endpoint of the assessment.

RBE and CBE factors are typically obtained through *in-vitro* measurements, using tumor cell lines irradiated with photons and neutrons in presence of boron. The cell survival as a function of dose is assessed and a fixed endpoint is chosen to calculate the RBE and CBE [17]. RBE and CBE factors have been established for different cell lines, including rat skin and HeLa cells, using both BPA and BSH carriers at different neutron facilities, including the thermal neutron field at the TRIGA Mark II reactor in Pavia [39, 40, 41, 42].

The BNCT weighted dose is calculated as follows:

$$D_w = CBE \cdot D_B + RBE_{th} \cdot D_{th} + RBE_f \cdot D_f + RBE_\gamma \cdot D_\gamma. \tag{1.5}$$

The RBE-weighted dose model has been used in BNCT clinical applications so far.

However, some recent studies conducted at the Comisión Nacional de Energía Atómica (CNEA) in Argentina show that this model is affected by formal inconsistencies which gives dose values that are not equivalent to photon dose [43]. As a matter of fact, the RBE dose model is proven to overestimate photon-equivalent dose values, and it is thus non suitable to understand the observed clinical outcomes in terms of the BNCT dose values translated using fixed-RBE factors.

1.2.3 The Photon Isoeffective Dose model

The RBE-weighted dose model leads to unrealistic tumor photon-equivalent doses, and this prevents its use to predict the effectiveness of a treatment or to limit the side effects.

One of the most important objections to the conventional RBE-weighted dose model is that weighting factors are considered as dose and dose rate independent values. As a result, a new formalism called **photon-isoeffective dose model** has been proposed by González and Santa Cruz [43], which exploits the entire survival curve, instead of relying only in one endpoint.

The photon-isoeffective dose model relies on two main assumptions: the first order repair of sublethal lesions and the synergistic interaction between different radiations. As said in Section 1.2.1, during a BNCT treatment several reactions take place and the products of these interactions are responsible of DNA lesions. This effect is described by the *theory of dual radiation action* (TDRA) [44] whose primary tenet is that lethal lesions can arise either from the direct action of single events or by the incoherent action of independent events combining together. This expression is formally equivalent to the linear quadratic (LQ) formalism described by

$$S(D) = e^{-(\alpha D + \beta D^2)} \tag{1.6}$$

according to which the yield of lethal damage due to a single event is proportional to the dose, while the one due to the combination of events is proportional to the square of the dose deposited. In the previous Equation (1.6), $S(D)$ corresponds to the survival fraction of irradiated cells, while the parameters α and β are the constant terms linking cell lethality respectively to damage per single hit (linear term) or per sum of two separate damages (quadratic term).

Cellular damages that, instead, do not imply lethal lesions are known as sublethal lesions, or *sublesions*, and they can either combine together leading to lethal damages or be fixed by cellular repair mechanisms. First-order lesion repair was originally taken into account in the pioneering works of Lea and Catcheside in the 1940s [45] resulting in the generalized Lea-Catcheside time dependent factor $G(\theta)$. This term refers to the reduced probability of sublesions interactions during the irradiation time θ considering the presence of DNA repair mechanisms.

Mathematically, it can be expressed as the linear combination of two contributes, taking into account the double kinetics of slow (t_{0s}) and fast (t_{0f}) repair as follows [43]:

$$G_i(\theta, t_{0f}, t_{0s}) = p_{i_f} G(\theta, t_{0f}) + p_{i_s} G(\theta, t_{0s}) \ , \tag{1.7}$$

where p_{i_f} and p_{i_s} are the proportions of sublesions repaired by the fast and slow kinetics for the $i-th$ radiation (with $p_{i_f} + p_{i_s} = 1$), and

$$G(\theta, t_0) = \frac{2t_0}{\theta}\left[1 - \frac{t_0}{\theta}\left(1 - e^{-\frac{t_0}{\theta}}\right)\right] \ , \tag{1.8}$$

where t_0 is the fast or the slow characteristic repair times.

Lea-Catcheside factor affects and modifies the quadratic term, taking values from zero to one, with $G = 1$ when all damages have combined and $0 < G < 1$ when repair has taken place reducing cell killing [46]. As a result, the former equation becomes

$$S(D) = e^{-(\alpha D + G(\theta)\beta D^2)}. \tag{1.9}$$

$G(\theta)$ factor is directly dependent on irradiation time without regard to the spatial distribution of the initial lesions.

The other effect taken into consideration by the photon-isoeffective dose model is the *radiation synergism* [47]. It is indeed possible for coexisting sublesions produced by different radiations to interact with each others and produce a lethal damage to the target. Several experiments have demonstrated this phenomenon [48, 49, 50]. In particular, it has been shown that there is a synergistic behavior of high and low LET radiation and how this leads to an increased effect. In terms of linear quadratic formalism, the expression of the survival fraction $S(D)$ is modified to include mixed terms each one modulated by the respective $G(\theta)$ factor.

The formalism of photon-isoeffective dose model is, thus, based on the following assumptions [43]:

(i) the survival dose-response relationship is described by linear quadratic model that accounts for dose-rate dependent sublesion repair (MLQ);

(ii) if synergism is taken into account, the survival dose-response is adequately described including the additional mixed terms from TDRA, modulated by Lea-Catcheside factors G experimentally derived by Suzuki [51] for simultaneous mixed irradiation.

The photon-isoeffective dose formalism applied to BNCT translates the BNCT dose into the photon-equivalent dose. For simplicity, let D_1, D_2, D_3 and D_4 be respectively the boron, thermal neutron, fast neutron and γ absorbed dose components of the BNCT mixed field and let us consider the reference radiation R. It is possible to calculate the reference radiation dose D_R as a function of the four dose components, $D_R = D_R(D_1, ..., D_4)$ using two different approaches [43].

Independent Action

Let us consider the survival probability $S(D_i)$ given from Equation (1.6) for each i-th dose component $i = 1, ...,4$. It is, however, possible to allow the quadratic term only for the gamma dose component and neglect it for the other contributions. High LET radiations are far more likely to induce single hit lethal lesions than low LET ones and so the survival probability becomes

$$S_i(D_i) = \begin{cases} e^{-\alpha_i D_i} & i = 1,2,3 \\ e^{-(\alpha_i D_i + G_i(\theta)\beta_i D_i^2)} & i = 4 \end{cases} \tag{1.10}$$

where α_i and β_i are the coefficients of the single-fraction linear-quadratic survival model for the corresponding radiations and $G_{i=4}(\theta)$ is the generalized Lea-Catcheside time factor for the gamma dose component. Let $S_R(D_R)$ be the survival fraction for the reference radiation R and be written as a function of the BNCT dose, i.e.

$$S_R(D_R) = S(D_1, ..., D_4). \tag{1.11}$$

If no synergistic effects are taken into account and each dose component acts independently, the survival probability for the combination of the four contributions is just given

by the product of the single components $S_i(D_i)$ from Equation (1.10):

$$S(D_1, ..., D_4) = \prod_{i=1}^{4} S_i(D_i). \tag{1.12}$$

Considering Equations (1.10) and (1.11), the latter Equation (1.12) becomes

$$-ln(S_R(D_R)) = \sum_{i=1}^{4} \alpha_i D_i + G_4(\theta)\beta_4 D_4^2. \tag{1.13}$$

It is moreover possible to rewrite D_R as the sum of the four dose contributes weighted by RBE factors $r_i(D_R)$ which, unlike the fixed ones used in the RBE-weighted dose model, are reference-dose dependent. One can, thus, obtain

$$D_R(D_1, ...D_4) = \sum_{i=1}^{4} r_i(D_R) D_i. \tag{1.14}$$

Combining Equations (1.14) with (1.13), it is possible to get RBE values $r_i(D_R)$ as

$$r_i(D_R) = \begin{cases} \alpha_i \dfrac{D_R}{-ln(S_R(D_R))} & i = 1,2,3 \\ (\alpha_i + G_i(\theta)\beta_i D_i^2) \dfrac{D_R}{-ln(S_R(D_R))} & i = 4 \end{cases} \tag{1.15}$$

Now, let us suppose that the survival probability of the reference radiation $S_R(D_R)$ is given by the following single-fraction linear quadratic dose

$$-ln(S_R(D_R)) = \alpha_R D_R + G_R(\theta')\beta_R D_R^2 \ , \tag{1.16}$$

where α_R and β_R are the LQ model parameters and $G_R(\theta')$ the generalized Lea-Catcheside time factor for the reference radiation R.

Considering the Equations 1.16, 1.15 and 1.14, the appropriate expression achieved to evaluate the photon-isoeffective dose for the BNCT mixed radiation field in absence of synergistic actions is

$$\begin{aligned} D_R(D_1, ..., D_4) &= \sum_{i=1}^{4} r_i(D_R) D_i \\ &= \sum_{i=1}^{3} \left[\alpha_i \frac{D_R}{-ln(S_R(D_R))} \right] D_i + \left[(\alpha_i + G_i(\theta)\beta_i D_i^2) \frac{D_R}{-ln(S_R(D_R))} \right] D_4 \\ &= \sum_{i=1}^{3} \left(\frac{\alpha_i}{\alpha_R + G_R(\theta')\beta_R D_R} \right) D_i + \left(\frac{\alpha_4 + G_4(\theta)\beta_4 D_4}{\alpha_R + G_R(\theta')\beta_R D_R} \right) D_4 \ . \end{aligned} \tag{1.17}$$

This equation, though, cannot be analytically solved and does not allow the calculation of D_R, as an explicit expression of D_R cannot be simply obtained. Nevertheless, when cell survival experiments are carried out with a constant irradiation time (i.e. changing

the dose rate for each point) or if the irradiation time is shorter than the characteristic repair time, then $G_R(\theta') = G_R$ is constant. Therefore, for these cases, Equation 1.17 can be solved for D_R, obtaining

$$D_R(D_1,...,D_4) = \frac{1}{2}\frac{\left(\frac{\alpha}{\beta}\right)_R}{G_R} \times \left(\sqrt{1 + \frac{4G_R}{\alpha_R \left(\frac{\alpha}{\beta}\right)_R} \left(\sum_{i=1}^{4} \alpha_i D_i + G_4(\theta)\beta_4 D_4^2 \right)} - 1 \right) \ . \quad (1.18)$$

Synergistic Action

In this case, different radiations are able to actively interact with each other, producing an additional effect that enhances cellular death. Sublesions produced by the *i-th* radiation can combine with the ones produced by any other different *j-th* radiation and form lethal damage.

Let assume the survival probability from LQ formalism

$$S_i(D_i) = e^{-(\alpha_i D_i + G_i(\theta)\beta_i D_i^2)} \quad (1.19)$$

for every *i-th* radiation component alone. Synergism between component i and j is expressed through the following relationship

$$-ln(S_{ij}(D_i, D_j)) = G_{ij}(\theta)\sqrt{\beta_i \beta_j} D_i D_j \ , \qquad i \neq j = 1,...,4 \quad (1.20)$$

where $G_{ij}(\theta)$ is the Lea-Catcheside time factor accounting for first-order repair of sublesions produced by radiation i interacting with sublesions produced by radiation j during the irradiation time θ. The whole survival probability for the combination of all dose components is, thus

$$-ln(S(D_1,...,D_4)) = \sum_{i=1}^{4} \alpha_i D_i + \sum_{i=1}^{4}\sum_{j=1}^{4} G_{ij}(\theta)\sqrt{\beta_i \beta_j} D_i D_j \ . \quad (1.21)$$

It should be noted that for $i = j$, Equation 1.21 turns to the Equation 1.9. As seen before, it is possible to introduce a reference radiation R whose dose $D_R(D_1,...,D_4)$ produces the same survival level as the combination of all components. Neglecting any variation of $G_R(\theta')$ for the reference data, we can finally obtain the general expression for the photon-isoeffective dose as

$$D_R(D_1,...,D_4) = \frac{1}{2}\frac{\left(\frac{\alpha}{\beta}\right)_R}{G_R} \times \left(\sqrt{1 + \frac{4G_R}{\alpha_R \left(\frac{\alpha}{\beta}\right)_R} \left(\sum_{i=1}^{4} \alpha_i D_i + \sum_{i=1}^{4}\sum_{j=1}^{4} G_{ij}(\theta)\sqrt{\beta_i \beta_j} D_i D_j \right)} - 1 \right) \ . \quad (1.22)$$

Both reference and BNCT radiation parameters must be experimentally evaluated for different experimental conditions and different cell lines.

To study the reliability of these models, few experimental trials have been led and the final dose outcomes have been put in comparison with the number of cancer cells controlled

after the treatment. The photon isoeffective dose values obtained for brain tumors using the radiobiological experiments performed for the rat gliosarcoma by *Coderre et al.* [52] are lower than the dose obtained by the fixed-RBE dose model. This is shown in Figure 1.8, where it can be seen that an absorbed dose in tumor of 15 Gy, corresponds to photon isoeffective doses of about 28 Gy (IsoE) and 30 Gy (IsoE) without and with synergism respectively, and to 51 Gy (RBE) using the fixed RBE-weighted dose model.

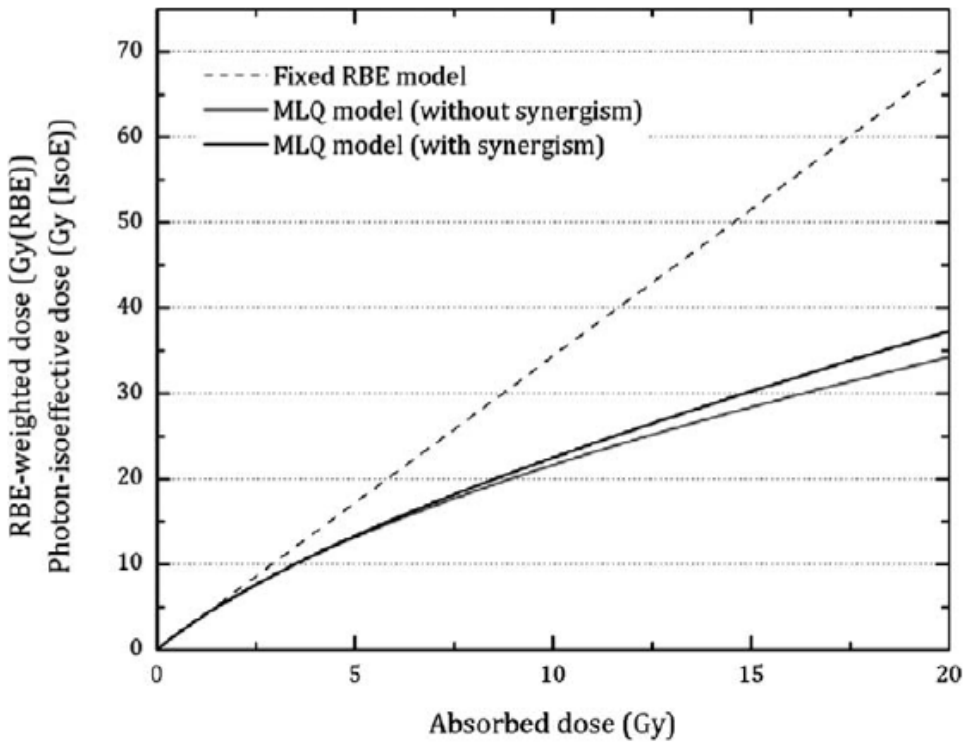


Figure 1.8: *RBE-weighted dose and photon-isoeffective dose as a function of the total absorbed dose computed with the fixed RBE model (dashed line), and the isoeffective dose model considering independent action (solid gray line) and synergistic action (solid black line). Parameters are taken from [52] and [43]. Illustration from [43].*

Figure 1.9, instead, shows the clinical outcomes of the cutaneous melanoma BNCT trial [53], comparing when the dosimetry is calculated with the RBE-weighted dose model and with the photon isoeffective dose model. The benchmark is the expected Tumor Control Probability.

This comparison shows that photon isoeffective dose values are lower than the fixed RBE-weighted dose ones. Moreover, it evidences how the traditional model is not adequate to compute photon-equivalent doses. In the region on the right of 95% tumor control probability curve almost all lesions should present a complete response, while it does not happen for the fixed RBE approach, where the large amount of open circles indicate the

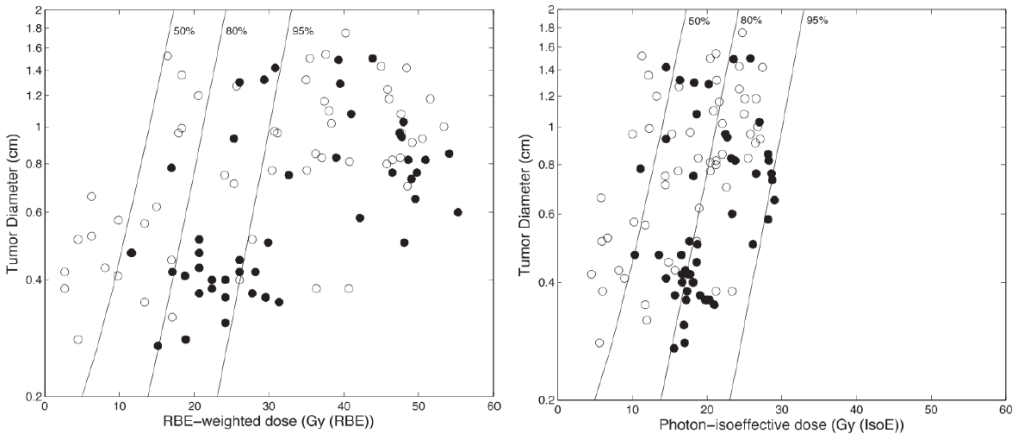


Figure 1.9: *Scatter plots showing the distribution of tumor diameters and minimum doses of the lesions with their response for fixed RBE-weighted (left) and photon isoeffective dose (right) models. Filled circles indicate complete response, and open circles indicate any other clinical outcome. Continuous lines represent the iso-tumor control probability curves computed with* TCP_{MLQ} *for* 50%, 80% *and* 95% *of tumor control. Illustration from [43].*

presence of other clinical outcomes. This therefore shows how fixed RBE-weighted doses are inconsistent with the standard radiotherapy clinical results. On the contrary, the photon isoeffective dose model gives a distribution that is consistent with the expected control of the lesions.

Chapter 2

Treatment Planning System

After the introduction to BNCT and photon-equivalent dose model in the previous chapter, it is now time to use those notions for clinical dose assessment. To do this it is necessary to analyse the functionality of a BNCT Treatment Planning System (TPS). This chapter will, therefore, outline the concept of TPS in all its components.

Treatment Planning Systems are indispensable tools used for radiation treatment planning as they are the means to calculate and optimize the dose distribution inside the patient in radiation therapy. It is through the use of these systems that specific treatment procedures are developed for individual patients. This includes the specification of beam energy, direction, field size and fluence to maximize the dose delivered to the target and to minimize the probability of normal tissue complications [54]. In clinics, the entire treatment planning process is overseen by a medical physicist, who is responsible for the overall process of the treatment planning to accurately and reliably produce dose distributions and associated calculations. The plan must be finally approved by a radiation oncologist who checks its accuracy and suitability before being actually implemented in patient treatments [55].

Treatment planning involves the following steps [56]:

(i) Identifying the shape and the location of the tumor (i.e. the *target*) and of the neighbouring *organs at risk* through the use of modern medical imaging (CT, MRI, SPECT, ...);

(ii) Choosing an appropriate patient positioning and immobilisation methods so that treatments will be reproducible;

(iii) Optimising and selecting a suitable beam arrangement;

(iv) Evaluating the resulting dose distribution in the selected volumes;

(v) Calculating the machine settings to deliver the required amount of absolute dose.

TPS are currently widely adopted in clinics worldwide and the technological improvements have been leading to the development of new computerized TPS relying on hardware and software able to generate accurate dose distributions.

2.1 Target Definition

The standard terminology adopted to describe volumes of the target organs and the other surrounding tissues is given by the ICRU Report 50 [57] of 1993, whose recommendations were integrated by the forthcoming ICRU Report 62 [58]. A common language in the definition of the volumes throughout the treatment planning process provides the basis for consistency within different centres worldwide. The main volumes taken into consideration during a treatment planning are shown in Figure 2.1 and are described as follows.

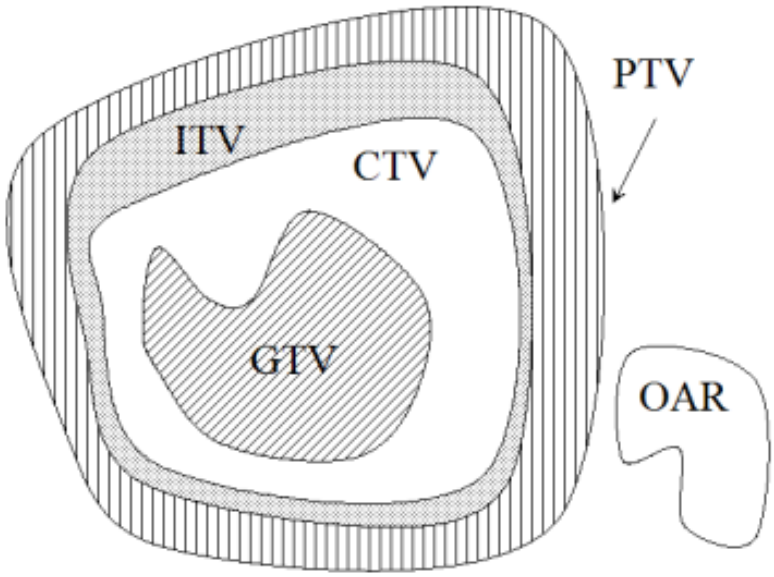


Figure 2.1: *Graphical representation of the volumes-of-interest, as defined by the ICRU 50 and 62 reports. Illustration from [55].*

Gross Tumor Volume

The *gross tumor volume* (GTV) is defined as the gross palpable, visible and demonstrated extent and location of the malignant growth. It can be identified and spatially localized inside the patient using modern and high resolution 2D and 3D imaging techniques, such as CTs and MRIs. However, GTV may be very ill-defined in case of very infiltrating tumors and high grade gliomas and its delineation may be subjected to large uncertainties [56].

Clinical Target Volume

The *clinical target volume* (CTV) is a tissue volume encompassing the whole GTV but with an additional margin. As a matter of fact, CTV takes into account possible microscopic extensions of the main gross volume which are not visible by imaging techniques due to their limited resolution. CTV thus has to be treated adequately to achieve the aim of therapy, cure or palliation.

Internal Target Volume

The *internal target volume* (ITV) was introduced by the ICRU Report 62 [58], and it consists of the CTV plus an internal margin that is designed to take into account the variations in the size and position of the CTV relative to the patient's reference frame. Any alteration due to organ motions such as breathing, bladder or rectal content are taken into account.

Planning Target Volume

The *planning target volume* (PTV) is a geometrical concept, and it is defined to select appropriate beam arrangements, taking into consideration the net effect of all possible geometrical variations, in order to ensure that the prescribed dose is actually absorbed in the whole CTV. In contrast with the GTV and the CTV, which are purely oncological concepts, the PTV does not necessarily correspond to the anatomy of the tumor target. It actually includes the internal target volume and an additional margin for set-up uncertainties, machine tolerances and intra-treatment variations.

Organ at Risk

Organs at risk (OAR) are non-target volumes adjacent to the PTV as they do not contain malignant cells. These organs have a key role in the treatment planning. As a matter of fact, it is fundamental to minimize the dose delivery to these organs as they are usually sensitive to the effects of ionising radiations and, if dose absorbed is higher than their tolerance, they can suffer from serious side effects and cause life-threatening injuries to the patient. Many OARs have a defined *tolerance dose* at which a given level of cell morbidity can be expected [59]. Specific attention should be paid to organs that, although not immediately adjacent to the CTV, have low tolerance dose (e.g. eye lens during nasopharyngeal or brain tumor treatments). In case the tolerance dose of a specific OAR is exceeded, it is necessary to modify some treatment features, such as the beam direction or the irradiation time. A complete review of the tolerance dose values in organs for photon therapy is provided in [60].

Radiation therapies are localized treatments as their anti-cancer effect is only achieved in the irradiated organs and tissues. The definition of target volumes, adjacent organs at risk and the other anatomical structures is therefore essential in regard to the development of an accurate treatment planning and to provide a basis for comparison of treatment

outcomes [55]. Future improvements in medical imaging techniques and their incorporation into the planning processes will significantly improve the correct delineation of both GTV and CTV, while an improved capacity to follow the internal organ movements with the beams will contribute to a better coverage of the ITV and thus also of the PTV.

2.2 Treatment Planning System for BNCT

Treatment planning for BNCT involves a computational analysis of the radiation dose distribution inside a patient. This evaluation is crucial for accurately identifying the most suitable neutron beam orientation, and determining the irradiation time (i.e. consequently the fluence), enabling the delivery of an optimized dose distribution limited by dose restrictions for normal tissues and OARs. Thus complying with dose prescriptions and optimized to maximise the dose to the target volume. The workflow of the treatment planning process is similar to other types of radiotherapy, such as external beams of photons, electrons, or protons. There are, however, some relevant differences between these approaches due to the different radiations involved and their physical properties, making BNCT treatment planning more complex [61].

As explained above, dose targeting is obtained through the biochemical tumor selectivity of the boron drug rather than through geometric targeting of highly conformal radiation beams. BNCT treatment planning system, therefore, must take into account the boron distribution at least at the tissue level, the injection route (for example if it is continuous or if the irradiation takes place after a while) and its retention inside the target. The shape of the beam becomes less relevant as the neutron beams are usually not highly collimated and their directionality is rapidly lost with depth in tissue due to the scattering.

The most important challenge is represented by the complex mixture of high- and low-LET dose components each with specific biological effectiveness, in contrast with the single low-LET dose component characteristic of conventional photon radiotherapy. As a consequence, the dose delivered by a photon beam can be easily calculated by both simple empirical algorithms and complex model-based algorithms, while BNCT treatment planning systems exclusively rely on Monte Carlo transport codes because of the variable, scatter-dominated nature of the radiation transport processes [62, 63, 64]. For this reason, the time required for planning may be longer than for conventional radiotherapy, depending on the available computational resources for dose calculations. Moreover, as a consequence of the complex mixed radiation field and the numerous dose components, dose prescriptions in BNCT treatment planning are defined by a limit on normal tissue dose rather than by delivery of a desired tumor dose. This is because the dose-effect relation in tumor for the mixed field of BNCT is not known yet in clinical application with the same precision as in photon-therapy.

2.2.1 Modeling patient geometry

The first step of the treatment planning is the computational construction of realistic individualized patient models for radiation transport calculations. Generally, the 3D models are constructed from the acquisition of the CT (Computed Tomography) of a patient.

Recently, the integration of these images with ^{18}F-BPA PET image data has been recommended, as it is useful for providing information on the boron accumulation inside both tumor and normal tissues and thus for enabling high-precision treatment planning [65]. The patient modeling relies on several different types of geometric 3D representations involving different approaches and software.

Voxel Models

One of the most widely and commonly used approaches is the voxel model [66], which consists of a 3D array of contiguous rectangular prisms called *voxels*. The construction of the patient's geometry consists of the adequate superimposition of a 3D rectilinear voxel-based grid on the 3D medical images. Each voxel, which is typically larger than the pixel of the reference computed tomography, is then assumed to be uniformly filled with materials corresponding to the image data related to Hounsfield units, as shown in Figure 2.2, where the CT image allows the corresponding 3D voxel-model to be reconstructed.

The NTCPlan is an example of voxel-based treatment planning system whose reconstruction technique relies on thin-slice CT image data to automatically create a heterogeneous multimaterial mathematical model of the relevant body part [67]. Other newly developed treatment planning systems for boron neutron capture therapy based on this patient modelling principle are the TsukubaPlan [68] developed by the University of Tsukuba and NeuMANTA [69]. For computational reasons, however, it is necessary to limit the number of materials and increase the voxel dimensions. As a matter of fact, if by shrinking the voxels the accuracy of calculated doses improves, on the other hand, in the case of excessively small voxels, the transport code has difficulty in computing the reactions occurring at each surface. The number of surfaces involved thus increases and it takes longer to calculate the trajectory of the particle. An alternative is to adopt a slightly different approach, called *multicell model*, where the voxel size varies as a function of the voxel location inside the body of the patient. It can be indeed advantageous to have finer resolution, corresponding to smaller voxels along the edges between materials, and lower resolution with larger voxels where the material is more homogeneous. This multi-voxel approach is currently employed in Argentina [70].

Univel Models

Univel models are based on a pixel by pixel uniform-volume element reconstruction of the patient geometry. From a geometric point of view, they are similar to voxel models as both use a uniform rectangular mesh to reproduce the patient geometry, but they differ in the univel size, which is smaller when compared to voxels. Each univel indeed represents a specific pixel from the corresponding medical image. Univel models have a high geometric accuracy, as they offer very high geometric fidelity, with the same spatial resolution as the medical images. However, the composition of each univel must be correctly identified and so each pixel must be computationally painted with a color that corresponds to a particular material composition, as shown in Figure 2.3.

This process may be tedious and very time-consuming and so it may therefore limit the

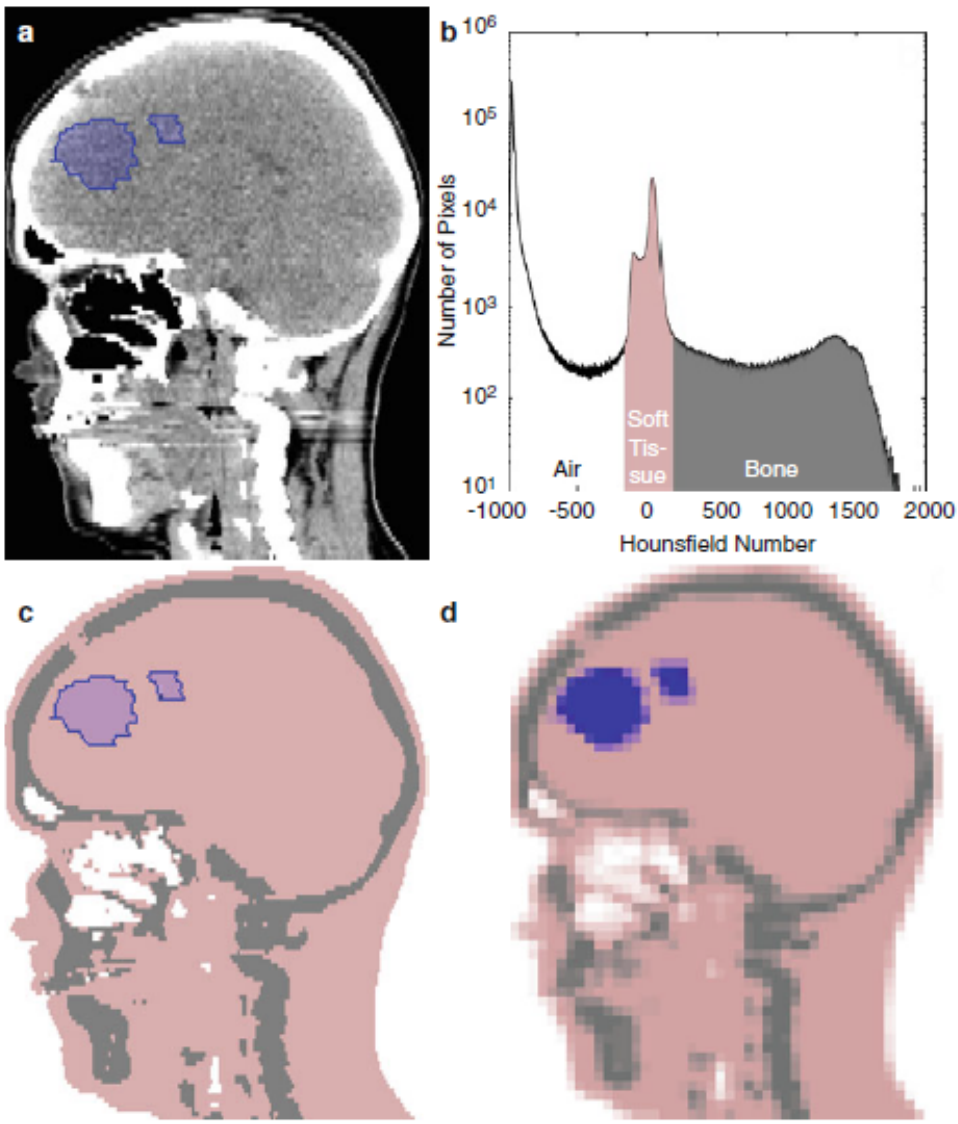


Figure 2.2: *Construction of a voxel model from CT image data in the MiMMC treatment planning system.* ***(a)*** *Sagittal CT image slice with contoured target volumes shown in blue.* ***(b)*** *Image intensity histogram showing the range of Hounsfield numbers for the selected tissues, i.e. air, soft tissue, and bone and* ***(c)*** *the resulting thresholded sagittal image.* ***(d)*** *The corresponding slice through a 4-mm voxel model. Illustration from [61].*

effort of the treatment planning. The SERA treatment planning system developed at Idaho National Laboratory (INL) and Montana State University (MSU) implements the univel model [71].

