Modelling of Brainstem Toxicity Including Variable Relative Biological Effectiveness in Paediatric Proton Therapy

Andreas Havsgård Handeland

Master's Thesis in Medical Physics

Department of Physics and Technology
University of Bergen
June 2021

Supervisors

Camilla Hanquist Stokkevåg
Department of Oncology and Medical Physics, Haukeland University Hospital /
Department of Physics and Technology, University of Bergen

Helge Egil Seime Pettersen
Department of Oncology and Medical Physics, Haukeland University Hospital

Acknowledgements

First, I would like to express my deepest gratitude to my supervisors Associate Professor Camilla Hanquist Stokkevåg and PhD Helge Egil Seime Pettersen. Thank you for continuously inspiring me with new and exciting ideas during this project, for providing superb feedback and for spending so much time on me and this project. Thank you, Camilla, for introducing me to the project and for helping me understand all the clinical aspects of the study, thus allowing me to understand the articles necessitated for this thesis. Thank you, Helge, for helping me along with all the programming performed for this thesis, for helping me with statistical and modelling approaches and for looking into every error message I could not fix by googling.

Further, I would like to thank PhD candidate (although not a candidate for long) Lars Fredrik Fjæra for performing the simulations necessary for this thesis, for explaining their implementation to me, and for insightful discussions on subjects related to both our projects. I also wish to thank MD Daniel J. Indelicato for providing the patient cohort utilised in this thesis and MD, PhD Yasmin Lassen-Ramshad for delineating the substructures that were so central for this thesis. Additionally, I wish to thank the ones mentioned above, and PhD Kristian S. Ytre-Hauge and Professor Ludvig P. Muren, for constructive feedback on my abstract submitted to NACP symposium 2021.

I would also like to thank all the colleagues I have worked with and the friends I have made at the department during my five years of study. Even this last year has not been too bad because of you. Thank you, Emilie Haugland Solheim, for spending so many of your hours assisting me on this project; I am eternally grateful. Thank you, Odin Alvestad, for teaching me topics our Master's projects shared. Thank you, Amalie Øie Hovland, Minh Chi To and everyone else who I also really wish I could mention by name.

Finally, I wish to thank my family for unconditionally supporting me all these years, despite me not visiting as often as I might have liked. Thank you to my mum, Mona Havsgård, my dad, Ståle Grude-Handeland, my siblings and everyone else.

Andreas Havsgård Handeland

Bergen, June 2021

Abstract

Brainstem necrosis is a rare but severe side-effect following paediatric proton therapy. Substructures of the brainstem may be associated with regional differences in radiosensitivity, but these are not accounted for clinically. The relative biological effectiveness (RBE), the ratio between a test dose and reference dose resulting in the same clinical endpoint, is also assumed to be constant for proton therapy. This may underestimate the biological effect of the radiation since the RBE is thought to be variable across the beam profile. Current dose constraints and normal tissue complication probability (NTCP) models for adult tissues are further developed than ones based on paediatric patients. However, paediatric tissue is associated with different radiosensitivity than adult tissues, and more data is required to quantify this. This study aimed to further explore the association between variable RBE, regional radiosensitivity of the brainstem and brainstem toxicity of paediatric proton therapy patients.

A cohort of 36 paediatric proton therapy patients that received significant dose to the brainstem, and were subsequently at risk of brainstem necrosis, were included in a case-control study. The patients had RBE-weighted dose distributions and dose-averaged linear energy transfer (LET_d) distributions recalculated with the FLUKA Monte Carlo code for variable RBE models. The brainstem was delineated into substructures. Dose-volume histograms and dose statistics of the cohort were used to fit Lyman-Kutcher-Burman (LKB) models to the data for different RBE-weighted dose distributions and substructures. Dose statistics were also used as a basis for cluster analyses to explore regional differences across the brainstem.

The results showed higher average variable RBE-weighted dose and LET_d observed for cases compared to controls, while this was not the trend for the constant RBE factor. This thesis shows the first fitting of LKB models to substructures of the brainstem. For the full brainstem structure, the tolerance dose (TD50) range was $61.7 - 68.6$ Gy(RBE) using RBE1.1 and $65.4 - 70.0$ Gy(RBE) based on the variable RBE models. The cluster analysis separated the data points into a small number of relatively solid clusters but overall did not show clear trends in sorting out cases from controls.

Contents

List of Figures

List of Tables

List of Abbreviations

3-D CRT Three-dimensional Conformal Radiotherapy

ALARA As Low as Reasonably Achievable

AUROC Area Under the Receiver Operating Characteristic

CI Confidence Interval

CNS Central Nervous System

CT Computed Tomography

CTCAE Common Terminology Criteria of Adverse Effects

CTV Clinical Target Volume

DBSCAN Density-Based Spatial Clustering of Applications with noise

DNA Deoxyribonucleic Acid

DSB Double-Strand Break

DVH Dose-Volume Histogram

gEUD Generalised Equivalent Uniform Dose

GTV Gross Tumour Volume

HU Hounsfield Unit

ICRU International Commission on Radiation Units and Measurement

IMPT Intensity Modulated Proton Therapy

IMRT Intensity Modulated Radiotherapy

LET Linear Energy Transfer

$\mathbf{LET_d}$ Dose-averaged Linear Energy Transfer

$\mathbf{LET_\Delta}$ Restricted Linear Energy Transfer

$\mathbf{LET_\infty}$ Unrestricted Linear Energy Transfer

$\mathbf{LET_t}$ Track-averaged Linear Energy Transfer

Linac Linear Accelerator

LKB Model Lyman-Kutcher-Burman Model

LQ Model Linear Quadratic Model

MC Monte Carlo

MCN McNamara Variable RBE Model

ML Machine Learning

MLC Multi-Leaf Collimator

MRI Magnetic Resonance Imaging

NMR Nuclear Magnetic Resonance

NTCP Normal Tissue Complication Probability

OAR Organ at Risk

PTV Planning Target Volume

RBE Relative Biological Effectiveness

RBE1.1 Relative Biological Effectiveness Factor 1.1

ROC Receiver Operating Characteristic

RORW Rørvik Weighted Variable RBE Model

SOBP Spread-Out Bragg Peak

SSB Single-Strand Break

TCP Tumour Control Probability

TD50 Tolerance Dose to 50% of the Sample

TRIPOD Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

TPS Treatment Planning System

UF University of Florida

VMAT Volumetric Modulated Arc therapy

Chapter 1

Introduction

Cancer is a collective term for several conditions all attributed to the uncontrolled, abnormal growth of cells [1]. 191 children under the age of 17 were diagnosed with cancer in Norway in 2019 and central nervous system (CNS) tumours were the most common form of cancer diagnosed [2]. Norwegian paediatric CNS tumours are associated with a five-year survival rate of over 80%, with paediatric cancer, in general, seeing a five-year survival rate of 86% [2]. Increasing survival rates are seen in most developed countries [3]. However, cancer survivors are at high risk of developing health problems later in life, largely attributed to the therapy they have undergone [4]. Increased focus is, thus, directed to quality of life of patients post-treatment and the mitigation of radiation-related side effects [5].

Paediatric CNS tumours are now commonly treated with proton therapy since this leads to increased sparing of healthy tissue and could potentially decrease the rate of both acute and late toxicity. However, there is a lack of clinical evidence, particularly for the latter, which emphasises the importance of further study [6, 7, 8]. The goal of radiation therapy as a form of cancer treatment is to maximise the dose to the tumour while minimising the dose to healthy tissue, thus achieving tumour control while avoiding adverse effects in healthy tissue [9]. For photons, this is achieved by irradiating from several angles with high energy X-ray beams. Treatment with photons has been the standard since radiation therapy was first used in 1896 [10, 11]. However, proton therapy is an alternative to the photon therapy which comes with certain key advantages [12]. The damage from radiation to biological tissue is described by the energy deposition from the radiation to the traversed mass, referred to as the dose. Protons can achieve significantly lower entry doses than photons, and the point of highest dose is deeper in the tissue and more localised. The point of maximum dose for a proton beam is referred to as the Bragg peak. The goal of proton therapy is for the position of the Bragg peak to fully cover the planning target volume (PTV), which is the volume to be treated and encompasses the cancerous tissue and error margins [9, 13]. This achieves increased sparing of healthy tissue while still achieving the target dose to the tumour. A spread-out Bragg peak (SOBP) is typically

utilised to ensure sufficient tumour coverage. Here, a beam configuration consisting of several different energies is used, resulting in several overlapping Bragg peaks that combine to a more extended range of consistent high dose deposition [12].

A disadvantage with proton therapy is the relative novelty of the technology compared to X-ray treatment technology and the long experience we have in treating with X-rays [10, 14]. A standard method of translating this experience to proton therapy is utilising the relative biological effectiveness (RBE) of a given modality, which relates the dose from the modality to the dose from photons. For proton therapy, this factor is generally set to 1.1 relative to photons [15] based on *in vivo* studies in the early days of proton therapy [16]. However, a constant factor of 1.1 is an oversimplified approach since *in vitro* studies have shown that not only is the RBE variable across the beam path, the maximum RBE in the distal end of the SOBP approaches a factor of 1.7 [16]. The variability of the RBE has been observed to correlate with other radiation and tissue properties, increasing with increasing linear energy transfer (LET) of the radiation and decreasing with physical dose [16] and $(\alpha/\beta)_{\mathrm{x}}$ of the tissue. The $(\alpha/\beta)_{\mathrm{x}}$-ratio of a tissue describes the response of the tissue to a reference radiation (i.e. photons), with high ratios being associated with tissues that respond early to radiation, while low ratios correspond to tissues that react later to radiation. This translates to higher predicted RBE for lower $(\alpha/\beta)_{\mathrm{x}}$, but uncertainties in experimental values of the tissues makes this trend difficult to quantify [16]. The variable RBE and LET might be especially critical for paediatric proton therapy patients since their tissue is associated with low $(\alpha/\beta)_{\mathrm{x}}$-ratios [17, 18], treatment is delivered in several fractions with low dose per fraction (≤ 2 Gy(RBE)) [19] and SOBPs with lesser modulation widths are associated with higher LET [16].

To account for variations in RBE, several phenomenological RBE models have been proposed [20, 21, 22], which are empirical formulas that approximate experimental *in vitro* results of RBE weighted dose. The LET is a particularly central parameter in describing RBE since the increase in RBE in the distal end of the SOBP is to a high degree attributed to the increase in LET [16, 21]. Additionally, most models account for increasing RBE with decreasing physical dose and $(\alpha/\beta)_{\mathrm{x}}$ [20]. Recalculating RBE with variable RBE models typically leads to higher average RBE-weighted doses and considerably altered dose distributions compared to the constant factor [20]. These are shown to better describe the RBE from experimental *in vitro* results [20, 16]. Currently, variable RBE is only accounted for clinically by avoiding the position of the SOBP falloff coinciding with vital organs close to the tumour site [23]. However, studies are increasingly investigat-

ing LET distributions and how variable RBE affects rates of toxicity, and while studies of brainstem necrosis have been inconclusive so far [24, 25], lesser degrees of brainstem injury have been associated with areas of increased LET and RBE [26, 27, 28].

An unfortunate side-effect of radiotherapy as cancer treatment is the associated risk of adverse effects to healthy tissue, referred to as toxicity. Toxicity spans a wide range of severities and can be both strictly asymptotic or potentially lethal. Limiting the risk of toxicity for vital radiosensitive organs close to the tumour volume is particularly concerning during radiotherapy treatment. These organs are referred to as organs at risk (OAR) [13] and are typically subject to strict dose constraints in the treatment planning phase to minimise the risk of complication [12].

In the treatment of CNS tumours, the brainstem is a critical OAR since the brainstem is responsible for vital body functions such as breathing [29], swallowing and regulating blood pressure [30]. The brainstem may also be associated with higher radiosensitivity than what is currently accounted for clinically [17, 31]. Brainstem dose constraints are not universal [32] and, apart from avoiding hot spots of excessive dose and the delineation of an outer 3 mm margin with slightly stricter constraints for the core [33], the further substructures of the brainstem are treated equivalently. However, studies have concluded that the brainstem's radiosensitivity is likely to be regionally variant due to the neuroanatomy of brainstem substructures [17]. Additionally, differences in fibre tract dose responses have been observed and given grounds for a further transversal delineation of the pons [34, 35], which could warrant refining the current dose constraints.

Dose constraints are upper limits of dose an OAR can be subject to from a given treatment plan in radiotherapy before the risk of toxicity becomes unacceptable. The constraints are based on experience from previous treatment and clinical studies [9, 32]. An alternative approach to utilising collected clinical data to guide future therapy is to construct normal tissue complication probability (NTCP) models. NTCP models are based on a cohort of patients where the prescribed treatment has resulted in a specific adverse clinical endpoint to an OAR [12]. The NTCP models are, thus, organ- and endpoint-specific. The model takes dosimetric data from the cohort as input, which is converted to a probability for the treatment plan of a given patient to result in the same clinical endpoint. NTCP models can then be utilised on future treatment schemes to associate the prescribed dose distribution with a probability of toxicity. The models can be used to guide treatment and to provide a second opinion on a specific dose plan. NTCP models have already been

utilised clinically for a time [36]. However, with increased amounts of data to build the models, increased hardware power and a heightened focus on machine learning (ML) and modelling in radiotherapy, NTCP models are likely to become an even more important aspect of modern treatment planning [37].

Cancer rates are observed to increase with age [38] and as a consequence NTCP models based on adult patients are typically further developed than peadiatric models [5, 39]. Paediatric tissue is associated with different radiosensitivity and clinical endpoints compared to adult patients, which need to be accounted for in their treatment [5, 39]. Hence, establishing new models based on paediatric data is crucial since paediatric patients are expected to live long lives following their treatment, and toxicity severely compromises the quality of life of these patients [39].

1.1 Objectives

The overall aim of this thesis was to investigate brainstem necrosis related to paediatric proton therapy in the interest of limiting incidence. A secondary goal was to explore if the parameters incorporated in phenomenological RBE models would yield significant changes in dose distributions compared to the constant RBE factor of 1.1 and if the constant RBE assumption could be underestimating the dose to the brainstem. A further aim was to study the regional differences in radiosensitivity across brainstem substructures to establish if certain volumes required prioritising over others. Additionally, constructing NTCP models for the brainstem and its substructures with brainstem necrosis as clinical endpoint based on RBE-weighted dose to brainstem substructures was another important objective of the thesis. Finally, the ability of cluster analysis to discover trends in the data not apparent through manual study of the cohort was another point of interest.

Chapter 2

Background

This chapter will cover the medical physics background important for understanding the method and results presented in this thesis. The interactions of photons and charged particles provide the basis of radiotherapy and will, thus, be recounted first (section 2.1). The description of radiation qualities, both physical and biological, will then be discussed (section 2.2), before moving on to how radiation interacts with and damages biological tissue, known as radiobiology (section 2.3). The subsequent section is dedicated to the principles behind radiotherapy, both with photons and protons (section 2.4). Further, the relative biological effectiveness is outlined, which is a means of quantifying the relative effect of two radiotherapy modalities (section 2.5). The treatment planning process that precedes the delivery of radiotherapy is the focus of the next section (section 2.6). Statistical methods useful for processing and presenting clinical data follows, which are necessitated in order to report results objectively (section 2.7). The processed clinical data can then be used to guide future treatment through incorporations in predictive clinical models, which are explained in the following section (section 2.8). Machine learning algorithms are increasingly being utilised within radiotherapy and will subsequently be covered (section 2.9), while the final section consists of a brief explanation of brainstem necrosis (section 2.10).

2.1 Radiation Interactions with Matter

The use of radiation in treating cancer is explained mainly by the mechanisms behind the energy deposition of ionising radiation. Thus, this section will describe the interactions of ionising particles with matter, both photons and heavier charged particles.

2.1.1 Photon Interactions with Matter

Photons interact with matter through several processes, but the most important contributors in the attenuation of a photon beam are the photoelectric effect, Compton scattering

and pair production, which are the processes dominant at energies relevant to radiation therapy [40, 41]. However, the range of energies is also characterised by different interactions, with the photoelectric effect dominating lower energies, Compton scattering dominating at intermediate energies, while pair production becomes important to account for at higher energies. The interaction processes reduce the intensity of the photon beam traversing matter, with radiation intensity at a point given as

$$I(x) = I_0 \exp(-\mu x) \tag{2.1}$$

where I_0 is the initial radiation intensity, μ is the linear attenuation coefficient of the matter the photon beam is passing through, and x is the penetration depth of the material. The linear attenuation coefficients for the different interaction processes outlined earlier vary depending on the incident photon energy and the electron density (Z) of the absorbing material. Hence, the linear attenuation coefficient describes what process contributes most significantly to the intensity loss of the beam given the current conditions [40, 41]. The relation between the linear attenuation coefficient and photon energy of the three processes is shown in figure 2.1.

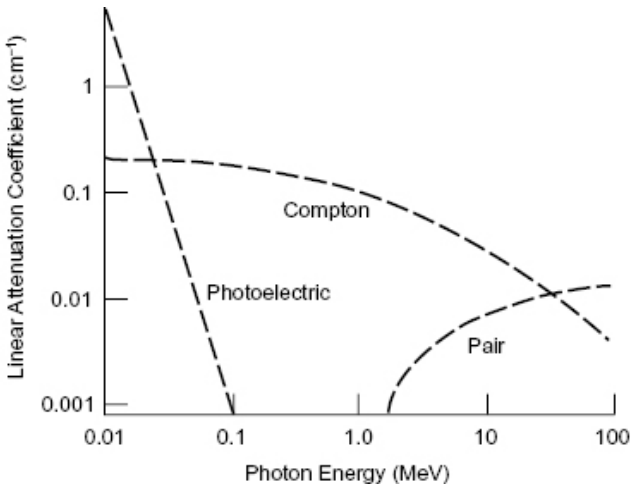

Figure 2.1: The dependence of linear attenuation coefficient on photon energy for photoelectric effect, Compton scattering and pair production. The figure is retrieved from [42].

The Photoelectric Effect

First described by Albert Einstein in 1905, the photoelectric effect entails a photon's absorption by an orbital electron and the subsequent ejection of the electron from its orbit. A schematic outlining the photoelectric effect is shown in figure 2.2.

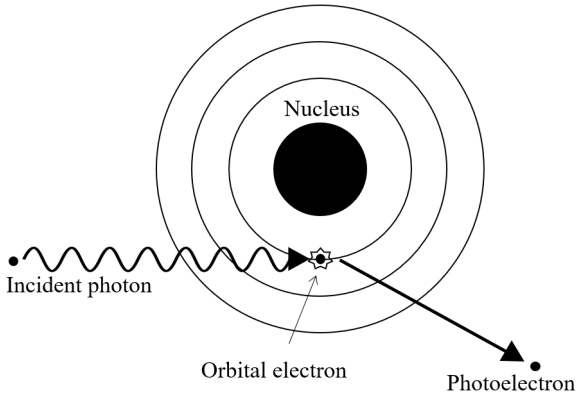

Figure 2.2: The photoelectric effect describes an orbital electron knocked out from its orbit by absorbing an incident photon. Modified figure from [42].

The electron is ejected with a kinetic energy equal to the difference between the incident photon energy and the electron's binding energy, given as

$$E = h\nu - \phi \tag{2.2}$$

where E is the energy of the electron, $h\nu$ is the energy of the photon, the product of Planck's constant $h = 6.63 \cdot 10^{-34}$ m^{2}kg/s and the frequency of the photon (ν), and ϕ is the binding energy of the electron. The probability of a particle interaction is commonly expressed as a cross-section. The cross-section of the photoelectric effect for photon energy exceeding 100 keV, which illustrates the rapid decline with increasing energy, is partially described by the proportionality given as

$$\sigma \propto \frac{Z^4}{E_\gamma^3} \tag{2.3}$$

where Z is the atomic number of the material and E_γ is the energy of the incident photon [9]. Thus, the process is dominant at lower photon energies, as can be seen from figure 2.1, and its probability increases with increasing atomic number (Z) of the material [9, 40, 41].

Compton Scattering

Unlike the photoelectric effect, Compton scattering does not result in the absorption of the photon, but the photon is instead scattered from its initial path. The process describes a photon colliding with an outer orbital electron, or a free electron, and transferring part of its energy to the electron. An illustration of the process is shown in figure 2.3.

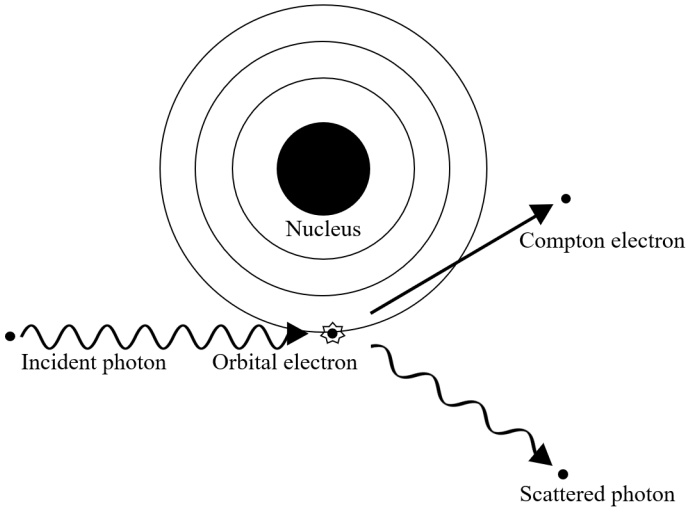

Figure 2.3: Compton scattering describes the process of an incident photon colliding with an electron and absorbing some of its energy, before both particles continue at an angle to each other. Figure modified from [42].

The particles are subsequently ejected at an angle between each other, with the angle dependent on the magnitude of energy transferred between the particles. The outgoing photon's energy is related to the incident energy of the photon by

$$E_{\gamma'} = \frac{E_\gamma}{\left[1 + \frac{E_\gamma}{m_e c^2}(1 - \cos\theta)\right]} \tag{2.4}$$

where $E_{\gamma'}$ is the energy of the outgoing photon, E_γ is the energy of the incident photon, m_e is the mass of the electron and, θ is the angle between the path of the incident photon and the outgoing photon [9]. Maximum energy occurs when the angle $\theta = \pi$, which describes a direct collision between the photon and the electron. The cross-section of Compton scattering has no direct energy proportionality, which makes Compton scattering dominate at intermediate energies, as can be seen in figure 2.1 [9, 40, 41].

Pair Production

Pair production is the final photon interaction process outlined and describes a photon affected by a nucleus' electromagnetic field and the subsequent spontaneous conversion into an electron-positron pair, as shown in figure 2.4.

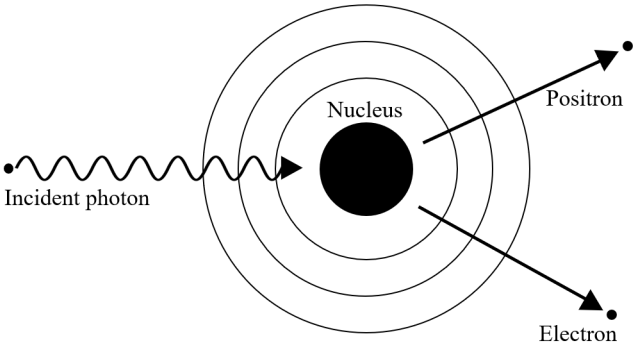

Figure 2.4: Pair production describes the spontaneous annihilation of a photon into an electron-positron pair. Modified figure from [42].

The particles are ejected at an angle θ to conserve momentum, but not necessarily at $\theta = \pi$ since the nucleus absorbs parts of the energy. Hence, conservation of energy is the reason pair production only occurs in the presence of a nucleus. Pair production is dominant at higher energies, with threshold energy equivalent to twice the rest mass of the electron, 1.022 MeV. The energy dependence of the process is illustrated in figure 2.1 [9, 40, 41].

2.1.2 Charged Particle Interactions with Matter

The mechanisms of energy deposition of photons and charged particles differ significantly, and as such, they are characterised by very different interaction patterns. A charged particle interacts with matter primarily through three processes: inelastic Coulomb interactions with atomic electrons, elastic Coulomb interactions with nuclei and inelastic nuclear reactions between heavier incident particles and nuclei. Several other processes are also partly responsible for the energy deposition of a charged particle, for example, bremsstrahlung and Cherenkov radiation. However, the former does not contribute much to the total loss for heavier particles but is integral to the energy loss of lighter particles, such as electrons and positrons. The latter, on the other hand, only becomes significant when the particle traverses a medium at velocities higher than the speed of light [41]. Thus, the following paragraphs will focus on the first three processes listed.

Unlike photon interactions where the photon is either absorbed, scattered or continues unhindered, the charged particle will instead lose small parts of energy in a series of inelastic interactions. As a result, the particle will eventually reach velocities where individual interaction events will alter its course to such an extent that the particle fluctuates ran-

domly in a localised space before finally stopping altogether [12]. For charged particles heavier than electrons, individual inelastic collisions with atomic electrons only contribute to a negligible part of kinetic energy lost. A particle will, however, interact with a large number of electrons through its path, which eventually adds up to a considerable average energy loss per unit length. The Coulomb interactions between a charged particle and atomic electrons, and the subsequent average energy loss of the charged particle (also called stopping power), is described by the Bethe-Bloch equation, which is given as

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{z^2}{\beta^2}\left[\ln\left(\frac{2m_e\gamma^2 v^2 W_{\max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{Z}\right] \tag{2.5}$$

where all constants are given in table 2.1 [41].

Table 2.1: Constants in the Bethe-Bloch equation for the average stopping power of radiation in a medium. Contents of table taken from [41].

	$2\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2 = 0.1535$ MeVcm^2/g
N_{A}	Avogadro's number $= 6.022 \cdot 10^{23}$ mol^{-1}
r_{e}	classical radius of the electron $= 2.817 \cdot 10^{-13}$ cm
$m_{\mathrm{e}}c^2$	rest energy of the electron $= 0.511$ MeV
ρ	density of the absorbing material
Z	atomic number of the absorbing material
A	atomic weight of the absorbing material
z	charge of the incident particle, in units of the elementary charge e
β	velocity of the incident particle in units of the speed of light, v/c
γ	Lorentz factor, $1/\sqrt{1-\beta^2}$
$W_{\max}$	maximum transfer of energy in a single collision
I	mean excitation potential
δ	density correction
C	shell correction

Of note from equation 2.5 is the energy loss being proportional to the atomic number of the absorbing material (Z) and the square of the charge of the incident particle (z^2), as well as the inverse proportionality to the square of the velocity of the incident particle (β^2). There are also two correction terms incorporated in equation 2.5, where the density

correction (δ) accounts for the offset in energy predicted by the Bethe-Bloch formula as the incident particle approaches relativistic velocities. The shell correction (C), on the other hand, becomes important at lower energies where the velocity of the incident particle approaches the orbital velocity of the bound electrons [41].

While inelastic Coulomb scattering largely accounts for the energy loss of the beam, the lateral spreading of the beam profile is instead caused by elastic Coulomb scattering with atomic nuclei. Elastic Coulomb scattering is often referred to as Multiple Coulomb Scattering. This process is also described by several small scattering events that individually deflect the beam only a slight amount, but which sum up to a significant widening of the beam. The spread of the particles in the beam can be approximated as a Gaussian distribution and is, unlike the total stopping power from equation 2.5, almost independent of the charge (z) of the traversing particle [12, 40, 43].

Lastly, there are nuclear reactions. These are inelastic and describe the absorption of the primary traversing particle by a nucleus and the subsequent ejection of secondary particles. The secondary particles are created with low energies, which leads to substantial localised energy deposition. Thus, the secondary particles cannot be neglected when considering the total energy loss of a particle beam. Furthermore, secondary particles are also responsible for the tails of energy deposition that can be observed following the peak of energy deposition of heavy particles [12].

The point of maximum energy deposition of charged particles is called a Bragg peak. As the velocity of the traversing particle decreases, the distance between subsequent inelastic Coulomb scattering events also decreases, which leads to increased localisation of the energy loss. The Bragg peak is associated with considerably higher energy loss compared to the previously traversed distance and is followed by a sharp fall-off in energy deposition, in what is called the range of the beam. The range is not fixed but instead normally distributed around a mean due to varying energy deposition along the tracks of individual particles. These individual variations are referred to as range straggling [12, 40, 44]. The low energy deposition of charged particles relative to the energy deposition in the Bragg peak, and the nearly complete fall-off following the Bragg peak, is one of the main arguments for the utilisation of charged particles in cancer treatment. This property can be utilised to maximise the dose to the tumour while minimising the dose to healthy tissue [12].

The Bethe-Bloch formula is an accurate approximation of the energy loss of heavier particles. However, for electrons and positrons, the assumption that only a small amount of energy is lost in any individual collision is no longer valid. Additionally, the indistinguishability of the electrons changes the calculations slightly. Finally, bremsstrahlung, which is energy lost in the form of a photon as the particle is decelerated by the electric field of nuclei and electrons in the medium, becomes relevant at a lower energy threshold for the lighter particles [41].

2.2 Formalism in Dosimetry and Biologically Modifying Factors

Ionising radiation is known to be hazardous to living tissue, and, in the interest of describing the damage induced by the radiation, several physical measures have been described. Furthermore, quantifying the effects of radiation are required both for radiation protection and for the planning of radiation therapy of cancers, a concept further described in section 2.4. This section will give an overview of units relevant in describing the physical effects of ionising radiation and factors used to quantify the biological effects of radiation, which are internationally agreed upon and outlined in the regularly updated *International Commission on Radiation Units and Measurements* (ICRU) reports [45].