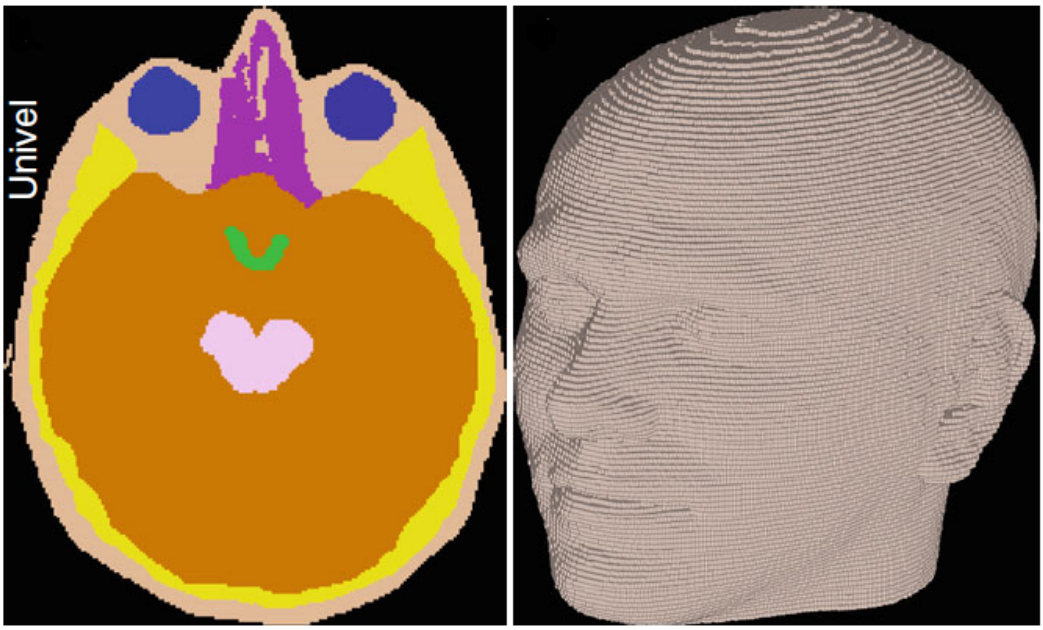


Figure 2.3: *2D slice Univel model where air, skin, skull, brain, brain stem, eyes, optic chiasm, and sinuses are defined. Each univel corresponds to a pixel in the MR images used to derive the model, which have a pixel size of* $1 \times 1 \times 2$ mm^3 *(left). Right: the corresponding 3D representation. Illustration from [61].*

NURBS Models

NURBS models are very flexible and powerful 3D modeling techniques already used in computer graphics. They work in a completely different way from the previous ones, as NURBS models define free-form curves and surfaces. Anatomical structures are indeed modelled and contoured by defining specific control points in the medical image that allow the reconstruction of a given curve or surface, as illustrated in Figure 2.4. Each surface area or curve refers to a specific volume which must be implemented in the treatment planning system.

These points are conventionally placed, either manually or using edge-detection algorithms, allowing an efficient and high geometric fidelity reconstruction of the patient. On the other hand, this technique may be tediously labor intensive and time consuming. One of the main disadvantages of NURBS models is given by the difficulty of selecting control points correctly, which can lead to possible overlapping of adjacent surfaces causing particles entering the region of overlap to be lost. Moreover, although the external contour of the patient is represented with high fidelity, some internal structures, such as sinuses and skull, may lack definition and accuracy because they are difficult to delineate [72]. The NURBS modeling technique was adopted in the BNCT_Rtpe planning system developed at INL [73].

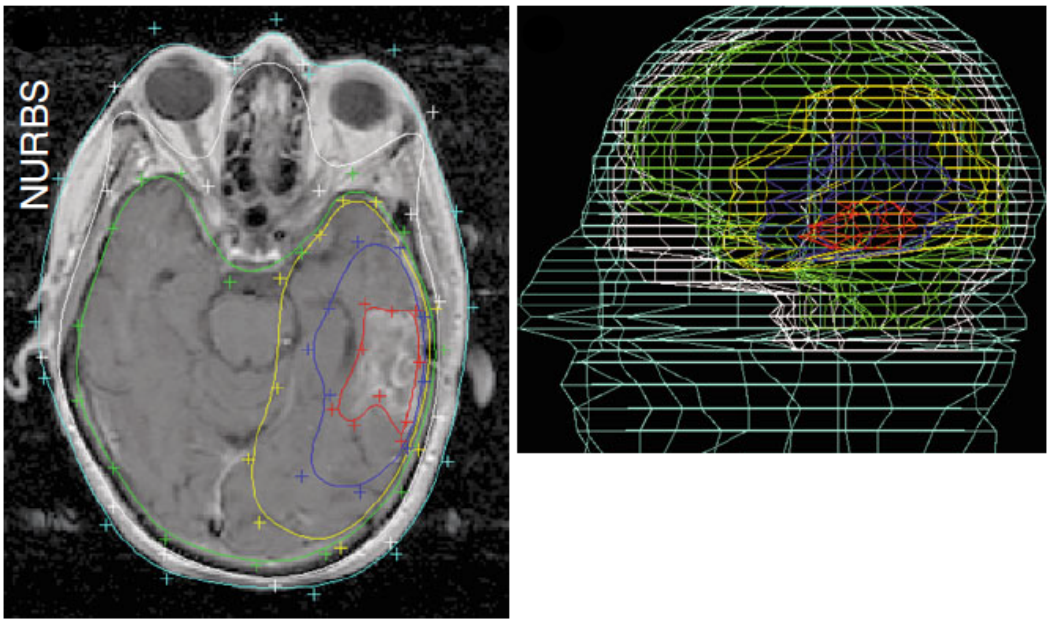


Figure 2.4: *Left: control points indicated by plus signs guide the placement of the spline in the NURBS model, which is overlaid on an MR image. Right: wire-frame rendering of the NURBS model shows the outer boundaries of skin (cyan), skull (white), brain (green), CTV (yellow), edema (blue), and GTV (red). Illustration from [61].*

2.2.2 Tissue Compositions and Boron Concentrations

The definition of the correct tissue composition in the patient is essential to design a proper treatment plan. As opposed to radiotherapy, where the beam photons and electrons only interact electromagnetically with the electrons of the target tissues and the dose calculation strictly depends on the electron density derived from the CT image data, in BNCT reactions take place with the different nuclei (see Section 1.2.1).

Modeling different tissue types is necessary because elemental composition and neutron cross sections vary significantly between different nuclides. As stated in ICRU Report 46 [74], uncertainties in the absorbed dose estimations in tissues arise from uncertainties both in the composition of the tissues and from radiation interaction coefficients. The former are usually larger. Treatment planning involving different nuclides composition of the same tissue can lead to completely different dose evaluations. Changes in dose of up to 9% when skull[1] is replaced by skin in the simulation of cranial irradiation were evaluated by *Wojnecki et. al* [75]. For some earlier clinical trials of BNCT Brooks brain composition [76] was adopted as the reference one. Anyhow, it was soon abandoned as it showed some experimental inconsistencies across trials. Currently, ICRU Report 44 [32] integrated with ICRU Report 46 [74] provides the adequate reference composition data

[1] which has a lower hydrogen density and a higher nitrogen concentration than soft tissues.

and densities for a wide range of tissues. The major difference between the Brooks brain and ICRU brain compositions is a 20% difference in nitrogen concentration, which leads to a similar discrepancy in thermal neutron dose [77]. From the graph in Figure 1.6 indeed, it can be appreciated how significant the dose component of nitrogen is, and how a large variation of this factor leads to a large variation of thermal neutron dose.

Because of the high thermal neutron cross section, also the ^{10}B concentration has to be correctly set in both tumors and healthy tissues using reasonable assumptions based on pharmacokinetic studies and with the support of PET images to have an adequate dose estimation [21].

2.2.3 Neutron Beam Source Definition

Once an adequate computational model of the patient has been generated, a proper neutron beam model has to be constructed to simulate a realistic clinical treatment. This is one of the most critical aspects involved in the treatment planning procedure, as it requires producing an accurate representation of the 5-dimensional probability distribution: 3 dimensions are related to spatial distribution, one for the energy spectrum, and another accounts for the angular characteristics of the beam [78]. Several methods for defining the radiation source have been used in BNCT treatment planning. In this chapter I will take into account the main techniques implied in accelerator-based BNCT, as the use of particle accelerators for this therapy is increasingly being employed on a large scale. The neutron beam source definition for the AB-BNCT relies indeed on the configuration of the accelerator components. In the course of this thesis work, I utilized a source containing both the BSA, the collimator, and the beryllium target to achieve enhanced source resolution.

This technique relies on the implementation of Monte Carlo simulations, whose mechanism is based on algorithms that sample a certain amount of random values that follow a specific probability distribution. The main Monte Carlo algorithms currently applied in BNCT treatment planning systems are MCNP [79], which will be discussed in more detail later (as used in NCTplan, Multicell, THORplan, JCDS), PHITS [80] (as used in TsukubaPlan, and in JCDS), and Geant4 [81] (as used in the dose engine implemented by Neutron Therapeutics). Typically, the TPS for the BNCT uses external Monte Carlo codes such as MCNP6, which is currently considered as the gold standard for particle transport, although some use specially developed Monte Carlo instead (Sera [71] and NEUMANTA [69]). What follows in this paragraph is specific to MCNP.

Surface Source Files

In the *surface source file* method of source definition, a detailed Monte Carlo simulation of the neutron beam is performed and the characteristics of all particles crossing a specific plane are recorded in a binary phase space file. This binary file contains all the track features, such as position, direction, energy, and the statistical weight of all particles located at a given position in the treatment geometry. The particle tracks from the phase space file are then sampled in subsequent simulations of particle transport through the

patient geometry.

The primary advantage of this method is that it does not introduce any approximation on the characteristic of the source. Moreover, as the transport does not take place from the actual radiation source, but from a surface closer to the geometry of interest, the speed of the dosimetry calculation is increased. On the other hand, the surface file method presents also few disadvantages. One being the size of the surface source file, that depends on the number of source points. Furthermore, the finite number of source points limits the statistics that can be achieved. This is partially overcome by using the same source point with a different random seed (i.e., changing the history of a single track).

Probability Density Function

To define a source in a volume or emerging from a surface, it is necessary to sample neutrons from the spatial coordinates, with appropriate energy and flight directions. *Probability density functions - PDF* are defined for the source particle energy, position and direction, and are then sampled in the simulation of each neutron generated in the source. As in surface source files, also the PDFs are usually constructed from previous Monte Carlo simulations or discrete ordinates transport calculations of the neutron beam [82], or they are known from experimental measurements.

Source PDFs may be defined in two ways in the patient transport simulation. Particles can be sampled in a plane located at the beam-port or located upstream, inside the collimator. In the latter method, a portion of the collimator must be explicitly modeled, to account for the interaction of neutron in the last part of the beam infrastructure. Defining the source at the beam exit is more efficient, but a greater amount of details may be required in the source distributions to achieve the desired level of accuracy in calculated doses. Finally, the PDF can be defined at the target, and in this case, the entire structure of the moderator, collimator and reflector must be modelled [27].

Benefits of source PDFs include a more compact representation and the ability to sample an unlimited number of particle histories, lowering the statistic uncertainty in calculated doses [78].

2.2.4 Dose Calculation

As a suitable neutron source has been defined and the patient geometry has been adequately modelled, it is possible to calculate the dose distribution inside the target volumes by Monte Carlo methods. The simulation calculates the absorbed dose in each voxel by considering all the possible neutron interactions inside the patient. To speed up the simulations, calculations are always performed to obtain the KERMA value, rather than the absorbed dose. As a matter of fact, *equilibrium of charged particles* CPE conditions ensure the equivalence between absorbed dose and KERMA, which is easier to calculate by using superimposed meshes for scoring neutron/photon fluence and coupling them to fluence-to-KERMA conversion tables [83]. For BNCT, CPE is surely satisfied in most of the patient volume due to the short track length of secondary charged particles. Thus the Monte Carlo run delivers as the output the KERMA value K, and consequently the

absorbed dose D_{abs} is given by the following Equation

$$D_{abs} = K = \varphi_n (F_n)_{E,A} \ , \tag{2.1}$$

where φ_n is the neutron fluence and $(F_n)_{E,A}$ are the tabulated KERMA factors, as a function of the neutron energy and the atomic weight of the target element.

However, this approach has some limitations as CPE conditions may not be verified in some volumes, such as skin, where secondary charged particles may escape the tissue and deposit their energy elsewhere. Skin is an important organ in BNCT, both for superficial and deep-seated tumors, and in some cases it is the OAR limiting the irradiation time [84, 85]. For this reason, a possible solution to correctly evaluate the absorbed dose is to dedicate part of the dose calculation to a detailed dose deposition simulation in those regions of the patient geometry where equilibrium cannot be assumed [86]. However, as the limiting organ defines the maximum fluence that is allowed, overestimating the absorbed dose through KERMA calculation provides a conservative approach.

In my thesis work I used MCNP [79] as the Monte Carlo engine. MCNP is a general-purpose, well validated, relatively easy to use Monte Carlo radiation-transport code designed to track many particle types over a broad range of energies. MCNP code provides seven instruments to calculate the quantities of interested, called *tallies*, which are described in the following Table 2.1.

Tally	**Description**
F1:N or F1:P or F1:E	Surface current
F2:N or F2:P or F2:E	Surface flux
F4:N or F4:P or F4:E	Track length estimate of cell flux
F5a:N or F5a:P	Flux at a point or ring detector
F6:N or F6:P or F6:N,P	Track length estimate of E deposition
F7:N	Track lentgh estimate of fission E deposition
F8:N or F8:P or F8:E or F8:P,E	Pulse height tally

Table 2.1: *Summary table of all standard tallies available in MCNP.*

In addition to the standard tallies, MCNP has one special tally type, the superimposed mesh tally, which can be invoked into the simulation by using the FMESH card. It allows the user to tally particles on a mesh, which can also be independent from the problem geometry. Currently, only track-length F4 mesh tallies have been implemented. MCNP provides various ways to calculate KERMA and consequently the absorbed dose in the desired target cells. In this thesis I chose to calculate KERMA by coupling the FMESH card with the dose energy DEn and dose function DFn cards, which specifies a list of KERMA factors as a function of energy displacement.

2.2.5 Plan Evaluation and Dose Prescription

The analysis of the treatment planning relies on standard quantities and ways to visualize the in-patient dose distribution. The *isodose* contour lines identify the regions within the

patient that absorb the same amount of dose. They are typically superimposed over the anatomical features displayed to locally indicate the effect of the neutron therapy on the different tissues of the patient. In general, the total weighted doses for tumor target and normal tissues are calculated separately, each one using different boron concentrations and weighting factors assumed for each tissue. Isodose are typically expressed by a simple percentage value related to the peak absorbed dose on the beam axis.

Dose-volume histograms (DVHs) are one of the most useful and commonly adopted tools relating radiation dose to tissue volume in radiation therapy planning for comparing alternative treatment plans for a patient. DVHs are plots of cumulative dose-volume frequency distribution, thus showing dose uniformity, minimum dose and maximum dose. They show the uniformity of the dose distribution, the minimum and the maximum value and they can thus give indication on the effectiveness of the plan. In general, they provide information on whether a sufficient volume of target tissue has absorbed sufficient dose to properly control the tumor. In BNCT this evaluation is more complex than in conventional photon treatment because a dose-effect relation in clinical treatment is not available yet. In fact, dose is never prescribed to the tumor but rather to the most radiosensitive healthy tissue involved to maximise the treatment effect.

These tools shown in Figure 2.5 are both useful with regard to further optimize the treatment plan. Optimizing a treatment plan involves choosing the orientation, the size, and the spectrum for a beam or set of beams to deliver radiation dose able to produce the highest dose in target volumes while complying with the dose prescription and respecting dose limits on normal tissues and organs at risk [61]. In the case of single-field treatment plans, optimization of the plan is usually relatively straightforward, as it only involves choosing a certain beam pattern among several candidates, comparing the delivered dose value. On the other hand, for multi-field treatment plans, the optimization is more complicated because it is necessary to choose few (usually non-overlapping) independent fields from a pool of candidates and to determine how they must be weighted to obtain an advantageous dose distribution. To solve this time-consuming process, a simplex search algorithm to optimize the boron dose to a large number of target volumes, subject to constraints on the organs at risk, has been implemented in software external to the treatment planning system [64, 88].

The estimation of dose in tumor suffers from some uncertainties, especially due to the knowledge of boron concentration and distribution in the volume. The typical choice is to assign to the whole volume a uniform boron concentration obtained by multiplying the boron concentration measured in blood by a fixed factor [89]. It is known, however, that this is only a rough estimation, as the tumor is an heterogeneous structure and boron is not uniform during irradiation. Finally, the lack of a robust method to calculate photon-equivalent dose in clinical application has prevented the use of the photon dose to predict the outcome of BNCT irradiation (see above). For this reason, several different methods of dose prescription mostly based on experience, have been used in BNCT clinical trials, most of which make use of the total weighted dose expressed in RBE-weighted dose units. Early methods were used to prescribe only the maximum weighted dose in the patient or particular organs, such as the brain for the treatment of intracranial tumors [90]. Other

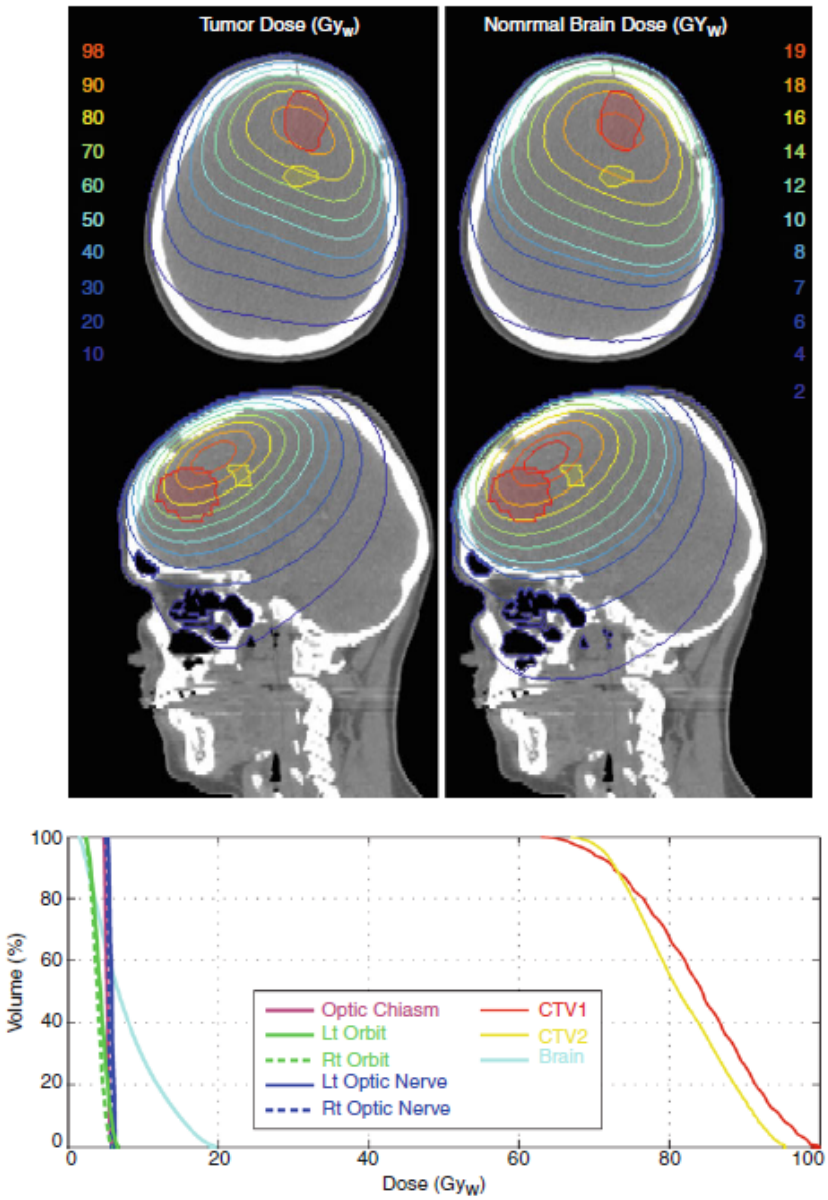


Figure 2.5: *Three-field treatment plan for a brain tumor (GBM) patient calculated using the MiMMC planning system. The prescription is a mean brain dose of 7.7 Gy-Eq. Isodose contours calculated for tumor and normal brain are shown on axial and sagittal sections through the target volumes. The integral dose-volume histograms (DVHs) summarize dosimetry for structures of interest including target volumes and organs at risk. Illustration from [87].*

extremely sensitive organs, for example, skin, oral mucosa, and spinal cord were also frequently applied as organs at risk, setting their tolerance dose as the prescribed one.

More recent protocols, instead, commonly employ limits on the mean weighted dose of the organs at risk [91].

2.2.6 Treatment Delivery

There are further aspects that are accounted in the TPS concerning the treatment dose delivery to the patient. In particular, in modern external radiation therapies, the setup precision, inter-treatment position reproducibility, and keeping of patient's intra-treatment position are important factors influencing the correct implementation of a treatment and influencing the treatment outcome [33]. Most BNCT facilities have a simulator that replicates the geometry of the treatment room for positioning and marking the patient in a convenient environment outside the treatment room [61]. Simulators have a dummy beam port mounted on the walls and the patient is positioned in relation to the dummy beam port by using a set of lasers and supporting devices like cushions, head cups, or, in some cases, a stereotactic frame. Reference ink marks are also applied to the patient. Furthermore, the patient must maintain the same body posture during the whole beam irradiation. This aspect may be particularly issuing, as a BNCT clinical treatment can take more than half an hour. Positioning patients for BNCT treatment is more difficult than for other external beam radiation modalities, because neutron beam ports at both nuclear reactors and accelerators are currently stationary and cannot be rotated around the patient. Moreover, an additional difficulty with patient positioning relates to the fact that it is not possible to observe the patient from the beams-eye view because of the need to place the patient as close as possible to the beam-port [92] to avoid out-of-beam organs exposure, since neutrons diverge out of the collimator.

Positioning must also consider the patient's posture and comfort. Several different patient setting systems are currently adopted at each BNCT facility to ensure the effective delivery of the established treatment plan [93, 94].

The knowledge of boron concentration in both tumor and normal tissues during irradiation is a crucial factor for the BNCT treatment planning implementation as it determines the boron dose component. In clinical BNCT dosimetry, estimations of ^{10}B dose in normal healthy tissues are generally based on the ^{10}B concentration in blood as a surrogate for normal tissue [89]. In glioblastoma treatment trials using the BPA-F complex as the boron delivery agent, a constant tumor-to-blood concentration ratio of 3.5–4 to 1 is typically assumed. This value, however, does not consider the dynamic ^{10}B loading and washout behavior expected in tissues nor the heterogeneity of its distribution in the tumor volume [95]. Pharmacokinetic modeling of the concentration of ^{10}B in blood following administration of the boron delivery agent is one of the first steps toward modeling the dynamic behavior of ^{10}B and providing more accurate dosimetry. A deeper understanding of the mechanisms and timing of boron retention in the blood and in tissues allows the selection of optimum infusion schedules and irradiation times.

Figure 2.6 shows how the blood boron concentration profile varies for an infusion of BPA-F as a function of the time, underlining a rapid rise in blood boron concentration during

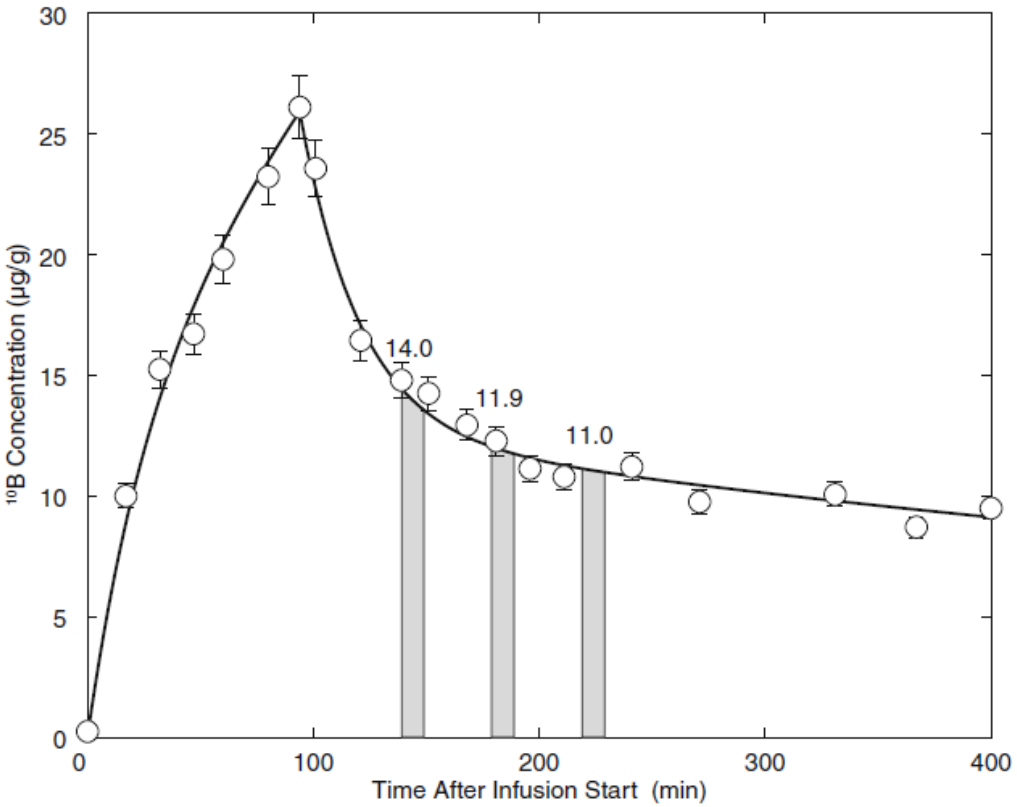


Figure 2.6: *Measured blood ^{10}B concentration profile and fitted pharmacokinetic model for a patient who received a 1.5 h infusion of 350 mg BPA/kg. Using only blood ^{10}B measurements available before the start of each irradiation (indicated by the three shaded areas), the pharmacokinetic model must accurately predict the mean ^{10}B concentration during each field so that the correct number of beam monitor units can be delivered. Mean ^{10}B concentrations are shown above each field. Illustration from [61], adapted from data in Ref. [95].*

BPA-F infusion followed by biphasic exponential decay [95]. Various schemes are available for blood boron concentration prediction. One of the simplest approaches involves measuring a blood sample drawn immediately before irradiation and another drawn at the (estimated) midpoint and extrapolating linearly to predict the mean concentration during irradiation. More sophisticated approaches for boron concentration prediction using multi-compartment pharmacokinetic models and exponential washout models are also available [96, 97].

In 1991, *Ishiwata* [98] developed the 4-^{10}B-borono-2-^{18}F-fluoro-L-phenylalanine (^{18}FBPA) as a positron emission tomography probe for imaging and evaluating the pharmacokinetics of BPA in vivo. This technique is still in use today, especially in Japan, and it represents one of the main methods to estimate the way boron accumulates inside both tumors and OARs. An example of application of this technique is shown in Figure 2.7, where the high

and selective boron accumulation inside the target is highlighted. The result is a three-dimensional boron distribution map carrying information about how BPA concentration varies over time. This has a key role in defining the right time to start the BNCT treatment after the drug injection [99].

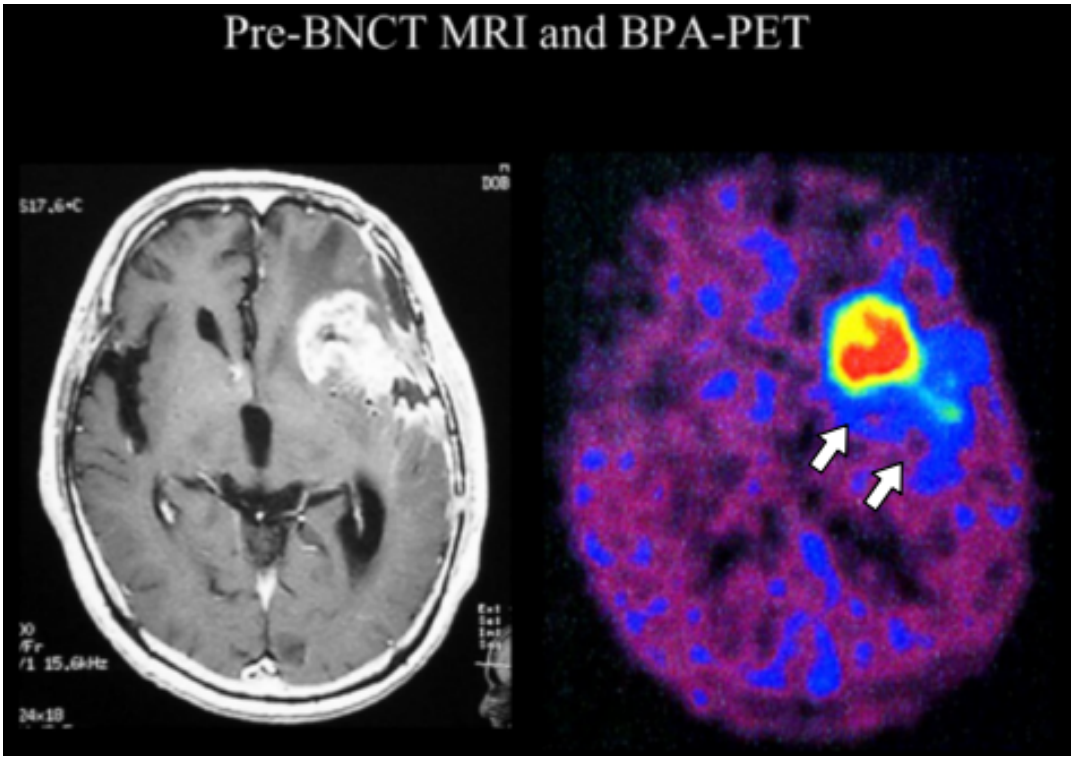


Figure 2.7: *Contrast-enhanced T1-weighted MRI of representative Glioblastoma patient and* ^{18}*F-labeled BPA-PET image after initial debulking surgery. Illustration from [99].*

2.3 IT_STARTS

The *Innovative Toolkit to Simulate neuTron cApture theRapy irradiaTion and doSimetry*, IT_STARTS, is a project funded by INFN in cooperation with the Argentinean Treatment Planning and Computational Dosimetry group at the Comisión Nacional de Energía Atómica, CNEA. This Argentinean group is the same who developed the Photon Isoeffective Dose formalism.

IT_STARTS is written in Python, which is an open source language, and the TPS code itself, after a validation process, will be made public through appropriate software sharing methods to extend its availability to different users of BNCT. It is based on the implementation of two successive modules for the creation of patient geometry and dosimetric evaluation, interposed by the application of a Monte Carlo transport code for the radiation transport.

The first module of IT_STARTS enables to transform the medical images of a patient into

a voxelised computational model. IT_STARTS works by executing several steps. Firstly, the code reads CT medical images uploaded in NIFTI format and the corresponding RTSTRUCT included, which are sets of anatomical structures identified on radiological images for the planning and delivery of radiotherapy. It is thus able to identify and define a *mask* relating to the GTV and all the other organs of interest within the RTSTRUCT. At this stage, IT_STARTS constructs a voxelised geometry of the patient, written in the syntax of the most common Monte Carlo transport codes, such as MCNP6 [79], PHITS [80] and GEANT4 [81]. Concerning the patient modelling, a voxel approach is adopted, different from the MultiCell model employed in Argentina [70].

This voxelised geometry is then provided to an external Monte Carlo transport code, such as MCNP6, in which also the source of the neutron beam is properly reproduced. In this configuration, the source is assumed to be fixed, so the correct positioning is achieved by orientating the patient's geometry in such a way as to optimise the treatment. A typical strategy applied is to place the patient in a position which minimises the distance between the beam-port and the centre of mass of the GTV. At the moment, IT_STARTS is not provided by a method able to automatically carry out this operation, therefore it is the user's task to find the right rotational and translational coordinates for a suitable positioning. This challenging and time-consuming procedure, however, can be supported with the use of external image visualisation programs, including Mango (`https://mangoviewer.com`) and 3D Slicer (`http://www.slicer.org`), both free and open source software for visualization, processing, segmentation, registration, and analysis of medical images. These tools provide a three-dimensional visualisation of the patient.

The input file is then run to calculate the absorbed dose distribution. As explained in Section 2.2.4, it is useful to superimpose a mesh on the patient's geometry, to have a precise voxel-by-voxel dose evaluation. The irradiation process can thus be simulated, providing the absorbed dose rate values resulting from each of the four dose components in each frame of the mesh as the output. The healthy organs and tumor segmentation is fundamental for IT_STARTS because it can calculate the spatial dose distributions only when the segmentation is supplied in input. Otherwise, there is no method for the toolkit to discern the respective volumes.

Finally, the output files containing dose rates in NIFTI format are uploaded into an Excel file, together with the dose prescription information. The second module of IT_STARTS thus reads this file and calculates photon equivalent dose in the segmented GTV and OARs. Starting with the dose prescription, it calculates the dose rate in the OAR and then establishes the treatment time by dividing the prescription value by the dose rate. There are generally two prescription options. One concerns the setting of an average dose on the volume of the OAR, D_{mean}, while the other a maximum dose, D_{max}.

In this work, which is focused on patients affected by Glioblastoma Multiforme, I have assumed an average dose limit of $D_{mean} = 2.5\ Gy(RBE)$ and a maximum dose of $D_{max} = 13\ Gy(RBE)$ in the brain [100] .

The most innovative characteristic of this TPS compared to the ones used in clinics, is the embedding of different models for the calculation of the photon equivalent dose, the

so-called *dose engine*. IT_STARTS indeed allows the user to load the radiobiological data, to fit survival curves for different cell lines and thus to obtain the RBE and the parameters for the isoeffective dose calculation.

As an output of the TP process, IT_STARTS displays the isodose curves superposed to the anatomy of the patient and builds dose volume histograms, DVHs, for the GTV and the OARs.

Another important characteristics is the inclusion of radiobiological figures of merits to evaluate the result of the treatment planning in terms of its therapeutic potential and safety. The *Tumor Control Probability*, TCP, the *Normal Tissue Complications Probability*, NTCP, and the *Uncomplicated Tumor Control Probability*, UTCP [101] are implemented in IT_STARTS for some types of tumors and healthy tissues. TCP, in particular, represents the probability that a tumor will be totally or partially controlled by a specific treatment. It is mathematically calculated by suitable models, including the linear-quadratic model, the Poisson model, and more complex ones incorporating factors such as tumor repopulation and repair. The possibility of implementing this probability function enables an assessment of the feasibility and applicability of the treatment plan for that specific clinical case, thus playing a key role in terms of plan evaluation.

Chapter 3

Artificial Intelligence in Cancer Treatment

Artificial Intelligence is a strongly developing and thriving field that studies how to build computer systems capable of simulating human thought. It is defined as *a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation* [102]. Artificial Intelligence (AI) is one of the most flourishing disciplines with many practical applications and active research topics, solving problems that are intellectually difficult for humans but relatively straightforward for computers. It is currently applied for many aspects of industry, science, and technology, as in many different areas in daily-life applications, such as speech and face recognition, and natural language processing, but also in vehicle driving and industrial activities.

Beyond these applications, another field where AI is introducing innovations is healthcare and medicine. It is more and more used for the interpretation of medical images, such as CTs and MRIs, in disease detection, and in the selection of morphological characteristics. The AI can also handle more challenging tasks like the accurate contour of tumors. These deep learning methods applied to clinical cases are still in the preclinical stage at the moment [103, 104]. However, through the relentless improvement of such automatic methods, there may be a systematic synergy in the future between clinicians and AI to find a new approach to handling cancer without overriding the physician assessment always needed in a clinical setup.

3.1 Applications in Cancer Imaging

AI excels at recognizing complex patterns in images and can quantify information from images that are not detectable by humans. Within the field of cancer imaging, the main contribution of AI focuses on the detection, characterization, and temporal monitoring of tumors [105].

Detection is defined as the capability to correctly localize and target the required objects of interest in radiographs or tomographies, and it is generally known as *Computer-Aided Detection* (CAD), which refers to the pattern recognition software that identifies suspicious features on the image and brings them to the attention of the radiologist [106]. AI-based detection tools can be used to reduce observational oversights and serve as a useful screen against misidentification errors by highlighting regions with suspicious characteristics.

Tumor characterization involves segmentation, staging evaluation[1] and its monitoring over time. Segmentation, in particular, helps in visualizing the extent of a specific abnormality. This is based on the contouring and isolation of the area to be treated. Such information can be implemented in diagnostic tasks as well as in therapy (i.e. dose calculations and prescription in treatment planning).

In current clinical practice, GTVs and OARs are typically manually defined by experienced radiologists, involving clinicians drawing contours by hand to delineate anatomical regions of interest (ROIs). While this practice is considered to provide good accuracy, on the other hand, it is operator-dependent and may be affected by the following types of uncertainties [107, 108, 109]:

(i) ***intra-observer variability***, which is quantified by the differences among contours delineated by a single observer on the same target volume at several different tries;

(ii) ***inter-observer variability***, which is quantified by the differences among contours delineated by multiple observers on the same target volume.

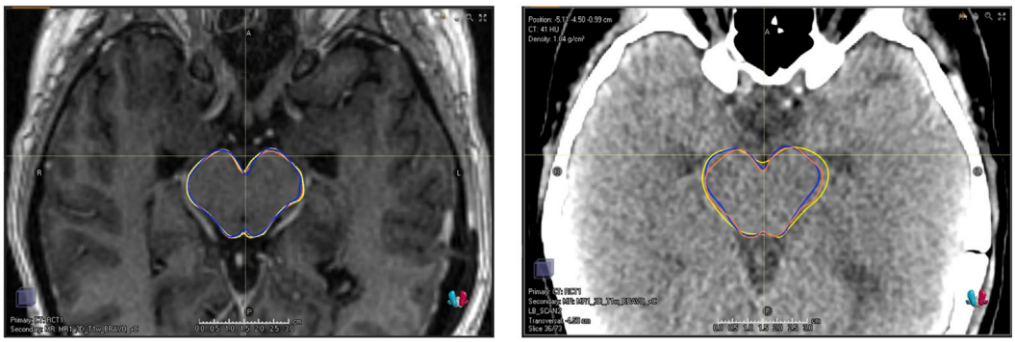


Figure 3.1: *Illustration of the effect of soft-tissue contrast on segmentation reliability. The brainstem was delineated with different colours by the same three clinicians on a magnetic-resonance scan (left) and a CT scan (right). The decreased soft-tissue contrast of the CT scan leads to larger inter-observer variability. (Based on data for a patient treated at Addenbrooke's Hospital, Cambridge, UK). Illustration from [110].*

[1]Staging the development of neoplastic tissue consists of categorizing tumors into predefined groups via genomics analyses,

The overall variability may be further intensified by the poor quality of the clinical images, as shown in Figure 3.1. As low-dose CT images allow for good distinction between materials with significantly differing densities, such as bone and air, on the other hand, they show limited contrast between soft tissues, causing a lower segmentation reliability. In addition, manual tumor contouring is a particularly tedious and time-consuming technique. This procedure may take several hours for the most complex cancer sites [111]. The long time spent by clinicians on contouring, and on associated review, may delay the start of the patient's treatment, and, consequently, it may lead to poorer tumor control and decreased survival probability [112].

AI has the potential to enhance the efficiency, reproducibility, and quality of tumor measurements by using automated segmentation techniques [105]. Artificial Intelligence has shown impressive accuracy and sensitivity in the identification of imaging abnormalities and tumors, trying to solve problems related to the wide variability associated with manual segmentation, as shown in Figure 3.2, where the multiple manual segmentations affected by unavoidable inter-observer variability are replaced by a single automatic contour. The development of Deep Neural Networks (DNN) has actively contributed to the implementation of algorithms able to adequately segment medical images to correctly identify and locate the regions of interest.

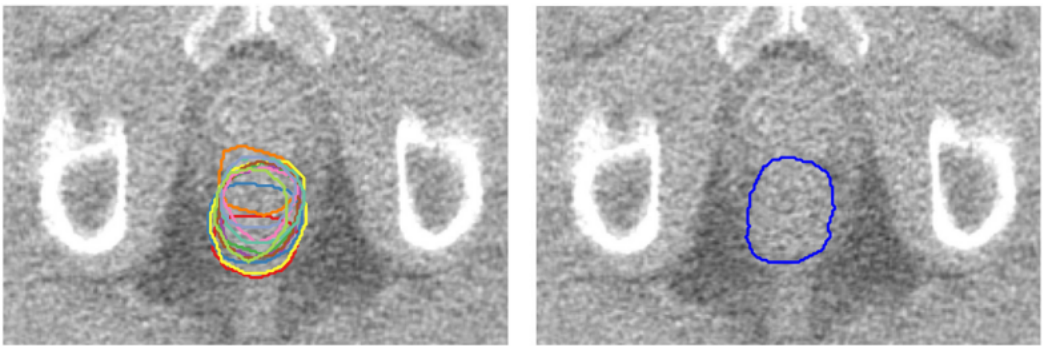


Figure 3.2: *Rectum contoured by multiple clinicians (left) and contours combined using the STAPLE algorithm (right). The images shown are from a megavoltage CT scan recorded to guide patient position, and have lower contrast than typical CT scan. Illustration from [110].*

3.2 Deep Learning and Artificial Neural Networks

3.2.1 The Perceptron

The idea of reproducing the human neural system has led to the modelization of the Perceptron [113], the first and the simplest example of an Artificial Neural Network. The Perceptron was designed to resemble the functional structure of a biological neuron, as shown in Figure 3.3. A biological neuron receives electrical signals as an input from its dendrites. These signals are then transmitted across the central body, where they

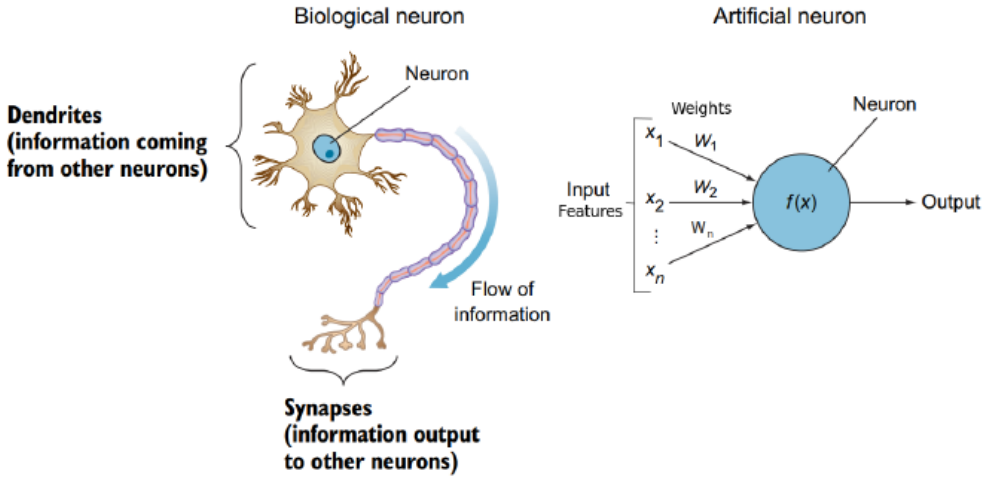


Figure 3.3: *Artificial neurons and biological ones share an analogous structure. The dendrites are emulated by the input connections, which collect input features instead of electrical signals. The function of the biological neuron central body is represented by the activation function and the activation state in the artificial neuron. Lastly the synapses is reproduced by the output connection. Illustration from [114].*

are modulated, and then the neuron fires an output signal through its synapses feeding another neutron, and so forth. This, however, happens only when the total amount of the input signals exceeds a certain threshold. The artificial Perceptron follows the same pattern.

The input signals are not electrical but are represented by the specific *features* x_i, which are computationally defined as *individual measurable properties or characteristics of an observed phenomenon*. Each feature comes to the neuron with a different importance, or *weight* w_i. The neuron thus modulates all these features obtaining an overall input z given by the weighted sum of each feature:

$$z = \sum_{i=1} w_i x_i + b \ , \tag{3.1}$$

where b is a general added bias term, which can be set and then modified during the learning process of the ANN.

In both artificial and biological neural networks, a neuron does not just output the bare input received. Instead, z is subjected to an additional step, called an *Activation Function*. This process corresponds to the decision-making unit of the brain, as the neuron emits the output signal only if the weighted sum is higher than a certain threshold provided by the activation function. There are many different activation functions currently adopted

in ANN systems based on the tasks required by the network, including

$$
\begin{aligned}
a(z) &= z && \text{Linear transfer Function} \\
a(z) &= \frac{1}{1+e^{-z}} && \text{Sigmoid Function} \\
a(z) &= \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} && \text{Hyperbolic Tangent Function} \\
a(z) &= max(0,z) && \text{Rectified Linear Unit (ReLU)} \\
a(z) &= max(0.01 \cdot z, z) && \text{Leaky ReLU}
\end{aligned}
$$

The simplest one used by the Perceptron algorithm is the step function which results in the following formula producing a binary output

$$\text{Activation Output} = \begin{cases} 0 & \text{if } z \leq \tau \\ 1 & \text{if } z > \tau \end{cases} . \tag{3.2}$$

If the total input signal passes a certain threshold τ, then the neuron is in an activated state and it is able to transmit the corresponding output, otherwise, the state remains 0 and the neuron is not fired.

3.2.2 Multilayer Perceptrons

ANNs made by single Perceptrons perform sufficiently well with simple datasets. However, for more complicated tasks, a more complex network is required. This means arranging Perceptrons in subsequent an interconnected layers [115, 113]. These artificial structures are so called *Multilayer Perceptrons* (MLP). A very common neural network architecture is to stack the neurons in layers on top of each others called *layers*, that are linked to each others by weight connections, as shown in Figure 3.4.

These additional layers are so-called *"hidden"* as it is impossible to see or control the input going into them and their output. Moreover, they are the only neurons with both incoming and outgoing connections, as their inputs are the activation states of the neurons from the previous layer, and their output is transmitted to the following layer of neurons. MLP is defined as a *fully-connected* or dense ANN, which corresponds to the fact that all the neurons inside a layer are connected to all the neurons in the previous one. Layers are structured in order to obtain increasingly complex features and thus each hidden layer builds new more complex features that will be analyzed by the next hidden layer. The first hidden layer finds patterns in the input data, the second one finds more complex patterns in the first discovered patterns, and so on. The number of hidden layers (i.e. depth of the ANN) is subject to variation depending on the complexity of the required network and the number of connections between each unit is related to the size of the ANN. Hypothetically, the deeper the ANN is, the more information about the data it is able to extract.

As stated before, each Perceptron is associated with a specific activation function. In general, the activation function usually adopted in the input and hidden neurons is the

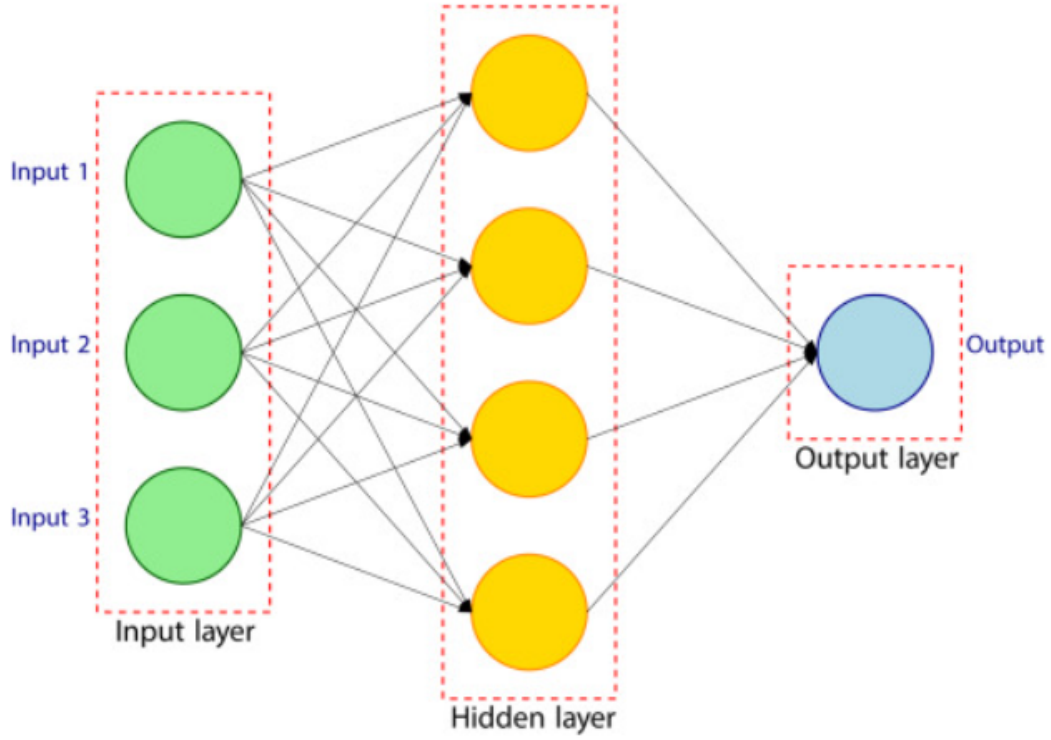


Figure 3.4: *Scheme of an artificial neural network with a single hidden layer. Input values are propagated through the network, with calculations performed at each Perceptron, to yield an output value. Illustration from [115].*

Sigmoid Function, while the Leaky ReLu is particularly suggested for images. For the output neutrons and for classification tasks, the most recommended one is the Softmax function, which is a generalization of the Sigmoid Function [116]. The process in which the information flows from the input layer through the hidden layers, all the way to the output layer is called *feedforward*. This process happens through the subsequent implementation of the weighted sum and the activation function between layers and can be schematized as in Figure 3.5. Starting the feedforward calculations from the input features, it is possible to calculate the first layer nodes as

$$\begin{aligned} a_1^{(1)} &= \sigma(w_{11}^{(1)} x_1 + w_{12}^{(1)} x_2 + w_{13}^{(1)} x_3) \\ a_2^{(1)} &= \sigma(w_{21}^{(1)} x_1 + w_{22}^{(1)} x_2 + w_{23}^{(1)} x_3) \\ a_3^{(1)} &= \sigma(w_{31}^{(1)} x_1 + w_{32}^{(1)} x_2 + w_{33}^{(1)} x_3) \end{aligned}$$

where σ corresponds to the applied activation function. Applying the same method for subsequent layers all the other nodes a_i^j can be evaluated until the output $\hat{y}$ is achieved as

$$\hat{y} = a_1^{(3)} = \sigma(w_{11}^{(3)} a_1^{(2)} + w_{12}^{(3)} a_2^{(2)} + w_{13}^{(3)} a_3^{(2)} + w_{14}^{(3)} a_4^{(2)}) \ .$$

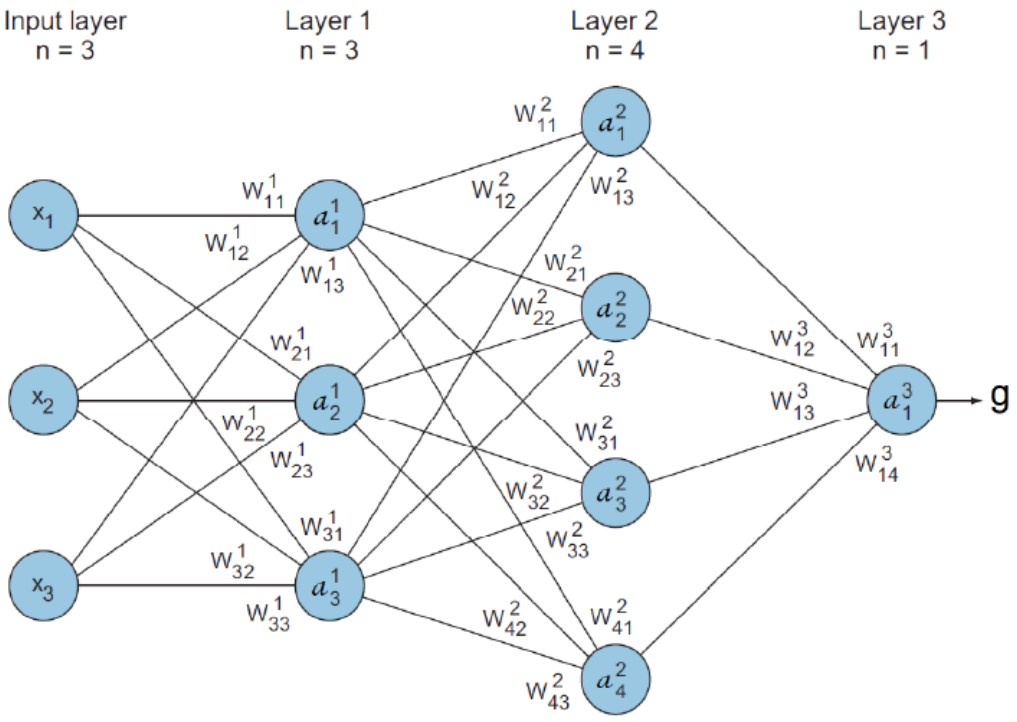


Figure 3.5: *Example of a simple three-layer neural network. The input features x_i are modulated and weighted by the respective values of w_i, obtaining the new node values a_i^j. Illustration from [114].*

Depending on the output prediction, the weights of the network can be tuned until the performance is satisfying. This iterative process is what network learning is based on.

Learning Process

There are various methods to train neural networks, and the one I will be addressing is Supervised Learning [117, 118]. One of the fundamental cores of ANN learning is the *Loss Function*, called also *Cost Function* or *Error Function*. It quantifies the discrepancy between the neural network prediction and the truth value. The Neural Network learning approach thus consists of adjusting the weight values in an iterative process based on the implementation of an *Optimization Algorithm* to decrease the difference between the predicted output and the real one. The smaller the Loss Function value, the better the ANN does its task. Optimizing the Cost Function in neural networks updates the connection biases and weights until a satisfactory loss value is achieved.

There are many types of Cost Functions that may be adopted for ANN. In general, the Cost Function is a weight-depending function having local and global minima, and few plateau regions, as shown in Figure 3.6 and Figure 3.7. As a result, an adequate optimization method is needed to obtain the global minima in the fastest and most

precise way. The first learning gradient-based algorithm introduced is called *Gradient Descent* [119, 120].

Gradient descent is one of the most popular algorithms to perform optimization and by far the most common way to optimize neural networks. It is an iterative method to minimize a differentiable function $f(x)$ by updating the parameter x, until reaching the local minimum. By taking the gradient of the chosen Loss Function concerning the connection weights, it is thus possible to update the connection weights making the ANN learn and improve its performance.

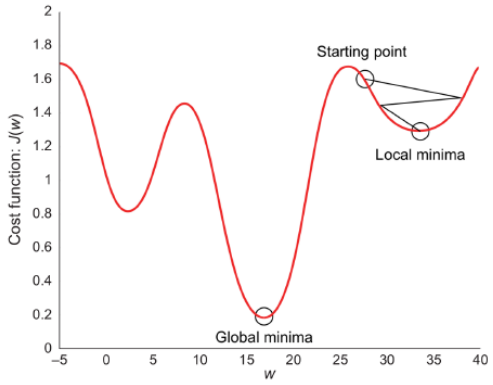


Figure 3.6: *The Cost Function with respect to its weight for a single Perceptron represented with a 2D curve. Illustration from [114].*

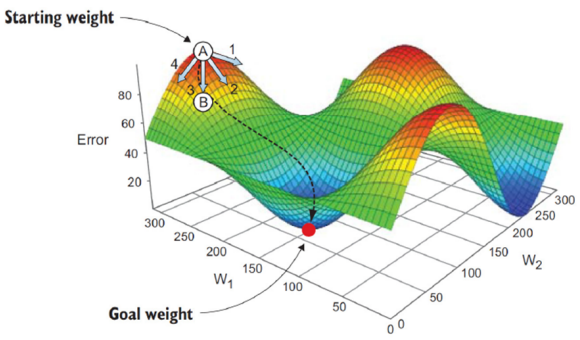


Figure 3.7: *3D graph showing the trend of the Cost Function in the presence of two weight factors. Illustration from [114].*

Although Gradient Descent is a very powerful method to get to the minimum error, it has some limitations, and not all Cost Functions fit this algorithm properly. Firstly, functions must be differentiable and the derivative must be calculable at every point. Moreover, as shown in Figure 3.6, the Gradient Descendent method allows to randomly start descending the curve and reach a minimum value. However, this local minima could not be the lowest possible error value for this Error Function, as it may not correspond to the global one. Certain Cost Functions may even present holes, ridges, and all sorts of irregular terrain that make reaching the minimum error very difficult and time-consuming. To overcome these drawbacks, other algorithms may be implemented, such as the *Stochastic Gradient Descent*, which provides many different weight-starting points, calculates their local minima, and finally, chooses the minimum one as the global minimum, and the *Mini-Batch Gradient Descent*, where the training sample is divided into mini-batches from which to compute the gradient [121].

An important requirement in the use of Deep Learning methods is the availability of a large number of data. These are needed for the Training Process in which the ANN learns how to do a task through the gradient-based algorithms just explained. However, not all the available data are involved in the training procedure. A part of them is applied to the ANN performance evaluation. These data must be different from the ones adopted in the

training and they should be completely independent from the model, not to bias the test evaluation. In addition, during the training process, which is a trial and error process, it is necessary to evaluate the model accuracy after each epoch to understand how it is performing and to tune its parameters. As the test data is only used to evaluate the final performance of the model after training is complete, it is also required to introduce a further additional dataset to evaluate the model during training.

The workflow for optimizing and evaluating the performance of a Supervised-Learning algorithm consists of training, validation, and testing steps, as shown in the scheme in Figure 3.8, and thus the data needs to be randomly split into three distinct sets:

(a) ***training dataset***, the sample of data used to tune weighting factors to train the model. These data must be representative in the best way possible of the real and complete phenomenon described.

(b) ***validation dataset***, the sample of data used to provide an unbiased evaluation of a model fit on the training dataset while tuning model *hyperparameters*. These are distinguished from parameters as they deal with the architecture design of the ANN. Moreover, unlike weights and biases, hyperparameters are tuned manually before the model training phase and comprise the number of hidden layers, the chosen activation and loss functions, the optimization algorithm, as well as the learning rate, and the number of epochs [114].

(c) ***test dataset***, the sample of data used to provide an unbiased evaluation of a final model fit on the training dataset.

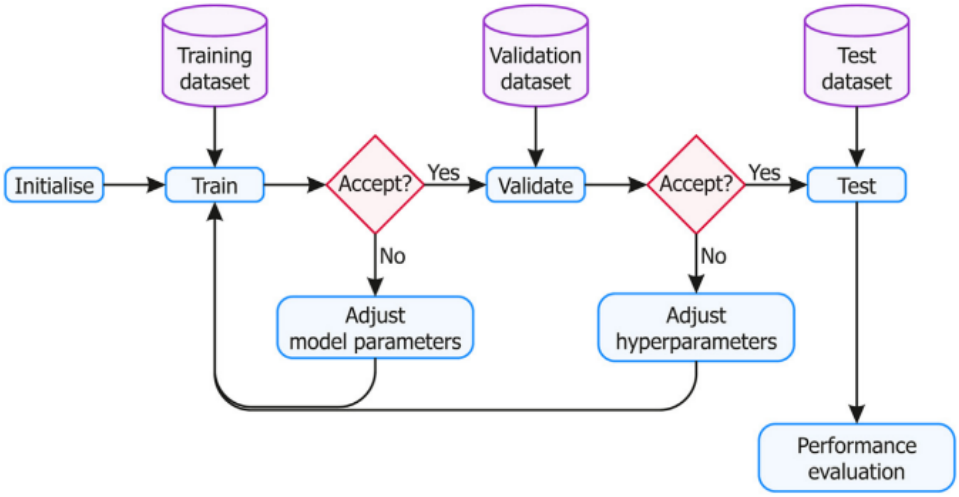


Figure 3.8: *Scheme of the workflow for optimising and evaluating the performance of a supervised-learning algorithm. Illustration from [110].*

Although Multilayer Perceptrons are powerful Deep Learning methods, they are not the best options when working with images. This is mainly due to the MLP architecture and

its connection network [114, 116].

As said before, the input layer neurons are fed by a 1D vector of features $(x_1,\ x_2, ..., x_n)$. However, images are at least composed of a 2D matrix, so they need to be flattened in a 1D vector to be analyzed by the MLP, losing all the spatial information in the pixel pattern of the image. Secondly, the MLP is a fully connected NN, hence the number of connection weights are closely related to the dimensions of the input features. In the case of large images composed of thousands of pixels, an extremely high number of neurons will be needed for each layer, and the number of connection weights would exponentially increase with the increasing number of hidden layers. All of this leads to a high inefficiency and computationally expensive process. The loss of information caused by flattening the 2D image matrix to a 1D vector and the computational complexity of fully connected layers implied the modeling of a new paradigm called *Convolutional Neural Network* (CNN) [122, 123].

3.2.3 Convolutional Neural Network

Convolutional Neural Network is a widely applicable model in supervised learning tasks involving image data processing [110, 124]. CNNs were inspired by studies of the visual cortex. CNNs employ filters to extract input image features in multiple ways, elaborating information at different resolutions and modifying images with different pixel values. The filters used, also referred to as *kernels*, are typically 3×3 matrices. This kernel "slides" over the original image pixel-by-pixel. At each point, a portion of the same size of the kernel is selected centered on the selected pixel. The dot product between this product and the kernel is calculated. It results being the values of the new "convolved" image at the next layer, as shown in Figure 3.9.

By applying different filters, it is possible to select different features and thus obtain different output images, depending on what is required to display. In Figure 3.10 it is shown how a CT image is modified after the application of different 3×3 filters. If edge information is required from that image, it is more suitable to apply a certain kernel (in this case the kernel **(f)**), while, on the contrary, if a sharper image or an image with a certain type of blurring is needed, further filters are available.