2.2.1 Absorbed Dose

Absorbed dose, or physical dose, is the simplest and most objective measure of the damage from radiation. ICRU report 33 [45] defines absorbed dose as

$$D = \frac{\Delta E}{\Delta m} \tag{2.6}$$

where ΔE is energy deposited by ionising radiation, and Δm is the mass that absorbs the energy. Its unit is the Gray [Gy], equivalent to Joules per kilogram [J/kg] [46].

2.2.2 Radioprotection

Absorbed dose gives a physical description of the energy deposition of radiation, but in the interest of radioprotection, one must also account for the biological damage induced

by radiation. Radioprotection aims to protect the population from adverse effects caused by radiation based on the ALARA principle. ALARA spells out *As Low as Reasonably Achievable* and is based on limiting exposure to the greatest extent possible since the risks involved in particularly low dose radiation are difficult to quantify [9, 47, 48]. However, adverse effects from radiation is known to be linked not only to the physical dose deposition of the radiation, but also on other parameters, such as the type of radiation received and the radiosensitivity of the irradiated biological tissue [40]. Equivalent dose is a measure that accounts for the former by including a weighting factor, w_{R}, accounting for the degree of damage typically induced by the specific type of radiation received. The equivalent dose is described by

$$H_{\mathrm{T}} = \sum_{\mathrm{R}} w_{\mathrm{R}} D_{\mathrm{T,R}} \tag{2.7}$$

where H_{T} is the equivalent dose received by a tissue T, R describes the different type of radiation, w_{R} are the aforementioned weighting factors, while $D_{\mathrm{T,R}}$ is the absorbed dose to a tissue T from radiation type R. The unit of equivalent dose is the Sievert [Sv] [40]. Weighting factors for common types of radiation are given in table 2.2.

Table 2.2: Weighting factors w_{R}, for common types of radiation. Contents of this table taken from Cerrito, L. [40].

Radiation type	Weighting factor, w_{R}
Photons	1
Electrons and muons	1
Protons and charged pions	2
α, fission fragments and heavy ions	20
neutrons	Between 2 and 22, depending on neutron energy.

By incorporating a weighting factor (w_{T}) accounting for the radiosensitivity of the different biological tissue in addition to w_{R}, one receives the effective dose (E). The effective dose is given as

$$E = \sum_{\mathrm{T}} w_{\mathrm{T}} H_{\mathrm{T}} = \sum_{\mathrm{T}} w_{\mathrm{T}} \sum_{\mathrm{R}} w_{\mathrm{R}} D_{\mathrm{R}} \tag{2.8}$$

where E is the effective dose which also has the unit Sievert [Sv], w_{T} is the weighting factor for a specific tissue type, T, and H_{T} is the equivalent dose as found from equation 2.7 [40]. Commonly used estimates for weighting factors, w_{T}, are given in table 2.3.

Table 2.3: Commonly used tissue weighting factors, w_T. Contents of table retrieved from Cerrito, L. [40].

Tissue	Weighting factor, w_{T}	$\sum_{\mathrm{T}} w_{\mathrm{T}}$
Bone marrow (red), colon, lung, stomach	0.12	0.72
Breast, remainder tissues	0.12	0.72
Gonads	0.08	0.08
Bladder, oesophagus, liver, thyroid	0.04	0.16
Bone surface, brain, salivary glands, skin	0.01	0.04

2.2.3 Linear Energy Transfer

An alternative physical quality of the radiation used to describe the degree of damage induced is the linear energy transfer (LET). LET was first described by Zirkle et al. in 1952 and describes the energy transferred to the traversed medium along a single particle track of the radiation [49]. LET of a charged particle is given as

$$\mathrm{LET} = \frac{\mathrm{d}E}{\mathrm{d}l} \tag{2.9}$$

where $\mathrm{d}E$ is the average energy loss of a charged particle over an average distance $\mathrm{d}l$. The unit most commonly used for LET is [keV/µm] [47]. One typically differentiates between the unrestricted LET (LET_∞) and restricted LET (LET_Δ). LET_∞ takes into account all energy deposited by the particles and is equal to the stopping power acting on the particle, as given by equation 2.5. However, a significant portion of the energy can potentially travel quite far from the local tissue in the form of delta rays. As such, LET_Δ, which excludes delta rays above a certain energy threshold, generally represents a more accurate depiction of the localised energy deposition, although for clinically relevant proton energies there is little difference between the two definitions of LET [9].

Since LET is a property associated with individual particles, a given position is described by several different LET values that combined make up an LET spectrum. Additionally, for massive particles, the LET varies along the radiation beam path, with LET increasing

as the particle is slowed down. Thus, a maximum LET is located in the distal end of the Bragg peak [12]. However, a spectrum of LET values is difficult to work with, so several methods are used to calculate average LET values for beam positions instead. The dose averaged LET ($\mathrm{LET_d}$) gives a relationship dependent on both LET and absorbed dose and is defined as

$$\mathrm{LET_d}(z) = \frac{\int_0^\infty S_{\mathrm{el}}(E) D(E,z) \mathrm{d}E}{\int_0^\infty D(E,z) \mathrm{d}E} \tag{2.10}$$

where z is the depth in the tissue, S_{el} is the stopping power, or LET_∞, for a particle of energy E, and D is the dose at a specific depth z from particles with energy E [50]. Track averaged LET ($\mathrm{LET_t}$), also sometimes referred to as fluence averaged LET, incorporates the particle flux of the radiation and is defined by

$$\mathrm{LET_t}(z) = \frac{\int_0^\infty S_{\mathrm{el}}(E) \Phi(E,z) \mathrm{d}E}{\int_0^\infty \Phi(E,z) \mathrm{d}E} \tag{2.11}$$

where z is the depth in the tissue, S_{el} is the stopping power, or LET_∞, for a particle of energy E, and Φ is the particle flux at a certain depth z of particles with energy E. LET cannot be directly measured, so Monte Carlo (MC) methods are commonly used to estimate the LET from a given radiation description, a concept that will be described in further detail in section 2.7.3 [50].

2.3 Radiobiology and Quantifying Biological Effects

Radiobiology is the study of ionising radiation's effect on living organisms and is primarily described by damage to deoxyribonucleic acid (DNA), which disrupts the cells' ability to reproduce. On a molecular level, the damage from ionising radiation can be categorised as direct action or indirect action. A schematic of both processes is given in figure 2.5.

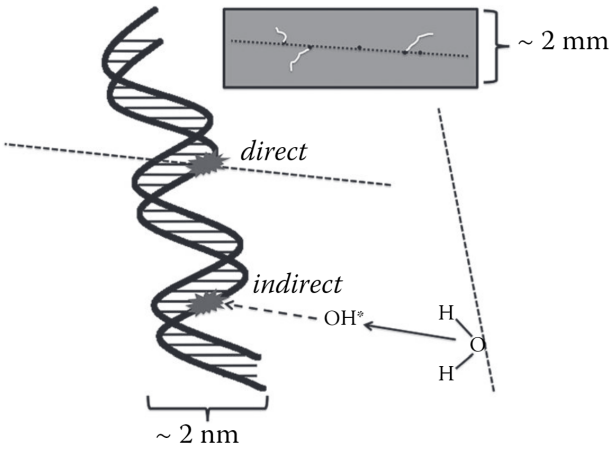

Figure 2.5: Figure outlining damage directly from the radiation (direct action) and subsequent damage from a radical produced by radiating a water molecule (indirect action), both on a nm and mm scale. Figure taken from Paganetti, H [12].

Direct action refers to the local damaging of the molecules in the cell, such as a protein molecule or nucleic acid, from the radiation itself or from δ-rays created from particle interactions. Damage from direct action is localised, which means only cells in the immediate vicinity of the radiation beam path are affected. Direct action is associated with high-LET radiation, such as heavier particles and ions, and also typically induces a high number of double-strand breaks (DSB) to the DNA molecular structure. DSB is the breaking of both DNA strands by a single particle and is more challenging for the cell to repair than its counterpart, the single-strand break (SSB). Thus, DNA damaged by DSB is more likely to lead to cell-killing, but is also at higher risk of leading to faulty repair of the cell. Faulty repair can result in mutations of the cell or even carcinogenesis, the development of cancer in the irradiated tissue [46, 47].

Indirect action, on the other hand, describes the ionisation of water molecules into chemically toxic radicals, which can then travel a distance from the site of the radiation before interacting with biological tissue. Indirect action will eventually damage DNA strands through chemical interactions between the molecules in the DNA and the radicals, which, unlike direct action, is more likely to induce an SSB. SSBs are typically easier to repair compared to DSB, which means most of the biological damage of radiation is associated with the DSB. Indirect action is generally associated with low-LET radiation, such as low energy X-rays [46, 47].

2.3.1 Linear Quadratic Model

While parts of our knowledge regarding the biological effects of radiation come from *in vivo* data collected from, for example, survivors of atomic disasters, most of our understanding of radiative biological effects are discovered through *in vitro* experiments of irradiated cell cultures. From *in vitro* studies, the linear-quadratic (LQ) model has been developed, a mathematical model describing radiation damage in cell cultures. The surviving fraction of cells predicted by the model coincide well with experimental results, particularly for lower dose regions [47]. The LQ model describes the surviving fraction of cells after being irradiated with a dose D as given by

$$\mathrm{SF} = \exp\left(-(\alpha D + \beta D^2)\right) \tag{2.12}$$

where α is a factor that can be thought of to describe cell-killing induced by a single particle, while β represents cell-killing caused by several separate particles [51]. Cell survival, as described by equation 2.12, is often characterised by survival fraction curves, which is a graphical representation of the surviving cells in a cell culture radiated by a dose D. An example of a cell survival curve is given in figure 2.6.

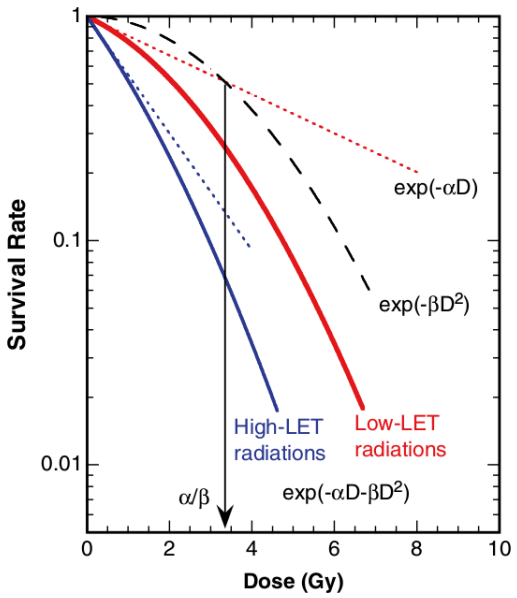

Figure 2.6: Example of a survival fraction curve showing the contribution of both the linear and quadratic term on the resultant survival fraction. The effect LET has on survival fraction is also outlined, and the α/β-ratio is indicated. Figure taken from Ridolfi, R [52].

In addition to showing the effect of the linear and the quadratic terms from equation 2.12, figure 2.6 also shows how LET relates to cell killing, with high LET radiation leading to a higher degree of cell killing than low LET radiation. The point where the linear and

quadratic components contribute to the same amount of cell killings is referred to as the α/β-ratio. The α/β-ratio of tissues and their radiosensitivity typically correlate, with early-responding tissues having high α/β-ratios, while late-responding tissues have lower α/β-ratios [9].

2.3.2 Fractionation

A radiotherapeutic approach to exploiting the different response of cells is to split radiotherapy treatment into several fractions. The initial rationale behind fractionated treatment was mainly empirical. However, the LQ model provides a generally accepted explanation for its success. Late responding-tissue (low α/β-ratio) is more sensitive to dose within a fraction than early-responding tissue (high α/β-ratio). Thus, by irradiating a cell culture in several iterations of low dose, one would see a higher degree of cell-killing to the latter than the former. This can be interpreted from figure 2.6 by low doses maximising the initial linear slope of the curves while not achieving a high enough dose for the quadratic slope to become significant. Human tumour tissue is typically associated with high α/β-ratios, while healthy tissue is typically late-responding. Thus, fractionation achieves maximum cell-killing of tumour tissue while minimising the damage to healthy tissue [9].

2.4 Radiation Therapy

Radiation therapy, or radiotherapy, is a central technology within cancer treatment that utilises the damage from ionising radiation to biological tissue to kill malignant tumour growths. Radiotherapy aims to achieve sufficient tumour control probability (TCP) by maximising the dose to the tumour volume while simultaneously avoiding excessive normal tissue complication probability (NTCP) by minimising the dose to healthy tissue. The TCP is defined as the dose required to achieve a certain probability of killing the tumour cell, whilst NTCP is the dose corresponding to a probability of causing a specific adverse effect to healthy tissue [9, 12, 53]. Adverse effect to healthy tissue from radiation is referred to as toxicity and is graded from 1 to 5 according to a universal set of guidelines referred to as the *Common Terminology Criteria for Adverse Events* (CTCAE) [54]. Hence, ideal radiotherapy treatment minimises the risk of toxicity by achieving a sufficient balance between TCP and NTCP, which is referred to as the therapeutic window and is shown graphically in figure 2.7.

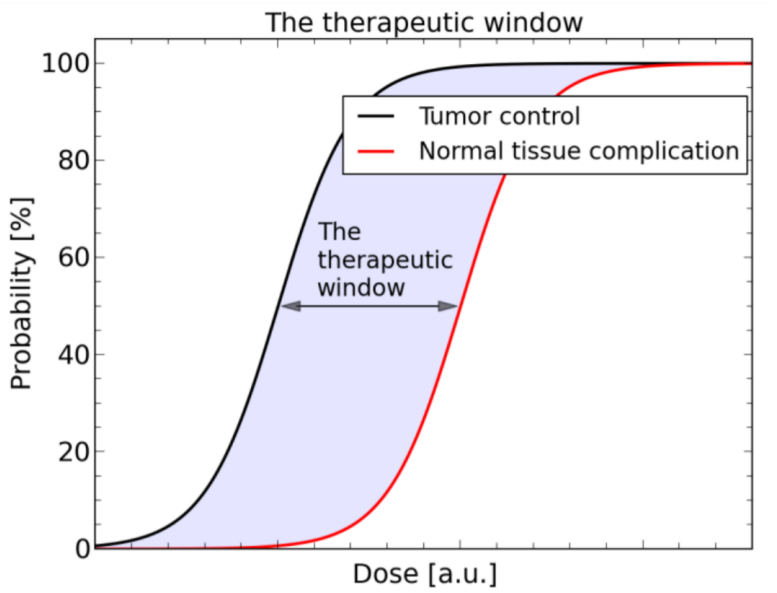

Figure 2.7: Graph outlining the relation between tumour control, normal tissue complication and the therapeutic window. Figure taken from Ytre-Hauge, K [55]

Treatment within the therapeutic window is achieved through careful treatment planning and quality assurance. First, the patient's anatomy is imaged with diagnostic imaging modalities, which is used to contour the volumes in the area of the tumour site. The dose plan is then calculated to cover the tumour volume while avoiding unnecessary dose to vital organs near the tumour [9]. The general workflow of the treatment is further explained in section 2.6, while this section focuses on the basic concepts of radiotherapy modalities. First, photon therapy will be discussed, its delivery techniques and the production of the X-rays that are central for the treatment. The basics of proton therapy will subsequently be explained, along with its beam modalities and the production and acceleration of the protons.

2.4.1 Photon Therapy

Radiotherapy using photons was, for a long time, the only modality available. Following Wilhelm Röntgen's discovery of X-rays, photons with energies between 124 eV and 124 keV, in 1895 [56], the particles were quickly used to treat breast cancer [10]. Our understanding of ionising radiation improved the following years, with a large part of the credit attributed to the efforts of Becquerel and the Curies [10]. Common for most X-ray production is the acceleration of electrons that in turn create X-rays through interacting with a target. Hence, X-ray beams are typically characterised by the voltages to which the electrons were accelerated before reaching the target and will consist of a range of X-ray energies peaking at the energy of the accelerated electrons [9].

As can be seen from figure 2.8 higher voltage X-rays have deeper peaks of dose deposition than lower voltage X-rays, and achieving sufficiently high voltage was historically an obstacle of radiotherapy. Damage to healthy tissue was then minimised by irradiating from several angles, which is a central principle within radiotherapy for higher voltage X-ray beams, as well. The 1950s to 1980s saw improved options for achieving higher voltage X-rays with the development of cobalt therapy machines, which use gamma-ray emitting ^{60}Co sources as the radiation source, and the linear particle accelerator (linac). The linac continues to be the primary source of X-rays for radiotherapy in modern times [57, 58].

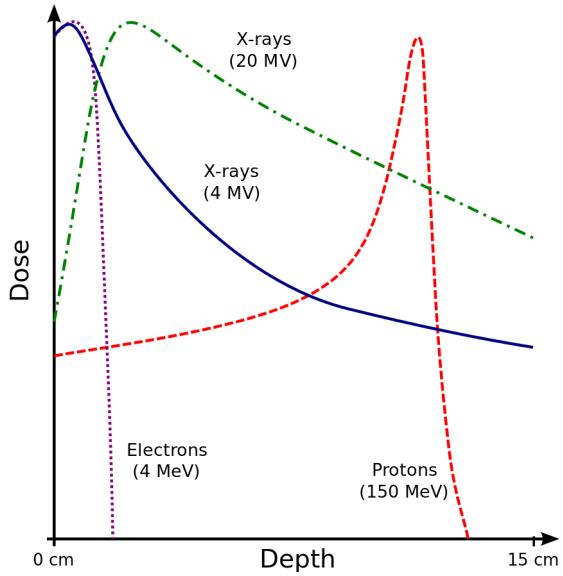

Figure 2.8: Depth-dose curve in an arbitrary medium for 4 MV and 20 MV energy X-ray beams, a 4 MeV electron beam and a 150 MeV proton beam. Modified figure from [59].

Linear Accelerator

The linear accelerator (linac) is the primary mechanism behind the modern production of high energy photons for clinical purposes, with clinical linacs typically able to accelerate electrons to energies up to 22 MeV [9]. Shallow tumours can be treated directly with electrons due to their high entrance doses, as seen from figure 2.8. However, more commonly the electrons are converted to X-rays by interacting with a target. Clinical X-rays are typically in the voltage range 6-20 MV [9].

An electron gun, usually a heated cathode releasing electrons through thermionic emission, is powered by a voltage pulse. The electrons then propagate through a waveguide

of alternating electric potential, accelerating the electrons in a linear beam path. The alternating potentials and the voltage to the electron gun are controlled by a radiofrequency pulse, which ensures the electrons are constantly accelerated towards a positive potential. Due to their large size, the linac is constructed horizontally, which means that the electron beam must be redirected approximately 90° before reaching the target. This is achieved by using magnetic fields since this will induce a Lorentz force on the moving charges and, subsequently alter the propagation path. The electrons are then directed towards a high Z target, typically tungsten, which will convert the electrons into X-rays [9]. A schematic of the linac system is given in figure 2.9.

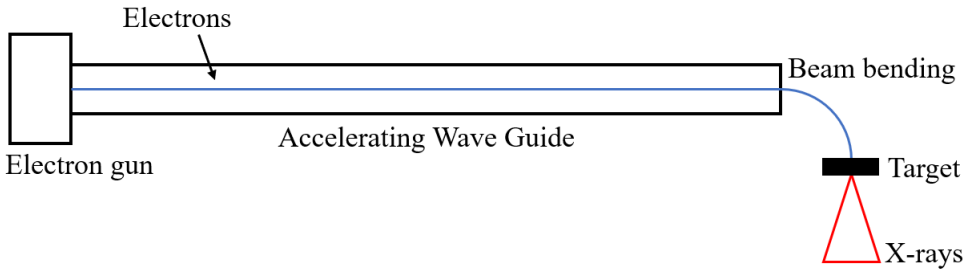

Figure 2.9: Schematic of a linac. The electrons are produced in the electron gun, accelerated and steered to a target where they produce electrons, which are subsequently delivered to the patient. Modified figure from Paganetti, H [9]

The X-rays are mainly produced through two mechanisms, bremsstrahlung and the excitation of orbital electrons. Bremsstrahlung describes the photons emitted as the electrons lose energy being slowed down by the electric field of orbital electrons or nuclei in the medium, as described earlier in section 2.1. The result of bremsstrahlung is a continuous energy spectrum of X-rays, and production is inversely proportional to the energy of the electrons. Unlike bremsstrahlung, the spectrum of X-rays released from the excitation of orbital electrons is not continuous. The orbital electrons inhibit quantised energies, and as they absorb energy, they ascend to a higher quantised potential. The subsequent de-excitation of the electrons will release the energy difference between the states as an X-ray. Due to the specific energy states these X-rays are emitted with, they are referred to as characteristic X-rays [60].

Beam Shaping and Delivery

The target is located in the beam head in the treatment room, and the next step in the process is for the X-ray beam profile to be contoured to match the prescribed treatment before being delivered to the patient. This is achieved by utilising different forms of physical blocking of the beam. The beam escaping the target does not have a homogeneous

beam profile. Instead, most of the X-rays are concentrated centrally in the beam profile, and to even out the beam profile a flattening filter is used. The flattening filter is a physical shape that is thicker in the centre than at the edges, which results in a larger number of X-rays absorbed from the centre of the beam profile compared to the edges. The beam escaping the flattening filter will, thus, have a significantly more homogeneous beam profile. The beam profile must also be limited to only irradiate the volume of the patient that was prescribed dose, which is achieved through collimation. The primary collimator shapes the beam to a circular beam profile, while the secondary collimator attains a rectangular field. Both are physical blockings made of high Z material, typically lead. Further shaping of the beam is achieved by more flexible collimators referred to as multi-leaf collimators (MLC) [60], and while the primary collimator is always present in the system, the MLC can replace the secondary collimator [9]. The MLC is a system of independently moving leaves that can flexibly shape the beam, resulting in a beam accuracy only limited by the width of the leaves [9].

The flexibility of shaping the beam developing alongside more advanced imaging methods and more powerful computing power has led to significant advancements within radiotherapy in the last decades [57]. Oncologists would previously plan treatment by hand with inflexible beam shaping mechanisms and limited knowledge of the actual dose distribution [9], but the application of MLCs and computer algorithms to achieve highly conformal and homogenous dose to the tumour volume has significantly improved the quality of the treatment. This form of treatment is commonly referred to as three-dimensional conformal radiotherapy (3D-CRT) and allows improved sparing of healthy tissue surrounding the tumour volume [11].

Further improvement in technology has also allowed the intensity of the beam to be adjusted during treatment to achieve even higher conformity to the tumour volume, in what is known as intensity-modulated radiotherapy (IMRT) [9]. The beam intensity is altered by using segmented beam apertures, where each segment has different beam-on times. Beam intensity can be changed during beam-on time, but one can also periodically turn off the beam and adjust the intensity [11]. Furthermore, volumetric Modulated Arc Therapy (VMAT) is a more recent development within photon therapy, and while IMRT utilises a fixed gantry, VMAT uses a 360° rotating treatment system. By delivering the treatment in arcs, while dynamically altering the intensity and collimation of the beam during therapy, the treatment time can be significantly shortened [61].

2.4.2 Proton Therapy

While there have been significant developments within photon therapy, proton therapy, and heavier particle therapy, have developed in parallel and inhibits certain advantages over their photon counterpart. The main benefit of particle therapy over photon therapy is shown in the depth-dose profile of protons in figure 2.8. Unlike photons that have considerable dose deposition along their full beam paths, protons accelerated to sufficiently high energies, typically 200-250 MeV [9], will have low entrance dose, high and localised maximum dose deposition in the Bragg peak, and a nearly complete fall-off following the Bragg peak [12]. To ensure the Bragg peak covers the entire tumour volume, a beam profile consisting of several different proton energies, and accordingly several differently positioned Bragg peaks, is often utilised. This is known as a spread-out Bragg peak (SOBP) [12], with the depth-dose profile of an SOBP shown in figure 2.10.

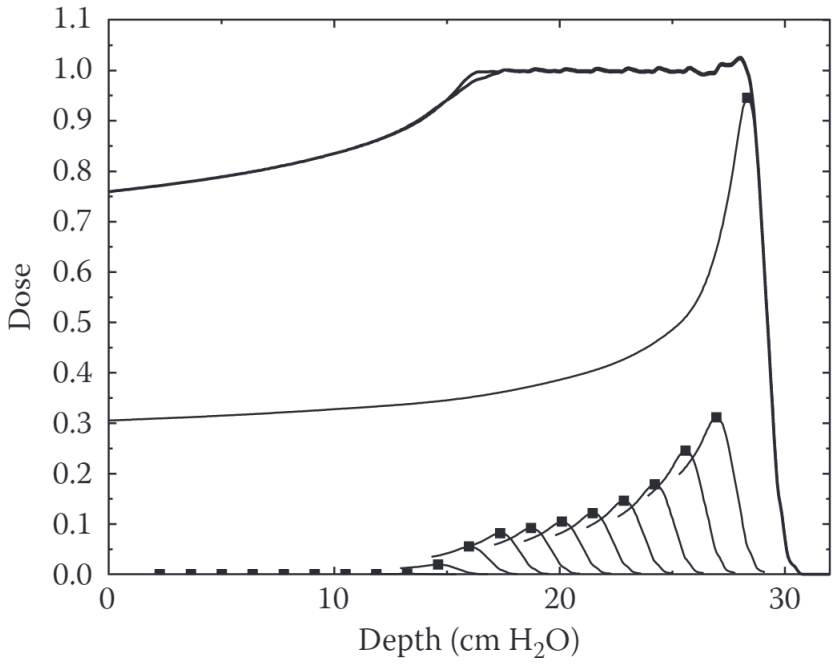

Figure 2.10: Example of a spread out Bragg peak curve of a radiation beam traversing a water medium. The wide Bragg peak on the top is the sum of all the smaller curves. Figure taken from Paganetti, H [12].

With improvements in imaging modalities and the ability to more accurately position the Bragg peak to cover the tumour volume, proton therapy and heavier particle therapy has been getting increased focus the last decades. However, the preferable depth-dose profile of particles has been known, and the subject of study, for considerably longer [62]. The use of protons for cancer treatment was first proposed by Robert R. Wilson in 1946 [63] and was first used clinically to treat a case of breast cancer in 1954 [14]. While early studies of the treatment method showed promising results, the expense of maintaining

cyclotrons to accelerate the protons and the lack of public support for the subject meant that the development stagnated for a while, until 1990 saw the first clinical synchrotron installed at the Loma Linda University Medical Center [14]. Since then, several proton centres have been established globally, and per April 2021, 97 proton centres were in operation, with both cyclotrons and synchrotrons being used as accelerators [64].

Cyclotron

A cyclotron utilises the Lorentz force set up by combined electric and magnetic fields to accelerate charged particles. The Lorentz force acting on a particle is described by

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{2.13}$$

where q is the charge of the particle, $\vec{E}$ is the electric field acting on the particle, $\vec{v}$ is the velocity of the particle, and $\vec{B}$ is the magnetic field acting on the particle. The mechanism of a typical cyclotron is given in figure 2.11.

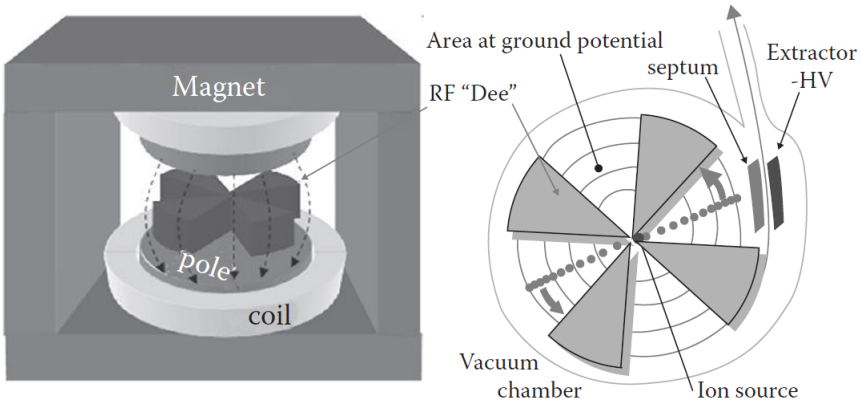

Figure 2.11: Schematic of a cyclotron. Left shows the magnets vertically over the vacuum chamber, while the right side shows the circular orbits of the charged particles in the vacuum chamber, and the extractor. Figure retrieved from Paganetti, H [12]

Two magnets, called "Dees", are positioned vertically over a hollow chamber, with the particles to be accelerated in the centre of the chamber. The resultant magnetic force imposed on the particles is, thus, pointing towards the centre. This would merely maintain the particles enclosed in a circular orbit, but, in addition to the magnetic field, there is an alternating electric field that changes polarity depending on the current position of the particles. Furthermore, a radiofrequency pulse alters the charge of the two sides of the system, constantly accelerating the particles towards the side of opposite charge. Hence, the radius of the particles will increase since the time taken to complete the circle remains constant, while the particle's velocity is increasing. Finally, when the particles have been

accelerated to sufficient energy, the particles are extracted from the circular orbit [40].

Synchrotron

A synchrotron also utilises the Lorentz force to accelerate and steer the particles, but through a slightly different approach. A depiction of a synchrotron given in figure 2.12.

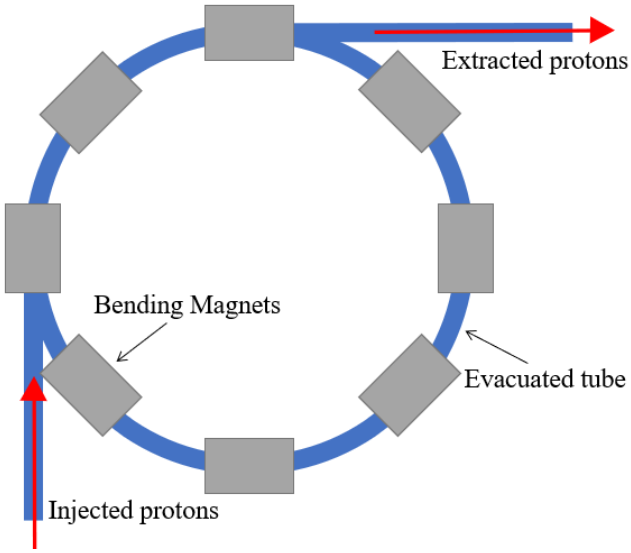

Figure 2.12: Schematic of a synchrotron. The particles are accelerated prior to entering the tube. In the tube their position is maintained by magnets, before they are extracted when the target particle energy is achieved. Modified figure from Khan et al. [60].

In synchrotrons, the particles are accelerated with electric fields, while several magnets keep them bound within a large, evacuated circular tube. The magnetic field is increased as the velocity of the particle increases, and this, combined with the considerable size of the main tube, results in greater flexibility in extracted particle energy from the synchrotron compared to the cyclotron. Before entering the main tube, the particles must first be accelerated to high velocity to maintain their rotational period despite the increased radius. This initial acceleration is generally achieved using a linac. The particles are subsequently accelerated by altering the electric fields in the main tube. Finally, when sufficient particle energy has been achieved, the magnetic fields are altered to bend the path of the particles and extract them from the circular orbit [40].

Beam Shaping and Delivery

The particles are extracted from the accelerator and delivered to the patient through one of two main categories of particle therapy delivery systems, passive scattering and active

scanning [12]. Passive scattering is the oldest particle delivery technique and uses physical blocks to shape the beam. The simplest technique utilises a single scattering foil called a flat scatterer. However, the flat scatterer is an energy inefficient approach since the dose profile is heterogeneous and a considerable fraction of the beam profile must be blocked to achieve clinically acceptable homogeneity of the treatment. A means of achieving more energy efficient treatment is to implement a contoured scatterer in the system. The heterogeneous beam profile escaping the flat scatterer is homogenised by blocking different regions of the beam profile to different degrees by the contoured scatterer. This achieves improved energy efficiency of the treatment by creating a wider beam profile of clinically acceptable homogeneity. However, while the contoured scatterer achieves improved energy efficiency, a high number of protons are still blocked in the process. This is not only energy inefficient, but the blocked protons also produce neutrons that can reach the patient and deposit unwanted dose [12].

In more recent years active scanning is instead more commonly utilised to shape the proton beam. Active scanning, also called spot scanning or pencil beam scanning, primarily uses magnets to steer the beam. Active scanning is an iterative process that irradiates the tumour volume spot-wise within a layer before adjusting the energy to repeat the process but contoured to a different volume in a deeper layer. Hence, active scanning can achieve a highly conformal treatment to the tumour. Still, due to the sequential nature of the treatment, the process is time-consuming and sensitive to organ motion. Since physical barriers are avoided, however, higher energy efficiency is achieved, and fewer secondary neutrons are produced [12, 65].

Active scanning forms the basis of the technical implementation of proton therapy, but equivalently to IMRT for photon therapy, each field's intensity and dose distribution is also commonly altered during treatment. The result is a dose distribution summed over all beams that fits the intended treatment but comprise several individually varying dose distributions [12, 66]. This is referred to as intensity-modulated proton therapy (IMPT). Additionally, equivalently to VMAT, proton arc therapy has been proposed as an alternative treatment modality in later years where the gantry is rotated in an arc around the patient. This can improve dose conformity to the target volume but gives less control over the distal ends of the Bragg peaks of the individual beams. The consequence of this, as will be explained further in the next section, is possibly excessive radiation damage to sensitive organs close to the target volume [67].

2.5 Relative Biological Effectiveness

Proton therapy is a less explored treatment modality than photon therapy, and our experience of the body's response to photon therapy is difficult to directly translate to proton therapy. The standard method utilised clinically to account for the different qualities of the radiotherapy modalities is to incorporate the relative biological effectiveness (RBE) of the radiation. RBE is defined as the ratio between the dose from a reference radiation and the dose from a test radiation that both result in the same clinical outcome. Mathematically, the RBE factor is defined as

$$\mathrm{RBE} = \frac{D_{\mathrm{x}}}{D_{\mathrm{test}}} \tag{2.14}$$

where D_{x} is the physical dose, as outlined by equation 2.6, from a reference radiation, while D_{test} is the physical dose from the radiation utilised, to which the RBE factor corresponds [47]. The physical dose is then recalculated to yield the RBE-weighted dose [Gy(RBE)], by rearranging equation 2.14 to give

$$D_{\mathrm{RBE}} = \mathrm{RBE} \cdot D \tag{2.15}$$

The RBE factor is often used to correct dose calculations for non-photon beams, such as proton and carbon ion therapy. For proton therapy, the ICRU report 78 recommends the utilisation of a constant RBE factor of 1.1, from here referred to as RBE1.1, compared to the photon dose from a ^{60}Co source resulting in similar cell-kill as the proton beam [12, 15, 68], while carbon ion therapy instead utilises different RBE models based on *in vitro* data [69]. However, the RBE of protons has been shown from *in vitro* experiments to vary in the interval 1.1 to 1.7 across the SOBP, resulting in an average value of approximately 1.15 in the centre [16]. RBE has been observed to increase with increasing LET of the radiation, up to a maximum at approximately 100 keV/µm before decreasing for further increasing LET [16, 47]. Additionally, RBE has been shown to increase with decreasing physical dose and $(\alpha/\beta)_{\mathrm{x}}$-ratio of the tissue, which describes the sensitivity of the tissue to the reference radiation [47]. The peak RBE is observed in the distal end of the Bragg peak, since this is associated with high LET values [16].

An alternative expression of RBE as a function of the test dose and the LQ parameters

of both the test and the reference radiation is given by

$$\mathrm{RBE}(D_\mathrm{p}, \alpha, \beta, \alpha_\mathrm{x}, \beta_\mathrm{x}) = \frac{1}{2D_\mathrm{test}} \left(\sqrt{\left(\frac{\alpha_\mathrm{x}}{\beta_\mathrm{x}}\right)^2 + 4D_\mathrm{test}\frac{\alpha_\mathrm{x}}{\beta_\mathrm{x}}\frac{\alpha}{\alpha_\mathrm{x}} + 4D_\mathrm{test}^2\frac{\beta}{\beta_\mathrm{x}}} - \frac{\alpha_\mathrm{x}}{\beta_\mathrm{x}} \right) \tag{2.16}$$

where D_test is the dose from the test radiation, α and β are the parameters for the LQ model (equation 2.12) for the test radiation, and α_x and β_x are the parameters for the reference radiation [16]. The minimum and maximum RBE values are defined as the RBE in the high dose limit and low dose limit, respectively, and are described by

$$\mathrm{RBE}_\mathrm{min} = \lim_{D\to\infty} \mathrm{RBE} = \frac{\alpha}{\alpha_\mathrm{x}} \tag{2.17}$$

and

$$\mathrm{RBE}_\mathrm{max} = \lim_{D\to 0} \mathrm{RBE} = \sqrt{\frac{\beta}{\beta_\mathrm{x}}} \tag{2.18}$$

This shows mathematically how the linear component dominates the low dose region of the LQ model, while the quadratic component dominates the high dose region [20].

2.5.1 Modelling Relative Biological Effectiveness

The increase in RBE in the distal end of the SOBP is accounted for clinically by avoiding the position of the distal end coinciding with sensitive organs at risk (OAR) [23], but otherwise variable RBE is not considered in treatment planning. However, several approaches to modelling the variations of RBE on dose distributions have been developed in recent years, with varying degrees of complexity [20]. The following paragraphs will outline a selection of RBE models utilised in this thesis accounting for several combinations of the previously outlined parameters. Phenomenological RBE models are a subcategory of RBE models that approximate experimental results and are often based on the formulation of the LQ model as shown in equation 2.16. What separates the phenomenological RBE models are their definitions of the extremal RBE values, as given by equations 2.17 and 2.18 [20].

Unkelbach et al. established what has come to be known as the LET weighted dose as a means of accounting for LET in dose calculations for proton therapy, to spare healthy

tissue [70]. The original description used dose averaged LET and is given by

$$D_{\mathrm{LET}} = D \times (1 + c\mathrm{LET_d}) \tag{2.19}$$

where D is the physical dose, c is a weighting factor, and $\mathrm{LET_d}$ is the dose averaged LET. Unkelbach et al. used a factor $c = 0.04$ since this achieved an RBE = 1.1 in the centre of an SOBP of 5 cm modulation and 10 cm range. Further, McMahon et al. empirically established a weighting factor $c = 0.055 \pm 0.003$ by fitting the curve to points from clonogenic cell survival experiments [71].

McNamara et al. [21] established a phenomenological RBE model based on 285 *in vitro* cell survival experiments previously assessed by Paganetti, H [16]. The RBE model is based on the LQ-model in equation 2.12, and predicts an increasing RBE with increasing $\mathrm{LET_d}$ and increasing RBE with decreasing absorbed dose and $(\alpha/\beta)_{\mathrm{x}}$. The extreme RBE values of the model are given by

$$\mathrm{RBE_{min}}\left(\mathrm{LET_d}, \left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}\right) = 1.1012 - 0.0038703\ \mathrm{Gy}^{-\frac{1}{2}} \left(\frac{\mathrm{keV}}{\mu\mathrm{m}}\right)^{-1} \sqrt{\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}}\,\mathrm{LET_d} \tag{2.20}$$

and

$$\mathrm{RBE_{max}}\left(\mathrm{LET_d}, \left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}\right) = 0.99064 + \frac{0.35605\ \mathrm{Gy}\left(\frac{\mathrm{keV}}{\mu\mathrm{m}}\right)^{-1}}{\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}}\mathrm{LET_d} \tag{2.21}$$

where $\mathrm{LET_d}$ is the dose averaged linear energy transfer, and $(\alpha/\beta)_{\mathrm{x}}$ is the (α/β)-ratio of the tissue to the reference radiation [21]. Furthermore, Rørvik et al. [22] established two phenomenological RBE models; the unweighted and the weighted Rørvik model. The models are based on 85 *in vitro* cell survival experiments and differ in whether weighted or unweighted regression was utilised to determine a model fit to the database. Both models assume a constant $\mathrm{RBE_{min}} = 1$, with the unweighted $\mathrm{RBE_{max}}$ given by

$$\mathrm{RBE_{max}}\left(\mathrm{LET_d}, \left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}\right) = 1 + \frac{0.645\ \mathrm{Gy}\left(\frac{\mathrm{keV}}{\mu\mathrm{m}}\right)^{-1}}{\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}}}\mathrm{LET_d} \tag{2.22}$$

Thus, incorporating $(\alpha/\beta)_{\mathrm{x}}$ and $\mathrm{LET_d}$ in the extreme value, while equation 2.16 also shows the model predicting increased RBE with decreasing absorbed dose. On the other hand,

the weighted function has a definition of $\mathrm{RBE}_{\max}$ dependent on the full dose weighted LET spectrum (L), incorporated as a biological weighting function $r_{\max}(L)$. The $\mathrm{RBE}_{\max}$ of the Rørvik weighted model is given by

$$\mathrm{RBE}_{\max}\left(\mathbf{d}(L)\right) = \int_0^\infty r_{\max}(L)\mathbf{d}(L)\mathrm{d}L \tag{2.23}$$

where L is the range of LET values, and $\mathbf{d}(L)$ is the local dose weighted LET spectrum at a given position in the patient. The biological weighting function $r_{\max}(L)$ incorporates L and $(\alpha/\beta)_{\mathrm{x}}$ as parameters [22].

2.6 Treatment Planning

As mentioned in section 2.4, radiotherapy treatment is preceded by a careful treatment planning process, which will be described further in this section. The first step in treatment planning is to diagnose the patient, where anatomical information is acquired through diagnostic imaging tools. Anatomical images are central through all steps in the treatment planning process. When a diagnosis has been established, the cancerous tissue is defined and delineated according to a universal set of volume definitions [13]. Further, the beam setup must be decided, with the current standard being inverse planning. In inverse planning, dose goals and constraints are set for the delineated volumes, and the beam configuration is subsequently optimised by a computer algorithm. Finally, the planned dose must be evaluated before the patient can be treated [9].

2.6.1 Image Acquisition

The first step in the treatment planning of tumours is to diagnose the patient, where the acquisition of anatomical information of the patient is essential. Several imaging modalities have importance within cancer diagnostics, but they also form the basis of delineating clinical volumes used to define prescribed doses and beam configurations. Magnetic Resonance Imaging (MRI) is commonly used for diagnostics, while computed tomography (CT), in addition to diagnostic use, is also widely used after the diagnosis has been given to assist in the calculation of a dose plan for the treatment [9].

Magnetic Resonance Imaging (MRI) utilises nuclear magnetic resonance (NMR) to image the protons in human tissue. NMR is a means of quantifying a spin system's alignment

with an external magnetic field, and the hydrogen atom is the atom that gives the strongest NMR signal. Thus, an MRI scan effectively maps the hydrogen composition of water and lipid molecules in the body which make up most of the body's soft tissues. Hence, an MRI scan yields images with good soft-tissue contrast, but an MRI can be time-consuming [72].

Computed tomography (CT) is another imaging modality commonly used in cancer diagnostics. CT is an X-ray scan performed 360° around the patient, and the individual X-ray images can then be digitally reconstructed to give three-dimensional anatomical information. The interactions of photons are highly dependent on the atomic number (Z) of the tissue, so CT is well suited for imaging bone structures, as their Z-values are significantly higher than for soft tissues. The soft-tissue contrast, however, while sufficient for most purposes is not as good as for an MRI. Another disadvantage of the CT is the use of ionising radiation. However, the CT is typically quicker than the MRI, and for radiotherapy the CT allows the acquisition of images in treatment position [9, 72].

In CT each voxel is assigned a Hounsfield unit (HU), which is defined by

$$\mathrm{HU} = 1000 \times \frac{\mu_{\mathrm{tissue}} - \mu_{\mathrm{water}}}{\mu_{\mathrm{water}}} \tag{2.24}$$

where μ is the linear attenuation coefficient of the given material, as shown in equation 2.1. By definition, water has HU $= 0$, air has HU $= -1000$, while bone typically has HU $= 1000 - 3000$. The HUs are translated to graphical information as shades of grey [72, 60]. Additionally, since the HU is a direct measure of the dose deposition in different media, the CT scan is central in dose planning for radiotherapy [9].

2.6.2 Volume Definitions

After the image acquisition, the patient's anatomy is contoured, either manually or through auto-segmentation by computer algorithms [73]. The volumes that are important for the delivery of radiotherapy are delineated. They are divided into several clinical subvolumes outlined in the ICRU 50 report and, thus, agreed upon by oncology clinics globally [13]. First, there is the gross tumour volume (GTV) which only contains the volume of the initial solid tumour. The clinical target volume (CTV) includes the GTV and also accounts for subclinically diseased tissue, not visible as a physical tumour. Including error margins around the CTV to account for organ motion and inherent uncertainties in the treatment setup yields the planning target volume (PTV). Additionally, depending

on the position of the tumour, there may be organs nearby that are especially sensitive to radiation and are, thus, often subject to dose constraints that should ideally not be exceeded. These are referred to as organs at risk (OAR) [9, 60]. A schematic showing the clinical volumes with relative position and size is given in figure 2.13.

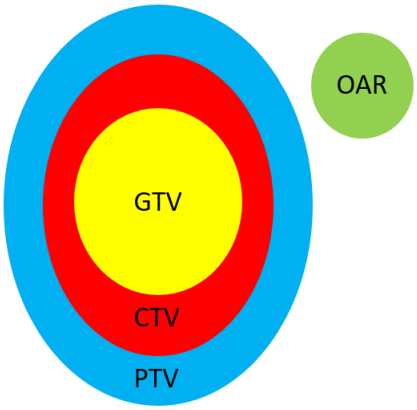

Figure 2.13: Clinically used volume definitions shown with relative sizes to each other. Gross tumour volume (GTV) in the center, surrounded by all subclinically diseased tissue in the clinical target volume (CTV). Error margins give the planning target volume (PTV). Adjacent to the target volume is an organ at risk (OAR). Modified figure from Mayles, P et al. [9].

2.6.3 Dose Optimisation

When the details of the anatomy are known the dose planning is the next step in the treatment planning. In modern treatment planning, this is usually supported by a treatment planning system (TPS). The current preferred method is to plan treatment inversely, which means a dose distribution is used as the basis, and a radiation field setup is planned from this basis. The contoured volumes and their corresponding dose goals and dose constraints are all given manually to the TPS. The TPS will then find beam setup and energies that meet the dose criteria to the greatest extent possible. The goal is to meet the required dose to the entire PTV while avoiding exceeding dose constraints set for proximate OARs. The treatment plan is typically found by setting up a function that penalises underdosage of treatment volume and overdosage of OARs and then attempting to minimise this equation. The function is typically a weighted sum of several cost functions, as given by

$$f(x) = \sum_i w_i c_i \tag{2.25}$$

where c_i are i cost functions, one for each dose goal to be met, while w_i are i corresponding weighting factors. The cost functions are only unequal to zero when their corresponding

dose criteria are not met. By attempting to minimise the given function, the TPS will find a dose regime as close as possible to fulfilling all dose criteria. The weighting factors ensure that if one dose criterion is more important to meet than another, this can be taken into account [9, 74].

The TPS also calculates the dose of each individual voxel. The HUs of the CT images form the basis of the dose distributions, and different algorithms are used to make the simulated doses as true to the physical dose distribution as possible. Currently, the pencil beam algorithm is the most used algorithm for these calculations. The pencil beam algorithm utilises point dose kernels to calculate dose distributions. Point dose kernels are localised descriptions of energy deposition from secondary electrons and scattered photons in water given by a predefined set of photon tissue interactions originating from a single point coordinate [53]. The pencil beam algorithm incorporates point dose kernels along an infinite ray of photons, and the localised dose to a point or voxel in the patient is then given by the dose contribution of each photon ray integrated over the area of the treatment field [9]. The pencil beam algorithm was initially developed for electrons but was quickly adapted for treatment planning with photons [9], and have also been translated to implementation for protons [75, 76]. An alternative approach to dose calculations are Monte Carlo based algorithms, which are generally agreed to be the most accurate means of simulating particle interactions, but was prohibitive for treatment planning purposes due to long computing times. With increased modern computing power, however, the methods have become the focus of development for treatment planning as well [77]. The concept of Monte Carlo methods will be further described in section 2.7.3.

2.6.4 Dose Plan Evaluation

Following the dose calculation, the planned treatment needs to be manually assessed before delivery. This is also often done in the TPS itself. There are several tools implemented in the system to help visualise the planned treatment. Colour mapped isodose curves are visualisations of the beams, where areas receiving lower dose will be coloured in cooler colours, such as blue. In contrast, areas receiving high dose will be coloured with warmer colours such as yellow and red. Isodose curves contoured over a CT image in the Eclipse TPS are shown in figure 2.14 [9].

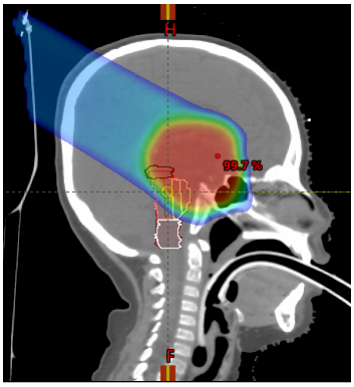

Figure 2.14: CT picture of vertical view of the head of a brain tumour patient with isodose curves showing the dose distribution of the treatment. Image is created with Eclipse TPS.

A dose-volume histogram (DVH) plots the relative irradiated volume against radiation dose and is another valuable tool for evaluating a dose plan. DVHs are divided into differential DVHs and cumulative DVHs, but the latter are used significantly more often than the former. The cumulative DVH considers all volume irradiated by a specific dose or higher, which means that as the dose approaches the maximum dose, the volume percentage will approach 0% [9]. An example of a cumulative DVH is shown in figure 2.15.

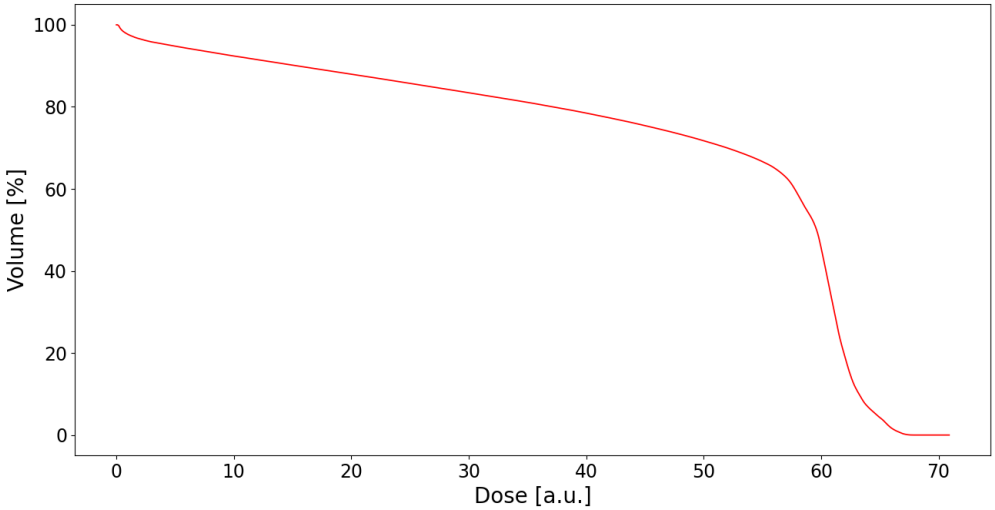

Figure 2.15: Example of a cumulative DVH for a given clinical volume with dose on the x-axis and percentage volume on the y-axis. DVH inspired by Mayles P. et al. [9]

From the DVH, dose statistics can be extracted to describe the dose given to a structure. Maximum dose (D_{max}), minimum dose (D_{min}), mean dose, modal dose and median dose give simplified views of the dose within a structure. Hence, they can be useful for comparing dose plans and clinical results, despite not providing direct anatomical information of the dose distribution. Additionally, they can be used as basis for dose constraints to organs since they give a concrete point of reference and can be compared directly to earlier

cases leading to adverse toxicity [32, 33]. The volume corresponding to a reference dose D_{ref}, $V_{D_{\text{ref}}}$ is also an often-used statistic that gives the percentage of the total volume that received a dose equal to or greater than the user-specified D_{ref}. Additionally, the reverse is also an often-used statistic, the dose corresponding to a reference volume V_{ref}, $D_{V_{\text{ref}}}$, which is the dose that means the specified volume V_{ref} received a dose equal to or larger than the given dose [9].

2.7 Statistical Methods

A clinical study of a patient cohort is a necessary tool to pinpoint the current treatment situation to determine whether altercations to the treatment regime are necessary. This section will outline statistical approaches central for the development and presentation of the results in this thesis. First regression will be outlined, along with methods that can be used in conjunction with regression to estimate the uncertainty in the model. Further, the Wilcoxon signed-rank test is outlined which tests the statistical significance of a matched database. Finally, the Monte Carlo (MC) method will be outlined, which is an algorithm that utilises randomised, statistical approaches to simulate particle behaviour.

2.7.1 Regression

Regression is used to find a relationship between two (or more) variables in the interest of using information from one variable to deduce further trends in the other(s). This is commonly done by fitting a model, a mathematical function, to a set of data points to minimise the error between predicted values and the true data points while preserving the general trend in the data [78].

Data points are sampled from a total population, and as such, the true distribution can only be determined with a certain accuracy. This is accounted for in regression by defining the model's parameters within a confidence interval (CI). A CI estimates how likely one is to find the true mean close to the predicted mean from a sample. The CI is associated with a given percentage, with 95% and 68% being common values, corresponding to the percentage of points within the first and second standard deviations of a perfectly normally distributed sample, respectively. The CI is then interpreted as the interval containing the given percentage of results after an experiment has been repeated several times [78].

The CIs can be estimated with several different approaches. The profile-likelihood method finds the log-likelihood of the parameters, which are then adjusted until the log-likelihood decreases below a pre-determined threshold. The threshold values will then yield the CI [78]. A mathematically easier but more computationally intensive method is the bootstrap. Bootstrapping was first developed in 1979 by Bradley Efron [79] and is a relatively brute force approach to establishing a CI. For non-parametric bootstrapping, a random sample is chosen with replacement from the data set and a model is fitted based on this random sample. The process is repeated several times, typically 1000-2000 iterations, which ultimately yields a distribution of parameters. The CI is determined at the relevant percentile cutoff of these parameters. Parametric bootstrapping differs from non-parametric bootstrapping by assuming the sampled data is a good representation of the population the sample is extracted from. Thus, synthetic patients derived from the distribution of the model are used to build the bootstrapped models instead of random selections of data points from the original data set [78, 79, 80].

In addition to estimating CIs, the bootstrap can also be used for bias estimation. Bias describes the difference between the predicted and true value of a data point, which bias correction can then help correct for. Median bias correction is used to shift the calculated median towards the true median of the population, which can be done with the bootstrapped median by

$$M_{\text{Bias Corrected}} = 2M_{\text{Sample}} - M_{\text{Bootstrap}} \tag{2.26}$$

where M is, from left to right, the bias-corrected median, the median from the original sample and the average median from the bootstrapped samples. Median bias correction shifts the sample median in the opposite direction of the bootstrapped median, which will move the median closer to the true population median, but at the risk of overshooting and underestimating the true value [81].

2.7.2 Wilcoxon Signed-Rank Test

A Wilcoxon signed-rank test is a non-parametric test used to determine whether the difference between the means of two groups of data is statistically significant [78]. The test applies to datasets where the data points are structured in matched pairs. The absolute differences between the matched pairs are found and ranked from lowest to highest. The sum of all positive ranks and negative ranks are then calculated, with the absolute lowest of

these two sums being called the test statistic. Suppose the null hypothesis (the hypothesis that there is no noticeable difference between the two means) is true. In that case, all rank distributions are equally likely, which means that the most probable test statistic would be close to the mean of the total sum of ranks. If the test statistic is small, the sum is less likely to be explained by the null hypothesis, with the probability of this test statistic translating to a p-value. The p-value can be found mathematically by first finding a z-score through

$$z = \frac{s - \mu}{\sigma} = \frac{s - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \tag{2.27}$$

where s is the test statistic, μ is the mean, and σ is the standard deviation of the values, calculated by the respective formulae in the equation, where n is the number of matched pairs in the test [78]. The z-value gives the number of standard deviations between the null mean and the new mean. The p-value is then given by the area under the z-curve outside the standard deviations between the two means. For example, $z = 2$ translates to 95% of the area covered by the standard deviations and would give a p-value of 0.05. A p-value of 0.05 is a common threshold to use for statistical significance from the test. One commonly distinguishes between one-tailed and two-tailed Wilcoxon tests, where a two-tailed test has been described so far. In contrast, the one-tailed test only accounts for one side of the normal distribution and can be utilised if one knows the nature of the shift and the only interest is in either a positive or a negative difference [78].

2.7.3 Monte Carlo Method

Monte Carlo (MC) methods use random number generators to simulate complex, stochastic systems. Due to the stochastic nature of particle behaviour, particle tracking is a common use of MC simulation [12]. Particles traversing a medium are simulated on a particle-by-particle basis. Since many of the interaction processes a moving particle is involved in produces secondary particles, these are generally simulated following the simulation of the primary particle. The knowledge of a particle's trajectory is known as the particle history, and the accuracy of the simulation is inversely proportional to the square root of the number of particle histories, $1/\sqrt{N}$. The simulation requires a well-defined geometry for the particle to traverse, with the particles simulated one at a time and with small step sizes. The step size is limited by the resolution of the underlying CT image since the probability of particle processes is mainly decided by the properties of the medium. Averaging material composition of adjacent voxels is a possible solution, but

this limits the accuracy of the simulation. Small step sizes and a large number of particles are ideal for accurate simulations but come at the cost of computational intensity and, hence, time-consuming simulations [12, 77].

2.8 Predictive Models

Predictive Models are models designed to take specific input parameters and predict the probability of any given output value. The models have importance within predicting the probability of NTCP from a given treatment scheme, as will be further elaborated during the course of this section. Predictive models are trained on a certain subset of data, such as a patient cohort for NTCP modelling, and will attempt to generalise the information from this subset to a population [82]. Thus, the ideal predictive model will have good accuracy on a large and inhomogeneous set of data, which is characterised by a high rate of correctly labelling positives, true positives, while also correctly labelling negatives, true negatives.