CNNs work the same way as Multilayer Perceptrons, as they are made of successive layers that learn increasingly complex features. The difference is that convolutional layers are employed instead of fully connected layers for the feature extraction task. In these layers, one or more filters are applied to the input. The numbers in the filter grids and the dimension of the filter that best suits the task are treated as weights that need to be adjusted during the training process. The CNN architecture is thus made of an input layer, followed by a succession of convolutional layers for feature extraction, fully connected MLP for image classification and, finally, an output prediction.

However, a large amount of convolutional layers may increase the number of weights of the network that need to be optimized through training. Furthermore, the size of the input image plays a role on the dimension of the CNN. For the same network structure, a larger image will have a larger number of weight. This will increase the computational size

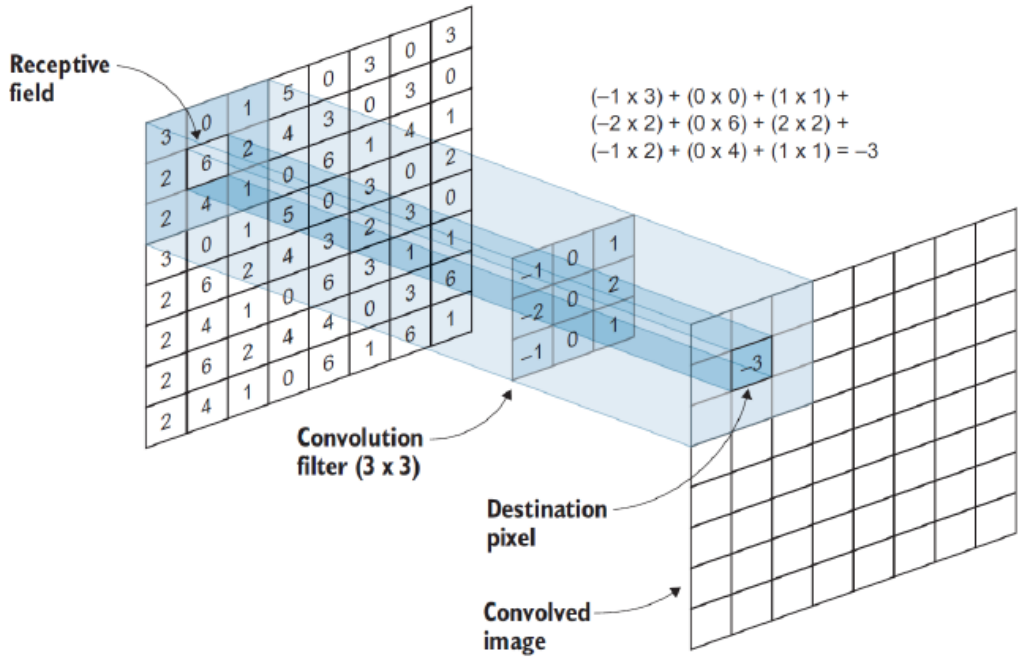


Figure 3.9: *Schematic representation of the convolutional process. Illustration from [114].*

of the problem, which if not taken into account may exceed the available computational memory.

To overcome this drawback, CNN layers are alternated by *pooling layers* [114, 116], or *subsampling*. As convolutional ones, also pooling layers are made of kernel "sliding" over the input images and applying some mathematical operations. However, these kernels are not made of weights. What pooling kernels effectively do is examine the feature map created by the previous convolutional layer and implement a specific function to pick only certain pixel values to be transmitted to the next layer, ignoring the remaining values. In this way, input values for a set of pixels give rise to a single output value, for example, their maximum or mean, depending on the applied kernel. The goal is thus to reduce the number of parameters of the network and so to lighten the computational weight, getting smaller feature maps and a shrunken image, but preserving the network depth. The diagram in Figure 3.11 represents a classification CNN. Convolutional layers are followed by pooling layers which decrease the dimension of the hidden layers. The last pooling layer is the flattened and fed into a MLP to produce the final image classification.

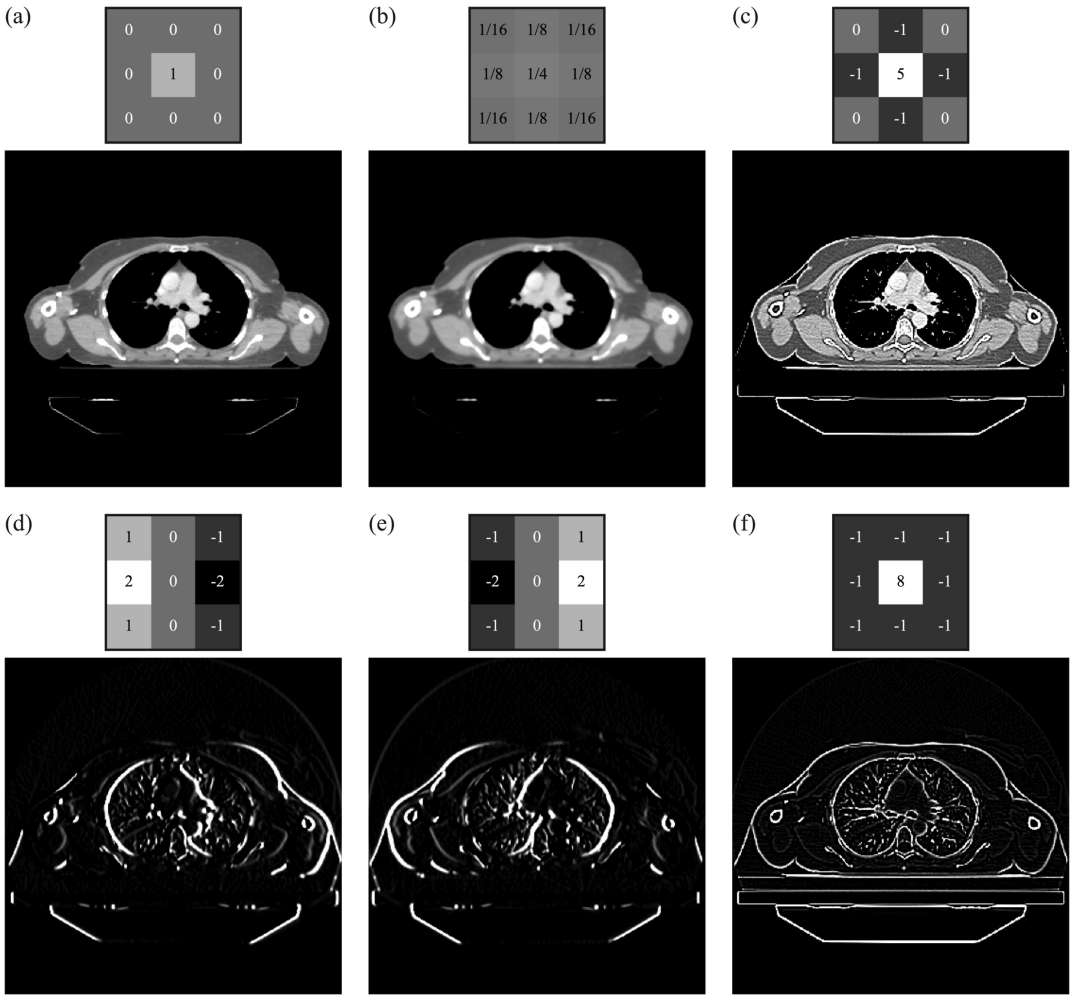


Figure 3.10: *Examples of 3×3 filters and the results of their convolutions with a CT image:* ***(a)*** *identity, which reproduces the original image;* ***(b)*** *Gaussian blur;* ***(c)*** *sharpening;* ***(d)*** *right-to-left gradient;* ***(e)*** *left-to-right gradient;* ***(f)*** *edge finding. Illustration from [110].*

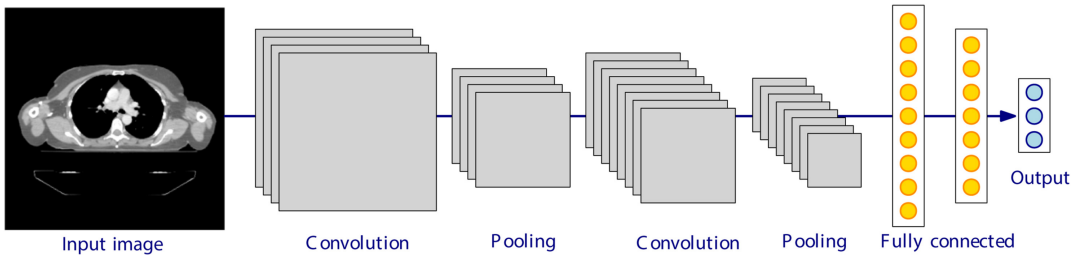


Figure 3.11: *Example of convolutional neural network (CNN). Illustration from [110].*

3.3 The U-Net

Image segmentation is achieved by replacing CNN's fully connected layers with convolutional ones, creating a *Fully Convolutional Neural Network* (FCNN) [125, 126]. This can be modeled by modifying and extending the architectural structure of Convolutional Networks to supplement a usual contracting network by successive layers, where pooling operators are replaced by upsampling operators.

The U-Net is one of the most popular Fully Convolutional Neural Networks. It was proposed by *Ronneberg et al.* in 2015 [127] and has now become the backbone and standard in medical image segmentation thanks to its very good performance. The U-Net architecture is shown in Figure 3.12. It is composed of two main parts having opposite roles, the Contracting Path, or *encoder*, and the following Expansive Path, or *decoder*. Every path is the repetition of the same architecture module repeated with different characteristics.

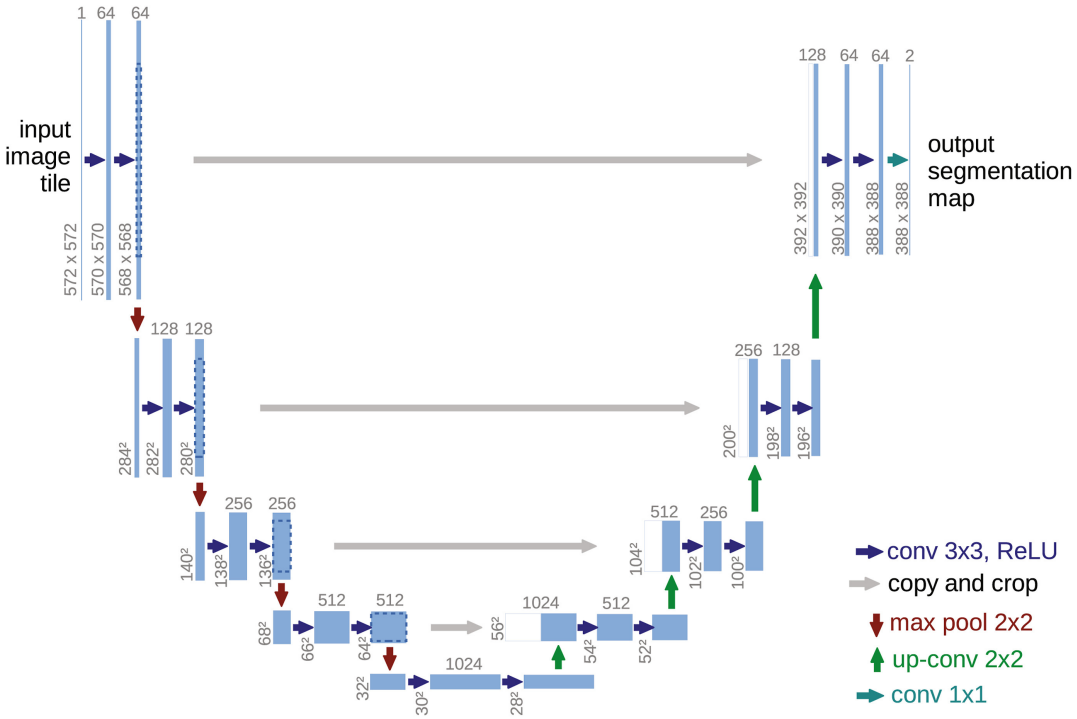


Figure 3.12: *U-Net architecture (example for 32×32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel image. The number of channels is denoted on top of the box, while the x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps, that are concatenated to the feature maps represented by the blue box. The arrows denote the different operations. Illustration from [127].*

The contracting path in U-Net is responsible for identifying the relevant features in the input image and it follows the typical architecture of a convolutional network. It consists of the repeated application of two 3×3 convolution layers, each followed by a ReLU activation function, leading to a deeper network able to extract more and more complex

features from the input image. Subsequently, a 2×2 max pooling operation is applied. It consists of selecting the maximum pixel value to pass along to the next layer, ignoring the remaining pixels. This sequence leads to an optimal feature extraction although the downsampling effect due to the application of pooling layers which lower the spatial resolution. The contracting path in U-Net is therefore responsible for identifying the relevant features in the input image.

On the other hand, the expansion path works on decoding the encoded data. It spatially locates the features and generates an output map with the same dimensions as the input image. Max pooling layers are substituted by *unpooling* [128] layers able to gradually restore the spatial dimensions of the feature maps while reducing the number of channels or features. The expansive module is thus repeated until the desired image size is reached. The segmentation is finally obtained through a 1×1 convolution layer associated with a Softmax activation function, that gives the probability of each pixel to belong to a class. Other ANNs are born from the adaptation of the U-Net to 3D image segmentation problems, such as 3D U-Net [129] and V-Net [130].

3.3.1 The nnU-Net

As U-Net is an extremely efficient tool, it has quickly evolved to a commonly used benchmark in medical image segmentation. However, for the network to function properly, it must be set up correctly according to what has to be analysed and processed. As said in Section 3.2.2, there are several benchmarks that are taken into account when building a network, including hyperparameters such as the learning rate or the batch size, which must be adequately tuned to improve the performance of the model. The model designing thus requires competence and experience, since its performance could be heavily compromised even by a small error, leading the network to underperform. This is the case of biomedical imaging. Moreover, the specificity of the ANN design is so deep that the same model could perform very well on a dataset but worse on another one. By contrast, the ideal would be the implementation of a model that would fit a greater variety of data. The critical decisions to be made during the shaping and the training of the DL method are in every phase, including the model architecture, how to do the training, and the methods for data pre- and post- processing. To overcome these issues, new models have been tested by adopting different values for each of the parameters involved and verifying the improvement of the performance.

This procedure led to the creation of a specific neural network model, the **nnU-Net** [131], an innovative segmentation method capable of performing automated configuration for new arbitrary datasets, without the occurrence of any manual setting. It is able to configure itself, including preprocessing, network architecture, training and post-processing for any new task in the biomedical domain in a very fast and data efficient way. As its name *"no new U-Net"* highlights, the nnU-Net does not introduce any new architecture of the NN, but applies the same architectural design as the standard U-Net and its adaptation 3D U-Net. The strength of this new network, indeed, relies on automatically setting network parameters according to the type of dataset fed and depending on its features in order to optimise the segmentation task.

The automatic configuration of nnU-Net is based on classifying domain knowledge into three parameter groups [132]:

(i) ***Fixed parameters***, which are parameters and design decisions that do not require adaptation between datasets. The values corresponding to these benchmarks are set and kept constant independently of the datasets employed. All those inherent to the *architecture model* and the *training schedule* are included in the fixed parameters. The architecture template adopted closely follows the original U-Net [127], its 3D counterpart [129] and the 3D U-Net Cascade [133] and the best suited one is automatically implemented. The only variations with respect to the original architecture are a smaller batch size of the networks and the replacement of the ReLU activation function with the Leaky ReLU [134]. Training instead is configured to end after 1000 epochs, each one defined as 250 iterations on the training set. The learning algorithm designed is the batch stochastic gradient descent and the loss function is the sum of the Dice loss function and Cross entropy loss function [135].

(ii) ***Rule-based parameters***, which have a dependency on specific dataset properties. It is therefore possible to establish heuristic rules that connect these design choices with the characteristics of the dataset and thus whenever the network is presented to an input having those qualities, it knows exactly how to autoset. The nnU-Net thus crops the implemented training data and is able to create a *dataset fingerprint* including parameters related to the image size, the number of voxels, the voxel size (imaging spacing), the number of modalities (CT, MRI or PET) along with the number of classes to segment and the total size of the training set [131]. According to the values of these parameters, the network generates the *pipeline fingerprint*, which is defined as the entirety of choices being made during method design. The nnU-Net reduces the design choices to the very essential ones and automatically infers these choices using a set of heuristic rules that operate on the above-described data fingerprint and project-specific hardware constraints [132].

(iii) ***Empirical parameters***, which need to be chosen empirically after the evaluation of the validation results after the training phase. Belonging to the latter category are the remaining design choices, that is, *model selection* and *post-processing*. This latter has a key role in organ image segmentation, as it often helps to remove some spurious false positives that may have been formed during network processing [132].

Once the fixed parameters have been set and the heuristic rules have been correctly established, a new dataset can be administered to the network.

The nnU-Net selects the image features and, exploiting the rules, sets the most adequate design choices. At the end of the process, the nnU-Net generates three different U-Net configurations: a two-dimensional (2D) U-Net, a 3D U-Net and a 3D U-Net Cascade. nnU-Net thus empirically selects the best performing configuration and through the application of post-processing techniques improves the final output.

This method has been performed on 11 international biomedical image segmentation challenges comprising 23 different datasets and 53 segmentation tasks, testing a variety

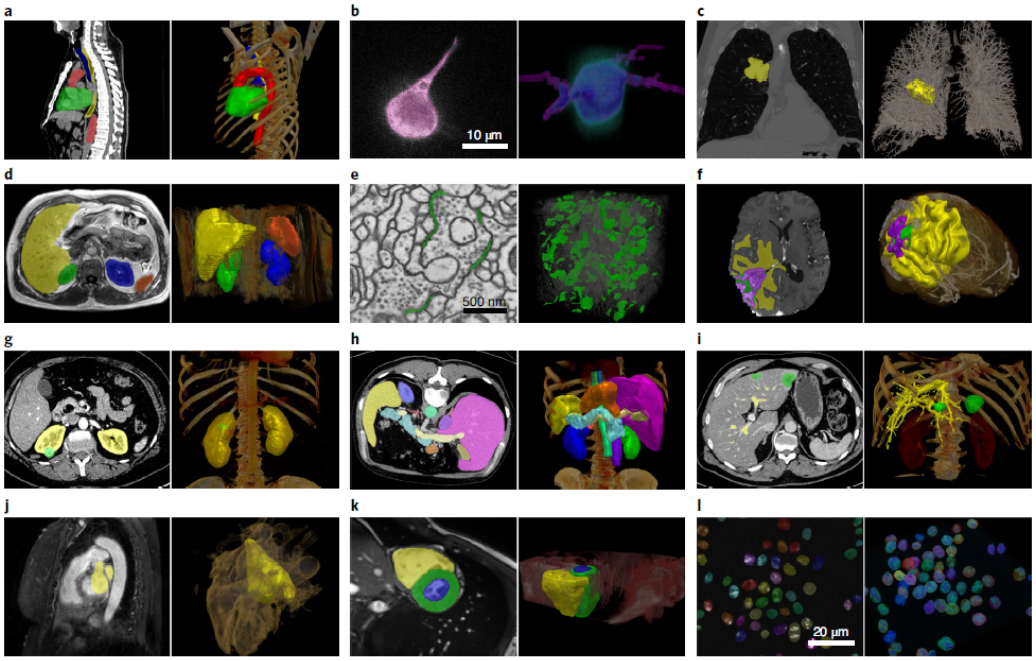


Figure 3.13: *A set of medical images segmented by the nnU-Net. Target structures for each dataset are shown in 2D projection (left) and in 3D (right).* ***(a)*** *Heart (green), aorta (red), trachea (blue) and esophagus (yellow) in CT images (dataset D18);* ***(b)*** *A549 lung cancer cells (purple) in FM (D22);* ***(c)*** *Lung nodules (yellow) in CT images (D6);* ***(d)*** *Liver (yellow), spleen (orange), left and right kidneys (blue and green, respectively) in T1 in-phase MRI (D16);* ***(e)*** *Synaptic clefts (green) in EM scans (D19);* ***(f)*** *Edema (yellow), enhancing tumor (purple), necrosis (green) in MRI (T1, T1 with contrast agent, T2, FLAIR) (D1);* ***(g)*** *Kidneys (yellow) and kidney tumors (green) in CT images (D17);* ***(h)*** *Thirteen abdominal organs in CT images (D11);* ***(i)*** *Hepatic vessels (yellow) and liver tumors (green) in CT (D8);* ***(j)*** *Left ventricle (yellow) in MRI (D2);* ***(k)*** *Right ventricle (yellow), left ventricular cavity (blue) and myocardium of left ventricle (green) in cine MRI (D13);* ***(l)*** *HL60 cell nuclei in FM (D21). Illustration from [132].*

of organs, tumors and cellular structures in 2D as well as 3D images acquired by different medical imaging techniques. Some results are shown in Figure 3.13.

These trials have evidenced the capability of the nnU-Net to adequately identify and segment a large variety of datasets relating to different organs and injuries, turning out to be an efficient and excellent tool. Therefore this DL algorithm has been employed for the development of a tool for medical imaging of patients undergoing BNCT treatment. This research work is framed in a more comprehensive project named **AI_MIGHT**, funded by the National Institute of Nuclear Physics (INFN) in the scheme of the Young Researchers Grants in 2021.

3.4 AI_MIGHT Project

The *Artificial Intelligence methods applied to Medical ImaGes to enHance and personalize BNCT Treatment planning*, AI_MIGHT project concerns the implementation of a Deep Learning software capable of obtaining segmented images of tumors from CTs and MRIs. This tool could provide an optimal support for researchers in the realisation and improvement of BNCT Treatment Planning Systems as it is able to quickly segment a wide dataset of images.

The prime aim of AI_MIGHT is to speed up the preliminary clinical evaluation of medical images performed by medical physicist, thus enabling them to start to work on the treatment planning of new patients without awaiting all the clinicians completed segmentation. As stated in Section 3.1, manual segmentation is a rather time-consuming process and possible delays in the TPS implementation may have consequences on the success of the therapy. On the contrary, in this way the analysis of patient positioning, the irradiation time and the beam configuration can be set-up in advance. Then, the clinicians can approve or re-countour the images and the dose distribution can be re-calculated in the final ROIs. Since the time-consuming procedures in the TPS are the patient positioning and the Monte Carlo calculations, the preliminary contouring done by AI_MIGHT could allow a faster Treatment Planning task.

Moreover, AI_MIGHT would be particularly useful for BNCT researchers looking to improve their Treatment Planning Systems. In order to do it, an extremely large amount of segmented data is required to test their computational instruments and enhance their performances. Therefore, the efficient and rapid segmentation of the tumors by AI_MIGHT would allow for a wide dataset of segmented data useful for research purposes without having to add a cumbersome task to the physicians and medical physicists. Moreover, the AI_MIGHT project could be coupled with the IT_STARTS TPS, discussed in Section 2.3 to study the efficacy of the automatic segmentation process and to provide a complete TPS tool. As the starting point, AI_MIGHT has created a standardized dataset of public Head & Neck and Brain tumors CTs and MRIs. To train the nnU-Net, open access databases have been used for both tumors and both images modalities. Focusing on CT images with corresponding RTSTRUCT files the database contains overall about 1934 H&N cases and 230 brain tumor cases. The RTSTRUCT files are manually contoured by physicians and are required for the training step since they create the ground truth of the segmentation.

In a precedent work the medical images have been cropped to optimise the NN performance by removing non useful information and, thus, considerably reducing the training time. Then, the dataset has been automatically separated by the nnU-Net into a set for training (and validation) and a set for testing. Finally the performance of the segmentation network was evaluated using a series of indexes commonly used in field.

In this thesis work I considered the CT images of 18 Glioblastoma Multiforme cases. For each case the BNCT dosimetry was obtained using IT_STARTS. The dosimetric study was conducted by using at first the ground truth segmentation present in the RTSTRUCT file and subsequentlly the automatic segmentation mask obtained by AI_MIGHT. The

dosimetry obtained in each ground truth case was compared to the relative one obtained by automatic segmentation to evaluate the performances of the algorithm and to test the possible coupling of AI_MIGHT with the IT_STARTS TPS.

Chapter 4

Treatment planning and dosimetrical analysis

This Chapter delves into the objective of this thesis, i.e., the evaluation of the application of nnU-Net in the automatic segmentation of medical images for BNCT applications. In particular, the accuracy of the segmentation achieved by nnU-Net and its ability to provide reliable dosimetric information for treatment planning of GBM have been examined. The potential benefits and challenges associated with the use of nnU-Net in dosimetry were also explored.

Specifically, I focused on the application of the nnU-Net trained in AI_MIGHT for the segmentation of GBM tumors. As explained in the previous Chapter, the use of nnU-Net for GBM segmentation involves training the network on a dataset of medical images provided with the corresponding manual definition of the tumors. The trained network is then used to automatically segment GBM tumors in the medical images of the selected set. The BNCT treatment planning was performed on the manually segmented volumes. The dosimetry was then analyzed in the manual and in the AI segmented volumes, using the same irradiation configuration. Therefore, using the same beam setup, I subsequently assessed the treatment by solely altering the irradiation time. This approach allows to study how the treatment might differ if the target volume is not in the expected location. The aim is thus to compare the doses when simulating treatment with the GTV contoured by a clinician to the results obtained when the GTV is contoured by an ANN, and how the possible differences affects the tumor control.

4.1 Dataset Creation

The creation of an accurate and representative dataset is crucial for the effective training of a GBM segmentation model using nnU-Net. The images used were downloaded from a free public archive, *The Cancer Imaging Archive* (TCIA) (`https://www.cancerimagingarchive.net/`), which offers medical images of different tumors. All imaging modalities are available such as CT, PET and MRI, and some

of them come with a corresponding manual segmentation image. These images are in DICOM format, that is the standard format for medical imaging. DICOM files are made by two parts, the image itself and the header, which stores all the information about the patient and the image.

Here is the full list of the datasets including brain tumors chosen in the TCIA and the corresponding number of patients available:

- **ACRIN-DSC** (123 patients);
- **ACRIN-FMIS** (45 patients);
- **IvyGAP** (39 patients);
- **GLIS-RT** (230 patients);
- **TCGA** (262 patients);
- **CPTAC** (66 patients).

These datasets should ideally cover a wide range of patients, imaging protocols, and disease stages. For this thesis, the CT images from the GLIS-RT dataset have been chosen, as it is the only one provided with the RTSTRUCT files, which are sets of anatomical structures identified on radiological images for the planning and delivery of radiotherapy. This collection consists of 230 cases of GBM and low-grade Glioma patients treated with surgery and adjuvant radiotherapy at Massachusetts General Hospital. All images are provided with the radiotherapy targets, the gross tumor volume (GTV) and the clinical target volume (CTV) manually delineated by the radiation oncologist. Of the 230 cases displayed, 198 are GBM patients. These 198 cases have been automatically sorted by the network, allocating 152 (76.7%) for the training and validation processes and the remaining 46 (23.3%) for the testing phase.

The training of nnU-Net for GBM segmentation operates as described in the previous Chapter. However, as the nnU-Net needs the images in the NIFTI format file, a specific tool is used to convert all the CTs and DICOM segmentations in NIFTI format. The GTV segmentation file obtained is hence a *mask*, a tensor whose voxels are 1 when the tumor is present and 0 everywhere else. The network architecture is thus configured to suit the characteristics of CT imaging, while the training and the validation stages are closely monitored adjusting weights and hyperparameters as necessary. The model evaluation relies on the application of the network on the test set. Thanks to a scientific collaboration with the Taipei Veterans General Hospital (Taiwan) and the Xiamen Humanity Hospital (China) the testing dataset was enriched with two more GBM patients who actually underwent BNCT.

4.1.1 Evaluation Coefficients

The evaluation process involves quantitative metrics and indexes to assess the segmentation accuracy and to compare the given ground-truth volume (G) and the NN predicted volume (S). The metrics used in this thesis are shown in Figure 4.1 and are defined as:

(i) ***Dice Coefficient (DC)***, also known as the *Dice similarity coefficient (DSC)* [136], which measures the similarity between two volumes. It ranges between 0 and 1, where 0 is obtained when there is no similarity, and 1 means perfect superposition. Dice score indicates the overlap volume of two structures relative to their average volume, and it is defined as

$$DC = \frac{2 \cdot (G \cap S)}{G + S}$$

(ii) ***Geometrical Miss Index (GMI)*** [137], which corresponds to the fraction of the ground-truth volume G that is not predicted. Also GMI ranges between 0 and 1, where a higher GMI indicates a greater geometric discrepancy between the real and the predicted volumes. A $GMI \simeq 0$ indicates instead a good geometric accuracy of the NN segmentation process. Geometrical Miss Index is defined as

$$GMI = \frac{G - (G \cap S)}{G}$$

(iii) ***Discordance Index (DI)*** [138], which is the fraction of the predicted volume S that does not belong to ground-truth volume G. It measures the over-contouring of the NN respect to the manual segmented image. Also DI ranges between 0 and 1, where a higher Discordance Index value indicates a larger disagreement of the images. It is defined as

$$DI = \frac{S - (G \cap S)}{S}$$

Dice Coefficient is thus an evaluation of the overlapping between the two masks, while the Geometrical Miss Index and the Discordance Index measure the dissimilarity between volumes.

After the training session, all the cases from the testing set have been evaluated: the coefficients distribution is shown in the box-violin diagram shown in Figure 4.2. The distribution of the Dice Coefficient has an average value of approximately 0.7, while the one related to the Geometrical Miss Index and the Discordance Index ranges between 0.2 and 0.3. “Violins” also denote a high concentration of data having high DC values and medium-low GMI and DI levels. This indicates a good level of overall overlapping between the two segments and that the network is performing well on the tested cases.

To obtain an analysis regarding the dosimetry due to BNCT, the medical images and the ROIs obtained by RTSTRUCT and by automatic segmentations were provided to the IT_STARTS TPS, for the simulations of the neutron irradiation.

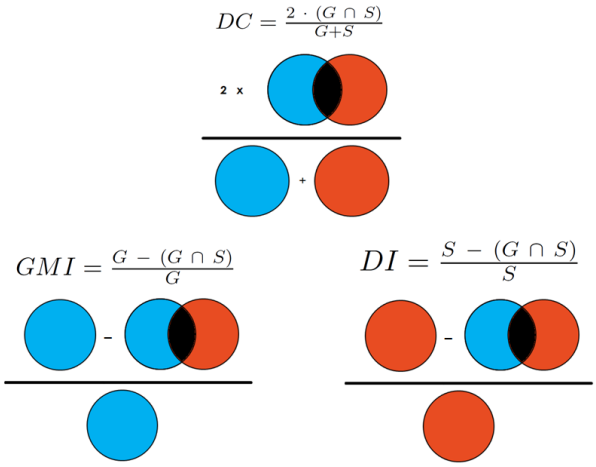


Figure 4.1: *Schematic illustration of the Dice Coefficient, the Geometrical Miss Index and the Discordance Index, respectively. The blue circle represents the ground-truth volume G, the red one is the simulated S, and the black part is the intersection between the two volumes $G \cap S$.*

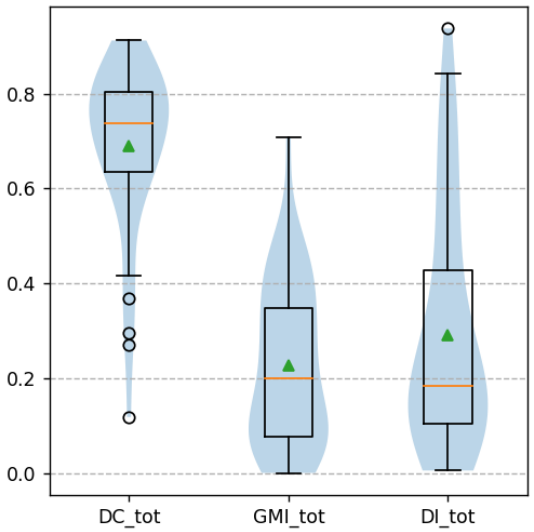


Figure 4.2: *The box-violin plot of the Dice Coefficient (DC), the Geometrical Miss Index (GMI), and the Discordance Index (DI) values regarding the whole testing dataset. The mean value is the green triangle and the median is graphically represented by the orange line. The box represents the interquartile range (IQR), in which the middle 50% of the data lies. The whiskers (or extension lines) extend from the box to the minimum and maximum values of the data within a certain range, typically corresponding to 1.5×IQR; empty dots represent the outliers. The violins depict the distribution of the data. The wider the violin, the higher the density of data in that region.*

4.2 Image Segmentation

To import a CT image and the corresponding ROIs from the RTSTRUCT into IT_STARTS as an overlaid mask over the CT a specific procedure is needed. The objective is to obtain a voxel-based geometry of the patient in which the different areas of interest can be distinguished. However, not all the required ROIs are included in the RTSTRUCT. Nevertheless, these can be derived, as discussed below. The same method is repeated for both the manually segmented GBM image and those obtained from the network.

In total, within my thesis work, I analyzed 18 patients from the test dataset, applying identical procedures until obtaining all necessary information regarding patient dosimetry. The volumetric characteristics of each patient's GTV are documented in the Table 4.1, alongside the DC, GMI, and DI values resulting from comparison with GTVs obtained from the neural network. A representative example (i.e. Patient_0022) of the application

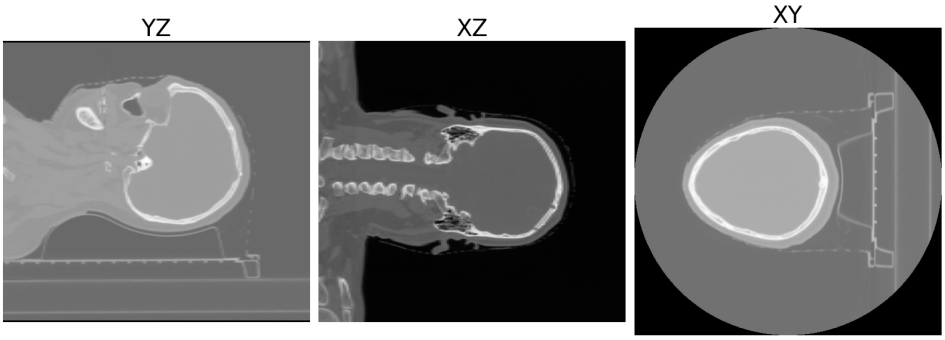
	DC	**GMI**	**DI**	**Volume true** (cm^3)	**Volume NN** (cm^3)
GBM_Taiwan	0.118	0	0.937	17.994	287.187
GBM_0006	0.271	0.700	0.699	53.434	323.650
GBM_0071	0.296	0.002	0.774	33.681	32.832
GBM_0057	0.369	0.038	0.734	29.267	128.957
GBM_0137	0.417	0.044	0.725	31.274	113.013
GBM_0202	0.427	0.105	0.574	11.446	39.766
GBM_0096	0.577	0.391	0.078	124.71	261.85
GBM_0217	0.688	0.097	0.333	253.558	152.465
GBM_0001	0.767	0.078	0.306	75.297	101.889
GBM_0020	0.777	0.039	0.348	30.747	42.127
GBM_0157	0.777	0.176	0.237	107.539	158.435
GBM_0192	0.793	0.231	0.150	34.354	37.096
GBM_0225	0.807	0.199	0.152	202.057	182.820
GBM_0022	0.824	0.137	0.177	61.243	57.871
GBM_0129	0.843	0.073	0.214	134.324	140.839
GBM_0144	0.851	0.093	0.173	89.635	105.703
GBM_0060	0.853	0.076	0.100	104.607	114.451
GBM_0128	0.912	0.042	0.842	207.673	213.312

Table 4.1: *Volumetric dimensions and values of DC, GMI, and DI for each analyzed patient. Volumes are referenced for both manually segmented GTVs and those artificially generated by the the neural network.*

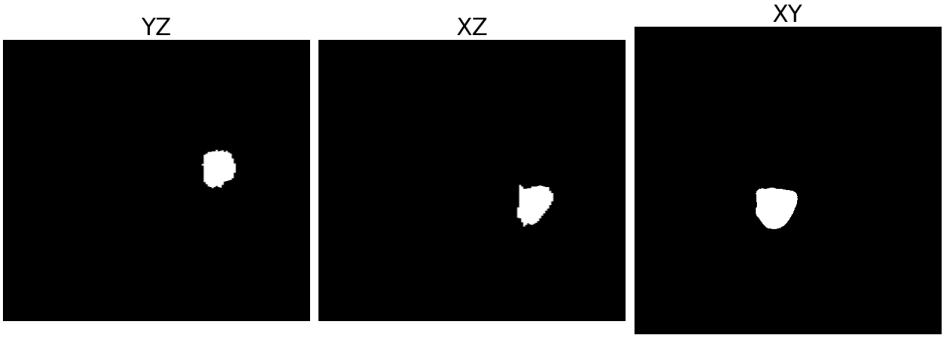
of IT_STARTS on a patient belonging to the test set is described step by step in the following section.

Patient_0022

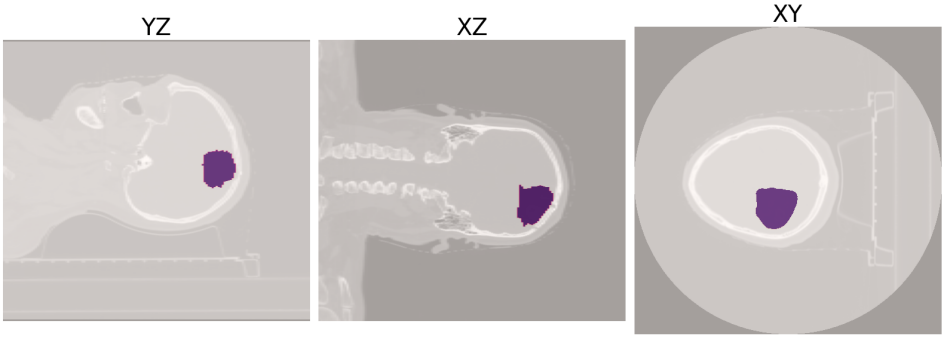
The first step of IT_STARTS treatment planning simulation is the loading of the CT image and the corresponding GTV mask provided by the RTSTRUCT in NIFTI format. IT_STARTS allows the display of the CT images in the three different planes: sagittal (YZ), longitudinal (XZ), and transversal (XY), as shown in Figure 4.3. The same can be applied to the mask of the GVT.



(a) *CT image.*



(b) *Manually segmented GTV.*



(c) *GTV mask (purple) overlaid to the CT.*

Figure 4.3: *From left to right: plot of the YZ, XZ, and XY slices.*

To obtain a meaningful dosimetric analysis, it is necessary to "build" masks corresponding to the other ROIs, including the brain, the bones and the soft tissue. This step is possible even if these anatomical regions are not present in the RTSTRUCT file, and therefore need to be created from scratch. IT_STARTS has the capability of exclusively selecting specific pixels within the CT image and map them to a particular anatomical structure, based on their HU numbers. The procedure employed to simplify this step involves defining as soft-tissue the voxels having HU in range of [-450 ; 150], as bones the voxels with HU greater than 150 and the rest as air. Regarding the creation of the *brain mask*, the process has been more complex, as its HU are in the range of [-10 ; 180], comprised in the range of values relative to soft tissue. In this case, I have applied to the soft-tissue mask the following additional filter:

$$\text{Kernel} = \left[\left(\left[\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right], 6 \right) \right]$$

The specified filter is a kernel for the erosion operation in a binary image processing context. Erosion is a morphological operation that is used to reduce the size of objects in the image while maintaining their general shapes. The kernel is a three-dimensional matrix that describes the shape of the structure to be used for erosion. In this case, the kernel contains two elements:

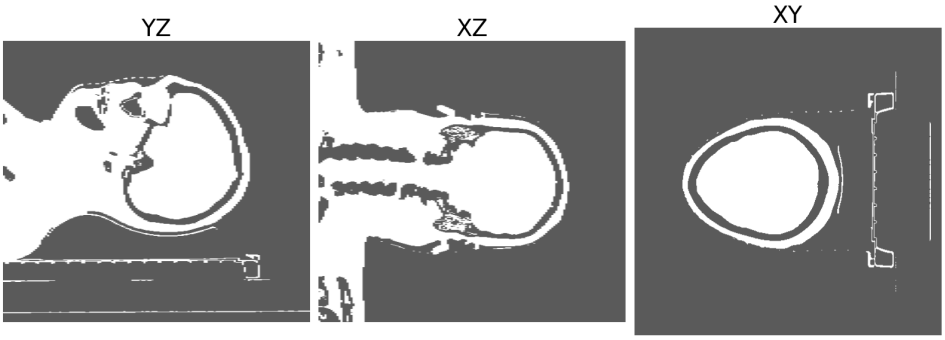
- A three-dimensional NUMPY [139] matrix representing the kernel itself. This matrix defines the shape of the structure used for erosion, has a $3 \times 3 \times 3$ shape and contains values 0 and 1.
- The number of iterations to be performed during erosion, which is six in this case.

The kernel is applied to each pixel in the image and the six iterations indicate that erosion will be repeatedly applied to the image six times.

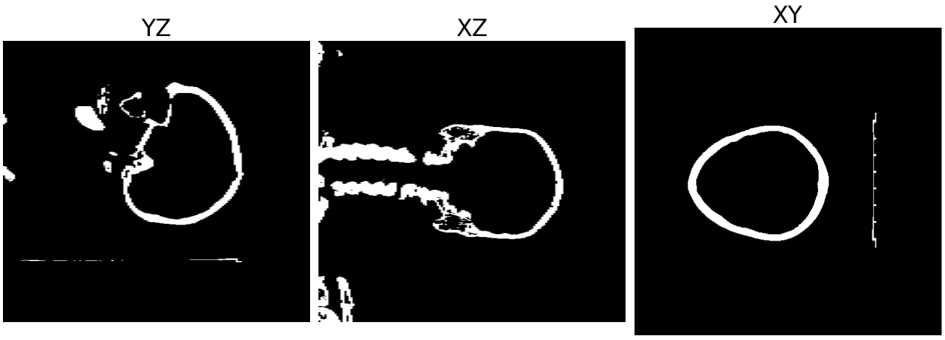
The OAR for the GBM treatment is the brain, excluding the GTV region. Therefore, it is necessary to create an additional mask representing the OAR, starting from the brain mask and excluding the GTV, i.e., setting the value to 0 for all its voxels (see Figure 4.4d). From this procedure I obtained the masks shown in Figure 4.4. The masks created are then assembled in such a way as to recreate the geometry of the entire patient and reproduce it as a set of individual templates, each depicting a different anatomical site (see Figure 4.4e).

A further step consists of the *voxelisation* of the total geometry, for which IT_STARTS requires to set a specific value for the size of the individual voxel; then it automatically converts the three-dimensional medical image into a set of voxels as shown in Figure 4.5. In this work, I set a voxel size of 4 mm for each dimension, to achieve good accuracy and at the same time avoid excessive computational time.

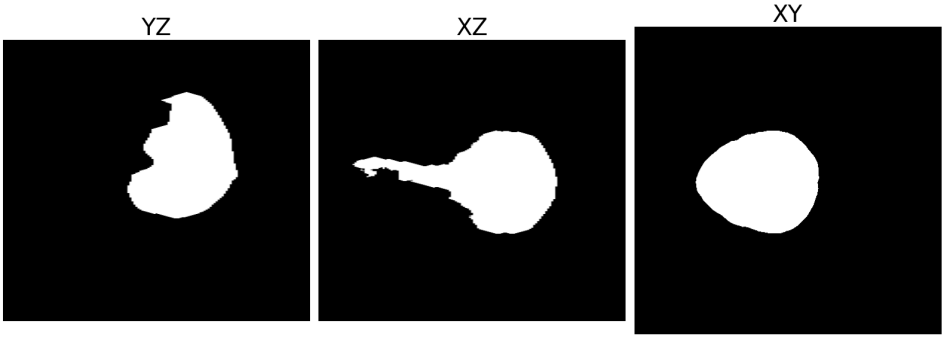
As the result of these procedures, IT_STARTS provides a voxel-based geometry of the patient, written in the syntax of the Monte Carlo transport codes, in this case, MCNP6.



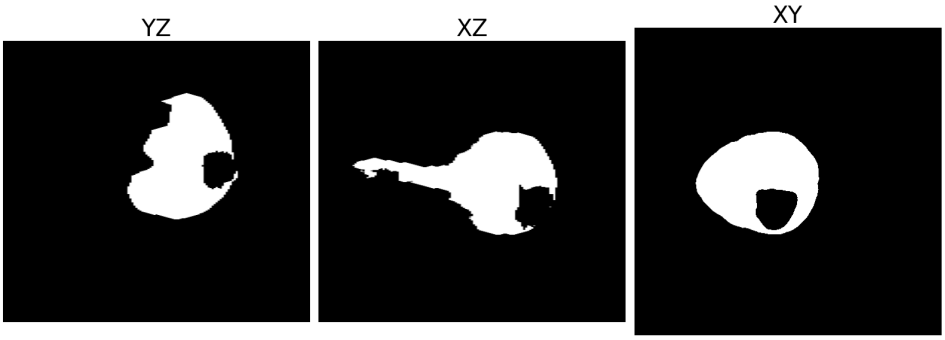
(a) *Soft tissue mask.*



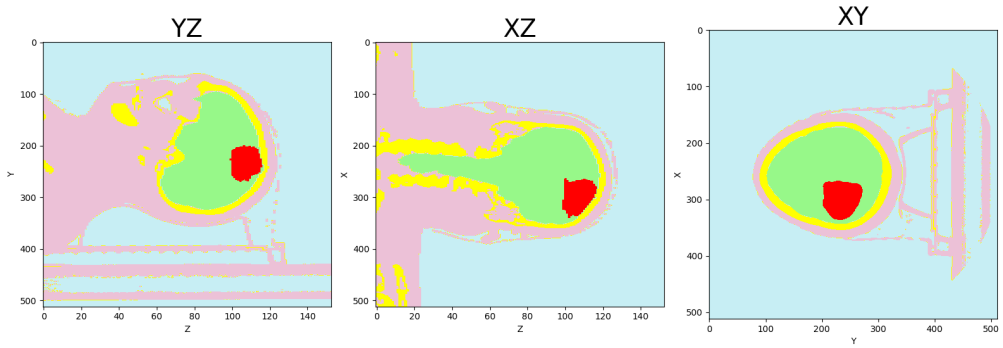
(b) *Bone mask.*



(c) *Brain mask.*



(d) *OAR mask.*



(e) *Assembly of all the masks created. The anatomical structures reproduced correspond to the soft tissue (pink), the bones (yellow), the organ at risk (green), the GTV (red), while the rest is the air (light blue).*

Figure 4.4: *From left to right: plot of the YZ, XZ, and XY slices.*

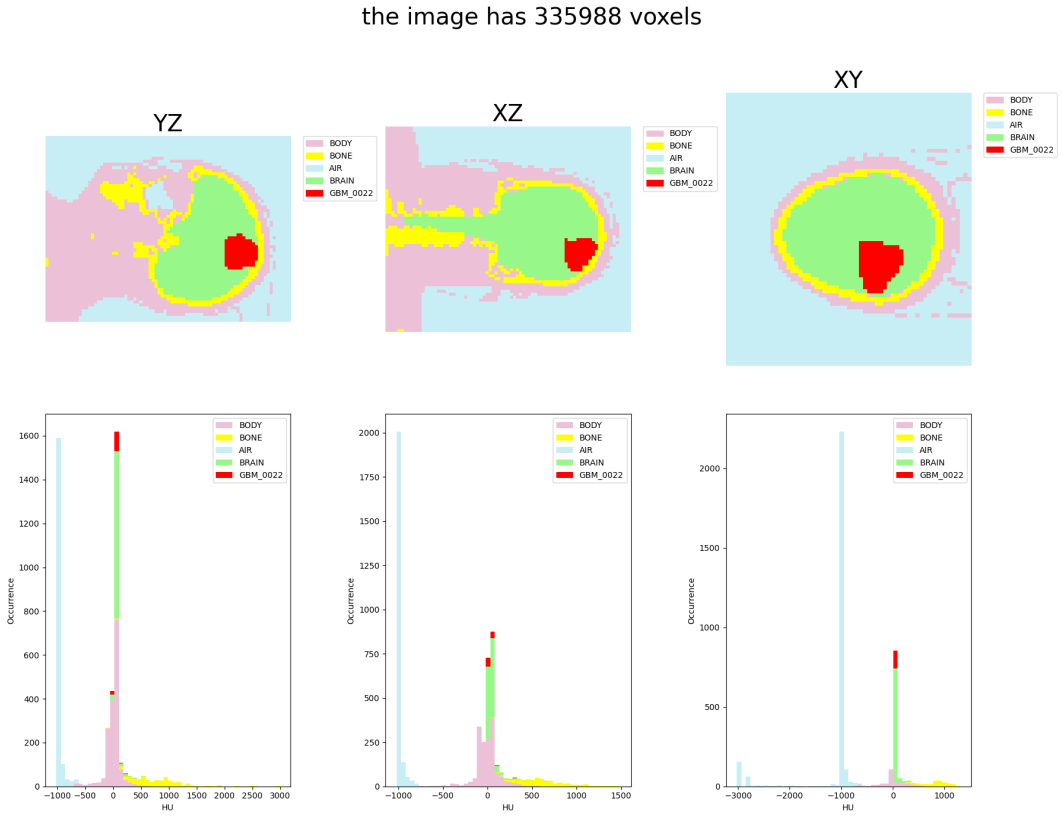


Figure 4.5: *Top: voxelised geometry of the patient, where the side of each voxel has been set to a size of 4 mm. The image has been also previously cut to remove the superfluous sites, including the CT scanner table and the headrest. Down:histograms relative to the occurrence of voxels having specific values of HU for each plane.*

4.3 Treatment Simulation

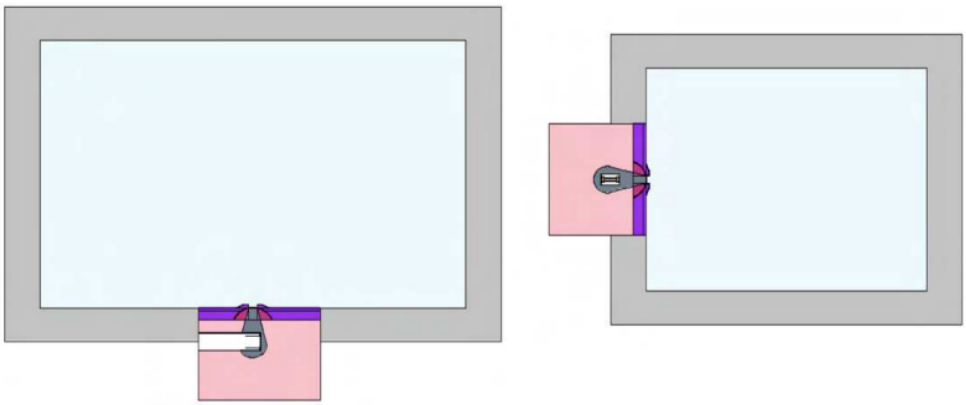
For the simulation of a BNCT treatment, it is necessary to position the voxelised patient inside the irradiation room, in the most advantageous orientation with respect to the beam port, where the neutron source is defined as described in Chapter 2.2.3.

The accelerator generating the beam employed in the simulation refers to the Radiofrequency Quadrupole accelerator [140] designed and built in Italy by INFN, delivering 5 MeV protons at 30 mA of continuous current. The protons impinge on a target of beryllium producing neutrons through the ^{9}Be(p,n)^{9}B reaction, which produces a neutron intensity on the target of approximately 10^{14} s^{-1}, with a maximum energy of 3.2 MeV. The beam is then moderated in the Beam Shaping Assembly (BSA), resulting in an epithermal spectrum (peak energy between 1 and 10 keV), suitable for deep-seated tumors. In Chiara Magni's doctoral thesis [141, 142], the treatment room was simulated taking into account radiation protection aspects with regard to air activation, materials in the BSA and walls, and dosimetry to the patient's organs and the environment, as shown in Figures 4.6a and 4.6b.

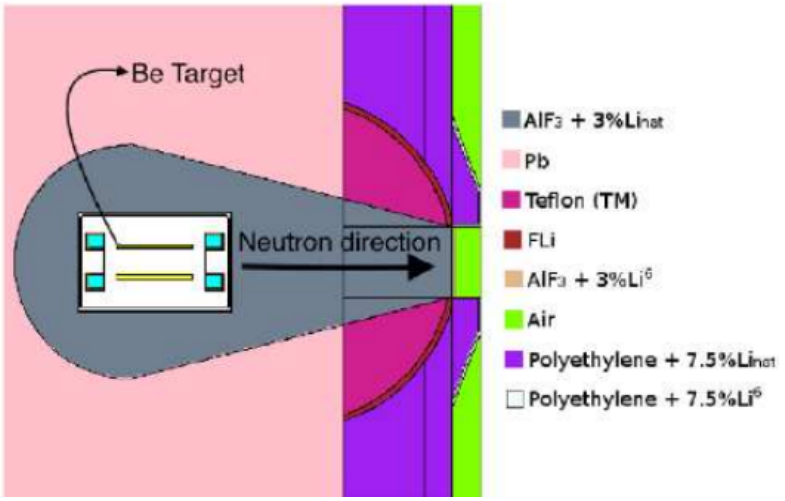
4.3.1 Patient Positioning

Positioning a patient for Boron Neutron Capture Therapy requires careful consideration to ensure effective treatment delivery while minimizing radiation exposure to healthy tissues. In order to achieve an optimal dose distribution, several strategies can be applied, and medical considerations must be taken into account regarding the protection of sensitive structures. In this work, where I needed to standardize a criterion to inter-compare different cases on common bases, I minimised the distance of the centre of mass of the GTV from the beam-port. This corresponds to maximise the irradiation of the GTV, regardless possible limitations of surrounding structures. This choice is justified because the goal of this study is not the evaluation of BNCT effectiveness in specific cases, but rather to evaluate the relative differences once fixed the irradiation position. I used the *3D Slicer* software to accomplish this criterion. Once uploaded the patient's CT and the masks in NIFTI format, the program enables to locate the centre of mass of the GTV and to generate a 3D geometry of the patient. This three-dimensional image can be turned and moved, making it easier to establish the rotation angles. Patient_0022 has undergone a rotation of 170° around the x-axis and 33° around the y-axis, and a translation of 3 cm along x, 1 cm along y and 66.5 cm along z. This results in a patient position in the treatment room as shown in Figure 4.7.

The next step is the simulation of the patient irradiation with the neutron beam. To assess the dose absorbed in each voxel, it is necessary to set the tallies as outlined in Chapter 2.2.4. In this work, I implemented the F4 tally in conjunction with the FMESH card. Meshes are three-dimensional grids superimposed on the patient; in this case, the grid elements were aligned with the voxels. KERMA factors corresponding to the neutron energy values reported in [83] are then coupled with the fluence tally for the four radiation components of BNCT. As in the inner volumes the charged particles equilibrium is ensured, KERMA can be used to calculate dose, and this method yields a single file



(a) *Horizontal (left) and vertical (right) sections of the MCNP6 model of the treatment room. The simulated treatment room has dimensions of 3.30 × 4× 6 m^3 with walls of 50 cm in borated concrete.*



(b) *Representation of the simulated BSA. Neutrons are produced via nuclear reaction of protons in the beryllium target and are moderated along the lithiated aluminum fluoride structure (grey).*

Figure 4.6: *Treatment room for the simulation of patient irradiation. The neutron beam is directed along the z axis. Both illustrations are from [141].*

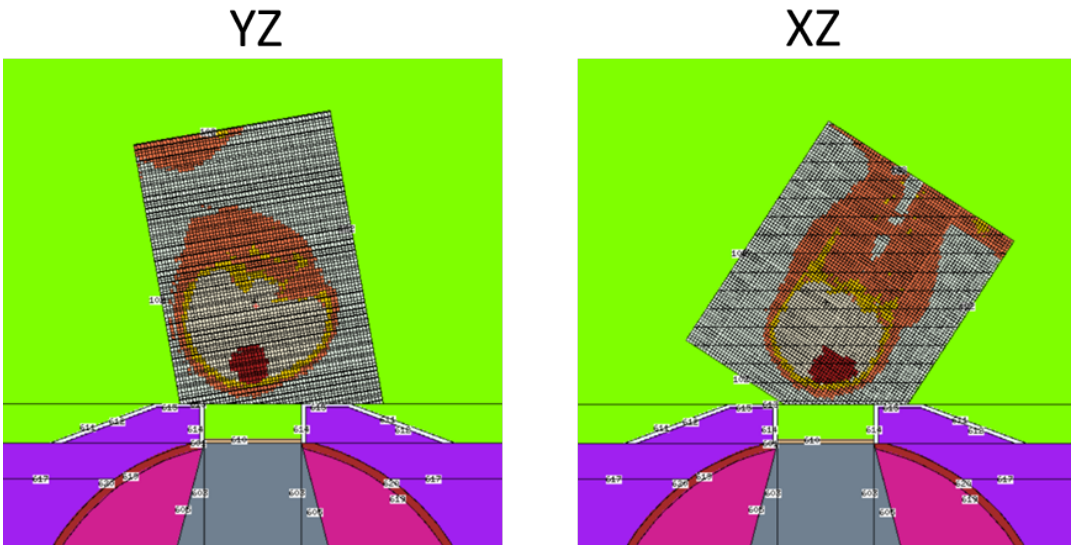


Figure 4.7: *Patient positioning in the treatment room along the YZ and XZ planes.*

recording the calculated dose rates for each of the four components involved in BNCT treatment. The irradiation simulation consisted in launching 10^8 source particles, resulting in a relative uncertainty of dose rate in the GTV voxels lower than 2%.

To reduce the statistical variance and speed up the Monte Carlo simulations, *weight windows* were employed in the particle transport [79]. Weight windows are defined by assigning and redistributing weights to particles throughout the simulation. Weight is a characteristic of particles that transforms *analog* Monte Carlo into *non-analog* Monte Carlo, where each particle can represent more than one, or a fraction of one, analog particle. Weights can change in the simulation as particles cross more important parts of the geometry, as labelled by the user. This produces the sampling of a higher number of particles with lower weights in these volumes, improving the statistics. However, an over-population may result in large computational time, thus weight windows are set to control and make the weights uniform along the particle transport. With windows, particles with higher weights are sampled more frequently (while their weight is lowered accordingly), and those with lower weights are compacted into fewer particles with higher weights. This concentrates the majority of Monte Carlo samplings in regions of greatest interest, such as those corresponding to the GTV, without losing efficiency.

4.4 Photon-Equivalent Dose calculation

At this stage, the dose rates extracted from the output file generated by the particle run are fed into a spreadsheet. Subsequently, this file is read by IT_STARTS to compute the total absorbed dose and provide essential details for clinical treatment. In addition to dose rates, the spreadsheet also contains the CT scan location in the file system, the boron concentration in the GTV, the ROIs, and OAR, as well as the prescribed dose

limit values for the patient's treatment to calculate the irradiation time. As stated in Chapter 2.3, the treatment time is determined using the dose limiting values, that can be the maximum dose value, D_{max}, or the mean dose calculated within 50% of the OAR, D_{mean}, depending on the tissue. In this work, albeit the limitation in brain is set on the mean value, we adopted the conservative approach to calculate both dose rates and to select the most restrictive one (i.e. the one that leads to the lower irradiation time).

For photon-equivalent dose calculation, IT_STARTS allows the input of radiobiological data for the RBE-weighted dose model and the photon isoeffective dose model. The absorbed dose in the OAR is calculated applying the RBE-weighted dose model described in Chapter 1.2.2 and assuming the RBE and CBE parameters given by *Coderre* [143], as it is the model applied in clinics, i.e. $CBE = 1.35$, $RBE_{th} = RBE_f = 3.2$, $RBE_\gamma = 1$. Even if the GTV dosimetry is obtained using the photon isoeffective dose model, for the healthy brain the model is not yet available. However, there are some reasons why the calculation of fixed-RBE weighted dose in healthy tissues is justified. First, the limiting criteria are available from the clinical BNCT experience using RBE/CBE. Moreover, literature demonstrated that at low dose, typical of healthy tissues, the RBE-weighted dose and the photon isoeffective dose do not differ significantly.

For this patient, the dose limit values of $D_{mean} = 2.5$ Gy-Eq and $D_{max} = 13$ Gy-Eq in the brain, generate the irradiation time of 39.08 minutes and 37.30 minutes, respectively. As said above, to protect as much as possible the normal brain the software selects the lower time: 37.30 minutes.

For the GTV, the photon isoeffective dose model is adopted. However, as an example of dose calculation with the traditional model, I report also RBE-weighted dose for one patient. To this end, it was possible to use original data, coming form radiobiological experiments conducted in Pavia on the human immortalised GBM cell line U87. The radiobiological parameters were derived from experimental measurements conducted at the TRIGA Mark II reactor of the University of Pavia. Flasks containing BPA-treated U87 cells were subjected to a thermal neutron beam characterized by low gamma and fast neutron contamination in the thermal column of the reactor. Other flasks were irradiated without boron treatment (beam-only experiments). Moreover, reference irradiation was carried out with a ^{60}Co photon source available at the reactor. All irradiation experiments were performed with dose escalation, in order to obtain dose-survival curves following the same procedure as described in [42]. The survival curves, are illustrated in Figure 4.8

The curves represent the *in vitro* GBM survival-dose relation in BNCT, allowing the translation of BNCT dose in photon-equivalent units using proper radiobiological parameters. In fact, the typical parameters used in clinics rely on the Coderre's experiments conducted with a Gliosarcoma experimental model. It was thus interesting to derive the parameters for the RBE-weighted dose model, defined in the first Chapter, using these curves. The CBE value is derived by comparing the dose associated with photons (depicted by the blue curve) to that of the boron capture reaction products, namely alpha and lithium (represented by the black dotted curve), for the same endpoint. A similar process is employed to determine RBE_{th} and RBE_f. Since protons are elicited in both

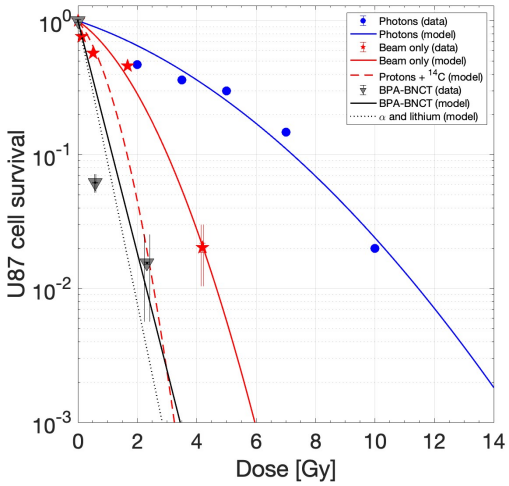


Figure 4.8: *Survival curves of U87 cells as a function of the absorbed dose. The blue dots and the corresponding blue line represent respectively the data from the reference radiation, ^{60}Co, and the fit. The red curve and the red stars are the fit and the experimental data for the beam-only irradiation. The red dashed curve is the contribution of protons. The black curve is the fit of the BNCT experimental points (black triangles). The dotted black line is the model considering only the boron component. Courtesy of Barbara Marcaccio, PhD Thesis, paper in preparation.*

reactions, their biological effectiveness is presumed to be identical, thus only one RBE_n value is considered, this time with reference to the dotted line representing survival just considering protons and ^{14}C ions. Applying Equation 1.4 and setting the endpoint as the 1% survival fraction of irradiated cell, the value of CBE was calculated to be 6.03 ± 0.03, and RBE_n was determined to be 4.49 ± 0.08. By definition, the RBE_γ for is 1.