2.8.1 Area Under the Receiver Operating Characteristic

Several operations can be utilised to rate a predictive model, with one example being the area under the receiver operating characteristic (AUROC). A receiver operating characteristic (ROC) curve is a curve that plots the rate of false positives against the rate of true positives for a predictive model. The true positive and false positive rates, or sensitivity and 1 - specificity (true negative rate), are defined as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad \text{and} \quad \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN} + \mathrm{FP}} \tag{2.28}$$

where TPR and FPR are the true positive rate and false positive rate, respectively, TP is the number of true positives, FN is the number of false negatives, TN is the number of true negatives, and FP is the number of false positives [83]. A depiction of a ROC curve is given in figure 2.16.

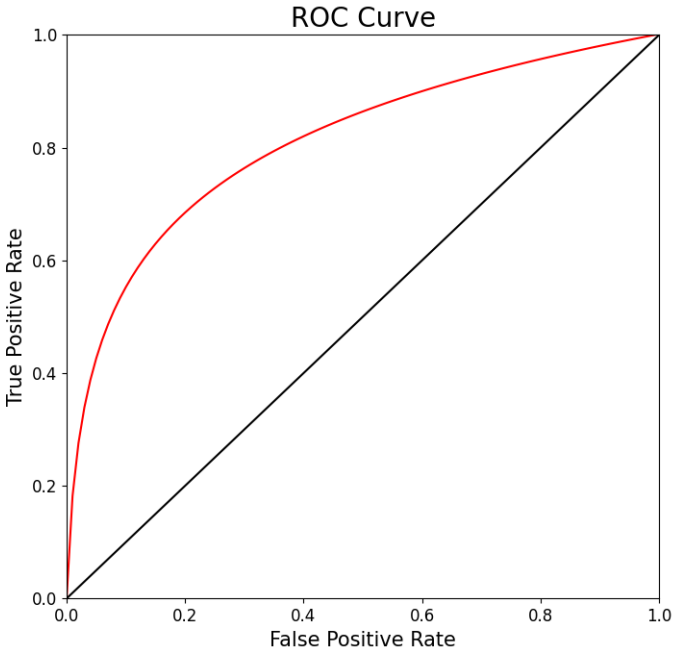

Figure 2.16: Example of a receiver operating characteristic (ROC) curve. The black line represents a model that performs equally well to a random guess, while the red line shows a typical shape of a ROC curve representing a better predictive model. Modified figure from Zou, K. et al. [83]

The black line in figure 2.16 represents the rate of true positives equal to false positives and describes a model that performs on par with a random guess. The red curve shows a typical curve comparing true positives to false positives, and the model's predictive ability improves as the curve approaches the upper left corner of the plot. The ROC curve can be generalised to evaluate predictive models by studying the area under the ROC curve, the AUROC. Thus, an AUROC-value of 0.5 describes a poor predictive model performing on par with a random guess, while an AUROC-value of 1 will describe a perfect model labelling all points correctly [83].

2.8.2 Logistic Regression Models

Logistic regression models, or logit models, are predictive models that utilise regression, as outlined in the previous section, to classify data points [82]. Logistic regression takes in data with at least one input variable (usually continuous) and discrete labels, with the simplest models having only binary labels. NTCP modelling is a possible use of logistic regression, where the data points are individual patients labelled either with toxicity or without toxicity, while the continuous input parameter is a dosimetric value [12]. A weighted linear sum is fit to the data points, which is then converted to a sigmoid function to yield an output between 0 and 1. In binary classification, each label has the binary value 0 or 1, with a threshold at 0.5 determining the label [84]. A sigmoid curve resulting

from logistic regression is illustrated in figure 2.17.

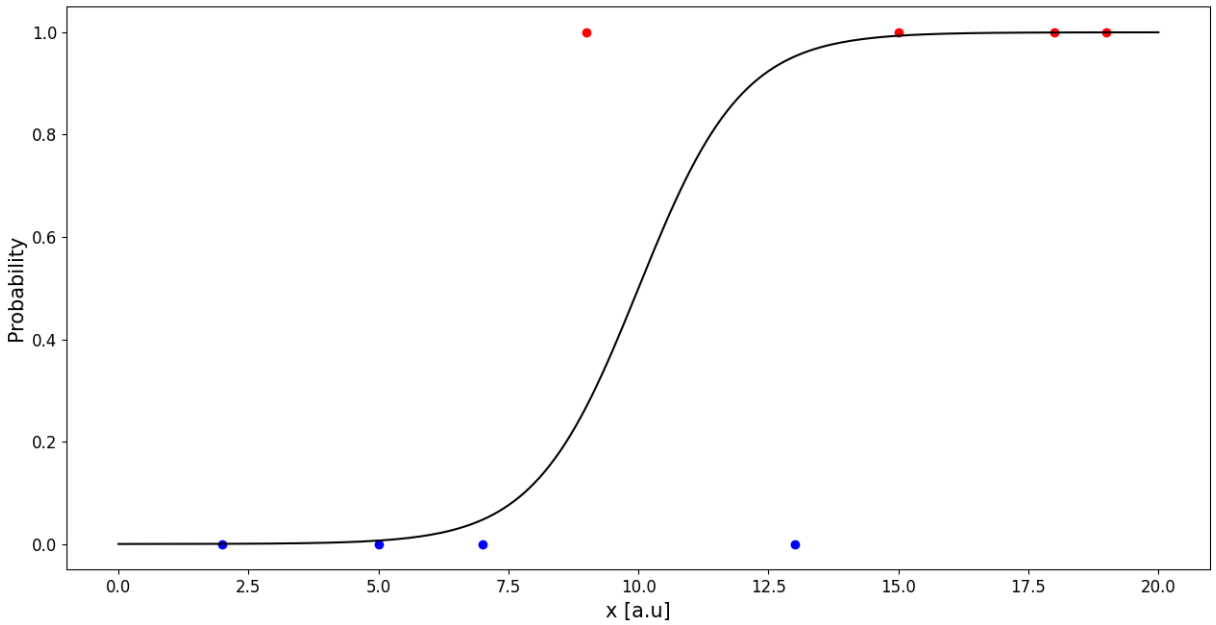

Figure 2.17: Example of a sigmoid curve for logistic regression. Coloured data points share a label of either 0 or 1. The given sigmoid curve would correctly label 6/8 points. Modified figure from Paganetti, H [12].

The example from figure 2.17 would correctly label 6/8 points for a standard threshold of 0.5, and the example had an AUROC-value of 0.94. The sigmoid function is given by

$$f(\mathbf{w^T x} + w_0) = \frac{1}{1 + e^{-(\mathbf{w^T x} + w_0)}} \tag{2.29}$$

where $\mathbf{x}$ is the set of input variables, $\mathbf{w^T}$ are corresponding weights for the input variables, and w_0 is a constant weight independent of all input variables. The summed weights are the parameters that describe the given model and are commonly optimised by maximising the likelihood, or more precisely, the model's log-likelihood, with respect to the fitted data points [82]. The likelihood, given an event, describes the probability distribution most likely to explain the event. In logistic regression, the event is the distribution of the data points' input variables and their corresponding labels, while the likelihood describes the weights resulting in the model with the best ability to fit the data points and label them correctly. The log-likelihood of a binary classifier is described by

$$\mathcal{L}(\mathbf{w}, w_0) = \sum_{i=1}^{N} y_i \log f_i(\mathbf{w^T x} + w_0) + (1 - y_i) \log(1 - f_i(\mathbf{w^T x} + w_0)) \tag{2.30}$$

where $\mathcal{L}$ is the log-likelihood, N is the total number of data points to be fitted, $f_i(\mathbf{w^T x} + w_0)$ is the solution of the sigmoid described in equation 2.29 for data point i, and y_i is the

label of data point i [78, 82]. Since the label, y_i, is discretely 0 or 1, equation 2.30 will have one term cancelled out for each data point. The log-likelihood can then be maximised by setting the first derivative equal to 0 and solving, which is generally done with iterative methods since the derivative of the log-likelihood does not have a closed-form solution [82].

One iterative approach that can be used to maximise the log-likelihood is gradient descent [84]. Gradient descent starts by calculating the weights at an arbitrary starting point and the derivative is subsequently found and set to 0 to find the point of maximum ascent. The derivative is then subtracted from the original weights. This process is repeated until a local minimum is found. Mathematically gradient descent is outlined as

$$w_{j,k} = w_{j,k-1} + \Delta w_{j,k-1}, \qquad \text{where} \quad \Delta w_{j,k-1} = -\eta \frac{\partial \mathcal{L}}{\partial w_{j,k-1}} \tag{2.31}$$

where $w_{j,k}$ is the weight corresponding to input variable j, after gradient descent iteration k, while $w_{j,k-1}$ is the weight from the previous iteration $(k-1)$. $\Delta w_{j,k-1}$ is the update to the weight, which is dependent on a manually determined step size, η, quantifying how much each iteration should update the weight, and the partial derivative of the log-likelihood ($\mathcal{L}$) with respect to the weight $w_{j,k-1}$ [84]. For binary classification, the weight updates, $\Delta w_{j,k}$, have the forms

$$\Delta w_{j,k} = \eta \sum_i (y_i - f_i(\mathbf{w^T x} + w_0)) x_{i,j} \tag{2.32}$$

and

$$\Delta w_{0,k} = \eta \sum_i y_i - f_i(\mathbf{w^T x} + w_0) \tag{2.33}$$

2.8.3 NTCP Models

Shackell et al. found as early as 1924 that biological effect relative to toxin dosage translates to a roughly S-shaped curve [85], which is similar to the biological effect of radiation. Dose-response curves are, thus, often given as sigmoid shapes, and logistic regression is a normal basis for what is often referred to as NTCP models [12]. NTCP models use patient data to determine the relationship between the treatment scheme's dosimetric (and non-dosimetric) characteristics and a given adverse clinical endpoint [12]. The *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis*

(TRIPOD) includes a set of guidelines used to guide the validity of multivariable NTCP models and universalise the models between clinics [86].

DVH Reduction Models

DVH-reduction models are NTCP models that compress information from DVHs into a singular representative metric. Since DVHs contain no spatial information of the radiation treatment, DVH-reduction models are based on estimated complication probability for uniform radiation. The difference in the volume of the organs can be considered by expressing tissue tolerance as a power law of fractional volume irradiated, as given in

$$D(V_{\text{irradiated}}) = D(V_{\text{reference}}) \left(\frac{V_{\text{reference}}}{V_{\text{irradiated}}} \right)^n \tag{2.34}$$

where $V_{\text{reference}}$ is a reference volume, $V_{\text{irradiated}}$ is a uniformly irradiated volume, D represents the tolerance doses of the respective volumes, and n accounts for how highly small irradiated volumes impact the final dose calculation [36]. Small n-values translate to changes in irradiated volume having a minor impact on the tolerance doses, while large n-values mean the tolerance dose is highly dependent on irradiated volume fraction [12]. Lyman used equation 2.34 as a basis to quantify risks associated with partial organ volume uniform irradiation in what would later come to be known as the Lyman-Kutcher-Burman (LKB) model [36, 87].

The Lyman-Kutcher-Burman Model

For uniform irradiation, the dose-response is described by a mathematical function with a minimum of two parameters. The LKB model, which is both one of the first and one of the most often used models [12], incorporates the parameters TD50, the dose corresponding to a 50% complication probability and m, a constant describing the slope of the response curve [36]. Partial volume irradiation is also accounted for by incorporating the volume parameter, n [87]. The n parameter is considered through the generalised equivalent uniform dose (gEUD), which is defined as the dose that, if given uniformly to the entire organ, would result in the same probability of complication [36]. Mathematically, gEUD is expressed as

$$\text{gEUD} = \left[\frac{1}{N_{\text{Voxels}}} \sum_{i=1}^{N_{\text{Voxels}}} d_i^{\frac{1}{n}} \right]^n \tag{2.35}$$

where N_{Voxels} is the number of equally sized voxels and d_i is the dose to voxel i. The LKB model is given in terms of gEUD as

$$\text{NTCP} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-x^2}{2}} dx, \quad \text{where} \quad t = \frac{\text{gEUD} - \text{TD50}}{m \times \text{TD50}} \tag{2.36}$$

The parameters of LKB models differ significantly based on both organ and clinical endpoint. The most apparent parameter difference is in the n parameter, which is highly dependent on the seriality of the organs. Parallel organs, such as the liver, lung and kidney, are organs where the components of the organ function reasonably independently of each other and, as such low irradiated volumes cause minor damage to the organ. Hence, complications in these organs are usually described by LKB models with large volume effects, i.e. $n \approx 1$ [36]. Serial organs, such as the rectum and the brainstem, on the other hand, are highly dependent on the entire organ to function and as such, only a minor volume of the organ irradiated will cause significant damage. Thus, they are described by LKB models with smaller volume effects, for example $n \approx 0.1$ for late rectal bleeding [36] and n-values of 0.05 [87] and 0.16 [88] for brainstem toxicity. Comprehensive lists of LKB parameters for several organs and endpoints have been collected and developed by Lyman in 1985 based on published data at the time [87] and Burman et al. in 1991 [88] based on tolerance doses compiled by Emami et al. [89] the same year. However, both parameter sets are limited by the data available at the time.

Functional Subunit-Based Models

Several NTCP models have been established based on the seriality of organs, with the initial three being the critical element model [90], the critical volume model [91] and the relative seriality model [92]. The critical element model assumes that dose to any subvolume of the organ will result in its failure and describes the response of serial organs. On the other hand, the critical volume model accounts for parallel organs by incorporating the organ's spare capacity. As long as the volume irradiated does not exceed the capacity, the dose will not result in organ failure. Finally, the relative seriality model incorporates an adjustable fitting parameter that results in a flexible description of the organ's degree of seriality [12].

Multiparametric Models

DVH reduction models compress the data from a DVH to a single metric, which is at risk of oversimplifying the distribution and yield similar metrics for significantly differ-

ent DVHs. An alternative could thus be multiparametric models that consider several parameters in predicting toxicity, typically a combination of both dosimetric and non-dosimetric parameters [36]. Logistic regression is a common method of fitting the models to the data, and there are several different techniques in establishing what parameters to utilise. Examples of parameter choice methods range from experienced manual choices, to determining the parameters that maximise the performance of the models through statistical methods or machine learning (ML) approaches [36, 93, 94].

2.9 Machine Learning

Machine learning (ML) has in recent years seen considerable development within the field of radiation oncology [95], and further improvements are likely to proceed. The regression and predictive modelling outlined in the previous sections are themselves examples of ML [84], which is, by definition, the optimisation of a computer algorithm based on a combination of past experience and example data [84]. ML can be roughly divided into supervised learning and unsupervised learning. Supervised learning has known labels for the computer to predict and a set of features for the computer to determine which have the most robust patterns with the labels. On the other hand, unsupervised learning provides the computer with a set of features as well but differs from supervised learning by relying on the computer to establish patterns in the data [84].

2.9.1 Clustering

Clustering is an unsupervised ML algorithm and describes the process of applying an algorithm to make a computer sort a set of data in groups based on specified parameters [84]. Clustering evaluates a set of data point coordinates, separates the data points tightly clustered together in dense areas and assigns the data points specific labels [84]. Examples of commonly used clustering algorithms include k-means clustering, hierarchical clustering and density-based clustering approaches.

k-means Clustering

As implied by the name, k-means clustering is a clustering algorithm that asks the user for manual input of the desired number of clusters, k, and the computer will then sort the data points into the most fitting of these clusters [82]. A cost function is defined as the

variance of the clusters multiplied by their size and is given by

$$C(D_1, ..., D_k) = \sum_{j=1}^{k} |D_j| Var[D_j] = \sum_{j=1}^{k} \sum_{x_i \in D_j} ||\vec{x} - \mu_j||^2 \tag{2.37}$$

where D_j represents the different clusters, $\vec{x}$ is a specific data point, and μ_j is the mean vector of points belonging to cluster j. The computer aims to find the cluster partition that minimises the value of this cost function, and a commonly used algorithm is Lloyd's algorithm, which is also referred to as Voronoi tesselation. Lloyd's algorithm chooses k random initial cluster centroids and clusters the data points accordingly. The centroids are subsequently updated to the centres of the new clusters, and the data points are relabelled. The process continues iteratively until the relabelling of the data points yields no changes in the clustering, which then gives the final clustering [82].

Hierarchical Clustering

Unlike k-means clustering, hierarchical clustering does not require the user to manually select the number of clusters one wants out of the clustering. Instead, in hierarchical clustering, the computer establishes a hierarchical order of differently sized clusters, and the user must then determine the ideal number of clusters based on different quality measures [82]. There are two main categories of hierarchical clustering, agglomerative clustering and divisive clustering.

Agglomerative clustering is a "bottom-up" approach where each data point is initially established as an individual cluster. The clusters are then consecutively merged based on proximity from a specified linkage criterion until all data points are gathered in one large cluster [82]. Divisive clustering describes the opposite approach where the basis is one large cluster containing all data points, which is then iteratively split apart until one has n clusters left with a single data point in each cluster. The clusters are formed based on a given linkage criterion, and there are several such criteria available, with certain examples being described here. Single linkage is defined as the distance from a data point to its closest neighbour. Alternatively, complete linkage is the distance from a data point to its furthest neighbour. Average linkage considers the average distance between all data points in the two clusters. Finally, centroid linkage is the distance from the centre of one cluster to the centre of an adjacent cluster [82].

Data from hierarchical clustering is often presented in a dendrogram, a graphical representation of the clusters as shown in figure 2.18.

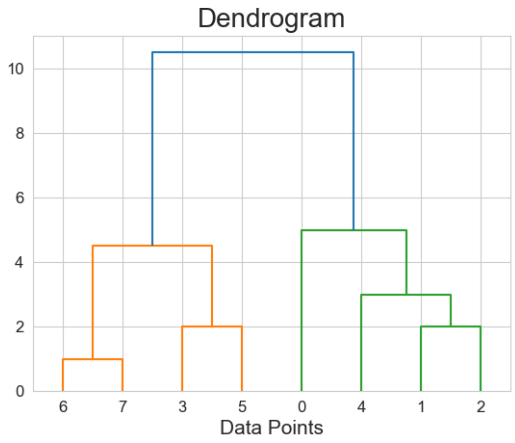

Figure 2.18: Example of a dendrogram. The y-axis shows radial distance between clusters, while the x-axis shows the data points in the clustering, referred to as the *leaves*. Figure inspired from Alpaydin E. et al. [84]

The dendrogram has labels for all data points on the x-axis, referred to as *leaves*. The y-axis gives the radial length between clusters. Hence, the dendrogram gives a visualisation of the distance where one cuts the tree in the horizontal direction, which yields the resultant clustering from the algorithm [82].

Density-Based Spatial Clustering of Applications with Noise

Density-based spatial clustering of applications with noise (DBSCAN) is an alternative clustering algorithm and, as the name implies, is prone to cluster dense regions of data points. In contrast, less dense regions are excluded from being clustered and are instead labelled noise. DBSCAN requires two parameters, the distance between two points that would qualify them as part of the same cluster, referred to as an Epsilon-neighbourhood, and a minimum amount of points required to form a cluster. These criteria exclude outliers from being included as clusters, which separates the algorithm from k-means clustering and hierarchical clustering and makes the algorithm a strong choice if one has a database with known outliers [96].

The choice of parameters for the DBSCAN is a central aspect of running the algorithm. A method of determining the ideal Epsilon-neighbourhood for the DBSCAN is to find the point of steepest slope on the k-distance graph of the dataset [96, 97]. k-distance refers to the distance from a point to its k-th nearest neighbour, and the point of inflection in the resulting k-distance graph will indicate the ideal Epsilon-neighbourhood. The k-distance graph plots distance versus data points with k neighbours within the specified

distance. Examples of k-distance graphs are shown in Appendix D. The choice of k has been suggested be set to twice the dimensions clustered minus one [98]. The minimum number of points within a cluster was observed to not change significantly over a value of 4 for 2-dimensional data in the original paper [96], which was generalised to a basis of $2 \cdot$ dimensions, or more precisely $2k + 1$ for higher dimensional data [98].

Silhouette Score

For all clustering algorithms described in this section, a metric to rate cluster quality is essential. Several metrics can be used for this purpose, with one example being the silhouette score of the clusters. The silhouette score is a measure that accounts for the similarity of data points within a cluster and the dissimilarity of data points between clusters. The silhouette score of a single data point is given by

$$s = \frac{B - A}{\max(A, B)} \tag{2.38}$$

where A is the average distance between the data point to the other data points in its cluster, and B is the distance between the data point and the average distance to each point within the closest neighbouring cluster. The silhouette score is normalised by the maximum of the two values, $\max(A, B)$. An average silhouette score between all the data points is calculated and is used to describe the total quality of the clustering. Hence, the silhouette score will always lie between -1 and 1, where a value below 0 signifies a poor clustering, while a score between 0 and 0.5 means the clustering is acceptable. Still, one usually wants to achieve a silhouette score of higher than 0.5, which translates to a solid clustering [99].

2.10 Brainstem Necrosis

Brainstem necrosis forms the clinical endpoint of the cases in this study, and this section is dedicated to a description of the condition. Necrosis refers to the irreversible injury and subsequent death of cells from external stimuli such as bacteria, toxins or radiation. Hence, necrosis describes the result of several different mechanisms of cell death [100]. Necrosis of brainstem tissue is a condition that can occur following radiotherapy treatment of tumours in the brain and has in certain studies been observed at greater rates for proton therapy compared to photon therapy [101]. Brainstem necrosis is graded from 1 to 5

according to CTCAE v5.0 guidelines for CNS necrosis, where grade 1 translates to a mild or asymptomatic condition, grade 2 to mild symptoms, grade 3 to moderate symptoms, grade 4 to a critical condition requiring urgent intervention, and grade 5 to death [18, 54, 102].

Chapter 3

Materials and Methods

This chapter will outline the acquisition of the data material that formed the basis for the work performed in this thesis, and the subsequent methods used to process the material. The origin of the patient cohort will be explained, as well as the recalculation of dose and contouring of the clinical volumes. Further, the investigation and statistical analysis of the cohort will be outlined, and the methods behind the NTCP modelling and the cluster analysis based on the cohort will also be covered.

3.1 Patient Cohort

The database used in this thesis consisted of DVHs and LET_d distributions from 36 anonymised paediatric patients part of a case-control study from an ongoing PhD project at the University of Bergen. The cohort was extracted from 954 paediatric patients treated with double-scattering proton therapy for posterior fossa tumours at the University of Florida proton centre between 2006 and 2017. Nine of the included patients developed symptomatic CTCAE grade 2+ brainstem toxicity following treatment and are, thus, the cases included in the study. These were each matched to three control patients based on diagnosis, adjuvant therapy, age ($\pm$1.5 years), as well as the maximum dose ($D_{0.1cc}$) and $D10\%$ to the brainstem (both $\pm$2 Gy(RBE)).

As part of the PhD project, dose distributions and LET_d distributions had been re-calculated using FLUKA MC code with a model of the double scattering nozzle used for treatments at the University of Florida proton centre [103]. Dose distributions were recalculated with LET_d-weighted doses [70], the McNamara RBE model [21], and the Rørvik weighted RBE model [22]. The former incorporated weighting factors $c = 0.04$ and $c = 0.055$ [70, 71]. The latter two models assumed constant $(\alpha/\beta)_\text{x}$-ratios for the full brainstem volume, and were recalculated with $(\alpha/\beta)_\text{x} = 2.1$ [17], $(\alpha/\beta)_\text{x} = 2.5$ [31] and $(\alpha/\beta)_\text{x} = 3.3$ [104], as $(\alpha/\beta)_\text{x}$-ratios associated with the brainstem from previous studies. The latter two models were chosen since they are based on a significant number of low $(\alpha/\beta)_\text{x}$ and LET values, while LET_d-weighted dose was chosen as a model independent

of the uncertain $(\alpha/\beta)_x$ of brainstem tissue [105]. 2–5 treatment fields were simulated per patient, with 600 million primary protons per treatment field, resulting in a statistical uncertainty of less than 1% in the PTV. The MC calculated doses were normalised to the same CTV coverage achieved during the initial treatment planning, which resulted in scaling factors for the total dose between -3.2% and 2.4% [103].

The substructures of the brainstem for all 36 patients had been contoured in relation with the PhD project. First, the brainstem was divided into its three primary substructures, the medulla oblongata, the midbrain and the pons. The pons was further delineated into four components; the posterior, the middle posterior, the middle anterior and the anterior, to approximate transversal fibre tracts. Additionally, the outer 3 mm of the brainstem was separated from the brainstem core, although the surface and core were not volumes focused on in this thesis, since the volumes are already clinically accounted for [33]. All delineated substructures are shown in figure 3.1.

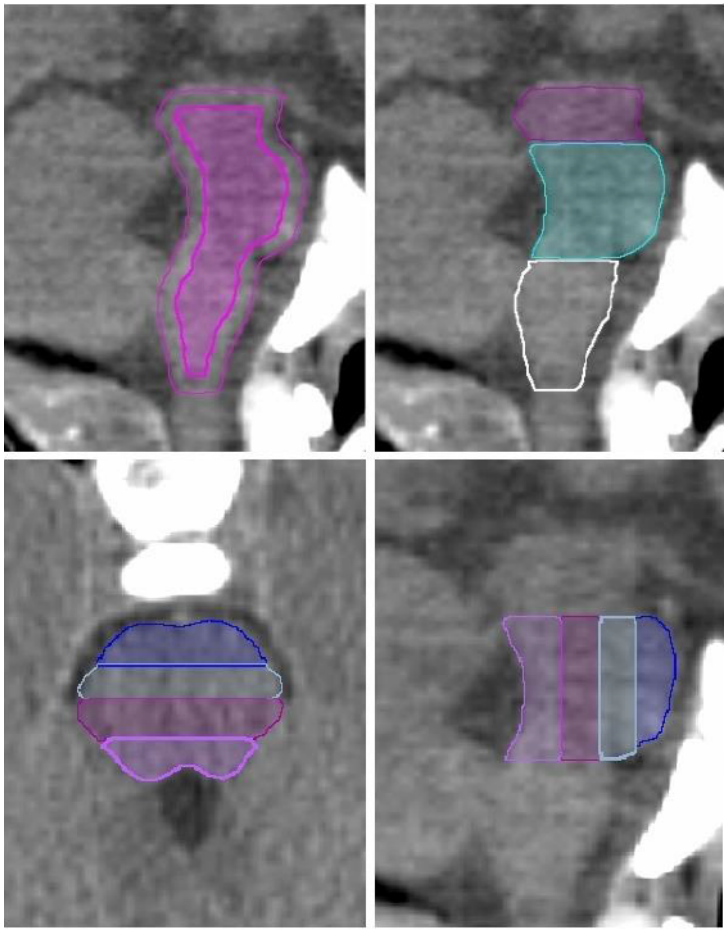

Figure 3.1: Brainstem substructures. Top left shows the brainstem core versus brainstem surface (inner versus outer structure). Top right shows the midbrain (pink), pons (turqoise) and medulla oblongata (white). The lower two images show a transversal view (bottom left) and a sagittal view (bottom right) of the four substructures of the pons: posterior (blue), middle posterior (white), middle anterior (dark pink) and anterior (light pink). Figure courtesy of Stokkevåg, C et al. [106].

3.2 Extracted Raw Data

The simulated dose distributions and LET_d distributions for all RBE calculations were extracted from FLUKA as DICOM files (.dcm) and exported to Eclipse TPS. The dose distributions of all patients were combined with the delineated volumes from CT and MRI images of the corresponding patients. DVHs and LET_d-volume were constructed with Eclipse TPS for all combinations of delineated subvolumes and recalculated RBE-weighted doses, and were then extracted from the system as text files (.txt). Isodose images were also studied in Eclipse TPS to illustrate the dose and LET_d distributions of the patients since, in the extraction of DVHs, the spatial information is limited to the delineation of the substructures. The LET_d had a cut-off threshold at 1 Gy(RBE) to avoid excessive impact from areas with insignificant dose, but moderate LET_d. This threshold was utilised in the study of isodose images, the LET_d-volume histograms and the LET_d statistics, which will be described later in this chapter.

3.3 Average Dose Volume Histograms

Average DVHs were constructed from the DVHs of the cohort through an in-house *Python* script [107], which takes as input the text files with DVH information from Eclipse and outputs a similar text file, but averaged over all cases and all controls. The text file was then plotted with a *Python* script made by the author of this thesis with imported functions from the Python library *Matplotlib* [108] for the plotting and the libraries *Pandas* [109] and *Numpy* [110] for data handling. Average LET_d-volume histograms were constructed to study the difference in LET_d distributions of the cases and controls. Additionally, DVHs comparing LET_d-weighted doses, McNamara RBE-weighted doses and Rørvik RBE-weighted doses with RBE1.1-weighted doses for cases and controls were also made. The full brainstem, the substructures medulla oblongata, midbrain and pons, and the further delineated pons were included in both LET_d-volume histograms and DVHs.