Regarding the photon isoeffective dose model, the characteristic repair time values for the U87 cell line and the relative percentage of cells repaired by both fast and slow kinetics for photons and high-LET radiations, which were used in the fits are:

- $\boldsymbol{t_{0f} = 91.42\ min}$, characteristic repair time considering the *fast* kinetics;
- $\boldsymbol{t_{0s} = 1238\ min}$, characteristic repair time considering the *slow* kinetics;
- $\boldsymbol{p_{\gamma_f} = 0.77}$, proportion of sublesions repaired by the *fast* kinetics for low-LET radiation;
- $\boldsymbol{p_{\gamma_s} = 0.23}$, proportion of sublesions repaired by the *slow* kinetics for low-LET radiation;
- $\boldsymbol{p_{BNCT_f} = 0.2}$, proportion of sublesions repaired by the *fast* kinetics for high-LET radiation;
- $\boldsymbol{p_{BNCT_s} = 0.8}$, proportion of sublesions repaired by the *slow* kinetics for high-LET radiation.

These parameters allow the calculation of the Lea-Catcheside time dependent factor $G(\theta)$ by applying the Equations 1.7 and 1.8.

It is also possible to derive the values of alpha and beta used in the calculation of the isoeffective dose for each dose contribute (see Equation 1.22). Some assumptions were made for the assessment of these parameters:

(i) For photons, $\alpha_4 = \alpha_R$ and $\beta_4 = \beta_R$, due to the similarities in the energy spectra of the reference photons (^{60}Co, about 1 MeV) and the ones from the neutron beam (typically around 2 MeV).

(ii) For neutrons, $\alpha_2 = \alpha_3 = \alpha_n$ and $\beta_2 = \beta_3 = \beta_n$, where α_2 and β_2 related to the dose contribution of thermal neutrons, and α_3 and β_3 to the one of fast neutrons. This is possible as the recoil hydrogen nucleus released during the scattering interaction of fast neutrons has an energy comparable to that of the proton generated by the reaction of thermal neutrons with nitrogen. As a consequence their biological effectiveness is the same, and so this assumption can be applied.

The number of parameters to be identified for the photon isoeffective dose model is thus reduced, and through the fitting experimental data in Figure 4.8, the following values listed in the Table 4.2 are obtained for the U87 cell line.

	$\boldsymbol{\alpha}$ $[\boldsymbol{Gy^{-1}}]$	$\boldsymbol{\beta}$ $[\boldsymbol{Gy^{-2}}]$
Boron	2.4 ± 0.5	0
Neutrons	0.6 ± 0.6	0.5 ± 0.9
Photons	0.17 ± 0.08	0.021 ± 0.009

Table 4.2: *Parameter values of the survival curves obtained by fitting the experimental data in Figure 4.8.*

IT_STARTS calculates the photon isoeffective dose distribution in each voxel inside the patient using the model fed with these radiobiological data, experimentally determined.

The software generates the corresponding DVH histograms for the dosimetric analysis. Figure 4.9 shows the dose-volume distributions inside the OAR and the GTV applying fixed-RBE and photon isoeffective dose models.

The DVH of the brain shows a rapid and steep decline: this is the curve that was used to determine the irradiation time and normalize all the histograms to obtain the dose. It is interesting to compare the curves regarding the GTV dosimetry calculated with the two models, echoing concepts introduced in the first chapter. The RBE-weighted dose model yields tumor photon-equivalent doses significantly higher than those derived from the photon isoeffective dose model [43, 144]. Table 4.3, reports the minimum, mean and maximum dose values. An important consideration concerns the minimum dose value: it is important because the probability to control a tumor critically depends on this. An adequate treatment, in fact, does not underdose any part of the tumor. On the other hand, the maximum dose value absorbed in tumor is not that significant for the clinical

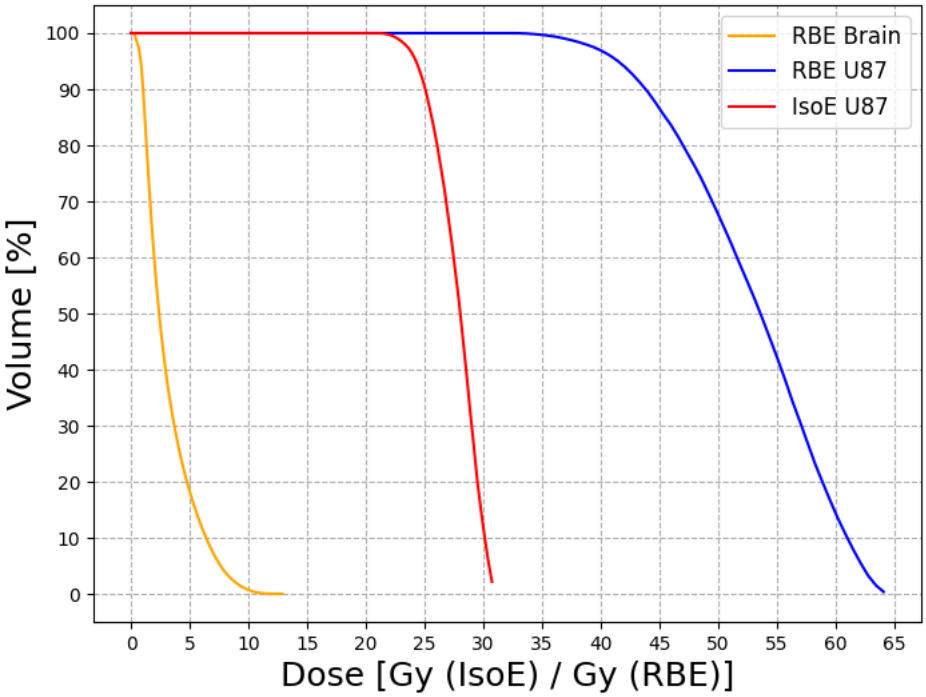


Figure 4.9: *DVHs of the brain (orange line), and the GTV calculated by applying both RBE-weighted (blue line) and the photon isoeffective dose (red line) models.*

outcome. It is from these considerations that it was possible to test whether the RBE or the isoeffective model was the most appropriate to explain retrospective clinical outcomes through TCP calculations [41, 43, 144]. In those works it was demonstrated that RBE dose model delivers unrealistic high dose in the tumor which is, in fact, not equivalent to photon therapy.

For this reason, in the rest of my thesis I will consider only the photon isoeffective dose model for the GBM dosimetry.

	RBE-weighted Dose	**Photon Isoeffective Dose**
Minimum Dose [Gy]	38.7	23.3
Mean Dose [Gy]	53.6	28.0
Maximum Dose [Gy]	63.2	30.7

Table 4.3: *Values of minimum, mean, and maximum dose calculated using the two equivalent dose models. The minimum dose refers to the dose value received by 98% of the volume, the mean corresponds to the dose received by 50%, and the maximum to 2% of the total volume of the GTV.*

The overlay of isodose curves computed applying the photon isoeffective dose model on the CT images shown in Figure 4.10 illustrates the distribution of radiation dose absorbed. They are calculated as a percentage of the maximum dose absorbed by the OAR. The fact that the maximum isodose comprises the GTV, demonstrates a good positioning of the patient.

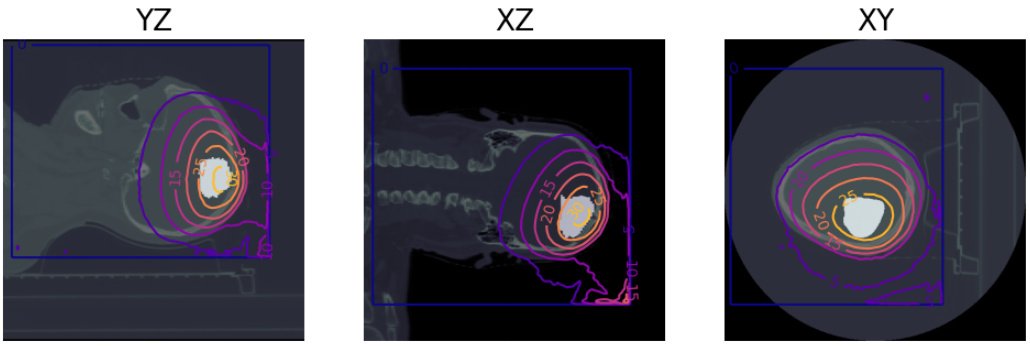


Figure 4.10: *Isodose curves computed using the photon isoeffective dose model overlaid on the patient's medical images.*

4.4.1 Tumor Control Probability

The DVH and the consideration on the difference in minimum and mean dose are indicators that can be used to evaluate the impact of the AI contouring. However, from a radiotherapy point of view, a fundamental criterion to evaluate a treatment planning is the *Tumor Control Probability* (TCP). This figure of merit, in fact, is a single number that summarises the effectiveness of a treatment in terms of its potential to achieve the tumor control. Of course, the purpose of this work is not to calculate TCP to make a clinical assessment of BNCT for GBM, because this would require an in-depth study of the TCP model itself, extended to the case of non-uniform BNCT dose distributions. In this work, the TCP has been used as a **figure of merit** to compare the impact of the difference in dose distributions obtained in automatic and human-based ROIs. As a matter of fact, it is very difficult to understand *a priori* the clinical implications of the variations in the minimum, mean and maximum dose on the treatment outcome. The TCP is an indicator closely related to the effect on the patient and therefore is more meaningful and easy to understand. IT_STARTS provides a module to calculate TCP, by implementing a chosen model with its radiobiological parameters.

Currently, there is no TCP model specifically tailored for the BNCT treatment of GBM. Therefore, for this study, I have adopted a TCP model derived from clinical photon treatments, which was then converted and adjusted for BNCT. This is justified because the use of photon isoeffective dose model translates the BNCT dose into photon-equivalent units; previous work has validated this approach with other tumors [144, 145].

The TCP model is based on the concept of cell survival fraction, previously introduced in the first chapter and defined as a function of dose according to Equation 1.6. Cell survival

follows the linear-quadratic (LQ) model, where the survival fraction $S(D)$ of cells in the tissue exposed to a total radiation dose $D = nd$ is defined as

$$S(D) = e^{-D(\alpha+\beta d)+\frac{\ln 2}{T_d}(T-T_k)} , \tag{4.1}$$

where n is the number of fractions and d is the dose per fraction. In this model α represents the intrinsic whole tumor radiosensitivity, and β its repair capability, while the time factors T_d instead indicates the repopulation doubling time, T_k the kick-off time for accelerated repopulation and T the treatment time.

The TCP model adopted here is based on the Poisson assumption, which states that cell death events follow a Poisson probability distribution, resulting in the Equation

$$TCP = e^{-NS(D)} = e^{-N\cdot e^{-D(\alpha+\beta d)+\frac{\ln 2}{T_d}(T-T_k)}} , \tag{4.2}$$

where N represents the number of clonogens for a given regimen of fractionation.

In the paper by *Pedicini et al.* [146], data concerning the conventional photon therapy of patients affected by GBM following a specific fractionation regimen were collected. These data were analyzed to obtain the parameters for Equation 4.2, in particular $N = (9.1 \cdot 10^3 \pm 1.2 \cdot 10^4)$, $\alpha = (0.12 \pm 0.02)\ Gy^{-1}$, and $\beta = (0.015 \pm 0.005)\ Gy^{-2}$.

The TCP model was then adapted to BNCT by setting a value of $n = 1$, as BNCT does not involve fractionation. To do this, the *biologically effective dose*, BED, is employed. Considering n fractions, BED is defined as

$$BED_n = -\frac{\ln(S)}{\alpha} = nd(\alpha + \beta d) - \frac{\ln 2}{T_d}(T - T_k) \tag{4.3}$$

This concept is particularly important in radiation oncology when comparing treatment plans or different fractionation schemes. By equating the BED values for different fractioned schemes ($BED_n = BED_1$), TCP equation for single fraction dose is thus obtained as

$$TCP = e^{-NS(D)} = e^{-Ne^{-\alpha BED_n}} = e^{-Ne^{-\alpha BED_1}} = e^{-Ne^{-D(\alpha+\beta D)}} \tag{4.4}$$

The available data of Tumor Control Probability as a function of the delivered dose, is then plotted in the case of single fraction dose ($n = 1$) as shown in Figure 4.11. To calculate the TCP value corresponding to a certain DVH it would be necessary to extend the calculation and include the information of dose distribution (which is not uniform as in conventional radiotherapy). However, for the purpose of this work, we rely on what demonstrated in [53]: the overall TCP is always comprised between the TCP due to the minimum and the TCP due to the mean dose values. Therefore, the minimum and mean dose values obtained from IT_STARTS into the Equation 4.4 give the range of values where the tumor control probability lies. In this specific case, the respective values of $TCP_{min} = (0.854 \pm 0.176)$ and $TCP_{mean} = (0.997 \pm 0.003)$ are obtained when the dose is calculated in the GTV contoured by the physician. The uncertainties have been calculated through error propagation of the fit described in the original work [146]. The two TCP values are very close and overlap within the uncertainties. It can thus be estimated that in this case the TCP is very close to 1.

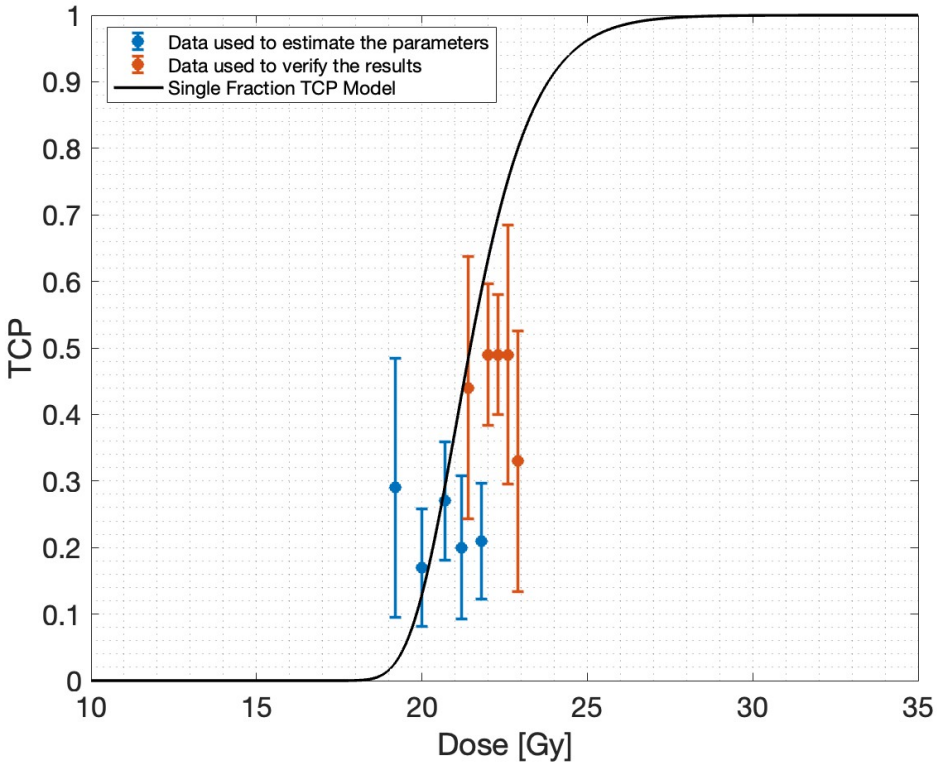


Figure 4.11: *Single Fraction TCP Model for GBM as a function of the delivered photon dose dose. Data taken from [146].*

4.5 Comparison with AI segmentation

The procedure was applied to the artificially segmented GTV image generated by the AI. After feeding the CT file into the nnU-Net algorithm (Chapter 3.3.1), the AI-based GTV mask is provided. By comparing the manually segmented tumor volume provided by the database with the one obtained from the neural network, a high Dice Coefficient value of DC=0.824 was achieved, along with low values of Geometrical Miss Index and Discordance Index, respectively GMI=0.198 and DI=0.152. The segmentation algorithm thus appears to perform well on Patient_0022's CT scan, with the two tumor regions largely overlapping, as shown in Figure 4.12.

To perform a dosimetric comparison, it is necessary to repeat some of the previously described operations, replacing the original GTV mask with the AI-based one. Consequently, a new voxelized patient geometry is obtained, where the new GTV and OAR differ from the previous ones. The calculated irradiation time in this case was 38.05 minutes.

The photon isoeffective dose values are collected in Table 4.4 and compared with those previously obtained. Figure 4.13 reports the DVHs in both cases, showing a minimal difference between the two dose distributions.

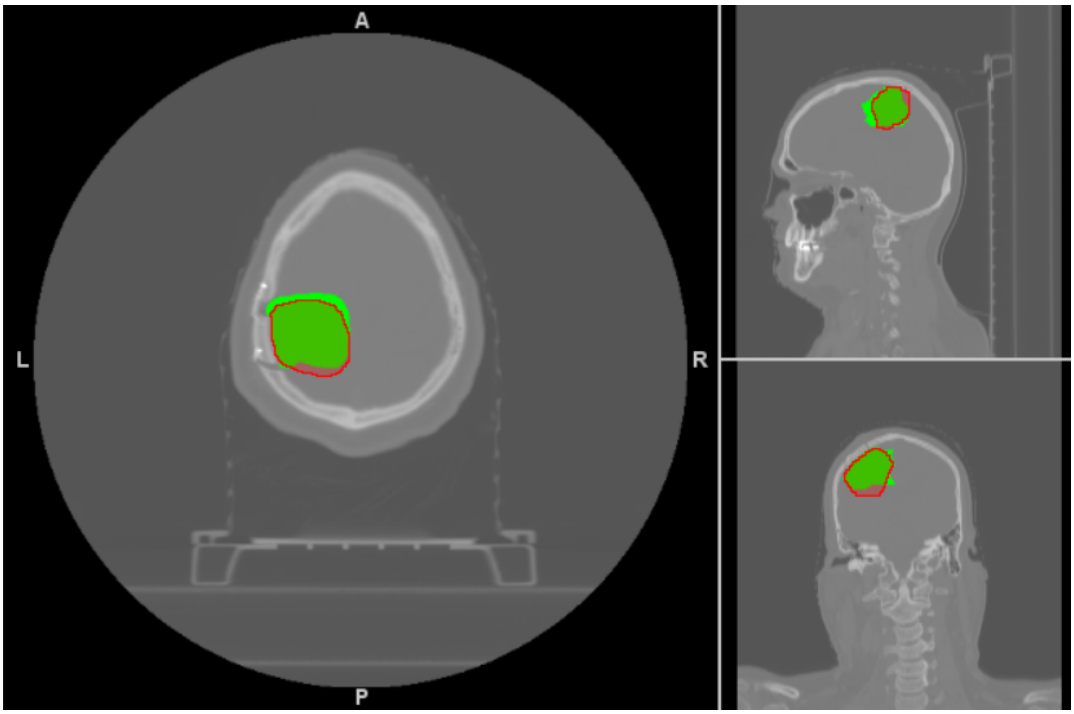


Figure 4.12: *CT of Patient_0022 overlaid with the manually segmented GTV (green) and the one segmented by the neural network (red).*

Regarding the uncertainty in dose values, this is linked to the error in Monte Carlo simulation. By simulating with 10^8 particles, it was possible to achieve an error in the GTV voxels even below 10% (in some instances, even below 1%). Considering the high number of voxels over which the dose measurement is averaged, the uncertainty is reduced to below 2%.

	Manual Segmentation	**AI Segmentation**
Minimum Dose [Gy]	23.3	23.1
Mean Dose [Gy]	28.0	28.4
Maximum Dose [Gy]	30.7	31.1

Table 4.4: *Photon isoeffective dose values delivered inside the GTV. The minimum dose refers to the dose value received by 98% of the volume, the mean corresponds to the dose received by 50%, and the maximum to 2% of the total volume. The uncertainty associated to these dose values is 2%.*

Regarding the calculation of TCP, the obtained range of values overlap with the case of manual contouring. By using the artificially segmented GTV, I obtained minimum and mean TCP values of respectively $TCP_{min} = (0.829 \pm 0.203)$ and $TCP_{mean} = (0.998 \pm 0.002)$. Therefore, even when contouring is performed automatically by the neural

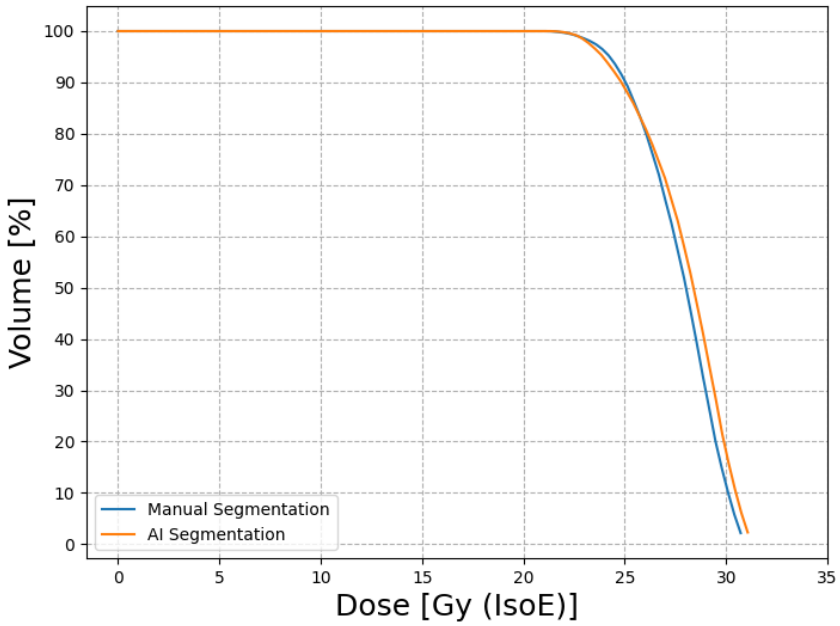


Figure 4.13: *Plot of DVHs for both manual segmentation and neural network segmentation. It reveals minor variations in the distributed dose values within the GTV volumes.*

network, the value of TCP is close to 1.

This entire procedure was then replicated for all 18 cases that I analyzed, each one having different values of DC, GMI, and DI, in order to extract information regarding a potential correlation between dose values and coefficients and to observe how the segmentation of GTV images affects the dose distribution within the examined patient. However, to give an overview of the situation encountered when the AI segmentation is not good, I describe below the analysis in Patient_0137.

Patient_0137

In this case, the manually contoured GTV and the AI-based GTV have significantly different dimensions, respectively 31.274 cm^3 and 113.013 cm^3. The Dice Coefficient value evaluated is lower, corresponding to DC=0.417, while GMI=0.038 and DI=0.734. Consequently, the correspondence of the two GTV segmentations is limited, as observed in Figure 4.14.

The low GMI value reflects the fact that the automatic segmentation almost entirely encompasses the manual one, resulting in $G \cap S \simeq G$, where G, the ground-truth volume represents the manually defined GTV and S the NN predicted one. Utilizing the

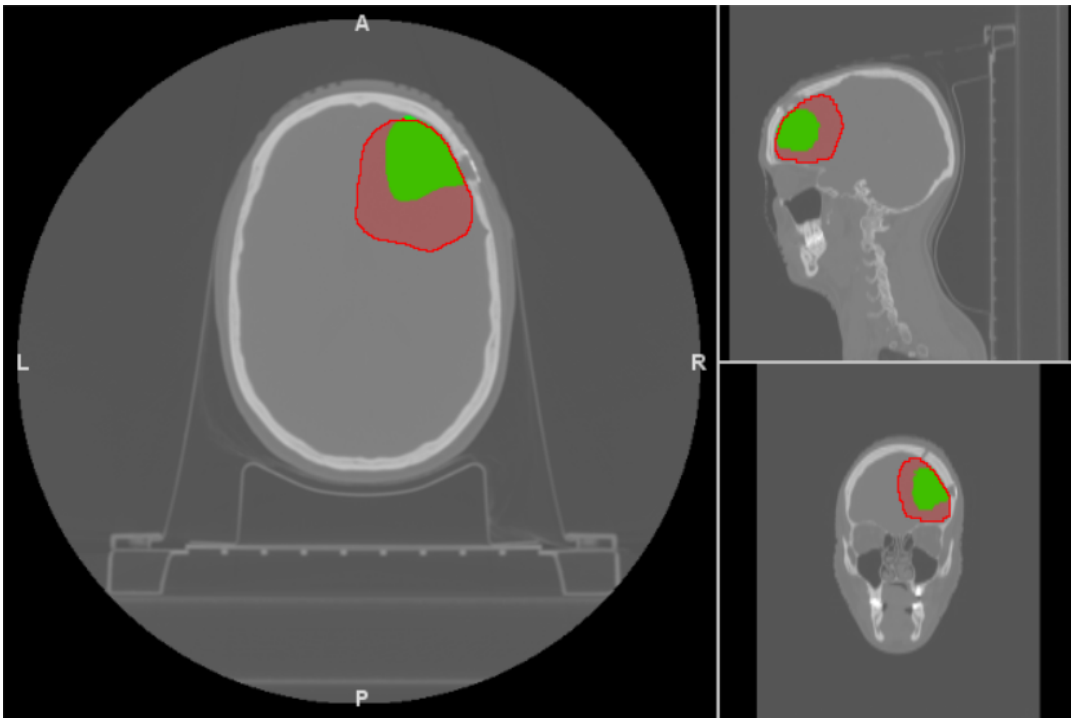


Figure 4.14: *CT of Patient_0137 overlaid with the manually segmented GTV (green) and the one segmented by the neural network (red).*

arithmetic definition of GMI, its calculated value tends to fall to zero, as

$$GMI = \frac{G - (G \cap S)}{G} \simeq \frac{G - G}{G} \simeq 0 \ .$$

On the other hand, the significant difference in volumes results in

$$S - (G \cap S) \simeq S - G >> 0 \ ,$$

leading to a high value of DI.

The dose distribution within the patient exhibits disparities in the two cases, as reported in Table 4.5, particularly concerning the minimum dose value. This is especially evident when observing the DVHs depicted in Figure 4.15.

This is because the artificially obtained GTV extends deeper, hence receiving a lower minimum dose value due to low-energy neutron beam penetration: the deeper the tumor, the more difficult is to ensure a suitable minimum dose.

As a result, TCP value represents a lower tumor control in the second case, as summarised in Table 4.6.

The reason why the neural network sometimes fails to segment the GTV is unknown, since, as illustrated in Chapter 3.2.2, what happens within the hidden layers of the algorithm cannot be observed or controlled by the user. Considerations on the performance of

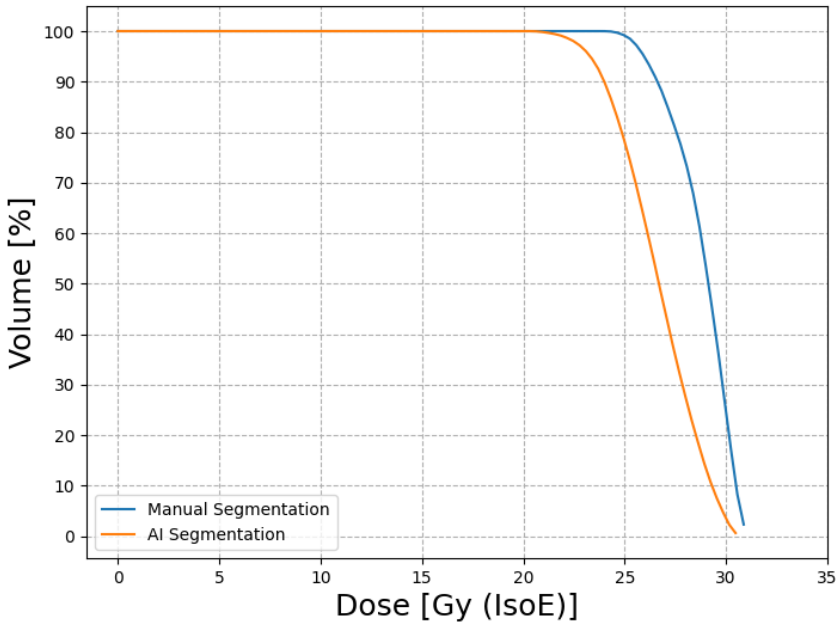


Figure 4.15: *Plot of DVHs for manual segmentation and neural network segmentation. Note the difference of the minimum dose values in the GTV volumes.*

	Manual Segmentation	**AI Segmentation**
Minimum Dose [Gy]	25.4	22.5
Mean Dose [Gy]	29.1	26.7
Maximum Dose [Gy]	30.9	30.2

Table 4.5: *Photon isoeffective dose values absorbed by the GTV. The minimum dose refers to the dose value received by 98% of the volume, the mean corresponds to the dose received by 50%, and the maximum to 2% of the total volume.*

the AI-based algorithm in segmenting the chosen cases are detailed below. ANN learns to contour the GTV region by minimizing a loss function. In this case Dice Similarity Index was used; in future work, results may be improved by introducing a new loss function.

	Manual Segmentation	**AI Segmentation**
$\mathbf{TCP}_{min}$	0.973 ± 0.035	0.734 ± 0.107
$\mathbf{TCP}_{mean}$	0.999 ± 0.001	0.992 ± 0.011

Table 4.6: *TCP values calculated considering the minimum dose absorbed by 98% of the GTV and the mean dose absorbed by 50% of the GTV for manual and automatic segmentations.*

4.6 Overall results

The procedure described so far was repeated on the remaining 16 cases selected from the test dataset. The dose and TCP values obtained for each of the 18 cases under consideration are compiled in Table 4.7.

For each case, nnU-Net processed the artificial segmentation of the GTV. Once this process has been successfully accomplished, I gathered the results and compared various parameters to obtain an in-depth analysis. It is worthwhile to assess the parameters while seeking potential correlations among them. To achieve this, I have implemented a ***correlation matrix***, as depicted in Figure 4.16, to show any existing correspondence between the evaluation parameters of the network (i.e., DC, GMI, and DI) and those characteristic of dosimetry obtained from the treatment plan.

4.6.1 Volume - Dice Correlation

A first consideration that can be drawn is that a clear correspondence exists between the volumes of manually segmented GTVs and the Dice Coefficient, as shown in Figure 4.17.

The majority of cases where the GTV has a volume above 100 cm^3 are concentrated in the region corresponding to a medium-high DC. Conversely, low DC values correspond to volumes below 50 cm^3. One possible explanation may lie in the training dataset used. In addition to being limited, since the small number of available GBM cases, it was probably composed mostly of large GTVs, making the network more efficient in recognizing large tumors.

A further observation was conducted on multi-volume GTVs. A notable case regards the patient from Taiwan (DC=0.118) shown in Figure 4.20. In this case, the network struggles to recognize and distinguish the separated volumes, but, instead, it creates a single, very large GTV that encompasses them all. Also this problem may be linked to the restricted training dataset. This issue leads to significant differences in the dose delivered in the tumor. There is a noticeable trend in the network to expand smaller volumes of GTV and narrow down larger ones, as depicted in the plot in Figure 4.18. The average volume value thus remains relatively constant and around 150 cm^3. Since I focused mainly on small tumors, the network generated GTVs with, on average, a larger volume compared to those manually contoured, as depicted in Figure 4.19.

Patient	$\mathbf{t_{irr}}$ **(min) True**	$\mathbf{t_{irr}}$ **(min) NN**	**D98% (Gy) True**	**D98% (Gy) NN**	**D50% (Gy) True**	**D50% (Gy) NN**	**D2% (Gy) True**	**D2% (Gy) NN**	**TCP min True**	**TCP min NN**	**TCP mean True**	**TCP mean NN**
Taiwan	33.06	34.90	17.7	14.8	27.8	23.8	30.7	31.1	0.000	0.000	0.997	0.902
$\mathbf{GBM_{0006}}$	31.63	33.99	23.4	15.8	28.3	24.4	30.8	31.4	0.862	0.000	0.998	0.940
$\mathbf{GBM_{0071}}$	29.60	29.58	18.3	18.6	20.7	21.0	23.0	23.4	0.001	0.004	0.293	0.379
$\mathbf{GBM_{0057}}$	32.06	34.77	21.7	20.9	26.6	26.2	29.0	30.2	0.567	0.341	0.991	0.987
$\mathbf{GBM_{0137}}$	39.20	38.30	25.4	22.5	29.1	26.7	30.9	30.2	0.973	0.734	0.999	0.992
$\mathbf{GBM_{0202}}$	30.92	31.67	27.1	24.0	29.3	28.7	30.5	30.7	0.994	0.911	0.999	0.999
$\mathbf{GBM_{0096}}$	32.58	35.35	11.5	9.0	19.0	18.2	26.7	28.3	0.000	0.000	0.016	0.001
$\mathbf{GBM_{0217}}$	36.98	36.35	13.3	16.5	21.8	23.5	30.0	29.8	0.000	0.000	0.576	0.876
$\mathbf{GBM_{0001}}$	36.42	35.65	20.2	18.8	27.4	26.0	31.1	30.7	0.167	0.008	0.996	0.985
$\mathbf{GBM_{0020}}$	36.09	36.09	25.1	23.4	29.0	28.4	31.3	31.3	0.966	0.861	0.999	0.998
$\mathbf{GBM_{0157}}$	39.93	39.81	19.9	17.6	26.6	26.0	30.9	30.8	0.111	0.000	0.991	0.984
$\mathbf{GBM_{0192}}$	30.71	30.79	24.0	25.5	27.6	27.9	29.4	29.4	0.912	0.975	0.997	0.997
$\mathbf{GBM_{0225}}$	33.83	33.86	15.5	15.5	22.8	24.1	30.5	30.7	0.000	0.000	0.784	0.920
$\mathbf{GBM_{0022}}$	37.30	38.05	23.3	23.1	28.0	28.4	30.7	31.1	0.855	0.829	0.998	0.998
$\mathbf{GBM_{0129}}$	35.58	35.39	20.5	20.5	26.4	26.4	30.8	30.7	0.241	0.251	0.989	0.989
$\mathbf{GBM_{0144}}$	33.50	33.94	22.1	21.9	26.8	26.6	29.7	29.9	0.662	0.611	0.993	0.991
$\mathbf{GBM_{0060}}$	31.37	31.37	20.9	22.3	26.6	26.5	29.8	29.7	0.346	0.693	0.991	0.990
$\mathbf{GBM_{0128}}$	36.96	36.96	14.4	14.6	22.7	23.0	30.6	30.6	0.000	0.000	0.765	0.807

Table 4.7: *Irradiation times, minimum, mean, and maximum dose values, and minimum and mean TCP values for the 18 analyzed patients. Values are referenced for both manually segmented GTVs and those artificially generated by the the neural network.*

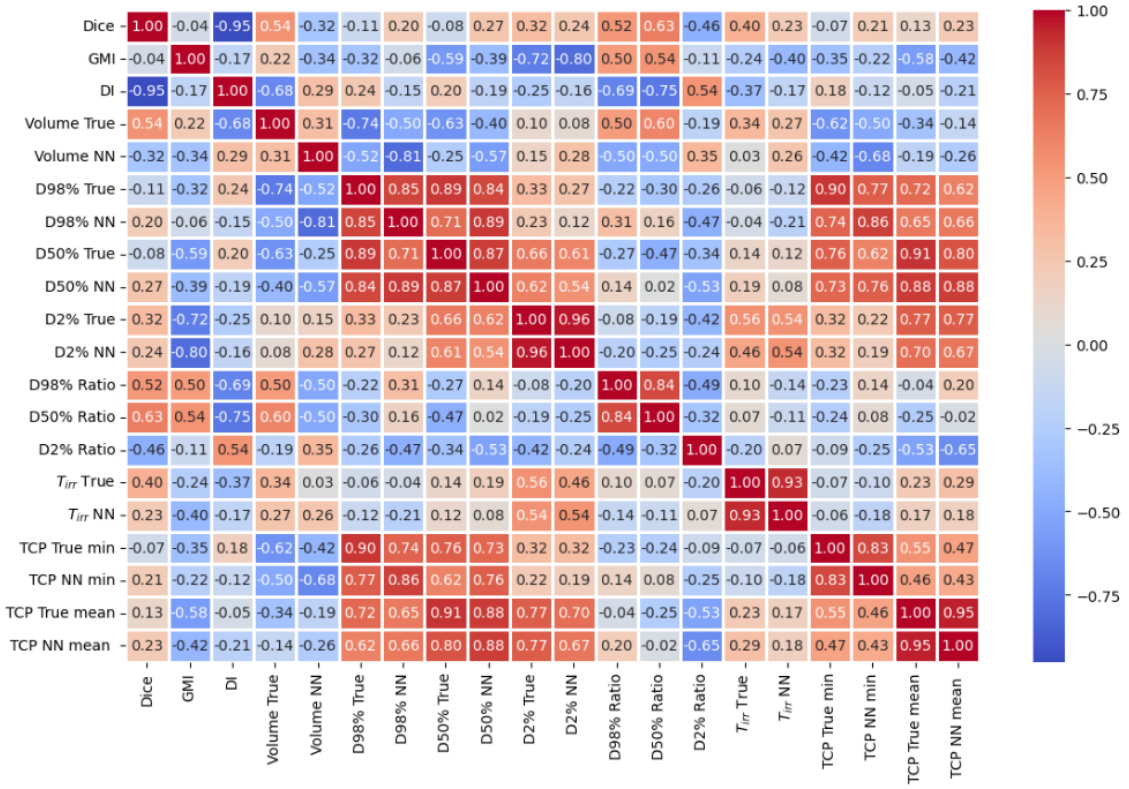


Figure 4.16: *Correlation matrix between all parameters obtained. The colours represent the correspondence value that lies in a range between 1 (maximum correlation) and -1 (maximum anticorrelation).*

4.6.2 Dose - Dice Correlation

An analysis was then conducted on how the absorbed dose distributions in patients change depending on whether the GTVs were manually segmented or by the neural network. Clearly, by altering the shape, size, and in certain cases, the position of the GTV, the OAR will also vary, consequently affecting irradiation times and doses delivered to the patient. Figure 4.21 shows a box-violin plot of the distribution of minimum, mean, and maximum doses in all the patients. A first consideration is that in AI-based GTVs the values of minimum and mean absorbed dose are lower. This tendency to underestimate the minimum and the mean doses may be a direct consequence of the different volume value, as shown in Figure 4.16, and a different tumor location predicted by the network. Moreover, the deeper the GTV develops, and therefore farther from the beam-port, the lower the minimum dose absorbed by 98% of the irradiated volume is. This difference tends to disappear regarding maximum dose values, which all remain slightly above 30 Gy(IsoE). The maximum dose absorbed by 2% of the GTV, in fact, is located closer to the patient's surface and is more closely related to the prescription, the irradiation time (see Figure 4.16) and dose limits, set equally for all patients under examination. Low correspondence between irradiated volumes and maximum doses are confirmed in

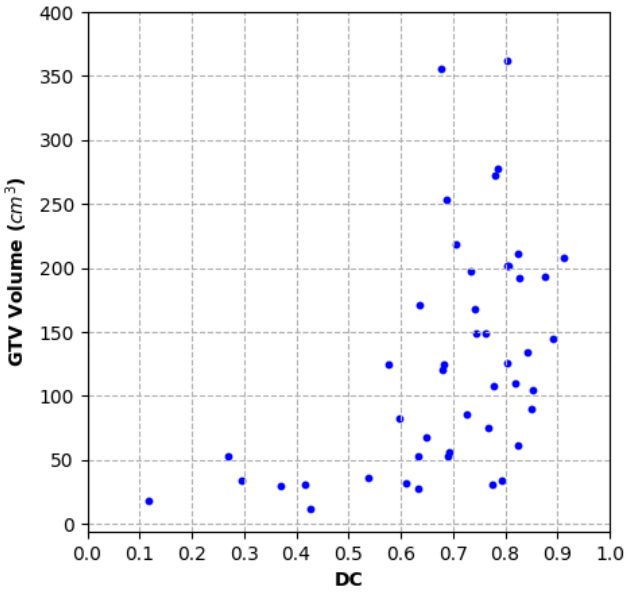


Figure 4.17: *Distribution of GTV volumes as a function of the Dice Coefficient.*

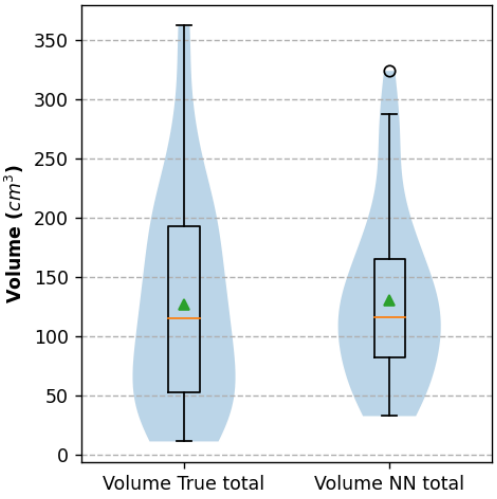


Figure 4.18: *Box-violin plot of the distribution of all the GTV volumes. The mean value is the green triangle and the median is graphically represented by the orange line.*

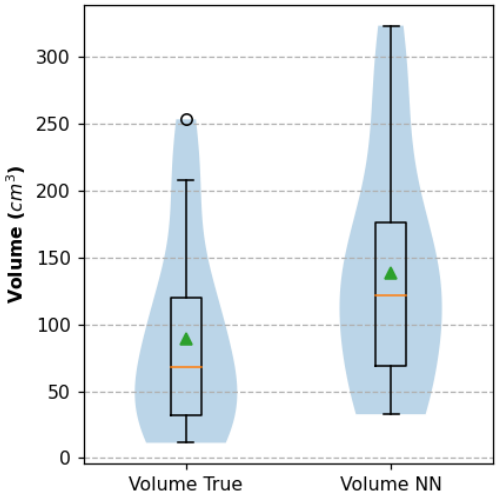


Figure 4.19: *Box-violin plot of the distribution of the GTV volumes analysed. The mean value is the green triangle and the median is graphically represented by the orange line.*

the correlation matrix.

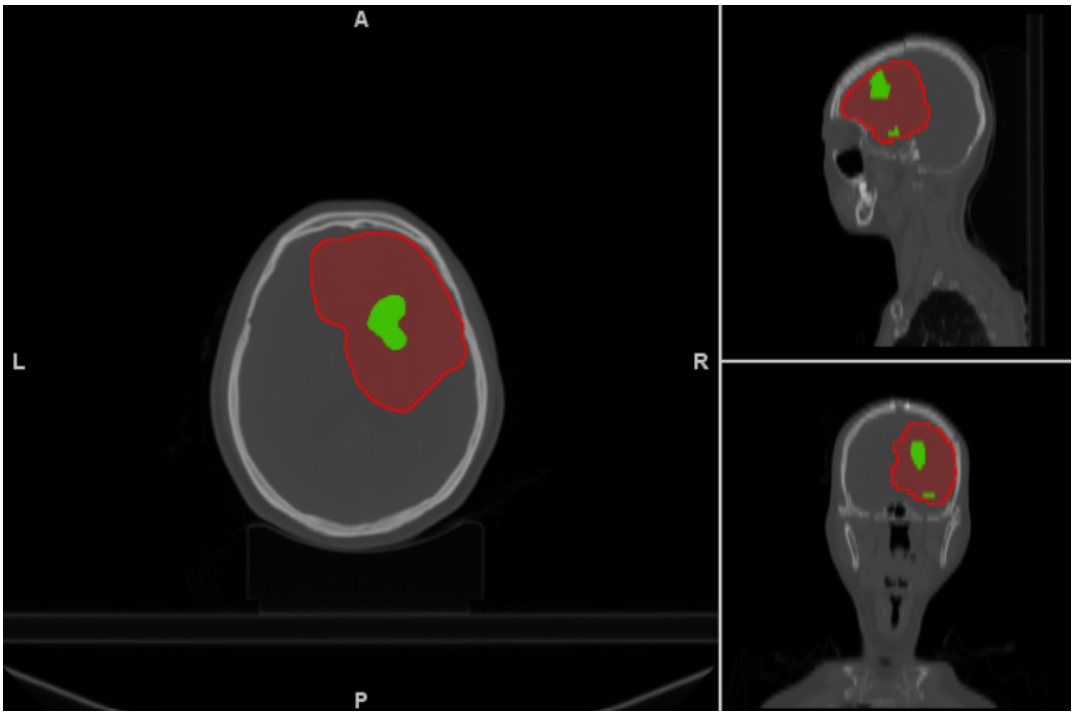


Figure 4.20: *CT of Taiwan Patient overlaid with the manually segmented GTV (green) and the one segmented by the neural network (red).*

It is interesting to observe how the difference in minimum doses varies depending on the evaluation parameters. The general trends are summarised in the 3D graph in Figure 4.22.

The cases where the ratio of minimum doses tends to 1 are mainly situated in the region of high Dice coefficient and low Geometrical Miss Index and Discordance Index. The reason is strictly related to the definition of these parameters. Higher DC as well as lower GMI and DI are attributable to a good correspondence and overlapping between the two GTVs, it is thus reasonable to assume that the distributions of doses in the two volumes are similar. As expected, a better overlap and a proper volume reconstruction lead to similar minimum doses. Focusing more on the relationship between the ratio of minimum doses and the Dice Coefficient, two different behaviors are observed in Figure 4.23 depending on the Dice value. For Dice coefficients greater than 0.7, dose values are the same within a 10% uncertainty. For lower Dice Coefficients, however, a random behavior is observed. Multiple assumptions can be made regarding this trend. Probably other factors come into play, including depth and volume. So, despite a poor overlap of the segmented volumes and despite a significant difference in GTV volumes, there is no significant disparity in minimum dose values. The fact that a similar dosimetry can be achieved in the two ROIs despite a low DC score suggests that while this coefficient may be a useful indicator in certain cases, it is sometimes insufficient for evaluating the performance of the neural network. Therefore, it is necessary to complement it with additional representative parameters to gain a more comprehensive understanding and to discern

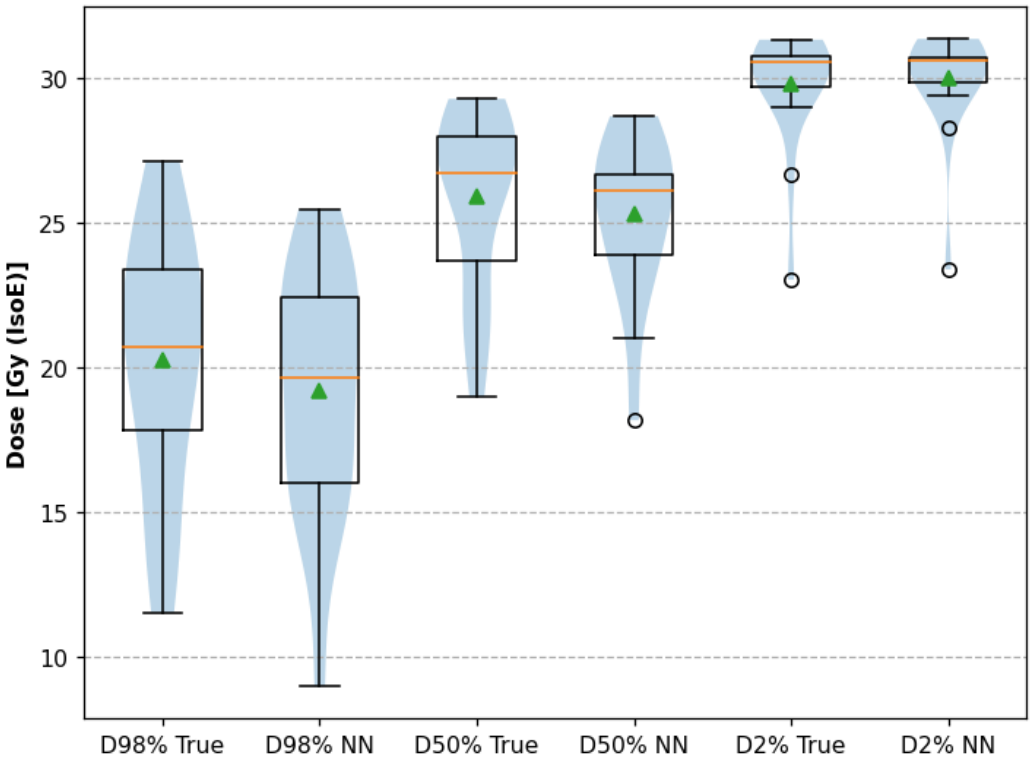


Figure 4.21: *Box-violin plot of the distribution of doses in the GVTs. From left to right, the values of minimum, mean, and maximum doses deposited in the GTVs are compared. The mean value is the green triangle and the median is graphically represented by the orange line.*

when the contouring obtained through the application of AI algorithms can be considered adequate.

4.6.3 TCP - Dice Correlation

An additional parameter worth analyzing in its correlation with the contouring performance is certainly the TCP. Minimum and mean dose values for each patient, and the corresponding artificial contouring, have been collected and incorporated into Equation 4.4, allowing the calculation of TCP values using the previously introduced model. TCP values of both manually and "artificially" segmented images are collected in the box-violin plot shown in Figure 4.24.

The TCP values calculated with the minimum dose associated with the GTV images produced by the network are on average significantly lower compared to those produced manually. Therefore, the dosimetry obtained in the GTVs generated by the network but optimised with the GTV manually contoured leads to lower probability of tumor control. One possible explanation can be sought by examining the correlation matrix in

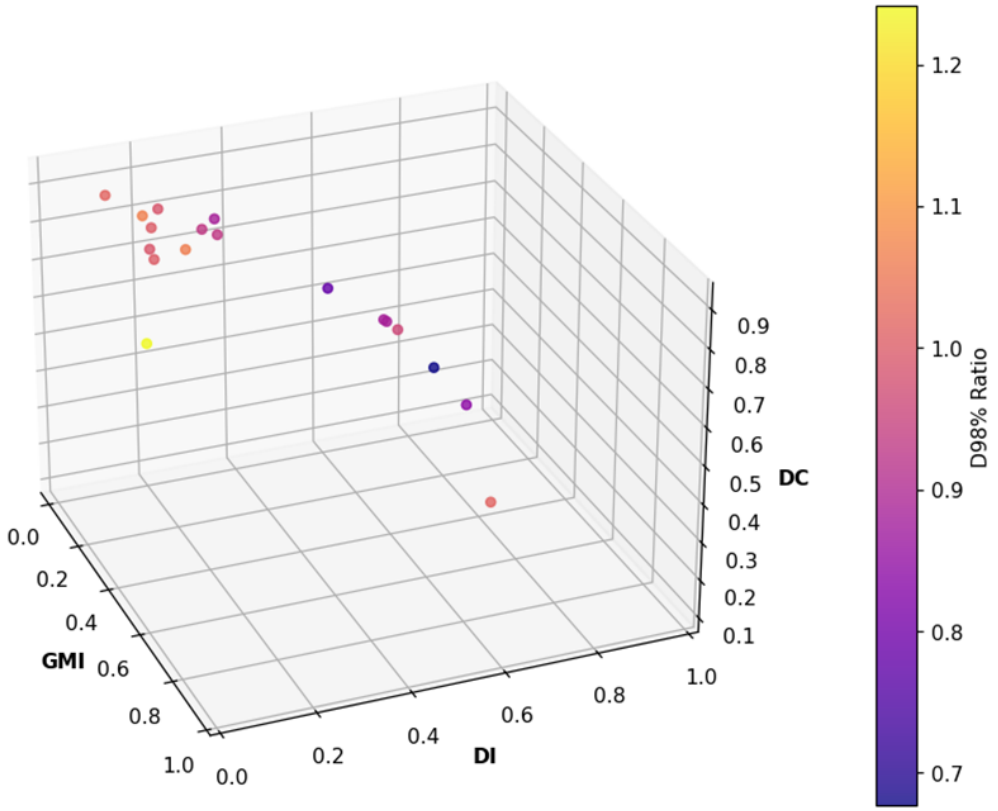


Figure 4.22: *3D representation of the cases under examination based on the parameters DC, GMI, and DI. The ratio is defined as the fraction between the minimum dose estimated on the GTV produced by the network and that manually obtained. The color of the dots is linked to the value of the minimum dose ratio, as indicated on the right.*

Figure 4.16. TCP_{min} is indeed highly correlated with both the volume of GTV (correlation index of -0.62) and, especially, with the value of the minimum dose (correlation index of 0.9), and this is what is expected. In fact, TCP is lower for big tumors and depends on the possibility to administer a sufficient dose to the whole GTV. Moreover, while the distribution of TCP_{min} shown in Figure 4.24 underlines a slightly lower average value, the median is markedly reduced, decreasing from 0.3 to a value ranging between 0.1 and 0.2.

The values of TCP_{mean} evaluated considering the mean doses are instead more similar and on average markedly high. In this case as well, the value of TCP is highly correlated with the absorbed dose, whereas it shows very little correlation with the volume of the GTV, unlike what occurs with TCP_{min}.

A further observation is the extremely low correlation between TCP and the Dice Coefficient, as evidenced by the correlation matrix in Figure 4.16. As a consequence, the

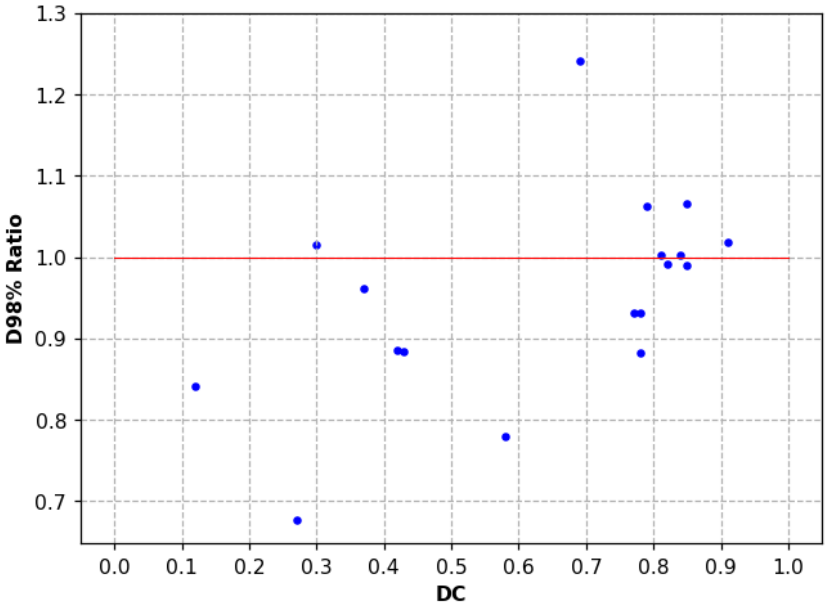


Figure 4.23: *Plot of the trend of the minimum dose ratio as a function of the Dice Coefficient.*

TCP may be used alongside the Dice Coefficient to enhance the performance of neural networks for medical image segmentation. Indeed, TCP and DC are two distinct indices, with different meanings and, importantly, uncorrelated and independent from each other. Therefore, incorporating both of them during the neural network training phase could lead to an enhancement in its performance.

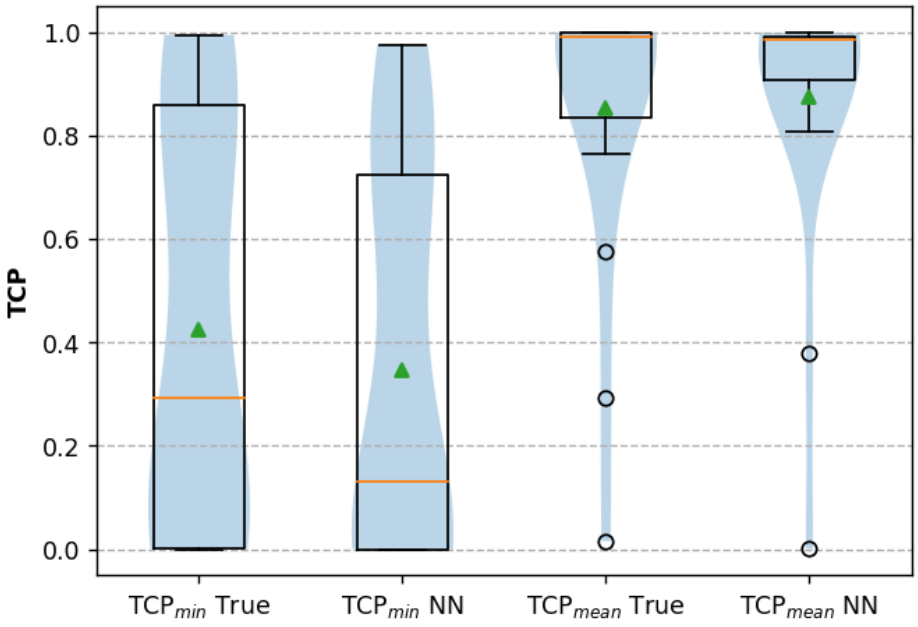


Figure 4.24: *Box-violin plot of the distribution of TCP_{min} and TCP_{mean} for both manually and artificially segmented GTVs. The mean value is the green triangle and the median is graphically represented by the orange line*

Conclusions and future work

The aim of this thesis was to assess the suitability of AI algorithms in contouring medical images for BNCT treatment of Glioblastoma Multiforme (GBM). The comparison between manually segmented CT images of GBM for dosimetry purposes and those segmented with the assistance of Artificial Intelligence provides valuable insights into the effectiveness and reliability of automated segmentation techniques in radiotherapy planning.

As manual segmentation provides more accuracy in the identification and the detection of the tumor volume, it allows for a much more precise treatment tailoring. However, this approach demands significant time and resources and it is subject to unavoidable variability. On the other hand, employing Artificial Intelligence methods for segmentation offers greater efficiency, since it reduces the human workload, but the outcome must be carefully tested. A preliminary evaluation of the reliability of such approach has been assessed in this study by observing how closely the automatic contouring matches the manual one using quantitative parameters such as DC, GMI, and DI, as well as conducting a dosimetric analysis. This work has revealed that the majority of studied cases showed good performance, as shown in the box-violin plot in Figure 4.21. In many instances, indeed, good overlapping contours and similar dose distributions were achieved. This behavior has been observed in those cases characterized by high Dice Coefficient. As shown in the plot in Figure 4.23, corresponding to high DC values, typically 0.7 and above, minimum dose values absorbed by the GTVs match within 10%.

This findings support the conclusion that the training of a neural network nnU-Net assuming the Dice Coefficient as the Loss Function can be efficient in achieving adequate segmented medical images. However, it is also observed that, for low Dice values, predicting correct dose distributions becomes more challenging. Clearly, additional factors come into play, such as, for example, the difference in depth-distribution of the GTVs. This is typical of BNCT, where the distribution of the total dose varies significantly as the neutron beam penetrates the patient's body.

Overall, some of these complications arise from the limited number of clinical cases used to train the nnU-Net. It is challenging for the neural network to generalize the obtained results, to take into account the complexity of the problem, and to achieve accurate outcomes with the number of cases analyzed in this work. However, even with these limitations, this study has demonstrated, for the first time, the possibility to apply Artificial

Intelligence algorithms for the BNCT treatment of patients affected by brain tumors. These artificial segmentations can indeed be incorporated into a treatment plan, and a corresponding dose distribution can be efficiently generated. This represents an initial approach and application of AI in BNCT, which has yielded encouraging results, upon which further research and development of increasingly efficient neural networks can be built.

Another novelty introduced in this work is the employ of the TCP as a method to evaluate the performance of an AI-based contouring. A strong correlation was found with the values of the minimum dose within the GTV, as well as the absence of correspondence with the DC values. For this reason, in the future, it may be considered to improve and optimize the training of the neural network not only by maximizing the Dice Coefficient, but also the TCP value. Clinical figures of merit linked to the outcome of the therapy could become a powerful tool to prepare the treatment planning itself. With the next developments dedicated to the optimization of TCP models for GBM, TCP can become a parameter, whose maximisation could establish the best irradiation position automatically.

Improvements in GTV contouring ANN could come from defining alternative loss functions, different from the dice similarity index. New loss functions may include more evaluation metrics, leading GTV prediction to ponder over multiple coefficients, increasing the ANN learning ability. Metrics that may be added to the loss function are GMI and DI. The former because, as observed in the correlation matrix in Figure 4.16, it is not correlated with DC. Therefore, incorporating GMI would provide additional information to the neural network, potentially benefiting its learning process. On the other hand, DI, being highly correlated with DC, would not provide any significant additional information to the ANN through the loss function. Hence, to enhance GTV segmentation, a function combining DC and GMI could be devised.

This study represents an initial approach in applying AI algorithms in the field of BNCT. The results obtained are encouraging and lead to further research in this sector. Indeed, one may wonder how to improve the network performance when applied to clinical cases. Firstly, as I have found differences in dose distributions within the GTV, it is difficult to ascertain whether these differences are clinically significant. In this work, I assumed the manually contoured structures provided as ground truth, as they were the only ones available. However, as stated in Chapter 3.1, manual contouring itself is subject to uncertainties and inter-observer variability. Contouring performed by a different radiologist would produce a different segmented GTV, resulting in a different dose distribution. It is thus interesting to compare the variability between different manual contours and the discrepancy observed between ground truth and AI.

Additionally, in this study, I analyzed the applicability of CNNs for tumor region contouring. However, other ROIs, particularly OARs, play a key role in BNCT treatment and therefore also require investigation. In this case, it would be useful to apply the Normal Tissue Complication probability (NTCP) models to the OARs, similarly to what was shown with TCP for GTVs. Currently, there is no established NTCP model for BNCT of GBM, but effeorts are being devoted to this end.

Finally, regarding the applicability of nnU-Net, the range of cases will be broadened by incorporating MRI or multimodal images. This approach would yield a more versatile ANN capable of recognizing a greater variety of different clinical images.

This field of development in BNCT represents an innovation that will yield positive effects both in research and clinical settings. In research, more patients can be studied and evaluated, and computational dosimetry can become more refined and robust, towards a better description of the dose distribution in the tissues. In clinics, a more efficient treatment planning procedure would enlarge the number of patients. This, in turn, would stimulate new research and a general improvement of the quality of the therapy for enhanced clinical applicability in the future.

Bibliography

[1] David Louis et al. “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A summary”. In: *Acta Neuropathologica* 131 (June 2016). DOI: 10.1007/s00401-016-1545-1.

[2] Shinji Kawabata et al. “Clinical Study on Modified Boron Neutron Capture Therapy for Newly Diagnosed Glioblastoma”. In: Nov. 2011. ISBN: 978-953-307-284-5. DOI: 10.5772/21989.

[3] Jeffrey A Coderre et al. “Boron neutron capture therapy: cellular targeting of high linear energy transfer radiation”. In: *Technology in cancer research & treatment* 2.5 (2003), pp. 355–375.

[4] Will Jin et al. “A Review of Boron Neutron Capture Therapy: Its History and Current Challenges”. In: *International journal of particle therapy* 9 (June 2022), pp. 71–82. DOI: 10.14338/IJPT-22-00002.1.

[5] Yi-Wei Chen et al. “Salvage Boron Neutron Capture Therapy for Malignant Brain Tumor Patients in Compliance with Emergency and Compassionate Use: Evaluation of 34 Cases in Taiwan”. In: *Biology* 10 (Apr. 2021), p. 334. DOI: 10.3390/biology10040334.