3.4 Statistical Analysis

The $D50\%$, $D10\%$ and D_{max} were the primary dose statistics studied since $D50\%$ represents a significant fraction of the total volume, $D10\%$ represented the approximate volume that maximised the difference between cases and controls as observed from DVHs, while D_{max} quantifies the effect of hot spots in the volumes. D_{max} is here defined as the dose to a volume 5 mm $\times$ 5 mm $\times$ 3 mm = 0.075 cm^3, as determined by the resolution of the images in Eclipse TPS. The volumes studied were the same that were outlined in the pre-

vious paragraph and RBE-weighted dose from McNamara, Rørvik, and RBE1.1 formed the dosimetric basis. A constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ was assumed for the full brainstem where applicable. From the DVHs made in Eclipse TPS the dose statistics $D50\%$ and $D10\%$ were calculated, while the $D_{\max}$ was provided in and read directly from the .txt files with the DVH information. gEUD was also studied with volume parameters n established previously by Lyman [87] and Burman et al. [88], $n = 0.05$ and $n = 0.16$, respectively. gEUD was calculated with equation 2.35, by using an in-house *Python* script [107]. $\mathrm{LET_d}$ statistics were also investigated, specifically the $L50\%$, the $\mathrm{LET_d}$ to 50% of the total structure volume, and the maximum $\mathrm{LET_d}$ ($L_{\max}$). The variation in dose statistics within and between individual case-control groups was studied and summarised as the average case and control patient with 95% CIs. The $D50\%$, $D10\%$ and $D_{\max}$ to the full brainstem volume were compared to the dose constraints set for the full brainstem at the University of Florida proton centre [33].

The cases were further compared to their average matched control by studying the difference between them, as given mathematically by

$$\text{tox } n - \frac{c_{n.1} + c_{n.2} + c_{n.3}}{3} \quad \text{for} \quad n \in [1, 9] \tag{3.1}$$

where tox n refers to the case of group n, while $c_{n.i}$ refers to its corresponding matched controls. Subsequently, an average was calculated for the differences of the nine case-control groups. A Wilcoxon signed-rank test was utilised to determine statistical significance of the differences, with a standard threshold of $p = 0.05$ used to signify statistical significance. Statistical significance is noted in tables where the differences are provided by an asterisk (*). The calculations and the Wilcoxon signed-rank test were both done in *Microsoft Excel*. The differences were found for the full brainstem, the medulla oblongata, the midbrain, the pons and the pons' substructures. The dose and $\mathrm{LET_d}$ statistics outlined in the previous paragraph include all the differences calculated, with the sole exception of $(\alpha/\beta)_{\mathrm{x}}$-ratio, where the effect of utilising a higher $(\alpha/\beta)_{\mathrm{x}} = 3.3$ was also investigated in these calculations.

3.5 Lyman-Kutcher-Burman Modelling

LKB models were fitted to the dataset through the use of an in house *Python* script [107], which processes DVHs, allows the extraction of dose statistics and using dosimetric patient data to establish parameters used for NTCP modelling. The models were used as binary classifiers with either toxicity or no toxicity as the outcome, and the clinical

endpoint was brainstem necrosis CTCAE grade 2+. LKB models were fitted based on the gEUD of the cohort with volume parameters $n = 0.05$, $n = 0.16$ and $n = 0.25$. The parameters $n = 0.05$ and $n = 0.16$ were established in previous studies by Lyman in 1985 [87] and Burman et al. in 1991 [88], respectively. The final parameter, $n = 0.25$, was chosen as an investigative value, since both $n = 0.05$ and $n = 0.16$ are relatively low estimates of the parameter.

The slope parameter (m) and the tolerance dose for 50% of the patients (TD50) were determined parametrically by establishing an initial model and using gradient descent, as outlined in section 2.8.2, to find a local maximum for the log-likelihood of the model. The gradient descent was performed using an algorithm called the Truncated Newton method. The method uses an iterated optimisation algorithm to minimise the log-likelihood of a function. The initial and subsequent parameter estimates are then updated by adjusting the parameters towards the gradient of decreasing log-likelihood error [111]. Additionally, to avoid local minima hiding a global minimum, a basin-hopping procedure was applied. Basin-hopping translates the identified local minimum in a random direction, before a new gradient descent search is started that might identify another lower local minimum [112]. The LKB parameters were constrained to relatively wide intervals encompassing parameters for previously established models, with m in the interval [0.03, 1.00] and TD50 in the interval [10 Gy(RBE), 200 Gy(RBE)]. Initially, the volume parameter n was also established by gradient descent. However, the log-likelihood consistently improved for n-values approaching 0, making the defined CIs insignificant since the parameter was overrepresented on the lower threshold. Hence, the choice was made to determine the parameter by literature instead.

The two parameters were defined within 95% CIs by performing non-parametric bootstrapping in 1500 iterations on the data, which was also performed using an in-house *Python* script [107]. The models were fitted based on the bootstrapped samples, as outlined in section 2.7.1, and each parameter associated with the model were subsequently assigned their respective CIs. The specific number of iterations was chosen since the parameters and CIs stayed reasonably consistent between runs, while the computing time was limited to around two minutes. Subsequently, the CIs were used to approximate the true parameters of the LKB models by applying a median bias correction. The bootstrapped parameters were compared to the original parameters based on the sample, and the parameters were shifted as outlined in section 2.7.1. However, the shift in the parameters after bias correction was minor so there was limited risk of significantly overshooting

the true value.

The gEUD of the cohort based on the variable RBE models by McNamara et al. [21], Rørvik et al. [22], and RBE1.1 were the primary basis for the modelling, but modelling solely based on the cohort proved difficult and yielded models with wide CIs. Since outliers were prone to account for a significant percentage of the patients in the bootstrapped cohorts, and the patient cohort was mostly concentrated in a limited dose range, synthetic constraints were introduced. The synthetic constraints consisted of controls at gEUD = 20 Gy(RBE) and cases at gEUD = 100 Gy(RBE) and were introduced to avoid unrepresented high and low doses too heavily impacting the model curvature around the TD50. The lowest weighting utilised was two controls at gEUD = 20 Gy(RBE) and two cases at gEUD = 100 Gy(RBE), while models with nine of each were also investigated. This was implemented in the *Python* script [107] as the ratio of the chosen weight divided by the total number of patients in the cohort. Thus, the first weight would correspond to $2/36 \approx 0.056$. The weights were minimised to the greatest extent possible to ensure the models were mostly fitted to the original patient cohort instead of the synthetic constraints. This resulted in the full brainstem, midbrain and pons using two patients at each constrained point. On the other hand, the medulla oblongata and the pons substructures used nine of each respective constraint.

The validation of the models was based on AUROC-values, with all toxicity thresholds along the sigmoid curves as the basis for the ROC-curve. The principles behind ROC-curves and AUROC are outlined in section 2.8.1, with the AUROC-value being a commonly used metric for evaluating the performance of a predictive model. The AUROC was found with a *Python* script plotting the ROC-curve, using the *Pandas* [109] and *Numpy* [110] libraries for data handling. The *Sci-kit Learn* library [113] was, subsequently, used to calculate the AUROC.

3.6 Cluster Analysis

Dose statistics of the patient cohort were also used as the basis of cluster analyses, to investigate if cluster analysis could discover systematic differences between the cases and controls of the cohort. The clustering was performed using *Python* scripts with imported packages from *Sci-kit Learn* [113]. The two clustering algorithms utilised were DBSCAN and hierarchical clustering and a range of dose statistics were clustered, with selected clusterings being presented in this thesis.

DBSCAN was chosen as a clustering algorithm because certain patients, or even whole case-control groups, differed significantly from the rest of the cohort. By labelling the outliers as noise, the hypothesis was that similar points could establish more robust clusters than the hierarchical clustering. The minimum points within an individual cluster were set to one more than the number of parameters tested, which does not follow the ideal number of clusters set by Sander et al. [98], but maintained the notion of the threshold being higher than the dimensions of the data, while allowing smaller clusters to form. This was considered beneficial since the database was of limited size. The minimum distance between two points within a cluster was determined by a k-distance graph, with k-distance graphs used for the clusterings presented in this thesis given in appendix D. A k-value of 2 was used, but the graphs and points of inflection stayed close to constant between different values of k. An interval within which the inflection point could be found was determined by studying the graph manually. Subsequently, the distances within the interval were explored until a distance yielding a maximum in silhouette score for the clustering was found. The silhouette score was also compared between the different clusterings to compare the quality of the clusters. Standard euclidean distance was the chosen distance metric utilised.

For the specific DBSCANs included in this thesis, the parameters established were as follows. The DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons used four minimum samples to form a cluster, one plus the number of dimensions clustered. The McNamara and Rørvik clusterings used a nearest neighbour distance, Epsilon, of 1.00 determined through studying the k-distance graph and the silhouette scores of the resultant clusterings. For the RBE1.1 clustering, a nearest neighbour distance, Epsilon, of 0.8 was chosen through the same method. The gEUD of the midbrain and pons with $n = 0.05$ used a minimum points to form a cluster of three. The nearest neighbour distance, Epsilon, of 0.80 was chosen for all three RBE weighted doses, decided from a combination of the k-distance graph and the silhouette scores.

The hierarchical clustering is not as robust to outliers as the DBSCAN. However, an advantage of the algorithm is less required input from the user and its ability to give an overview of all possible cluster combinations. The hierarchical clustering used the silhouette score to determine where to set the threshold for the result of the clustering, and the silhouette score also gave an estimate of the quality of the clustering. The clustering was an agglomerative approach and used the centroid linkage method to cluster the data

points.

Chapter 4

Results

The first results given are the isodose curves studied for selected example patients from the cohort. LET_d-volume histograms and LET_d statistics follow. Subsequently, average DVHs and dose statistics will be presented, before the study of the gEUD is provided. The next section is dedicated to LKB models fitted for the cohort before finally, the results from the cluster analyses are shown.

4.1 Isodose Curves

Both dose and LET_d distributions of the patients varied considerably across the case-control groups, mainly because individual case-control groups were matched based on different diagnoses. Notably, most of the patients in case-control groups one and two had a treatment field that did not cover the medulla oblongata, which can be seen from both dose and LET_d distributions of the case from group one in figure 4.1 (since the LET_d was subject to a cut-off at 1 Gy(RBE)). Otherwise, the LET_d was relatively low in large parts of the brainstem for the majority of the patients, which is a consequence of the treatment planning attempting to avoid overlap of the brainstem and the distal end of the Bragg peak. However, while average LET_d was reasonably low as shown in LET_d statistics presented in this chapter, certain outliers received considerable LET_d to parts of the brainstem volume. For the case from group six, the distal end was unable to avoid the brainstem altogether, and the patient ended up receiving LET_d in the $6-8$ keV/µm range to both the midbrain and the pons posterior. The patient also received $D_{\max}$ approaching 70 Gy(RBE), exceeding average $D_{\max}$ for the cohort to the full brainstem, which is also presented in this chapter.

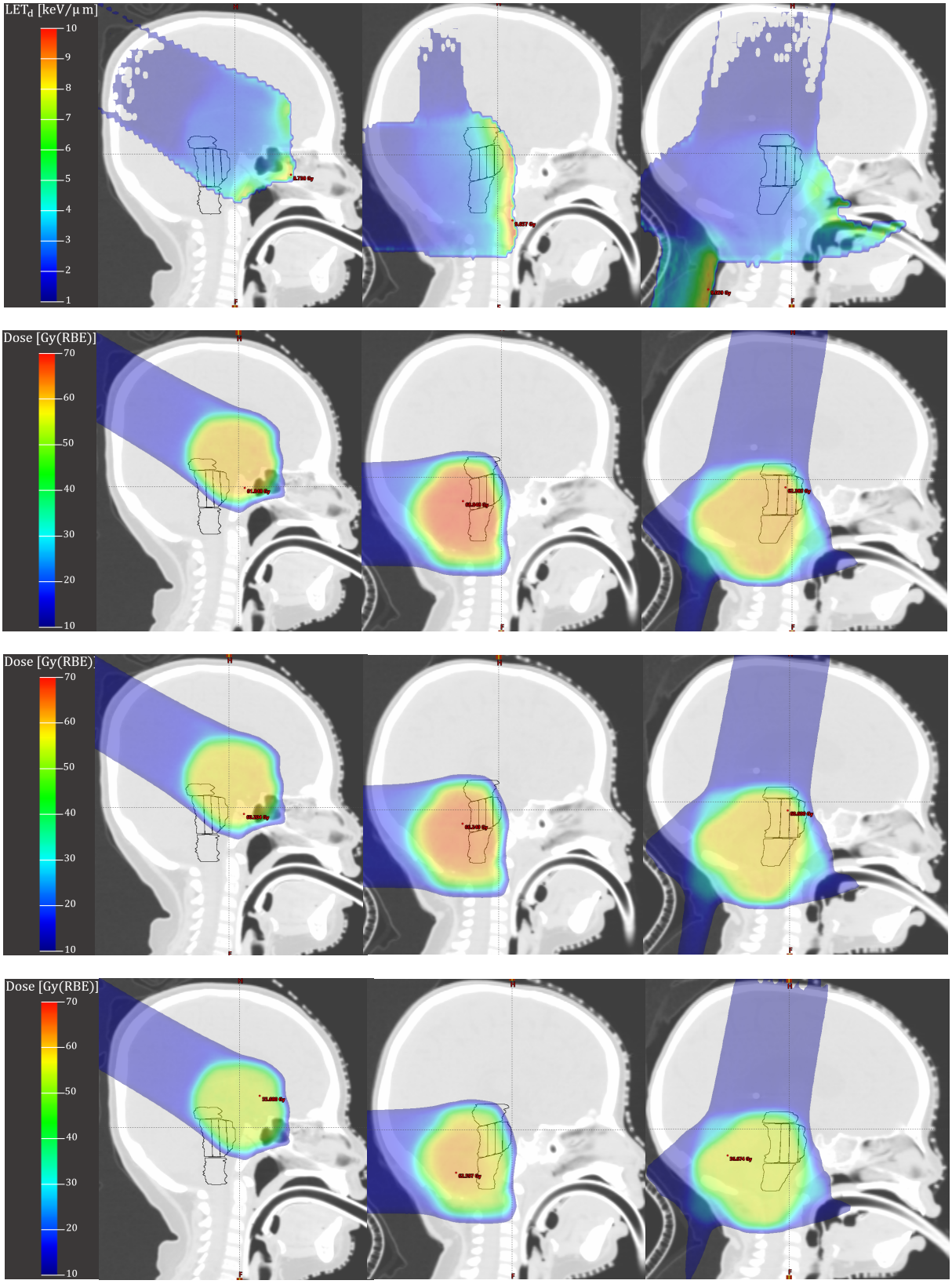

Figure 4.1: CT images showing dose and LET_d distributions from FLUKA MC recalculation using RBE1.1, and the McNamara and Rørvik RBE models, shown in Eclipse TPS. The left shows the case from group 1, the middle shows the case from group 6, and the right shows a control from group 9. From top to bottom row are LET_d distributions, dose distribution from the McNamara model, dose distribution from the Rørvik model and dose distribution from RBE1.1.

4.2 Linear Energy Transfer

Average LET_d-volume histograms constructed for the brainstem showed that, on average, cases received slightly higher LET_d than the controls, particularly in the $2-8$ keV/µm range. The brainstem is shown along with its substructures in figure 4.2. The medulla oblongata had a similar increase for cases compared to controls in the range $2-6$ keV/µm, while the midbrain also showed the same trend in the whole $2-8$ keV/µm range. The pons showed the smallest interval of increased LET_d to cases over controls at only $2-4$ keV/µm. The pons substructures, shown in figure 4.3, showed similar trends with increased LET_d to cases compared to controls for most LET_d. However, the values fluctuated more compared to the previously studied volumes.

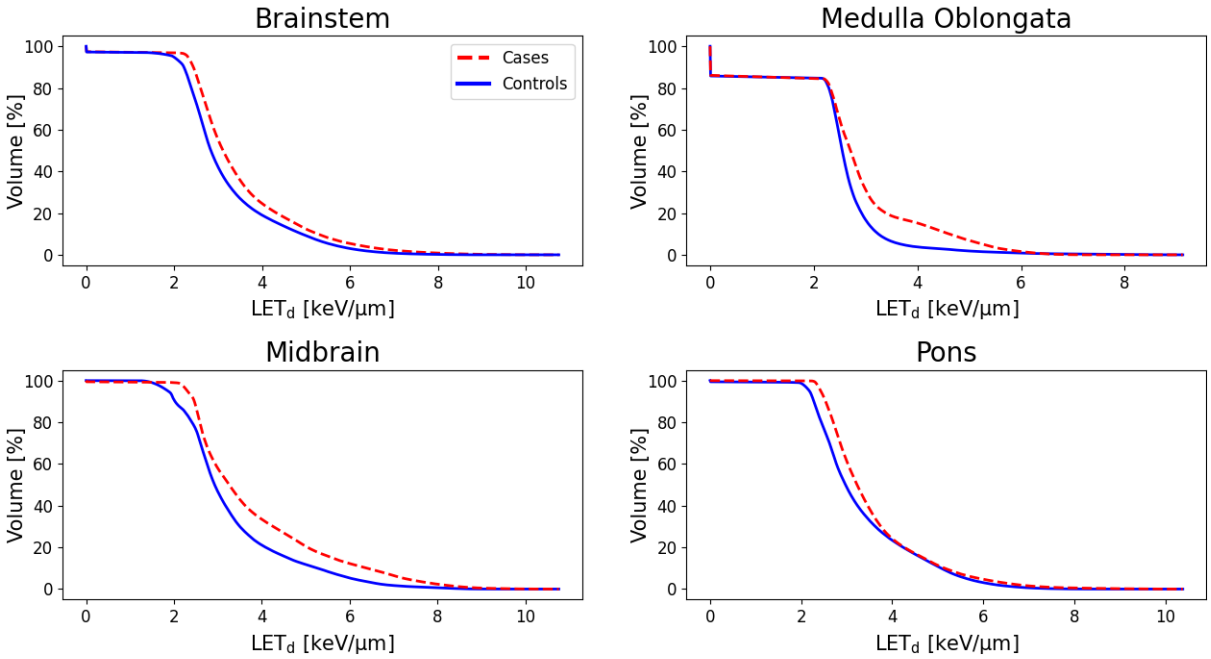

Figure 4.2: Average LET_d-volume histograms for cases versus controls for the full brainstem, the medulla oblongata, the midbrain and the pons.

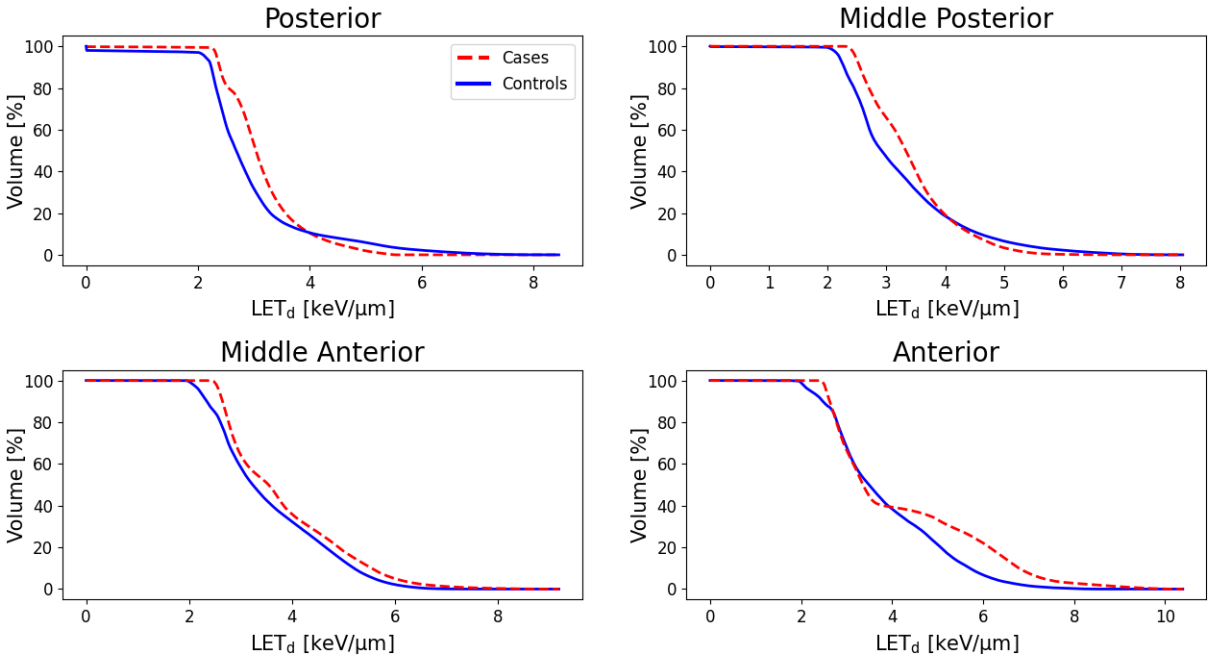

Figure 4.3: Average LET_d-volume histograms for cases and controls for the substructures of the pons.

Average $L50\%$ was seen to increase towards the anterior of the brainstem, as shown by the substructures of the pons in table table 4.1. $L_{\max}$ followed a similar trend, but also had high values for the midbrain. All subvolumes had larger standard deviations for $L_{\max}$ compared to $L50\%$.

Table 4.1: Average $L50\%$ and $L_{\max}$ for brainstem substructures for the whole cohort. The first standard deviation is included as an uncertainty estimate.

	$L50\%$ [keV/μm]	$L_{\max}$ [keV/μm]
Brainstem	3.08 ± 0.61	6.29 ± 2.10
Medulla Oblongata	2.34 ± 1.04	4.41 ± 1.78
Midbrain	3.34 ± 1.07	5.46 ± 2.29
Pons	3.26 ± 0.76	5.74 ± 1.96
Pons Posterior	3.00 ± 0.76	4.07 ± 1.57
Pons Middle Posterior	3.22 ± 0.78	4.66 ± 1.57
Pons Middle Anterior	3.56 ± 0.97	5.17 ± 1.75
Pons Anterior	3.90 ± 1.32	5.40 ± 1.97

From investigating the variation in $L50\%$ and $L_{\max}$ within and between case-control, as

well as average cases and controls, all subvolumes were observed to have higher values for cases than controls for both statistics. However, the differences were within the 95% CIs of both average case and control. This is depicted in figure 4.4 for the brainstem, the medulla oblongata, the midbrain and the pons, and for the pons substructures in figure 4.5. Similarly, $L_{\max}$ is given in figures 4.6 and 4.7. The full brainstem showed the least variation in $L50\%$, despite the medulla oblongata having most points more concentrated, but with notable outliers. The midbrain showed the greatest difference between cases and controls but with comparable CIs to the medulla oblongata. On the other hand, the pons showed a smaller difference between cases and controls than the other volumes and similar CIs to the full brainstem. The pons posterior and middle posterior exhibited similar minor variation and CIs for both cases and controls, while the middle anterior had wider CIs for controls. The anterior showed even greater CIs for both cases and controls.

The largest difference in $L_{\max}$ between cases and controls was observed for the full brainstem and the midbrain. However, these also yielded the widest CIs. The pons had slightly narrower CIs than the previous two structures, while the medulla oblongata yielded even narrower CIs. For the pons substructures, the CIs of the controls were similar across all four subvolumes, with the posterior showing the greatest consistency in the position of the data points. However, the CIs of cases for the two anterior subvolumes were significantly wider than their posterior counterparts.

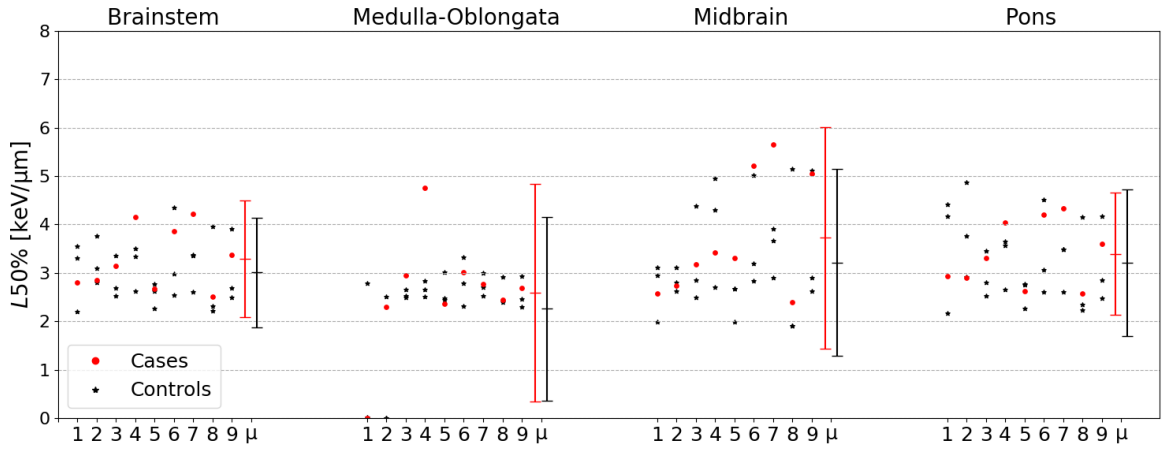

Figure 4.4: $L50\%$ of the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI.

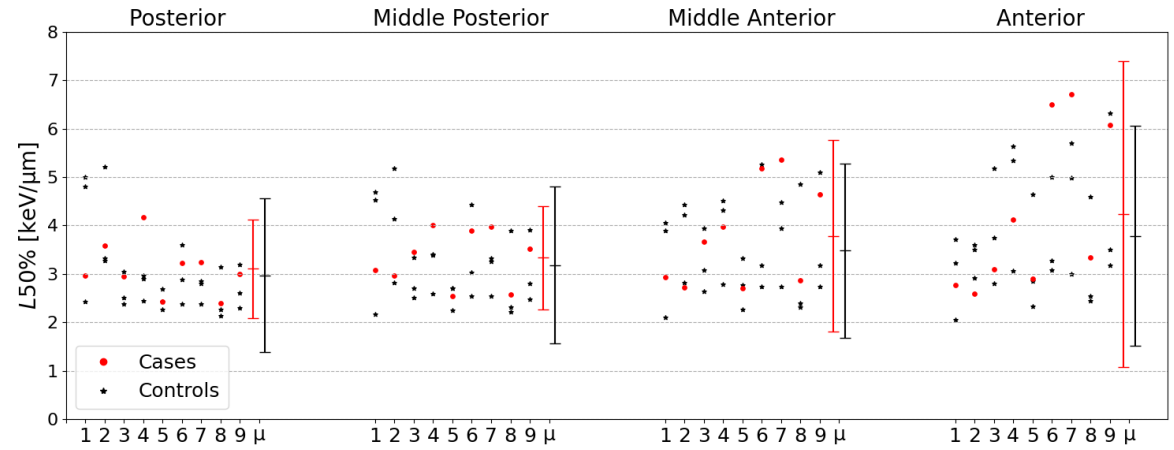

Figure 4.5: $L50\%$ of the substructures of the pons for all case-control groups, including average case and control with 95% CI.

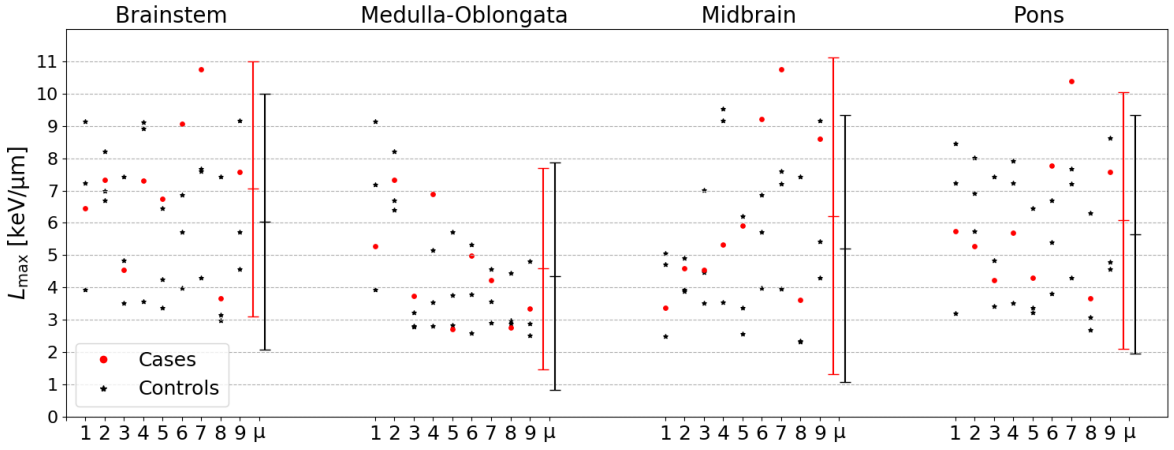

Figure 4.6: $L_{\max}$ of the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI.

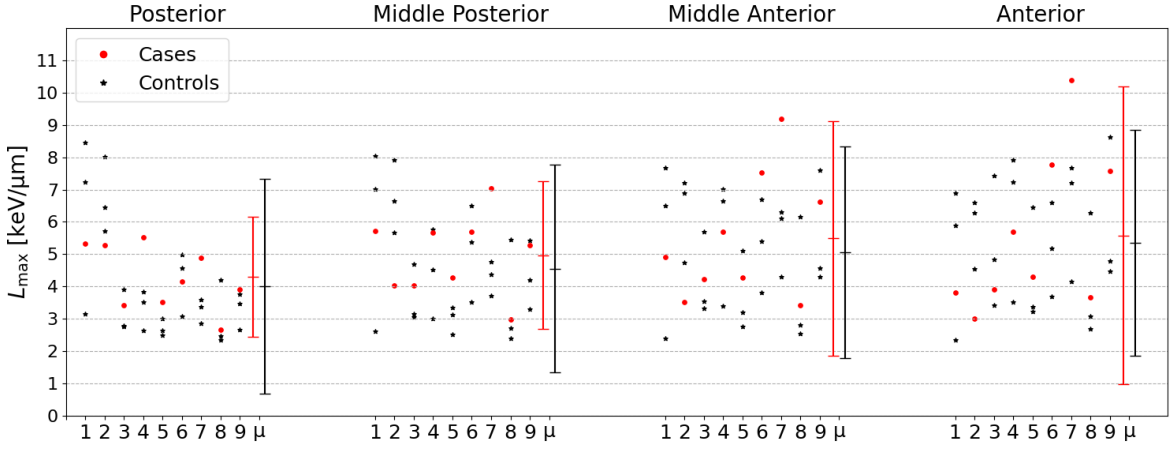

Figure 4.7: $L_{\max}$ of the substructures of the pons for all case-control groups, including average case and control with 95% CI.