[6] Heikki Joensuu et al. “Boron neutron capture therapy of brain tumors: Clinical trials at the Finnish facility using boronophenylalanine”. In: *Journal of neuro-oncology* 62 (Mar. 2003), pp. 123–34. DOI: 10.1023/A:1023293006617.

[7] GL Locher. “Biological Effects and therapeutical possibilities of neutrons”. In: *Am. J. Roentgenol. Radium. Ther.* 36 (1936), pp. 1–13.

[8] A. G. Vladimirov et al. “Synthesis of composite Si-B nanoparticles by the laser-induced pyrolysis method”. In: *Laser Physics Letters* (2014).

[9] Ys. Bae et al. “Advances of LINAC-based boron neutron capture therapy in Korea”. In: *AAPPS Bull. 32, 34* (2022). URL: https://doi.org/10.1007/s43673-022-00063-2.

[10] S. P. Soyland C. Hassfjell. “Survival of Human Lung Epithelial Cells Following In Vitro Alpha-particle Irradiation with Absolute Determination of the Number of Alpha-particle Traversals of Individual Cells”. In: *International Journal of Radiation Biology* (1991).

[11] L. E. Farr et al. “Neutron capture therapy with boron in the treatment of glioblastoma multiforme”. In: *The American Journal of Roentgenology Radium Therapy and Nuclear Medicine* (1954), pp. 279–293.

[12] R.F. Barth et al. “Boron Neutron Capture Therapy of cancer: current status and future prospects”. In: *Clin Cancer Res* 11 (2005), pp. 3987–4002.

[13] Paola Porcari, Silvia Capuani, and Francesco Saverio Pastore. “Novel Pharmacological and Magnetic Resonance Strategies to Enhance Boron Neutron Capture Therapy (BNCT) Efficacy in the Clinical Treatment of Malignant Glioma”. In: *Management of CNS Tumors* (2011). URL: https://api.semanticscholar.org/CorpusID:43939686.

[14] Y Mishima et al. “New Thermal Neutron Capture Therapy for Malignant Melanoma: Melanogenesis-Seeking 10B Molecule-Melanoma Cell Interaction From In Vitro to First Clinical Trial”. In: *Pigment Cell Research* 2.4 (1989), pp. 226–234.

[15] K. Yoshino et al. “Capture of p-boronophenylalanine in malignant melanoma cells by complex formation with melanin monomers, DOPA, DHI and DHICA. BPA trapping mechanism.” In: *Advances in Neutron Capture Therapy* (1997), pp. 234–238.

[16] J. A. Coderre et al. “The Therapeutic Ratio in BNCT, Assessment using the Rat 9L Gliosarcoma Brain Tumor and Spinal Cord Models”. In: *Y.Mishima ed, Cancer Neutron Capture Therapy* (1996), pp. 757–762.

[17] JA Coderre et al. “Boron neutron capture therapy of glioblastoma multiforme using the p-borophenylalanine-fructose complex and epithermal neutrons: tryal design and early clinical results”. In: *J. Neuro-Oncol.* 33 (1997), pp. 141–151.

[18] “Report on the first patient group of the phase I BNCT trial at the LVR-15 reactor”. In: *Sauerwein W, Moss R, Wittig A, editors. Research and Development in Neutron Capture Therapy, Bologna, Monduzzi Ed.* (2002), pp. 1107–1112.

[19] *Current Status of Neutron Capture Therapy*. TECDOC Series 1223. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2001. URL: https://www.iaea.org/publications/6168/current-status-of-neutron-capture-therapy.

[20] D.A. Brown et al. “ENDF/B-VIII.0: The 8th Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data”. In: *Nuclear Data Sheets* 148 (2018). Special Issue on Nuclear Reaction Data, pp. 1–142. ISSN: 0090-3752. DOI: https://doi.org/10.1016/j.nds.2018.02.001. URL: https://www.sciencedirect.com/science/article/pii/S0090375218300206.

[21] *Advances in Boron Neutron Capture Therapy*. TECDOC Series. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2023.

[22] O Harling and KJ Riley. “Fission reactor neutron sources for neutron capture therapy — a critical review.” In: *J Neuro-Oncol* 62 (2003), pp. 7–17.

[23] Andres J Kreiner. “Accelerator-based BNCT”. In: *Neutron Capture Therapy* (2012), pp. 41–54.

[24] Andres Juan Kreiner et al. “Present status of accelerator-based BNCT”. In: *Reports of Practical Oncology & Radiotherapy* 21.2 (2016), pp. 95–101.

[25] H Kumada, T Sakae, and H Sakurai. “Current development status of accelerator-based neutron source for boron neutron capture therapy.” In: *EPJ Techn Intrum* 10.18 (2023).

[26] S. Miyatake et al. “Boron Neutron Capture Therapy for Malignant Brain Tumors”. In: *Neurol Med Chir (Tokyo)* (2016), pp. 361–371.

[27] Ian Postuma et al. “A novel approach to design and evaluate BNCT neutron beams combining physical, radiobiological, and dosimetric figures of merit”. In: *Biology* 10.3 (2021), p. 174.

[28] Shinji Kawabata et al. “Accelerator-based BNCT for patients with recurrent glioblastoma: a multicenter phase II study”. In: *Neuro-oncology advances* 3.1 (2021), vdab067.

[29] Silva Bortolussi. “BORON NEUTRON CAPTURE THERAPY: a technological journey to bring neutrons in hospital”. In: *Il Nuovo Saggiatore* 39.3-4 (2023), pp. 66–72.

[30] M Comunian, A Pisent, and G.V. Lamanna. “TRASCO RFQ design”. In: *Proceedings of EPAC 2000, Vienna, Austria*. 2000.

[31] J Esposito et al. “Be target development for the accelerator-based SPES-BNCT facility at INFN Legnaro”. In: *Applied Radiation and Isotopes* 67.7-8 (2009), S270–S273.

[32] AA.VV. “ICRU Report 44: Tissue Substitutes in Radiation Dosimetry and Measurement”. In: *Journal of the ICRU* (1989).

[33] H. Kumada and K. Taka. “Treatment planning system and patient positioning for boron neutron capture therapy”. In: *Therapeutic Radiology and Oncology* (2018).

[34] D. A. Brown et al. “ENDF/B-VIII.0: The 8^{th} Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data”. In: *Nuclear Data Sheets* (2018).

[35] K.J. Riley, P.J. Binns, and O.K. Harling. “A state-of-the-art epithermal neutron irradiation facility for neutron capture therapy”. In: *Physics in Medicine and Biology* 49 (2004), pp. 3725–3735.

[36] J.A. Coderre and A.Z. Diaz. “The radiobiology of boron neutron capture therapy: Are "photonequivalent" doses really photon-equivalent?” In: *Current status of neutron capture therapy* (2001).

[37] Edward L. Alpen. *Radiation Biophysics (Second Edition)*. Academic Press, 1998.

[38] J.W. Hopewell et al. “The radiobiological principles of boron neutron capture therapy: A critical review”. In: *Applied Radiation and Isotopes* (2011), pp. 1756–1759.

[39] G.M. Morris et al. “Response of rat skin to boron neutron capture therapy with p-boronophenylalanine or borocaptate sodium”. In: *Radiotherapy and Oncology* (1994), pp. 144–153.

[40] M.A. Davis et al. “Relative biological effectiveness of the ^{10}B(n, α) ^{7}Li reaction in HeLa cells”. In: *Journal of Radiation Research* (1970).

[41] Giuliano Franco Perotti Bernardini et al. “Comparison of Photon Isoeffective Dose Models Based on In Vitro and In Vivo Radiobiological Experiments for Head and Neck Cancer Treated with BNCT”. In: *Radiation Research* 198.2 (2022). Cited by: 0, pp. 134–144. DOI: 10.1667/RADE-21-00234.1. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135449146&doi=10.1667%2fRADE-21-00234.1&partnerID=40&md5=1f99d2281462154edb42da0bc79fdabe.

[42] Ana Mailén Dattoli Viegas et al. “Detailed dosimetry calculation for in-vitro experiments and its impact on clinical BNCT”. In: *Physica Medica* 89 (2021). Cited by: 4, pp. 282–292. DOI: 10.1016/j.ejmp.2021.08.010. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114042359&doi=10.1016%2fj.ejmp.2021.08.010&partnerID=40&md5=ca96b979c99844ddea25326691277808.

[43] Sara J. González and Gustavo A. Santa Cruz. “The photon-isoeffective dose in boron neutron capture therapy”. In: *Radiation research* (2012).

[44] A.M. Kellerer and H.H. Rossi. “A generalized formulation of dual radiation action”. In: *Radiation Research 75* (1978), pp. 471–488.

[45] D. E. Lea and D. G. Catcheside. “The mechanism of the induction by radiation of chromosome aberrations inTradescantia”. In: *Journal of Genetics, Volume 44* (1942), pp. 216–245.

[46] D.J. Brenner et al. “The Linear-Quadratic Model and Most Other Common Radiobiological Models Result in Similar Predictions of Time-Dose Relationships”. In: *Radiation Research* 150 (1998), pp. 83–91.

[47] M. Zaider and H.H. Rossi. “The Synergistic Effects of Different Radiations”. In: *Radiation Research* 83 (1980), pp. 732–739.

[48] R. Railton, R.C. Lawson, and D. Porter. “Interaction of gamma-ray and neutron effects on the proliferative capacity of Chinese hamster cells”. In: *International Journal of Radiation Biology and related Studies in Physics, Chemistry, and Medicine* (1975), pp. 75–82.

[49] R.E. Durand and P.L. Olive. “Irradiation of Multi-cell Spheroids with Fast Neutrons Versus X-rays: A Qualitative Difference in Sub-lethal Damage Repair Capacity or Kinetics”. In: *International Journal of Radiation Biology and related Studies in Physics, Chemistry, and Medicine* (1976), pp. 589–592.

[50] S. Hornsey, U. Andreozzi, and P.R. Warren. “Sublethal damage in cells of the mouse gut after mixed treatment with X rays and fast neutrons”. In: *The British journal of radiology* (1977), pp. 513–5127.

[51] Shozo Suzuki. “A theoretical model for simultaneous mixed irradiation with multiple types of radiation”. In: *Journal of Radiation Research* 39 (1998), pp. 215–221.

[52] J.A. Coderre et al. “Derivations of relative biological effectiveness for the high-let radiations produced during boron neutron capture irradiations of the 9L rat gliosarcoma in vitro and in vivo”. In: *International Journal of Radiation Oncology Biology Physics* (1993), pp. 1121–1129.

[53] S.J. González et al. “Tumor control and normal tissue complications in BNCT treatment of nodular melanoma: A search for predictive quantities”. In: *Applied Radiation and Isotopes 67* (2009), S153–S156.

[54] *Specification and Acceptance Testing of Radiotherapy Treatment Planning Systems*. TECDOC Series 1540. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2007.

[55] E.B. Podgorsak. *Radiation Oncology Physics: A Handbook for Teachers and Students*. IAEA, Vienna, 2005. Chap. 11.

[56] P. Mayles, A. Nahum, and J.C. Rosenwald. *Handbook of radiotherapy physics: theory and practice*. CRC Press, 2007.

[57] D. Jones. "ICRU Report 50-Prescribing, Recording, and Reporting Photon Beam Therapy". In: *Journal of the ICRU* 21 (1994).

[58] A. Wambersie. "ICRU Report 62, Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU 50)". In: *Journal of the ICRU* 32 (1999).

[59] B. Emami et al. "Tolerance of normal tissue to therapeutic irradiation". In: *International Journal of Radiation Oncology, Biology, Physics* 21 (1991), pp. 109–122.

[60] Søren M Bentzen et al. "Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): an introduction to the scientific issues". In: *International Journal of Radiation Oncology, Biology, Physics* 76.3 (2010), S3–S9.

[61] W.A. Sauerwein et al. *Neutron capture therapy: principles and applications*. Springer Science & Business Media, 2012. Chap. 16.

[62] David W. David W. Nigg et al. "Computational dosimetry and treatment planning for Boron Neutron Capture Therapy". In: *Journal of Neuro-Oncology* 33 (1997), pp. 93–104.

[63] David W. Nigg. "Computational dosimetry and treatment planning for Neutron Capture Therapy". In: *Journal of Neuro-Oncology* 62 (2003), pp. 75–86.

[64] V.A. Nievaart, G.G. Dacquino, and R.L. Moss. "Monte Carlo based treatment planning systems for Boron Neutron Capture Therapy in Petten, The Netherlands". In: *Journal of Physics: Conference Series 74* (2007).

[65] N. Fukumitsu and Y. Matsumoto. "Development of an Imaging Technique for Boron Neutron Capture Therapy". In: *Cells* (2021).

[66] Sara J. González et al. "Voxel model in BNCT treatment planning: performance analysis and improvements". In: *Physics in Medicine & Biology* 50 (2005), pp. 441–458.

[67] R. Zamenhof et al. "Monte Carlo-based treatment planning for boron neutron capture therapy using custom designed models automatically generated from CT data". In: *International Journal of Radiation Oncology Biology Physics* 35 (1996), pp. 383–397.

[68] "Verification of nuclear data for the Tsukuba plan, a newly developed treatment planning system for boron neutron capture therapy". In: *Applied Radiation and Isotopes* 106 (2015), pp. 111–115. ISSN: 0969-8043. DOI: https://doi.org/10.1016/j.apradiso.2015.08.032.

[69] Jiang Chen et al. "Development of Monte Carlo based treatment planning system for BNCT". In: *Journal of Physics: Conference Series* 2313 (July 2022), p. 012012. DOI: 10.1088/1742-6596/2313/1/012012.

[70] R.O. Farias and S.J. González. "MultiCell model as an optimized strategy for BNCT treatment planning". In: (2012), pp. 10–14.

[71] David Nigg et al. "SERA – An advanced treatment planning system for neutron therapy and BNCT". In: *Transactions of the American Nuclear Society* 80 (Sept. 1999).

[72] W.F. Verbakel et al. “Towards in vivo monitoring of neutron distributions for quality control of BNCT”. In: *Phys Med Biol* 47 (2002), pp. 1059–1072.

[73] Nigg D.W. “Methods for radiation dose distribution analysis and treatment planning in boron neutron capture therapy”. In: *International Journal of Radiation Oncology Biology Physics* 28 (1994), pp. 1121–1134.

[74] AA.VV. “ICRU 46: Photon, electron, proton, and neutron interaction data for body tissues”. In: *Journal of the ICRU* (1992).

[75] C. Wojnecki, A. Wittig, and E. Bourhis-Martin. “Patient geometry modeling for treatment”. In: *Research and development in neutron capture therapy* (2002). Ed. by Bologna Monduzzi Editore.

[76] R.A. Brooks, G. Di Chiro, and M.R. Keller. “Explanation of cerebral white-gray contrast in computed tomography”. In: *Journal of Computer Assisted Tomography* 4 (1980), pp. 489–491.

[77] J.T. Goorley, W.S. Kiger, and R.G. Zamenhof. “Reference dosimetry calculations for neutron capture therapy with comparison of analytical and voxel models”. In: *Medical Physics* 29(2) (2002), pp. 145–156.

[78] J.R. Albritton. “Computational aspects of treatment planning for neutron capture therapy”. PhD thesis. Nuclear Science and Engineering Department, Massachusetts Institute of Technology, Cambridge, 2009.

[79] C. J. Werner (editor). “MCNP Users Manual - Code Version 6.2”. In: *Los Alamos National Laboratory, Los Alamos, NM, report LA-UR-17-29981* (2017).

[80] Tatsuhiko Sato et al. “Features of Particle and Heavy Ion Transport code System (PHITS) version 3.02”. In: *Journal of Nuclear Science and Technology* 55 (Jan. 2018), pp. 1–7. DOI: 10.1080/00223131.2017.1419890.

[81] Stefano Agostinelli et al. “GEANT4-a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment* 506 (July 2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[82] M.R. Palmer et al. “Treatment planning and dosimetry for the Harvard-MIT phase I clinical trial of cranial neutron capture therapy”. In: *International Journal of Radiation Oncology Biology Physics* 53 (2002), pp. 1361–1379.

[83] J.T. Goorley, W.S. Kiger, and R.G. Zamenhof. “Reference dosimetry calculations for neutron capture therapy with comparison of analytical and voxel models”. In: *Medical Physics* (2002).

[84] Silva Bortolussi et al. “Understanding the potentiality of accelerator based-boron neutron capture therapy for osteosarcoma: Dosimetry assessment based on the reported clinical experience”. In: *Radiation Oncology* 12.1 (2017). DOI: 10.1186/s13014-017-0860-6.

[85] Rubén O. Farías et al. “Exploring Boron Neutron Capture Therapy for non-small cell lung cancer”. In: *Physica Medica* 30.8 (2014), pp. 888–897. DOI: 10.1016/j.ejmp.2014.07.342.

[86] INTERNATIONAL ATOMIC ENERGY AGENCY VIENNA. *Advances in Boron Neutron Capture Therapy*. IAEA publications, 2023.

[87] Rolf Barth et al. “Current status of boron neutron capture therapy of high grade gliomas and recurrent head and neck cancer”. In: *Radiation oncology (London, England)* 7 (Aug. 2012), p. 146. DOI: 10.1186/1748-717X-7-146.

[88] S. Nievaart et al. “Use of linear programming to obtain an optimum, multi-beam treatment plan in NCT”. In: *Advances in neutron capture therapy. International Society for Neutron Capture Therapy* (2006).

[89] Leena Kankaanranta et al. “Boron neutron capture therapy in the treatment of locally recurred head-and-neck cancer: final analysis of a phase I/II trial”. In: *International Journal of Radiation Oncology* Biology* Physics* 82.1 (2012), e67–e75.

[90] W.S. Kiger et al. “Preliminary treatment planning and dosimetry for a clinical trial of neutron capture therapy using a fi ssion converter epithermal neutron beam”. In: *Applied Radiation and Isotopes* 61 (2004), pp. 1075–1081.

[91] R. Henriksson et al. “Boron neutron capture therapy (BNCT) for glioblastoma multiforme: a phase II study evaluating a prolonged high-dose of boronophenylalanine (BPA)”. In: *Radiotherapy and Oncology* 88 (2008), pp. 183–191.

[92] .L Wielopolski et al. “Patient positioning in static beams for boron neutron capture therapy of malignant glioma”. In: *Radiation medicine* 18 (Nov. 2000), pp. 381–387.

[93] H. Kumada et al. “Development of the patient setting system for BNCT at JRR-4”. In: *Proceedings of the ninth international symposium on neutron capture therapy for cancer, Osaka* (2000).

[94] L. Wielopolski et al. “Patient positioning in static beams for boron neutron capture therapy of malignant glioma”. In: *Journal of Radiation Medicine* 18 (2000), pp. 381–387.

[95] W.S. Kiger et al. “Pharamacokinetic Modeling for Boronophenylalanine-fructose Mediated Neutron Capture Therapy: 10B Concentration Predictions and Dosimetric Consequences”. In: *Journal of Neuro-Oncology* 62 (2003), pp. 171–186.

[96] W.S. Kiger et al. “A pharmacokinetic model for the concentration of 10 B in blood after boronophenylalanine-fructose administration in humans”. In: *Radiant Research* 155 (2001), pp. 611–618.

[97] Y. Shibata et al. “Prediction of boron concentrations in blood from patients on boron neutron capture therapy”. In: *Anticancer Research* 23 (2003), pp. 5231–5235.

[98] Kiichi Ishiwata et al. “Synthesis and radiation dosimetry of 4-borono-2-[18F]fluoro-d,l-phenylalanine: A target compound for PET and boron neutron capture therapy”. In: *International Journal of Radiation Applications and Instrumentation. Part A. Applied Radiation and Isotopes* 42.4 (1991), pp. 325–328. ISSN: 0883-2889.

[99] Trent Nichols et al. “Improved treatment planning for boron neutron capture therapy for glioblastoma multiforme using fluorine-18 labeled boronophenylalanine and positron emission tomography”. In: *Medical physics* 29 (Nov. 2002), pp. 2351–8. DOI: 10.1118/1.1507780.

[100] Young-soon Bae et al. “Advances of LINAC-based boron neutron capture therapy in Korea”. In: *AAPPS Bulletin* 32 (Oct. 2022). DOI: 10.1007/s43673-022-00063-2.

[101] L. Provenzano et al. "The essential role of radiobiological figures of merit for the assessment and comparison of beam performances in boron neutron capture therapy". In: *Physica Medica* 67 (2019). Cited by: 8, pp. 9–19. DOI: 10.1016/j.ejmp.2019.09.235. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073030518&doi=10.1016%2fj.ejmp.2019.09.235&partnerID=40&md5=6d66ef73f42b350aaea3879bc8592cbb.

[102] Andreas Kaplan and Michael Haenlein. "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence". In: *Business Horizons* 62 (2019), pp. 15–25.

[103] Anca Diaconu et al. "Expending the power of artificial intelligence in preclinical research: an overview". In: *IOP Conference Series: Materials Science and Engineering* 1254 (Sept. 2022), p. 012036. DOI: 10.1088/1757-899X/1254/1/012036.

[104] Alina Küper et al. "Is There a Role of Artificial Intelligence in Preclinical Imaging?" In: *Seminars in Nuclear Medicine* 53 (Apr. 2023). DOI: 10.1053/j.semnuclmed.2023.03.003.

[105] Wenya Linda Bi et al. "Artificial intelligence in cancer imaging: Clinical challenges and applications". In: *CA: A Cancer Journal for Clinicians* 69.2 (2019), pp. 127–157.

[106] Castellino R.A. "Computer aided detection (CAD): an overview". In: *Cancer Imaging* 5 (2005).

[107] Anne Meesters et al. "What Are the Interobserver and Intraobserver Variability of Gap and Stepoff Measurements in Acetabular Fractures?" In: *Clinical Orthopaedics & Related Research* Publish Ahead of Print (July 2020). DOI: 10.1097/CORR.0000000000001398.

[108] Jordan Wong et al. "Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning". In: *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 144 (Dec. 2019), pp. 152–158. DOI: 10.1016/j.radonc.2019.10.019.

[109] Hui Lin et al. "Deep learning for automatic target volume segmentation in radiation therapy: a review". In: *Quantitative Imaging in Medicine and Surgery* 11 (Dec. 2021), pp. 4847–4858. DOI: 10.21037/qims-21-168.

[110] K. Harrison et al. "Machine Learning for Auto-Segmentation in Radiotherapy Planning". In: *Clinical Oncology* 34 (2022), pp. 74–88.

[111] David Teguh et al. "Clinical Validation of Atlas-Based Auto-Segmentation of Multiple Target Volumes and Normal Tissue (Swallowing/Mastication) Structures in the Head and Neck". In: *International journal of radiation oncology, biology, physics* 81 (Oct. 2010), pp. 950–7. DOI: 10.1016/j.ijrobp.2010.07.009.

[112] Michael Sherer et al. "Metrics to evaluate the performance of auto-segmentation for radiation treatment planning: A critical review". In: *Radiotherapy and Oncology* 160 (May 2021). DOI: 10.1016/j.radonc.2021.05.003.

[113] Frank Rosenblatt. "The Perceptron: A Probabilistic Model for Information Storage and Organization (1958)". In: Feb. 2021, pp. 183–190. ISBN: 9780262363174. DOI: 10.7551/mitpress/12274.003.0020.

[114] M. Elgendy. *Deep Learning for Vision Systems*. Manning Publications, 2020. ISBN: 9781617296192.

[115] Christoph von der Malsburg. "Frank Rosenblatt: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms". In: *Brain Theory* (Jan. 1986), pp. 245–248. DOI: 10.1007/978-3-642-70911-1_20.

[116] Aston Zhang et al. *Dive into Deep Learning*. June 2021.

[117] Amir Ali and Muhammad Amin. *Hands-On Machine Learning with Scikit-Learn*. Mar. 2019. ISBN: 1090626797.

[118] Jeremy Watt, Reza Borhani, and Aggelos Katsaggelos. *Machine Learning Refined: Foundations, Algorithms, and Applications*. Jan. 2020. ISBN: 9781108480727. DOI: 10.1017/9781108690935.

[119] Eustace Dogo et al. "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks". In: (Dec. 2018). DOI: 10.1109/CTEMS.2018.8769211.

[120] Moritz Hardt, Tengyu Ma, and Benjamin Recht. "Gradient Descent Learns Linear Dynamical Systems". In: *Journal of Machine Learning Research* 19 (Sept. 2016).

[121] Sebastian Ruder. "An overview of gradient descent optimization algorithms". In: (Sept. 2016).

[122] Keiron O'Shea and Ryan Nash. "An Introduction to Convolutional Neural Networks". In: *ArXiv e-prints* (Nov. 2015).

[123] Saad Albawi, Tareq Abed Mohammed, and Saad ALZAWI. "Understanding of a Convolutional Neural Network". In: (Aug. 2017). DOI: 10.1109/ICEngTechnol.2017.8308186.

[124] Kunihiko Fukushima and Sei Miyake. "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position". In: *Pattern Recognition* 15.6 (1982), pp. 455–469. ISSN: 0031-3203. DOI: https://doi.org/10.1016/0031-3203(82)90024-3.

[125] Jonathan Long, Evan Shelhamer, and Trevor Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *Arxiv* 79 (Nov. 2014).

[126] Yabo Fu et al. "A review of deep learning based methods for medical image multi-organ segmentation". In: *Physica Medica* 85 (May 2021), pp. 107–122. DOI: 10.1016/j.ejmp.2021.05.003.

[127] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *LNCS* 9351 (Oct. 2015), pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28.

[128] Chao Zuo et al. "Deep learning in optical metrology: a review". In: *Light: Science & Applications* 11 (Feb. 2022), p. 39. DOI: 10.1038/s41377-022-00714-x.

[129] Özgün Çiçek et al. "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation". In: (Oct. 2016), pp. 424–432. DOI: 10.1007/978-3-319-46723-8_49.

[130] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation". In: (June 2016).

[131] Fabian Isensee et al. “nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation”. In: (Sept. 2018).

[132] Fabian Isensee et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. In: *Nature Methods* 18 (Feb. 2021), pp. 1–9. DOI: 10.1038/s41592-020-01008-z.

[133] Shuqing Chen et al. “Towards Automatic Abdominal Multi-Organ Segmentation in Dual Energy CT using Cascaded 3D Fully Convolutional Network”. In: (Oct. 2017).

[134] Andrew L. Maas. “Rectifier Nonlinearities Improve Neural Network Acoustic Models”. In: *Proceedings of the International Conference on Machine Learning* 3 (2013).

[135] Jun Ma et al. “Loss odyssey in medical image segmentation”. In: *Medical Image Analysis* 71 (2021), p. 102035. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2021.102035.

[136] V.W.J. Pelt et al. “Evaluation of delineating the target volume by radiation therapists in breast cancer patients”. In: *Technical Innovations & Patient Support in Radiation Oncology* 17 (Mar. 2021), pp. 78–81. DOI: 10.1016/j.tipsro.2021.01.001.

[137] Christina T. Muijs et al. “Consequences of additional use of PET information for target volume delineation and radiotherapy dose distribution for esophageal cancer”. In: *Radiotherapy and Oncology* 93.3 (2009), pp. 447–453. ISSN: 0167-8140. DOI: https://doi.org/10.1016/j.radonc.2009.08.030.

[138] Lucyna Kepka et al. “Delineation variation of lymph node stations for treatment planning lung cancer radiotherapy”. In: *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 85 (Jan. 2008), pp. 450–5. DOI: 10.1016/j.radonc.2007.10.028.

[139] Charles Harris et al. “Array programming with NumPy”. In: *Nature* 585 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[140] Chiara Magni et al. “Design of a BNCT irradiation room based on proton accelerator and beryllium target”. In: *Applied Radiation and Isotopes* 165 (July 2020), p. 109314. DOI: 10.1016/j.apradiso.2020.109314.

[141] Chiara Magni. “Experimental and computational studies for an Accelerator-Based Boron Neutron Capture Therapy clinical facility: a multidisciplinary approach - PhD thesis - University of Pavia”. PhD thesis. 2022.

[142] Chiara Magni et al. “Radiation protection aspects in the design of a Boron Neutron Capture Therapy irradiation room”. In: *Radiation Physics and Chemistry* 218 (2024), p. 111621. ISSN: 0969-806X. DOI: https://doi.org/10.1016/j.radphyschem.2024.111621.

[143] Jeffrey A. Coderre and Gerard M. Morris. “The Radiation Biology of Boron Neutron Capture Therapy”. In: *Radiation Research Society* 151 (1999), pp. 1–18.

[144] SJ González et al. “Photon iso-effective dose for cancer treatment with mixed field radiation based on dose–response assessment from human and an animal model: clinical application to boron neutron capture therapy for head and neck cancer”. In: *Physics in Medicine & Biology* 62.20 (2017), p. 7938.

[145] Ian Postuma et al. "Using the photon isoeffective dose formalism to compare and combine BNCT and CIRT in a head and neck tumour". In: *Scientific Reports* 14.1 (2024). Cited by: 0; All Open Access, Gold Open Access, Green Open Access. DOI: 10.1038/s41598-023-50522-5. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181255062&doi=10.1038%2fs41598-023-50522-5&partnerID=40&md5=556fc92bb9df93c0d69482ea588c61c3.

[146] Piernicola Pedicini et al. "Clinical radiobiology of glioblastoma multiforme Estimation of tumor control probability from various radiotherapy fractionation schemes". In: *Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft ... [et al]* 190 (Apr. 2014). DOI: 10.1007/s00066-014-0638-9.

Acknowledgments

Desidero esprimere un sincero ringraziamento a tutti coloro che hanno contribuito a rendere possibile questo straordinario viaggio accademico, accompagnandomi attraverso gli anni in un percorso che mi ha vista trasformarmi da semplice studentessa del liceo fino a diventare una persona sempre più consapevole di sé.

Innanzitutto, desidero rivolgere un caloroso ringraziamento a Silva, Ian, Seta e Barbara, senza i quali non avrei mai potuto raggiungere questo traguardo. La loro guida, il sostegno costante e la loro infinita pazienza nel corso di questi mesi sono stati fondamentali per me, dandomi la forza necessaria per affrontare questa sfida. Grazie per avermi accolta, ascoltata e per aver sempre creduto in me, anche nei momenti più difficili, quando io stessa dubitavo delle mie capacità.

Un ringraziamento speciale va ai miei genitori e a mio fratello Gigi, per il loro amore incondizionato e il costante supporto durante questi anni di università.

Devo anche ringraziare tutte le persone che sono entrate nella mia vita a Pavia. Grazie ai miei *amici dell'auletta*, con i quali ho condiviso i momenti più belli tra le mura del dipartimento. Grazie per le risate, i pranzi insieme e per aver reso più piacevoli le giornate universitarie, regalandomi momenti di leggerezza e serenità tra lezioni, esami e sessioni di studio matto e disperato. Un ringraziamento particolare va alla mia coinquilina Eleonora, per avermi sempre fatta sentire a casa e per avermi ridato il sorriso dopo lunghe giornate di studio. E non posso dimenticare Beatrice, con cui ho instaurato una forte amicizia, tra passeggiate in montagna e appuntamenti per bere il tè nel suo ufficio.

Un sentito ringraziamento alle mie amiche di sempre, Chiara e Valeria, che mi conoscono da una vita e continuano a sostenermi nonostante le poche occasioni di vederci e di stare insieme. Grazie per essere sempre state un punto di riferimento per me.

Questo lavoro di tesi segna la conclusione di un capitolo molto importante della mia vita, fatto di gioie, difficoltà e incertezze. Tuttavia, desidero che sia anche l'inizio di un nuovo percorso, fatto di continue scoperte e sfide, senza mai perdere la curiosità che mi ha spinta fin qui e senza dimenticare tutto ciò che ho costruito.