The average difference in $L50\%$ and $L_{\max}$ between the cases and their average matched control showed that cases received higher $\mathrm{LET_d}$, compared to controls, consistent for all

subvolumes studied. The results are provided in table 4.2. The greatest differences were observed in the $L_{\max}$ of the full brainstem and the midbrain. However, the differences were all less than one standard deviation. Additionally, none of the differences proved statistically significant from a Wilcoxon signed-rank test.

Table 4.2: Average difference in $L50\%$ and $L_{\max}$ between cases and their average matched controls for brainstem substructures. The first standard deviation is included as an uncertainty estimate. No difference was found to be statistically significant from a Wilcoxon signed-rank test

	$L50\%$ [keV/μm]	$L_{\max}$ [keV/μm]
Brainstem	0.28 ± 0.54	1.02 ± 1.87
Medulla Oblongata	0.33 ± 0.92	0.23 ± 1.40
Midbrain	0.52 ± 1.02	1.01 ± 2.21
Pons	0.18 ± 0.71	0.44 ± 1.84
Pons Posterior	0.14 ± 0.69	0.30 ± 1.17
Pons Middle Posterior	0.15 ± 0.71	0.42 ± 1.51
Pons Middle Anterior	0.30 ± 0.91	0.43 ± 1.81
Pons Anterior	0.45 ± 1.37	0.22 ± 2.15

4.3 Variable Relative Biological Effectiveness

Average DVHs showed that the full brainstem had slightly larger differences between cases and controls for the LET_d-weighted dose with $c = 0.04$ than for RBE1.1. A further increase was observed for LET_d-weighted dose with $c = 0.055$. The DVHs of the brainstem are shown in figure 4.8 along with the medulla oblongata, the midbrain and the pons. The further delineated pons is given in figure 4.9. The differences for the full brainstem, the medulla oblongata and the pons were greatest in the high dose range, between 55 Gy(RBE) and 62 Gy(RBE) for both RBE1.1 and LET_d-weighted dose. However, the midbrain differed by instead having the most significant difference between cases and controls in the middle to high dose range, at around 40 Gy(RBE) to 58 Gy(RBE).

An increase in the difference between cases and controls from RBE1.1 to LET_d-weighted dose can also be seen in the high dose range for all pons substructures except the anterior.

However, the pons anterior had a lower average dose to cases compared to controls, but with increased $\mathrm{LET_d}$ weighting, the difference was seen to decrease, consistent with the trend of the remaining substructures.

DVHs for the McNamara and Rørvik models are given in figures 4.10 and 4.11. The models yielded greater differences between variable RBE and RBE1.1 in the high dose range (55 Gy(RBE) and up) compared to RBE1.1 versus $\mathrm{LET_d}$ weighted doses. The two phenomenological RBE models had similar differences between cases and controls.

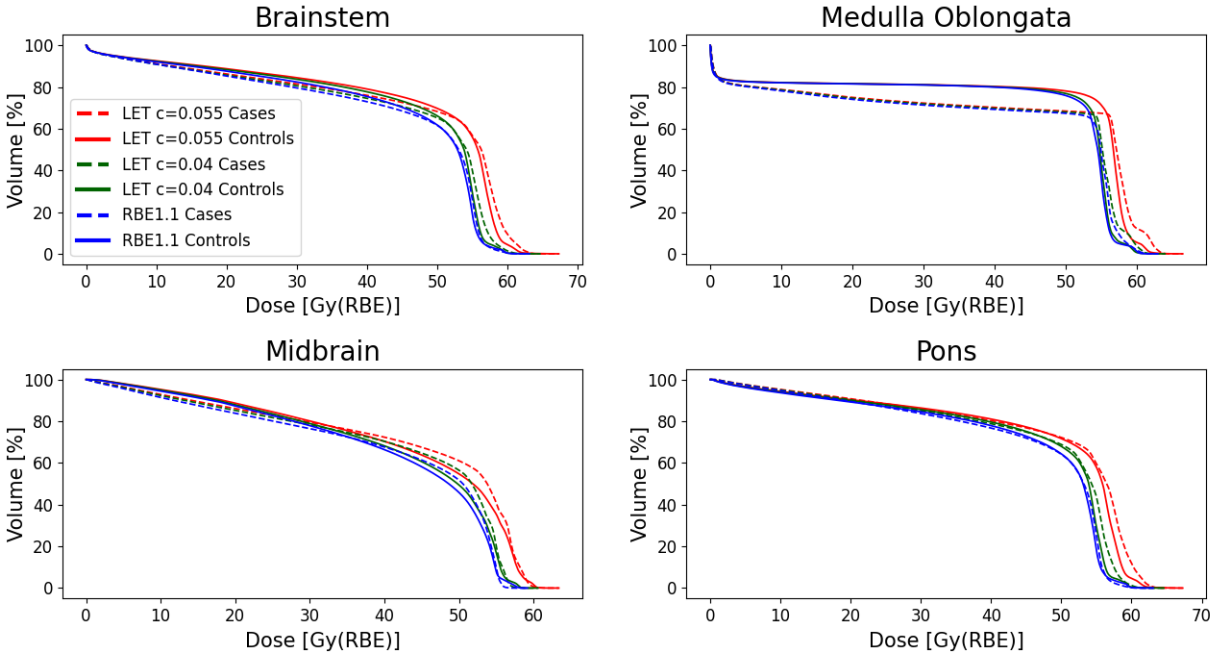

Figure 4.8: Average DVHs of cases and controls for the full brainstem, the medulla oblongata, the midbrain and the pons. RBE-weighted doses are found with RBE1.1, and $\mathrm{LET_d}$ weighted doses with $c = 0.04$ and $c = 0.055$.

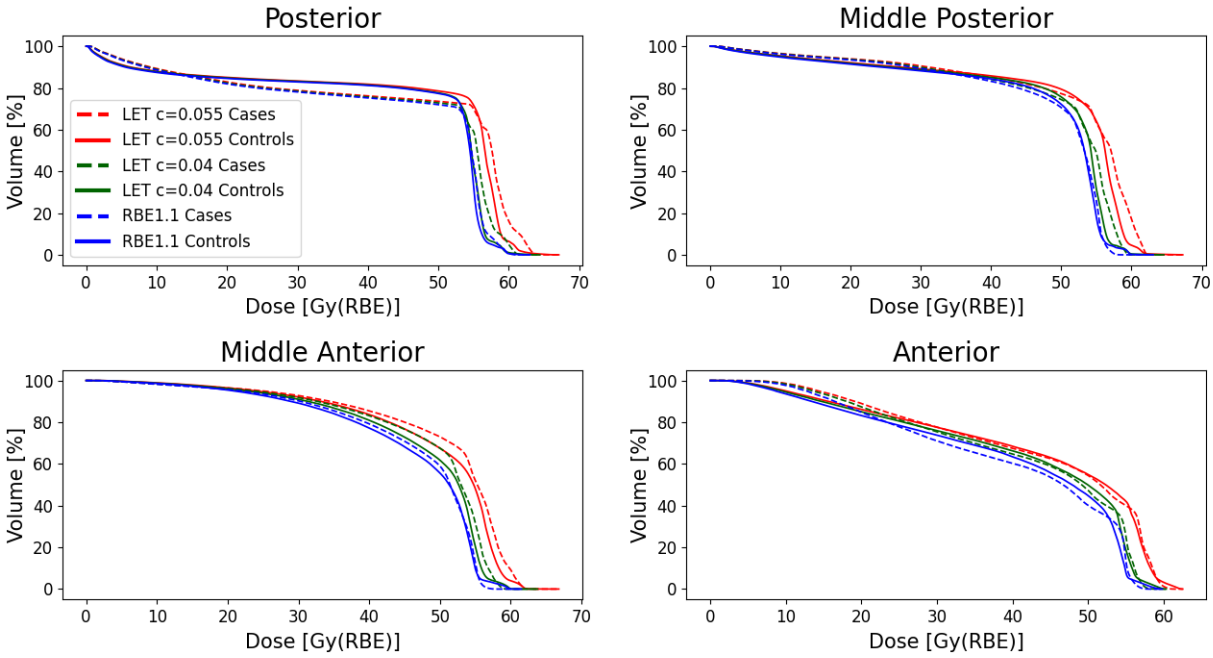

Figure 4.9: Average DVHs of cases and controls for the substructures of the pons. RBE-weighted doses are found with RBE1.1, and $\mathrm{LET_d}$ weighted doses with $c = 0.04$ and $c = 0.055$.

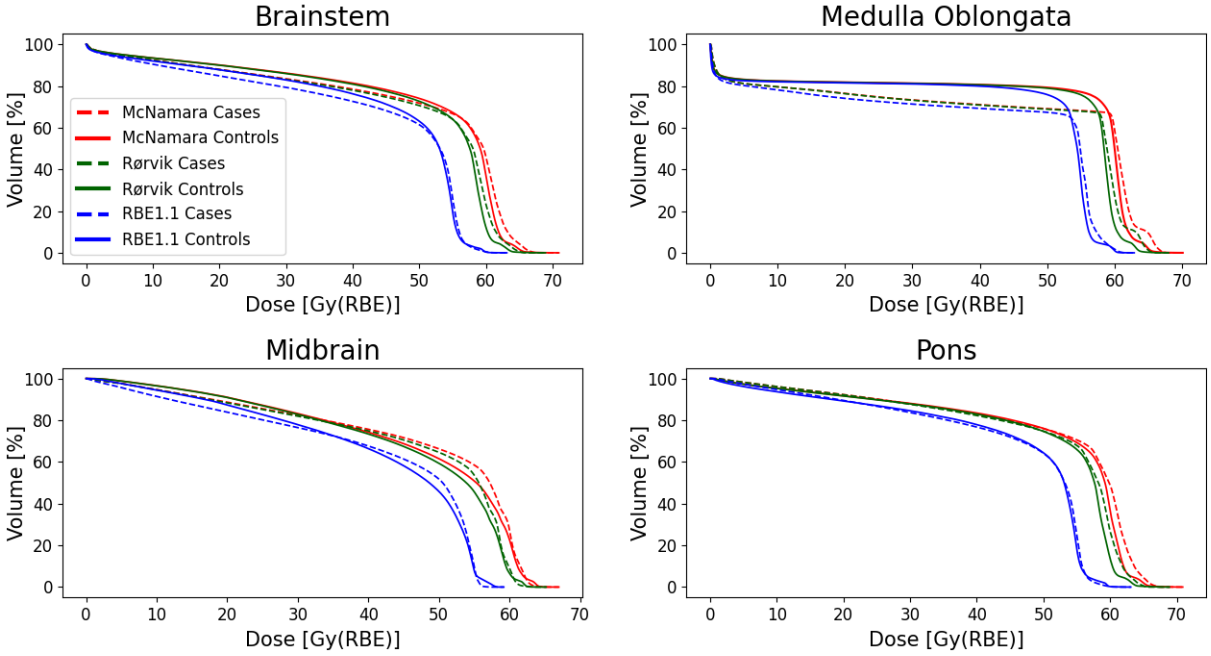

Figure 4.10: Average DVHs of cases and controls for the full brainstem, the medulla oblongata, the midbrain and the pons. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik variable RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume.

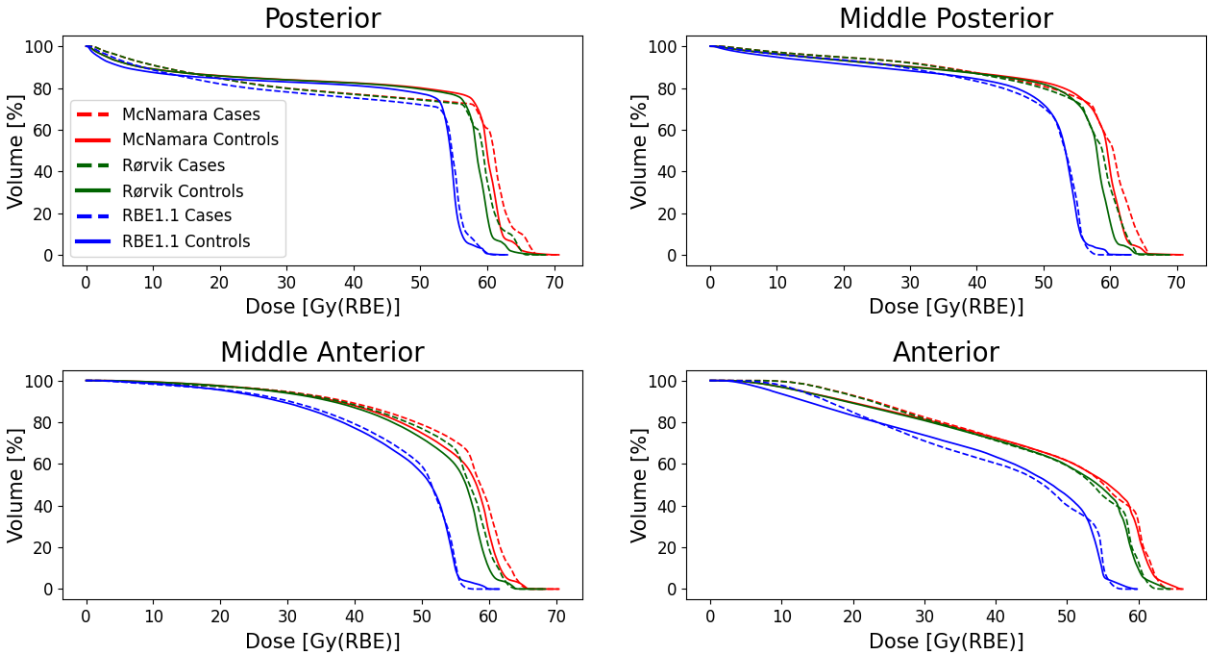

Figure 4.11: Average DVHs of cases and controls for the substructures of the pons. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik variable RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume.

The subsequent study of dose statistics showed that the variation in D50% exceeded the observations for $D10\%$ and $D_{\max}$, for both variable RBE and RBE1.1. Average $D50\%$ is given in table 4.3 for RBE1.1, as well as the Rørvik and McNamara RBE models, and showed structures with lower average $D50\%$ accompanied by high standard deviations. This can imply that the $D50\%$ was consistently high for most groups, but with a few outliers. The University of Florida proton centre practises an upper constraint of $D50\% = 54$ Gy(RBE) to the full brainstem, which the average patient met for RBE1.1, while both variable RBE models exceeded the limit [33].

Table 4.3: Average $D50\%$ to brainstem substructures for the whole cohort. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate.

Structure	$D50\%$ RBE1.1 [Gy(RBE)]	$D50\%$ Rørvik [Gy(RBE)]	$D50\%$ McNamara [Gy(RBE)]
Brainstem	49.1 ± 9.5	54.3 ± 9.1	55.7 ± 9.4
Medulla Oblongata	43.3 ± 22.1	46.8 ± 23.5	48.1 ± 24.1
Midbrain	44.7 ± 10.7	50.2 ± 10.2	51.5 ± 10.6
Pons	49.3 ± 9.6	54.6 ± 9.2	56.1 ± 9.5
Posterior	45.6 ± 18.9	49.8 ± 19.4	51.1 ± 20.0
Middle Posterior	48.5 ± 11.8	53.7 ± 11.6	55.1 ± 12.0
Middle Anterior	48.2 ± 7.6	54.3 ± 6.8	55.8 ± 7.1
Anterior	41.2 ± 15.3	47.3 ± 14.9	48.6 ± 15.4

The variation within and between case-control groups confirms that the low averages and large standard deviations were mainly caused by outlier groups, as shown in figures 4.12 and 4.13 for the McNamara variable RBE model. Appendix A provides additional plots for the Rørvik model and RBE1.1 in figures A.1-A.4. The variation in $D50\%$ was greatest for the medulla oblongata, but this was mainly caused by outliers. The average was higher for cases than controls for the full brainstem volume, the midbrain and the pons, while the medulla oblongata showed the opposite trend. However, the differences were well within the CIs of both cases and controls. The average dose received by both cases and controls to the full brainstem volume, as calculated by variable RBE, exceeded the University of Florida dose constraint, while both cases and controls met the constraint for RBE1.1.

The substructures of the pons showed wide CIs for the posterior and the anterior, with the posterior showing mostly compact values, bar a few outliers. The variable RBE models had higher means for cases compared to controls for all pons substructures. On the other hand, RBE1.1 showed higher averages for cases than controls for the two middle subvolumes but similar average case and control for the posterior and the anterior.

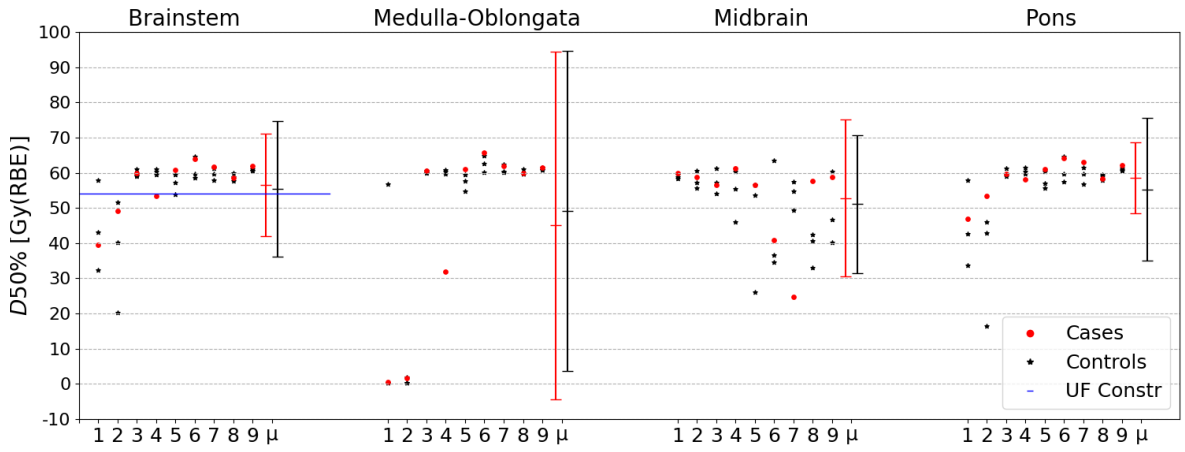

Figure 4.12: $D50\%$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. Dose constraint for $D50\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

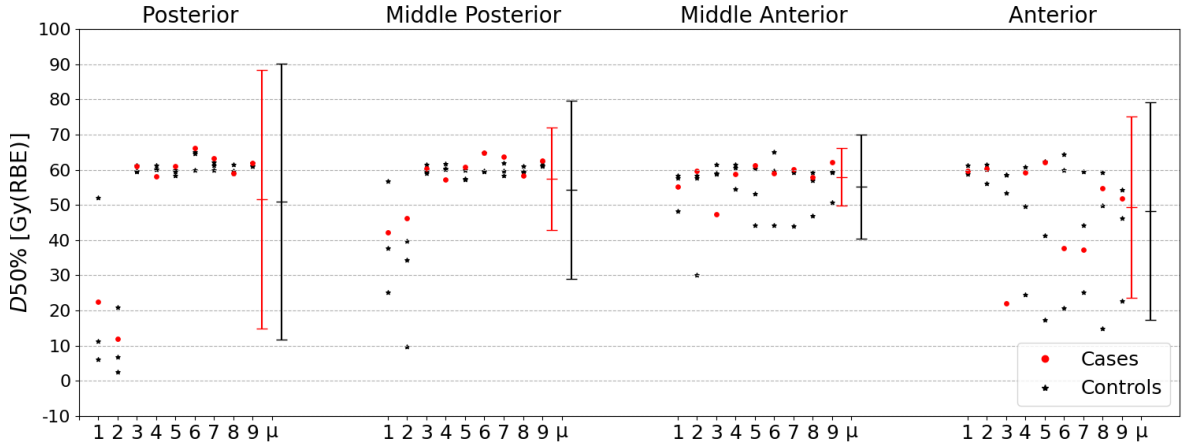

Figure 4.13: $D50\%$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

The average $D10\%$ to all brainstem substructures for the whole cohort are given in table 4.4 and generally showed smaller deviation compared to the $D50\%$ for most substructures. Additionally, the University of Florida proton centre limit of $D10\% = 56$ Gy(RBE) to the full brainstem was met for RBE1.1, but not for the McNamara and Rørvik RBE models [33],

Table 4.4: Average $D10\%$ to brainstem substructures for the whole cohort. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate.

Structure	$D10\%$ RBE1.1 [Gy(RBE)]	$D10\%$ Rørvik [Gy(RBE)]	$D10\%$ McNamara [Gy(RBE)]
Brainstem	55.4 ± 1.4	59.8 ± 1.8	61.4 ± 1.8
Medulla Oblongata	45.9 ± 20.7	49.7 ± 21.6	51.1 ± 22.2
Midbrain	53.5 ± 2.3	58.2 ± 2.5	59.7 ± 2.6
Pons	55.2 ± 1.6	59.8 ± 1.9	61.4 ± 2.0
Posterior	50.3 ± 13.0	54.7 ± 13.0	56.1 ± 13.5
Middle Posterior	53.6 ± 4.0	58.5 ± 4.6	60.1 ± 4.8
Middle Anterior	53.5 ± 2.9	58.8 ± 2.6	60.5 ± 2.7
Anterior	48.1 ± 10.3	54.0 ± 9.5	55.5 ± 8.3

The $D10\%$ to the full brainstem, midbrain and pons showed similar trends between cases and controls as the $D50\%$ to the same volumes. On the other hand, the $D10\%$ to the medulla oblongata showed very similar means between cases and controls, as can be seen for the McNamara model in figure 4.14. The $D10\%$ to the further delineated pons is shown in figure 4.15, while similar plots for Rørvik and RBE1.1 are provided in appendix A in figures A.5-A.8. The average $D10\%$ to the full brainstem for both cases and controls exceeded the University of Florida dose constraint for variable RBE, with all individual patients exceeding the constraint. However, RBE1.1 had both averages meeting the constraint, but with 33% of cases and a lower percentage of controls individually not meeting the constraint. The dose constraints not being met for certain patients for RBE1.1 is possibly a consequence of the FLUKA recalculation and rescaling to achieve similar CTV coverage as the original treatment. The groups had minor variance and narrow CIs for the full brainstem, the midbrain and the pons. On the other hand, the medulla oblongata had significant CIs, with case-control groups one and two deviating from the other groups. The narrow CIs can partly be explained by the fact that the case-control groups were directly matched based on $D10\%$ to the full brainstem volume.

The pons substructures posterior and anterior had relatively wide CIs, and all four pons substructures had a higher average for cases than controls for both variable RBE and

RBE1.1.

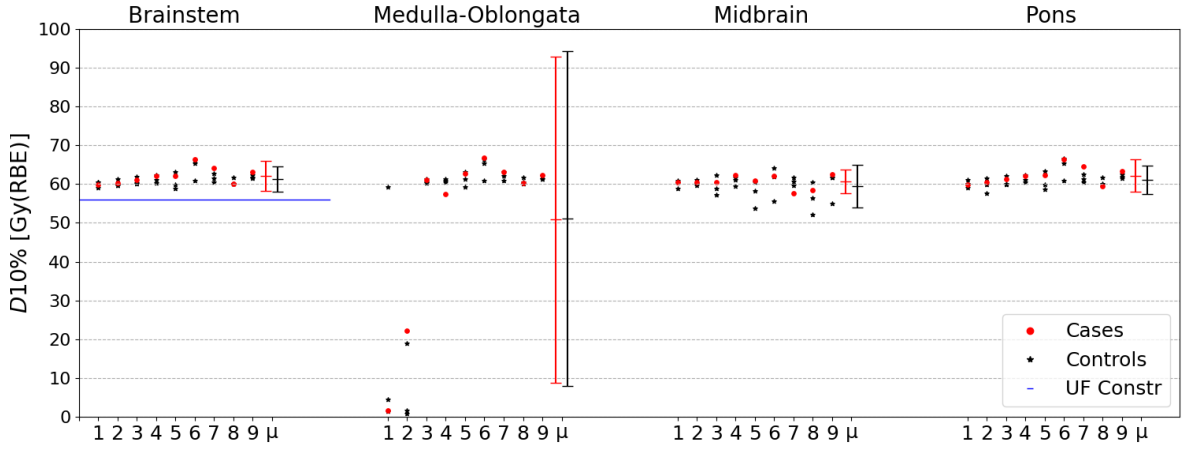

Figure 4.14: $D10\%$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. Dose Constraint for $D10\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

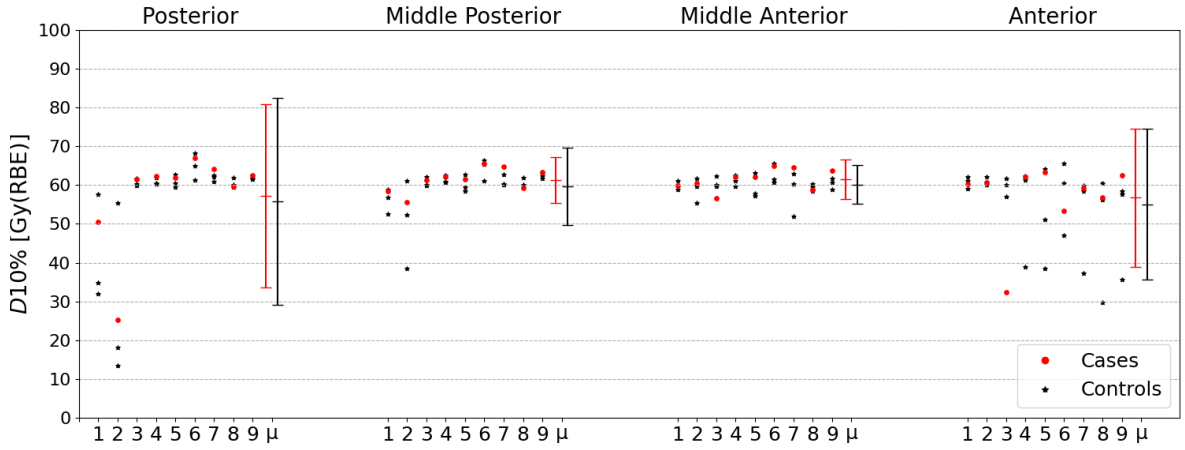

Figure 4.15: $D10\%$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

The average D_{max} is given for the whole cohort in table 4.5 and the standard deviations were minor, comparably to the variation in $D10\%$. Similarly to the other dose statistics, the University of Florida proton centre constraint of $D_{max} = 58$ Gy(RBE) was met for RBE1.1, but not for the variable RBE [33].

Table 4.5: Average $D_{\max}$ to brainstem substructures for the whole cohort. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate.

Structure	$D_{\max}$ RBE1.1 [Gy(RBE)]	$D_{\max}$ Rørvik [Gy(RBE)]	$D_{\max}$ McNamara [Gy(RBE)]
Brainstem	57.2 ± 1.9	61.7 ± 2.2	63.4 ± 2.3
Medulla Oblongata	49.6 ± 18.0	53.9 ± 18.2	55.4 ± 18.7
Midbrain	55.9 ± 1.5	60.3 ± 2.1	62.0 ± 2.2
Pons	57.0 ± 1.9	61.6 ± 2.2	63.3 ± 2.3
Posterior	54.2 ± 7.9	58.6 ± 7.8	60.2 ± 8.1
Middle Posterior	56.0 ± 2.5	60.7 ± 2.7	62.4 ± 2.9
Middle Anterior	55.6 ± 2.4	60.7 ± 2.5	62.4 ± 2.7
Anterior	52.6 ± 6.3	58.3 ± 5.4	59.9 ± 5.6

The variation between case-control groups and the average case and control are depicted for the McNamara RBE model in figures 4.16 and 4.17, while the Rørvik model and RBE1.1 are shown in appendix A in figures A.9-A.12. The full brainstem volume, the midbrain and the pons showed very narrow CIs, and all had slightly higher $D_{\max}$ for the cases compared to controls for variable RBE. In comparison, RBE1.1 showed very similar means between both cases and controls for all three volumes. On the other hand, the medulla oblongata exhibited wide CIs for both variable RBE and RBE1.1. The lack of intra-group variations for the full brainstem volume can partly be attributed to the matching of patients based on $D_{0.1\text{cc}}$. Additionally, the $D_{\max}$ with variable RBE models exceeded the dose constraint to the full brainstem for both average cases and controls, with all patients receiving excessive $D_{\max}$ to the organ. On the other hand, RBE1.1 had both mean case and control meeting the constraint, with a similar ratio of cases and controls not meeting the constraint as the $D10\%$.

The pons substructures mostly showed cases and controls with similar means but, for RBE1.1, posterior and anterior had the average case lower than the control. On the other hand, the variable RBE showed all substructures with slightly higher $D_{\max}$ to cases than controls.

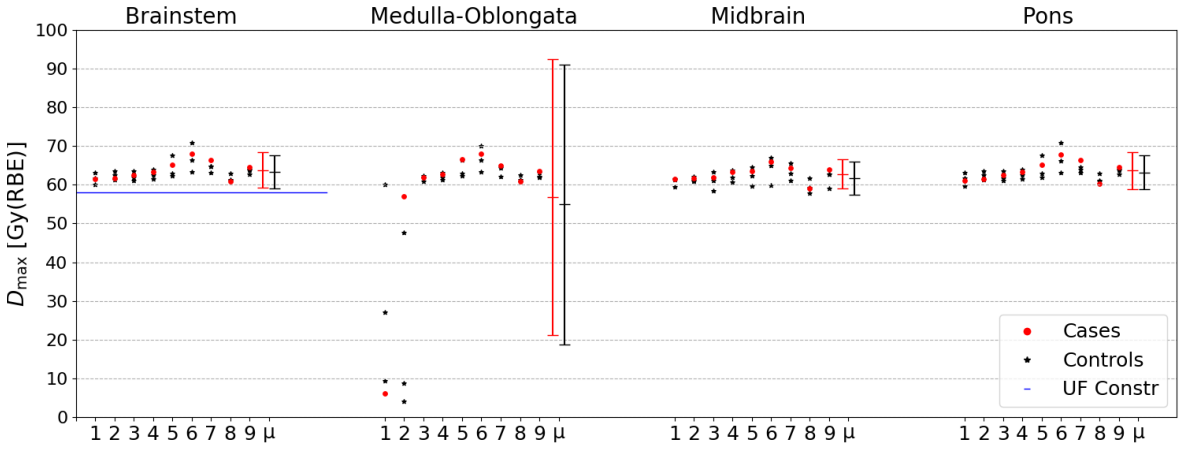

Figure 4.16: $D_{\max}$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. Dose constraint for $D_{\max}$ utilised clinically at the University of Florida is also given as a blue horizontal line [33].

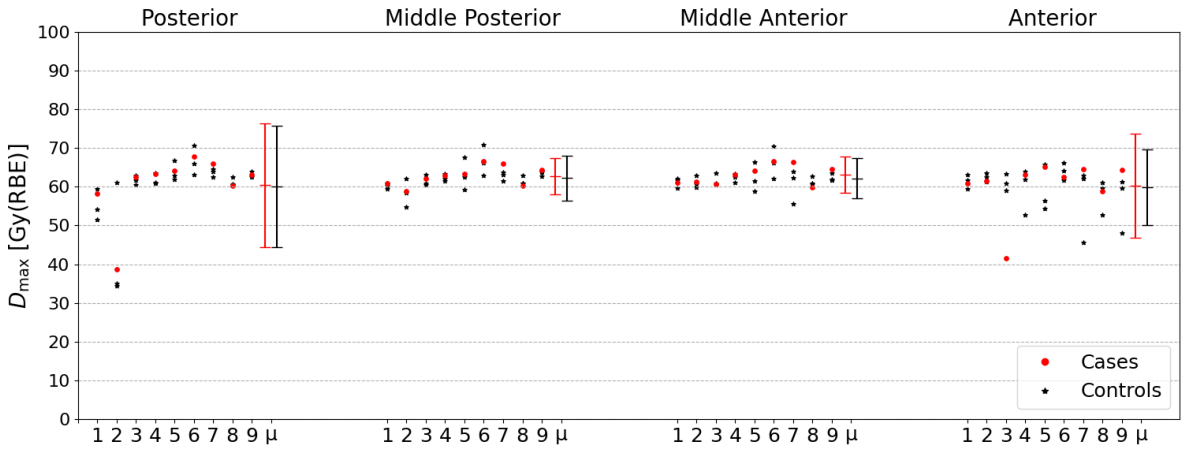

Figure 4.17: $D_{\max}$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

The difference in $D50\%$ between cases and their average matched controls are given in table 4.6 and shows an increased difference of 0.5 Gy(RBE) to 1 Gy(RBE) between cases and controls for variable RBE compared to RBE1.1, where a positive difference translates to a higher dose received by the case. This trend was observed for all substructures except the medulla oblongata, which was the only substructure without a positive difference for the variable RBE. On the other hand, RBE1.1 had negative differences for the medulla oblongata, the pons posterior and the pons anterior. However, the differences between cases and controls were minor compared to their first standard deviations. Additionally, no statistical significance was found from a Wilcoxon signed-rank test for any of the differences.

Table 4.6: Average difference in $D50\%$ between cases and controls for brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate. No difference was found to be statistically significant from Wilcoxon signed-rank test.

Structure	$D50\%$ RBE1.1 [Gy(RBE)]	$D50\%$ Rørvik [Gy(RBE)]	$D50\%$ McNamara [Gy(RBE)]
Brainstem	0.4 ± 6.1	1.1 ± 5.2	1.1 ± 5.4
Medulla Oblongata	-4.4 ± 11.6	-4.0 ± 11.0	-4.1 ± 11.3
Midbrain	0.7 ± 13.8	1.5 ± 13.2	1.7 ± 13.6
Pons	2.4 ± 6.9	3.1 ± 5.9	3.2 ± 6.0
Posterior	-0.3 ± 2.7	0.7 ± 1.7	0.7 ± 1.8
Middle Posterior	1.9 ± 6.9	2.9 ± 6.1	3.0 ± 6.2
Middle Anterior	1.7 ± 7.1	2.7 ± 6.4	2.8 ± 6.7
Anterior	-0.1 ± 16.0	1.0 ± 15.5	1.1 ± 17.1

The average differences in $D10\%$ between cases and their average matched controls are given in table 4.7 and showed increases between 0.5 Gy(RBE) and 1 Gy(RBE) from RBE1.1 to variable RBE for almost all substructures. The only substructure with higher $D10\%$ to cases than controls was the medulla oblongata, which also had the second largest first standard deviation, second to the pons anterior. Furthermore, The $D10\%$ to the full brainstem volume and the pons middle posterior were found to be statistically significant from a Wilcoxon signed-rank test.

Table 4.7: Average difference in $D10\%$ between cases and controls for brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate.
*Statistically significant from Wilcoxon signed-rank test

Structure	$D10\%$ RBE1.1 [Gy(RBE)]	$D10\%$ Rørvik [Gy(RBE)]	$D10\%$ McNamara [Gy(RBE)]
Brainstem	0.5 ± 0.7	0.9 ± 1.0*	1.0 ± 1.1*
Medulla Oblongata	-1.1 ± 7.9	-0.3 ± 8.8	-0.3 ± 9.1
Midbrain	0.3 ± 3.0	1.1 ± 1.9	1.2 ± 1.9
Pons	0.5 ± 0.9	1.0 ± 1.2	1.0 ± 1.2
Posterior	1.1 ± 3.6	1.4 ± 3.3	1.4 ± 3.4
Middle Posterior	1.1 ± 2.1	1.6 ± 1.7*	1.7 ± 1.8*
Middle Anterior	0.7 ± 2.8	1.3 ± 2.7	1.4 ± 2.9
Anterior	0.7 ± 11.5	1.6 ± 11.7	1.7 ± 12.3

$D_{\max}$ was subsequently studied, with the average difference between cases and their average matched controls provided in table 4.8. $D_{\max}$ showed more apparent systematic differences compared to the $D10\%$ and $D50\%$, where, for the variable RBE, all differences were positive. In contrast, for RBE1.1, the differences were minor and not systematic in either direction for all substructures, bar the medulla oblongata. The medulla oblongata was, however, the substructure with the greatest standard deviation. An increase from RBE1.1 to variable RBE of between 0.2 Gy(RBE) and 1.0 Gy(RBE) was observed for all substructures. The highest increase was seen for the midbrain at approximately 1 Gy(RBE) for both RBE models, and both were found to be statistically significant.

Table 4.8: Average difference in $D_{\max}$ between cases and controls for brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate.
*Statistically significant from Wilcoxon signed-rank test

Structure	$D_{\max}$ RBE1.1 [Gy(RBE)]	$D_{\max}$ Rørvik [Gy(RBE)]	$D_{\max}$ McNamara [Gy(RBE)]
Brainstem	0.2 ± 0.6	0.4 ± 0.9	0.5 ± 1.0
Medulla Oblongata	1.5 ± 14.2	1.9 ± 15.5	2.0 ± 15.9
Midbrain	0.1 ± 1.3	1.0 ± 0.9*	1.1 ± 0.9*
Pons	-0.0 ± 0.6	0.4 ± 1.1	0.5 ± 1.2
Posterior	-0.1 ± 2.2	0.4 ± 2.2	0.4 ± 2.3
Middle Posterior	0.0 ± 1.0	0.5 ± 1.1	0.5 ± 1.1
Middle Anterior	0.3 ± 1.8	0.9 ± 2.0	1.0 ± 2.2
Anterior	-0.2 ± 8.1	0.4 ± 8.0	0.5 ± 8.4

RBE-weighted dose from the Rørvik and McNamara models both depend on $(\alpha/\beta)_x$ of the irradiated tissue, so to study the effect of changing $(\alpha/\beta)_x$ on the dose distribution, average DVHs are shown in figures 4.18 and 4.19. The DVHs showed minor differences with altered $(\alpha/\beta)_x$-ratio of the tissue, with DVHs shifted slightly to the right with decreasing $(\alpha/\beta)_x$. Additionally, the difference between $(\alpha/\beta)_x$ did not reposition cases and controls to varying degrees, resulting in constant deviations between the curves.

The differences between cases and the average of their controls for the different $(\alpha/\beta)_x$-ratios are given in table 4.9 for the McNamara model and table 4.10 for the Rørvik model. However, the differences between altered $(\alpha/\beta)_x$ were minor compared to the effect of variable RBE to RBE1.1. Additionally, the $D50\%$ from RBE models with $(\alpha/\beta)_x = 3.3$ did not show statistical significance from a Wilcoxon signed-rank test for any of the substructures.

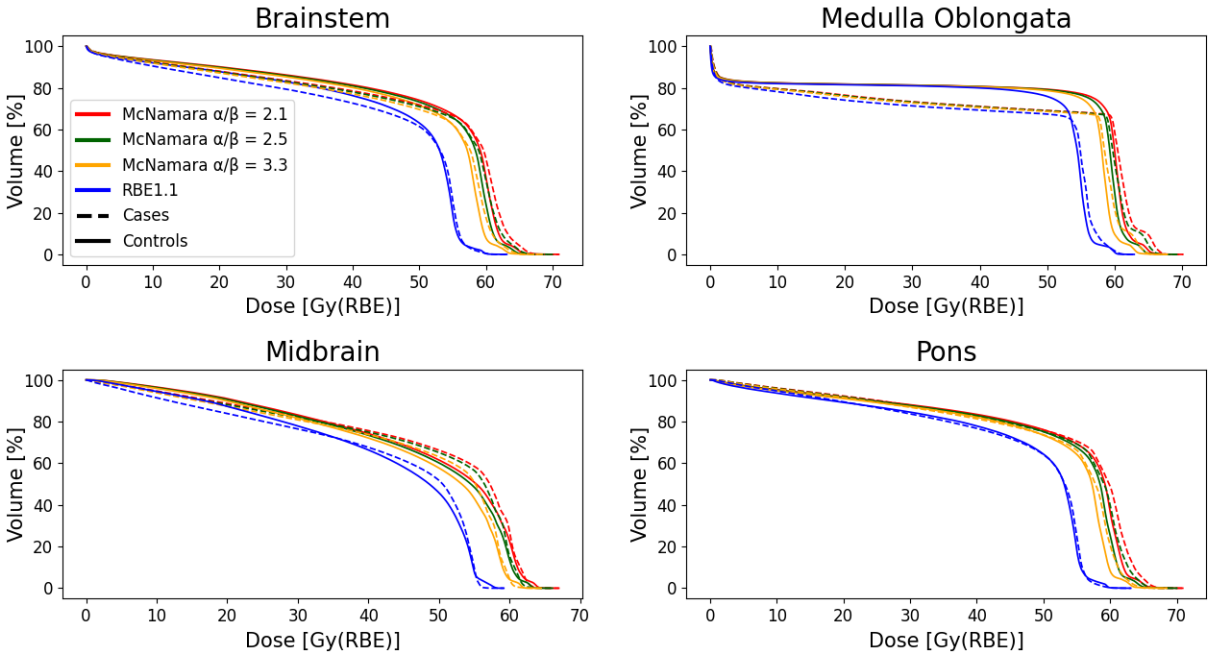

Figure 4.18: Average DVHs for cases (dashed lines) and controls (full lines) with RBE-weighted doses from RBE1.1, and the McNamara RBE model assuming constant $(\alpha/\beta)_x = 2.1$, $(\alpha/\beta)_x = 2.5$ and $(\alpha/\beta)_x = 3.3$ for the full brainstem.

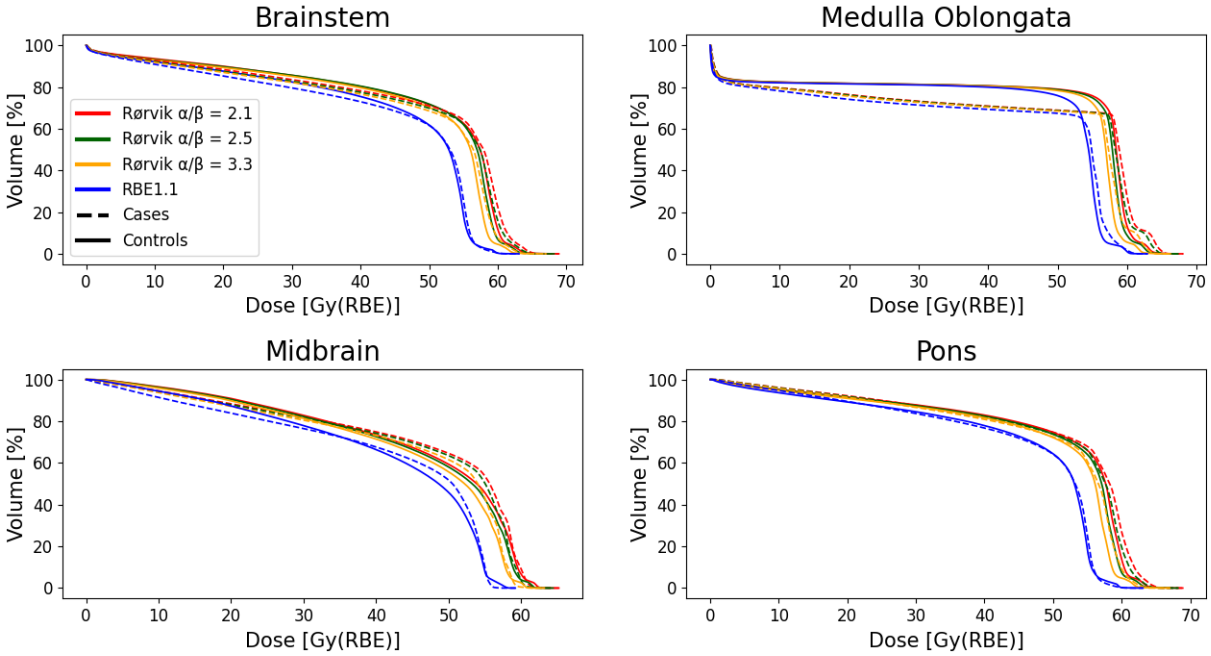

Figure 4.19: Average DVHs for cases (dashed lines) and controls (full lines) with RBE-weighted doses from RBE1.1, and the Rørvik RBE model assuming constant $(\alpha/\beta)_x = 2.1$, $(\alpha/\beta)_x = 2.5$ and $(\alpha/\beta)_x = 3.3$ for the full brainstem.

Table 4.9: Average difference in $D50\%$ between cases and controls for brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara RBE model assuming constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ and $(\alpha/\beta)_{\mathrm{x}} = 3.3$, for the full brainstem. The first standard deviation is included as an uncertainty estimate. No statistical significance was found from a Wilcoxon signed-rank test for any of the differences. †Variable RBE from McNamara model

Structure	$D50\%$ RBE1.1 [Gy(RBE)]	$D50\%^{\dagger}$ $\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}} =$ 2.1 [Gy(RBE)]	$D50\%^{\dagger}$ $\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}} =$ 3.3 [Gy(RBE)]
Brainstem	0.4 ± 6.1	1.1 ± 5.4	1.0 ± 5.5
Medulla Oblongata	-4.4 ± 11.6	-4.1 ± 11.3	-4.2 ± 11.3
Midbrain	0.7 ± 13.8	1.7 ± 13.6	1.5 ± 13.6
Pons	2.4 ± 6.9	3.2 ± 6.0	3.1 ± 6.1
Posterior	0.3 ± 2.7	0.7 ± 1.8	0.5 ± 1.9
Middle Posterior	1.9 ± 6.9	3.0 ± 6.2	2.8 ± 6.3
Middle Anterior	1.7 ± 7.1	2.8 ± 6.7	2.7 ± 6.7
Anterior	-0.1 ± 16.0	1.1 ± 17.1	0.9 ± 16.8

Table 4.10: Average difference in $D50\%$ between cases and controls to brainstem substructures. RBE-weighted doses are found with RBE1.1, and the Rørvik RBE model assuming constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ and $(\alpha/\beta)_{\mathrm{x}} = 3.3$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate. No statistical significance was found from a Wilcoxon signed-rank test for any of the differences. †Variable RBE from Rørvik model

Structure	$D50\%$ RBE1.1 [Gy(RBE)]	$D50\%^{\dagger}$ $\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}} =$ 2.1 [Gy(RBE)]	$D50\%^{\dagger}$ $\left(\frac{\alpha}{\beta}\right)_{\mathrm{x}} =$ 3.3 [Gy(RBE)]
Brainstem	0.4 ± 6.1	1.1 ± 5.2	1.0 ± 5.3
Medulla Oblongata	-4.4 ± 11.6	-4.0 ± 11.0	-4.1 ± 11.0
Midbrain	0.7 ± 13.8	1.5 ± 13.2	1.4 ± 13.3
Pons	2.4 ± 6.9	3.1 ± 5.9	3.0 ± 6.0
Posterior	0.3 ± 2.7	0.7 ± 1.7	0.5 ± 1.9
Middle Posterior	1.9 ± 6.9	2.9 ± 6.1	2.8 ± 6.2
Middle Anterior	1.7 ± 7.1	2.7 ± 6.4	2.5 ± 6.5
Anterior	-0.1 ± 16.0	1.0 ± 16.4	0.8 ± 16.3

4.4 Generalised Equivalent Uniform Dose

Trends in gEUD were similar for volume parameters $n = 0.05$ [87] and $n = 0.16$ [88], and no n-value was shown to systematically predict higher gEUD. The variations in gEUD between case-control groups are given in figures 4.20 and 4.22 for $n = 0.05$, and figures 4.21 and 4.23 for $n = 0.16$, all for the McNamara model. The equivalent figures for the Rørvik model and RBE1.1 are included in appendix A in figures A.13-A.20 For $n = 0.05$, the full brainstem, the medulla oblongata, the midbrain, and the pons all showed higher average gEUD to cases than controls, with a higher difference for variable RBE compared to RBE1.1, although well within the CIs of both cases and controls. Similar results were observed for $n = 0.16$, with the exception of the medulla oblongata, where the average control was higher than the average case. Additionally, the higher n-value showed wider CIs compared to $n = 0.05$, with the medulla oblongata having wider CIs than the full brainstem volume, the midbrain and the pons. ¨

The further delineated pons showed the variable RBE with a higher average gEUD for all substructures and both n values. On the other hand, RBE1.1 showed very similar values between cases and controls for the posterior and anterior, while the middle subvolumes had higher averages for cases than controls. The posterior and anterior also had the widest CIs, and the greatest intra-group variations, among the substructures.

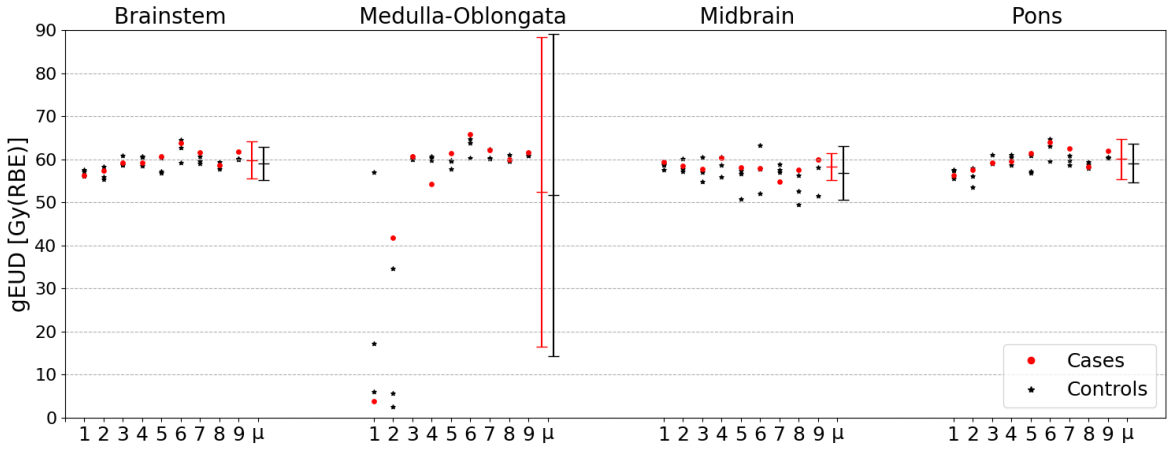

Figure 4.20: gEUD with $n = 0.05$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

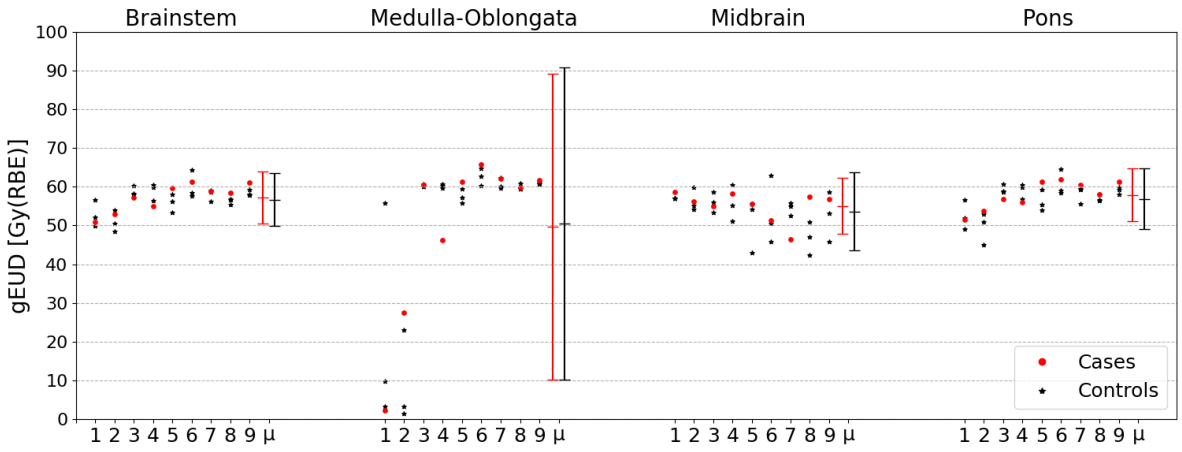

Figure 4.21: gEUD with $n = 0.16$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

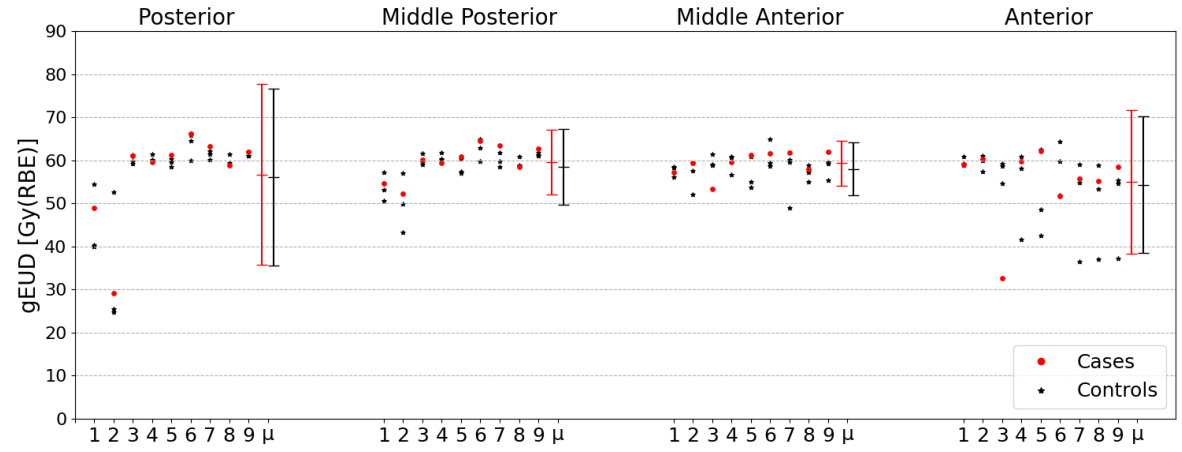

Figure 4.22: gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

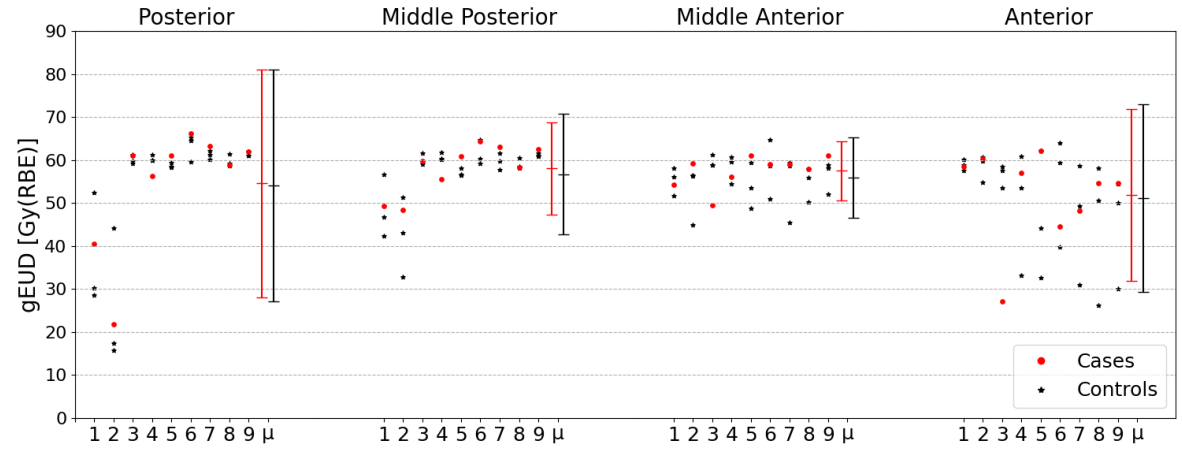

Figure 4.23: gEUD with $n = 0.16$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the McNamara model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

The average difference in gEUD between cases and the average of their matched controls are given in table 4.11 for $n = 0.05$ and table 4.12 for $n = 0.16$ and showed an increase of between 0.3 Gy(RBE) and 0.8 Gy(RBE) from RBE1.1 to variable RBE. Additional minor increases were observed from Rørvik to McNamara of about 0.1 Gy(RBE). For the variable RBE, all the differences were positive for all structures, except the medulla oblongata. However, the medulla oblongata and the pons anterior showed greater first standard deviations than the remaining volumes. Additionally, no statistical significance was found from a Wilcoxon signed-rank test.

Table 4.11: Average difference in gEUD with $n = 0.05$ between cases and controls to the brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem. The first standard deviation is included as an uncertainty estimate. No difference was found to be statistically significant from a Wilcoxon signed-rank test.

Structure	gEUD RBE1.1 [Gy(RBE)]	gEUD Rørvik [Gy(RBE)]	gEUD McNamara [Gy(RBE)]
Brainstem	0.3 ± 1.1	0.7 ± 1.2	0.8 ± 1.2
Medulla Oblongata	0.3 ± 11.5	0.7 ± 12.6	0.7 ± 12.9
Midbrain	0.5 ± 3.0	1.3 ± 2.3	1.3 ± 2.4
Pons	0.4 ± 1.4	0.9 ± 1.4	1.0 ± 1.5
Posterior	0.2 ± 2.6	0.6 ± 2.6	0.6 ± 2.7
Middle Posterior	0.5 ± 1.6	1.1 ± 1.6	1.1 ± 1.7
Middle Anterior	0.6 ± 3.6	1.3 ± 3.5	1.4 ± 3.6
Anterior	-0.0 ± 10.3	0.6 ± 10.6	0.7 ± 11.1

Table 4.12: Average difference in gEUD with $n = 0.16$ between cases and controls for brainstem substructures. RBE-weighted doses are found with RBE1.1, and the McNamara and Rørvik RBE models. Both variable RBE models assume a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. The first standard deviation is included as an uncertainty estimate. No difference was found to be statistically significant from a Wilcoxon signed-rank test.

Structure	gEUD RBE1.1 [Gy(RBE)]	gEUD Rørvik [Gy(RBE)]	gEUD McNamara [Gy(RBE)]
Brainstem	0.1 ± 2.5	0.5 ± 2.4	0.6 ± 2.5
Medulla Oblongata	-1.1 ± 10.2	-0.7 ± 10.8	-0.8 ± 11.1
Midbrain	0.6 ± 5.5	1.3 ± 5.0	1.4 ± 5.2
Pons	0.4 ± 2.9	1.0 ± 2.7	1.0 ± 2.8
Posterior	-0.1 ± 3.0	0.4 ± 2.8	0.4 ± 2.9
Middle Posterior	0.6 ± 3.2	1.3 ± 3.1	1.4 ± 3.2
Middle Anterior	0.7 ± 5.4	1.5 ± 5.2	1.6 ± 5.5
Anterior	0.1 ± 12.5	0.7 ± 13.0	0.8 ± 13.5

4.5 Lyman-Kutcher-Burman Models

LKB model curves for the full brainstem volume with $n = 0.05$, $n = 0.16$, and $n = 0.25$ showed that variable RBE gave steeper LKB curves compared to RBE1.1 and that the slope of the curves became even steeper for decreasing volume parameter n. The LKB models are shown in figures 4.24 and 4.25 with synthetic constraints in the form of two patients and nine patients, respectively. More heavily weighting the constraints also caused most models to become steeper. The TD50 generally decreased for increased weighting. Additionally, TD50 increased with increasing n for $w = 2$, but remained nearly constant with changing n for $w = 9$.

The LKB model parameters of the full brainstem volume for $n = 0.05$, $n = 0.16$ and $n = 0.25$, as well as both sets of synthetic weights, are shown in table 4.13. The models had TD50 values that fluctuated between 61 Gy(RBE) and 70 Gy(RBE), with CIs ranging from around 55 Gy(RBE) to approaching 80 Gy(RBE). Thus, the TD50 values with corresponding CIs are distributed around the TD50 established from previously established models [87, 88]. The m-parameter, on the other hand, ranged from approximately 0.150

to slightly exceeding 0.300, with CIs from 0.030 to 0.500, which consistently exceeded the previously established parameters. The CIs of both parameters were wider for RBE1.1 than the variable RBE and were narrowed slightly by decreasing n and more prominently by weighting the synthetic patients more heavily.

The AUROC-values decreased with decreasing n-values, and variable RBE had stronger AUROC-values to RBE1.1. However, the weighting had minimal impact on the AUROC. The AUROC-values ranged from slightly exceeding 0.50 to just surpassing 0.60, which are AUROC-values describing poor predictive models.

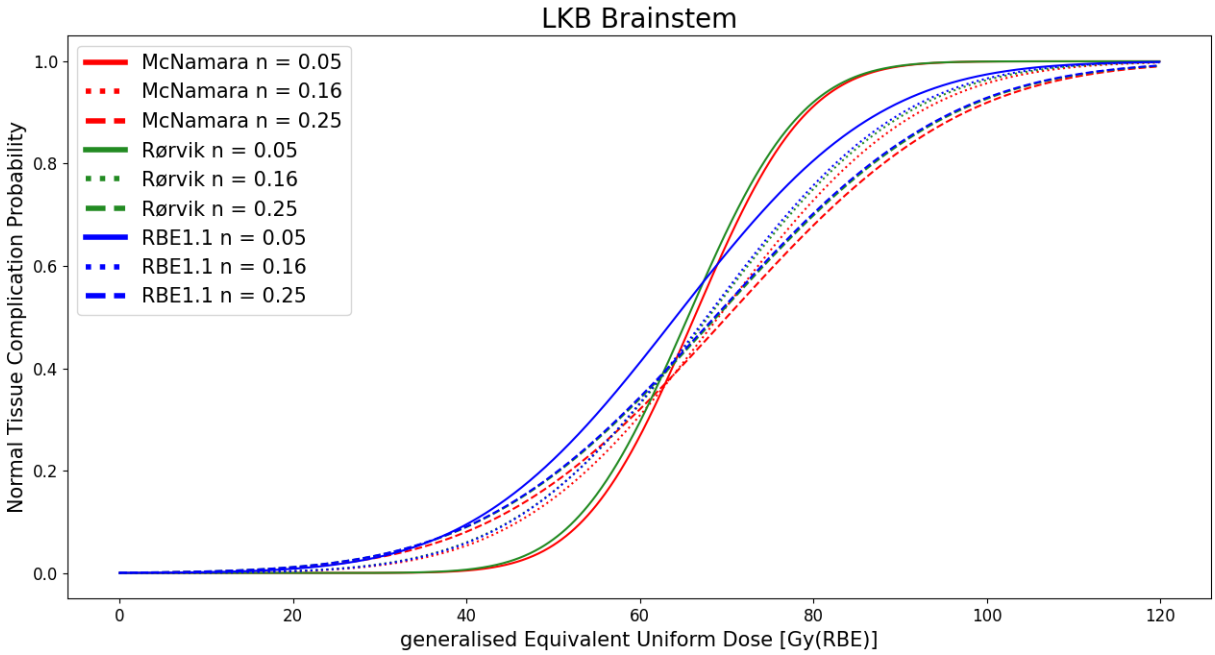

Figure 4.24: LKB models for the brainstem with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models. A synthetic constraint of two controls with gEUD = 20 Gy(RBE), and two cases with gEUD = 100 Gy(RBE) were included in the fitting of the model.

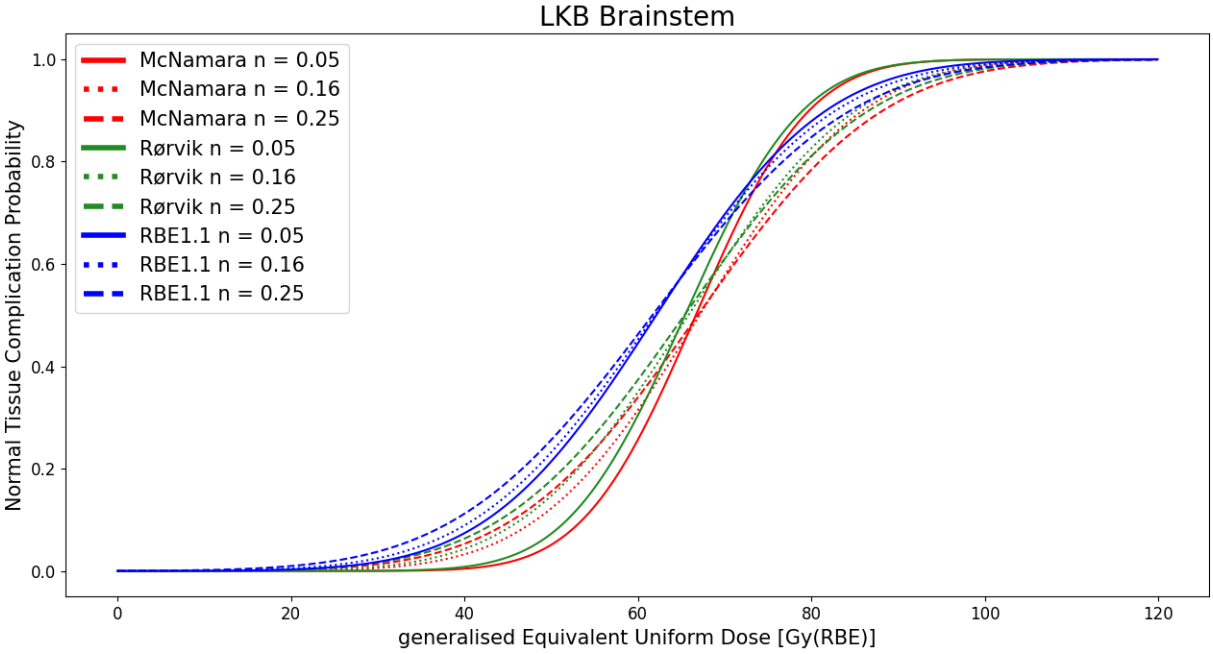

Figure 4.25: LKB models for the brainstem with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models. A synthetic constraint of nine controls with gEUD = 20 Gy(RBE), and nine cases with gEUD = 100 Gy(RBE) were included in the fitting of the model. Three different n-values are shown.

Table 4.13: Parameters of the LKB models for the full brainstem with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models, including 95% CIs of the parameters. A synthetic constraint of controls with gEUD = 20 Gy(RBE), and cases with gEUD = 100 Gy(RBE) were also included in the modelling, with w as seen in the table referring to the amount of patients introduced with each gEUD. Parameters established by Lyman [87] and Burman et al. [88] are included for comparison.
*Lyman [87], †Burman et al. [88]

RBE	w	n	m	TD50	AUROC
MCN	2	0.05	0.153 [0.030, 0.269]	66.28 [59.17, 73.04]	0.609
	9	0.05	0.153 [0.059, 0.252]	66.68 [59.82, 70.89]	0.609
	2	0.16	0.261 [0.099, 0.430]	69.04 [59.05, 76.97]	0.580
	9	0.16	0.219 [0.130, 0.358]	67.11 [60.86, 73.14]	0.580
	2	0.25	0.306 [0.115, 0.500]	70.02 [59.28, 78.79]	0.564
	9	0.25	0.249 [0.122, 0.368]	66.91 [60.61, 73.52]	0.560
RORW	2	0.05	0.157 [0.030, 0.287]	65.53 [57.78, 72.17]	0.609
	9	0.05	0.162 [0.067, 0.265]	65.43 [58.67, 69.88]	0.611
	2	0.16	0.264 [0.107, 0.444]	67.90 [57.77, 75.68]	0.582
	9	0.16	0.229 [0.112, 0.346]	65.83 [59.55, 71.95]	0.582
	2	0.25	0.312 [0.104, 0.521]	68.82 [56.66, 77.92]	0.564
	9	0.25	0.254 [0.141, 0.374]	65.36 [59.05, 72.15]	0.564

RBE1.1	2	0.05	0.287 [0.145, 0.473]	64.12 [54.75, 72.48]	0.553
	9	0.05	0.246 [0.130, 0.362]	62.12 [55.22, 68.66]	0.556
	2	0.16	0.261 [0.106, 0.443]	67.66 [57.60, 75.69]	0.543
	9	0.16	0.263 [0.141, 0.377]	61.94 [55.43, 69.04]	0.541
	2	0.25	0.312 [0.099, 0.512]	68.61 [57.57, 77.77]	0.506
	9	0.25	0.289 [0.201, 0.435]	61.69 [54.98, 69.03]	0.508
MCN*		0.05	0.1	64	0.609
RORW*		0.05	0.1	64	0.609
RBE1.1*		0.05	0.1	64	0.556
MCN†		0.16	0.15	65	0.580
RORW†		0.16	0.15	65	0.582
RBE1.1†		0.16	0.15	65	0.539

The LKB models for the midbrain and pons with RBE-weighted doses from the McNamara model and RBE1.1 are given in figure 4.26. The midbrain model had steeper curves than the pons, which increased for variable RBE compared to RBE1.1. Additionally, decreasing n led to increasing slopes of the curves. On the other hand, the medulla oblongata, which is given in figure 4.27, proved difficult to fit because most patients in case-control groups one and two received low dose to the substructure. Thus, the dosimetry did not correspond with representative NTCP of the cohort. Hence, the resulting model excludes these two case-control groups. Despite heavier weighting and the exclusion of a subset of the cohort, the medulla oblongata models were still more shallow than the other two substructures.

The parameters of the LKB models fitted for the three substructures are given in table 4.14. Decreasing n caused decreased m-values and increased TD50 values for all substructures, similar to the full brainstem. The pons had similar m-values to the full brainstem in the approximate range of 0.150 to 0.350. The TD50, however, had a narrow interval compared to the full brainstem, ranging from 63 Gy(RBE) to 67 Gy(RBE). The midbrain differed from the pons with greater variations in m and TD50 with altered n-values. The midbrain had m defined between 0.130 to 0.330 and TD50-values between 62 Gy(RBE) and 70 Gy(RBE). The heavier weighting and the exclusion of case-control groups in the fitting of the medulla oblongata models were found to have the effect of narrowing the CIs. However, the CIs of the resultant LKB models for the medulla oblongata were still in similar ranges to the other two substructures. The variations in m and TD50 parameters with changing n were also minor compared to the other substructures. The m-values were

similar at 0.190 to 0.300, while TD50 was higher at 67 Gy(RBE) to 73 Gy(RBE).

The models' AUROC-values are also included in table 4.14, with the highest AUROC-value of 0.64 observed for the midbrain with variable RBE. However, the AUROC-score fluctuated more for the midbrain and pons than for the medulla oblongata, with the medulla oblongata having consistent AUROC at around 0.63. The ranges of the midbrain and pons were similar, but the pons showed slightly lower AUROC-values. The lowest AUROC of all models in table 4.14 were seen for the pons with higher n-values and RBE-weighted dose from RBE1.1. The value was marginally lower than 0.5. AUROC was generally seen to decrease with increasing n parameter and was more robust for variable RBE compared to RBE1.1, similar to the models of the full brainstem.

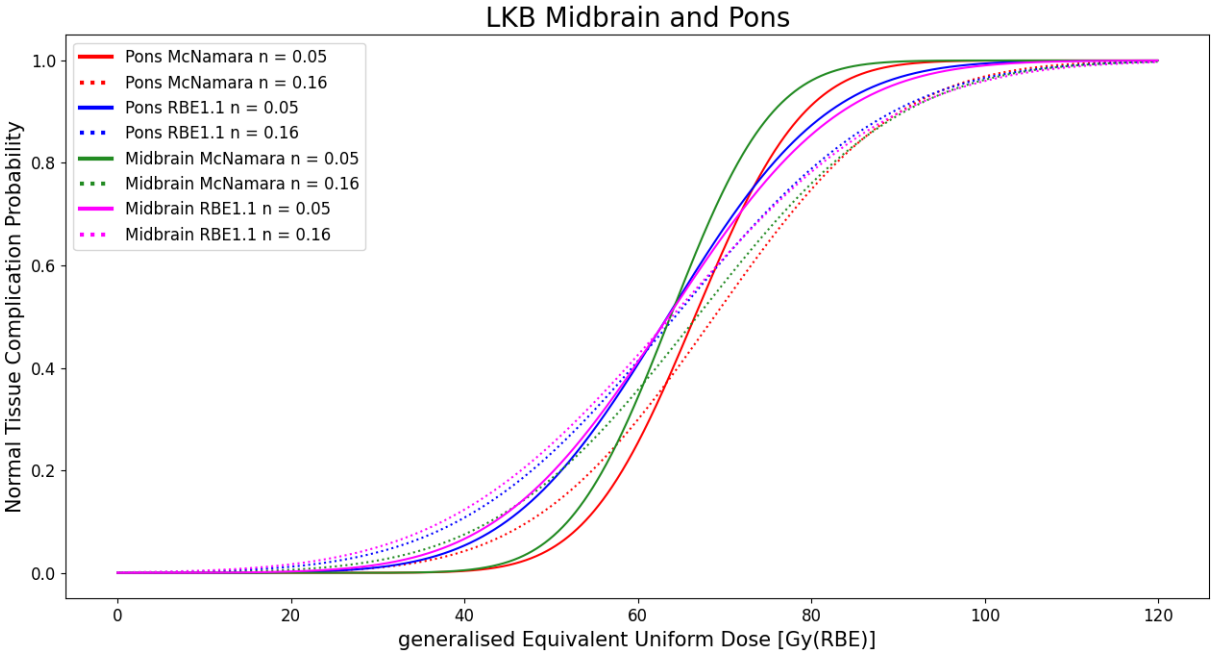

Figure 4.26: LKB models for the midbrain and the pons with RBE-weighted doses from the McNamara RBE model and RBE1.1 with $n = 0.05$ and $n = 0.16$. A synthetic constraint of two controls with gEUD = 20 Gy(RBE), and two cases with gEUD = 100 Gy(RBE) were also included in the modelling.

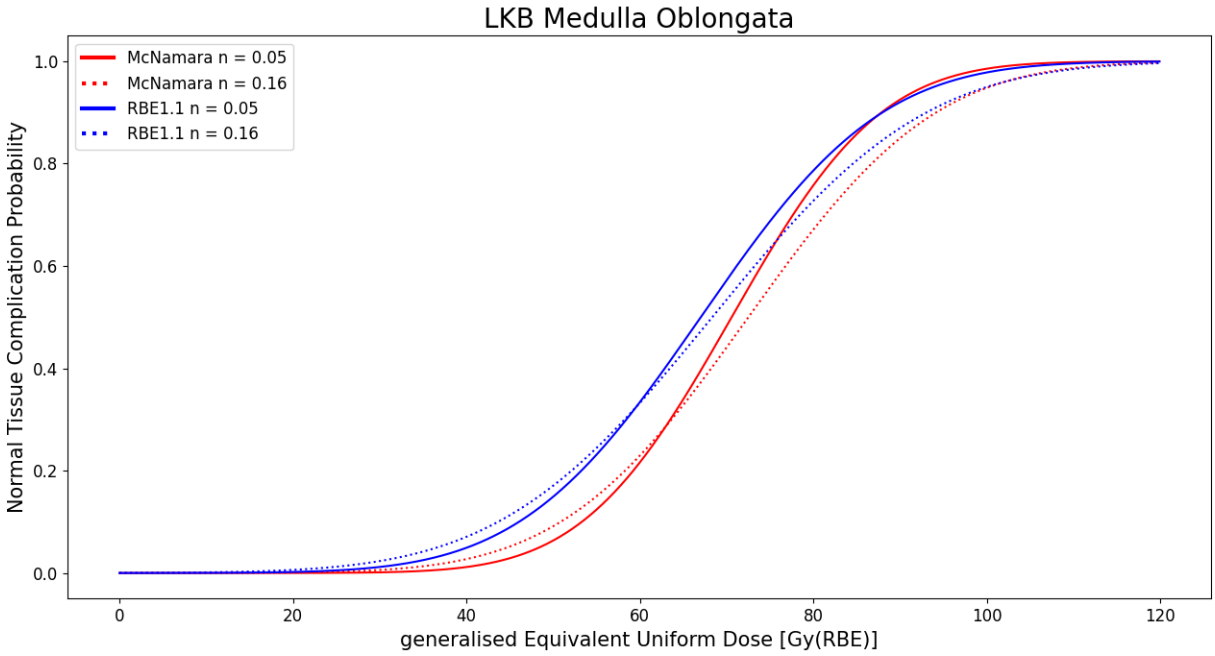

Figure 4.27: LKB models for the medulla oblongata with RBE-weighted doses from the McNamara RBE model and RBE1.1 with $n = 0.05$ and $n = 0.16$. A synthetic constraint of two controls with gEUD = 20 Gy(RBE), and two cases with gEUD = 100 Gy(RBE) were also included in the modelling.

Table 4.14: Parameters of the LKB models for the medulla oblongata, the midbrain and the pons with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models, including bootstrapped 95% CIs of the parameters. A synthetic constraint of two controls with gEUD = 20 Gy(RBE), and two cases with gEUD = 100 Gy(RBE) were also included in the modelling, two for the pons and the midbrain, and nine for the medulla oblongata.
*Excluding case-control groups one and two.

RBE	substructure	n	m	TD50	AUROC
MCN	Pons	0.05	0.149 [0.030, 0.279]	66.56 [59.73, 73.76]	0.613
		0.16	0.243 [0.099, 0.405]	68.81 [59.03, 75.80]	0.595
		0.25	0.285 [0.113, 0.462]	69.48 [58.85, 77.79]	0.576
RORW	Pons	0.05	0.151 [0.030, 0.280]	64.82 [57.90, 72.26]	0.613
		0.16	0.254 [0.100, 0.416]	67.38 [57.45, 74.47]	0.586
		0.25	0.297 [0.117, 0.480]	68.44 [57.63, 76.30]	0.572
RBE1.1	Pons	0.05	0.229 [0.110, 0.387]	63.42 [54.61, 70.00]	0.539
		0.16	0.305 [0.144, 0.488]	64.28 [54.71, 72.92]	0.498
		0.25	0.347 [0.163, 0.516]	64.84 [53.82, 74.54]	0.490
MCN	Midbrain	0.05	0.145 [0.030, 0.262]	63.74 [57.39, 71.01]	0.640
		0.16	0.279 [0.110, 0.446]	66.83 [55.92, 74.98]	0.601
		0.25	0.327 [0.096, 0.509]	69.93 [55.20, 77.07]	0.576

RORW	Midbrain	0.05	0.133 [0.030, 0.256]	61.97 [56.46, 70.11]	0.642
		0.16	0.289 [0.120, 0.454]	65.82 [54.64, 73.68]	0.609
		0.25	0.337 [0.115, 0.527]	66.35 [53.66, 75.61]	0.572
RBE1.1	Midbrain	0.05	0.246 [0.098, 0.419]	63.48 [53.10, 69.90]	0.574
		0.16	0.323 [0.133, 0.521]	63.90 [52.10, 72.60]	0.537
MCN	Med Obl*	0.05	0.191 [0.077, 0.310]	70.58 [63.77, 75.87]	0.629
		0.16	0.233 [0.104, 0.419]	72.52 [64.65, 78.65]	0.633
		0.25	0.253 [0.085, 0.456]	73.15 [64.79, 79.64]	0.633
RORW	Med Obl*	0.05	0.205 [0.100, 0.337]	69.42 [62.06, 74.78]	0.633
		0.16	0.244 [0.119, 0.436]	71.18 [62.81, 77.65]	0.636
		0.25	0.262 [0.106, 0.471]	72.21 [62.75, 78.47]	0.633
RBE1.1	Med Obl*	0.05	0.244 [0.129, 0.421]	67.02 [59.21, 73.37]	0.612
		0.16	0.282 [0.125, 0.513]	68.35 [59.93, 75.40]	0.629
		0.25	0.300 [0.129, 0.547]	68.85 [58.87, 76.53]	0.629

The LKB models for the substructures of the pons with $n = 0.05$ are shown in figure 4.28 and gave posterior, anterior, middle posterior and, finally, middle anterior in ascending order of steepness. RBE1.1 gave similarly shaped curves to variable RBE, but they were slightly shifted to the left, implying a more significant change in TD50 compared to the change in the m parameter. The model parameters for the pons substructures are given in table 4.15, and similarly to the previously studied structures decreasing n caused decreased m values, as well as slightly decreasing TD50 values. Posterior had high m and the highest TD50 among the pons substructures. Unlike the other three substructures, the pons posterior did not have models fitted for $n = 0.25$ since the variance in the data points made the fitting of a model with bootstrapped samples yield undefined CIs. Further, the middle posterior had significantly lower m and TD50 values than the posterior, while the middle anterior had the lowest range of m and similarly low TD50. The anterior had higher m-values than the middle anterior, but similar TD50.

The highest AUROC-value of all eight structures modelled for this thesis was observed at 0.667 for the middle anterior with variable RBE (both Rørvik and McNamara) and $n = 0.05$. In contrast, the RBE1.1 model for the same substructure barely exceeded 0.5. The middle posterior had AUROC-value around 0.57 for variable RBE, while RBE1.1 had lower values at around 0.53. Compared to the middle posterior, the posterior had lower AUROC-value for variable RBE but higher for RBE1.1. The anterior showed the consistently lowest AUROC-value, with the highest only slightly exceeding 0.5, while

many were below the 0.5 benchmark.

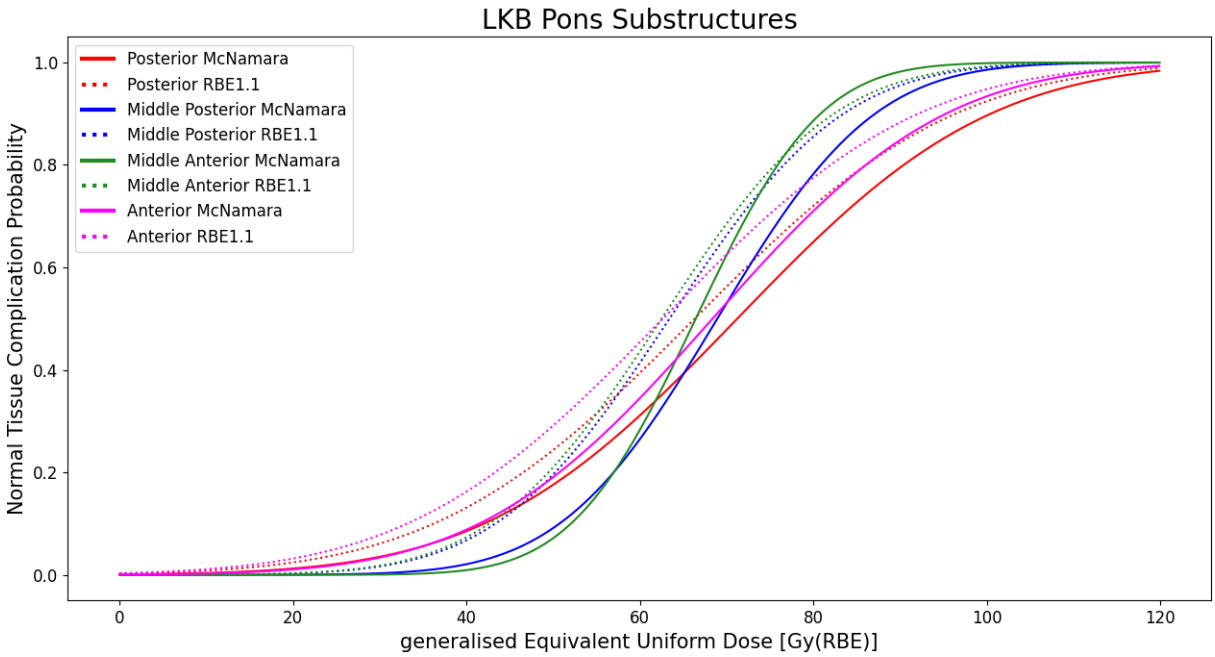

Figure 4.28: LKB models for the pons substructures with $n = 0.05$ with RBE-weighted dose from the McNamara RBE model and RBE1.1. A synthetic constraint of nine controls with gEUD = 20 Gy(RBE), and nine cases with gEUD = 100 Gy(RBE) and toxicity were also included in the modelling.

Table 4.15: Parameters of the LKB models for the pons substructures with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models, including bootstrapped 95% CIs of the parameters. A synthetic constraint of two controls with gEUD = 20 Gy(RBE), and two cases with gEUD = 100 Gy(RBE) were also included in the modelling.

RBE	substructure	n	m	TD50	AUROC
MCN	Posterior	0.05	0.320 [0.141, 0.565]	71.23 [63.41, 78.73]	0.566
		0.16	0.374 [0.165, 0.651]	71.33 [62.94, 79.75]	0.549
RORW	Posterior	0.05	0.328 [0.147, 0.569]	69.82 [62.56, 76.99]	0.564
		0.16	0.374 [0.139, 0.660]	70.16 [62.14, 78.03]	0.551
RBE1.1	Posterior	0.05	0.354 [0.181, 0.614]	66.31 [58.48, 74.46]	0.541
		0.16	0.400 [0.144, 0.488]	66.44 [58.28, 75.54]	0.549
MCN	Mid Post	0.05	0.206 [0.094, 0.318]	68.93 [62.91, 74.43]	0.570
		0.16	0.250 [0.105, 0.376]	68.94 [63.03, 75.18]	0.570
		0.25	0.263 [0.166, 0.446]	69.08 [62.69, 75.69]	0.572
RORW	Mid Post	0.05	0.328 [0.147, 0.569]	69.82 [62.56, 76.99]	0.570
		0.16	0.374 [0.139, 0.660]	70.16 [62.14, 78.03]	0.568
		0.25	0.273 [0.159, 0.435]	67.49 [61.20, 74.18]	0.568

RBE1.1	Mid Post	0.05	0.247 [0.155, 0.369]	63.42 [57.04, 70.18]	0.527
		0.16	0.289 [0.198, 0.454]	63.35 [56.50, 70.61]	0.516
		0.25	0.310 [0.210, 0.464]	63.44 [56.42, 70.79]	0.535
MCN	Mid Ant	0.05	0.169 [0.077, 0.279]	66.44 [59.60, 71.02]	0.667
		0.16	0.218 [0.107, 0.334]	66.62 [60.24, 72.44]	0.609
		0.25	0.237 [0.151, 0.361]	66.51 [59.97, 72.59]	0.569
RORW	Mid Ant	0.05	0.166 [0.073, 0.274]	65.07 [57.50, 69.46]	0.667
		0.16	0.225 [0.122, 0.343]	65.43 [58.80, 71.19]	0.611
		0.25	0.240 [0.150, 0.372]	65.07 [58.23, 71.59]	0.560
RBE1.1	Mid Ant	0.05	0.248 [0.123, 0.354]	62.51 [55.61, 68.86]	0.504
		0.16	0.278 [0.193, 0.401]	61.11 [54.28, 68.06]	0.516
		0.25	0.295 [0.207, 0.409]	61.01 [53.31, 68.43]	0.529
MCN	Anterior	0.05	0.307 [0.157, 0.469]	68.38 [61.13, 75.39]	0.527
		0.16	0.352 [0.186, 0.507]	66.72 [59.31, 75.72]	0.494
		0.25	0.371 [0.182, 0.540]	66.88 [59.29, 75.45]	0.494
RORW	Anterior	0.05	0.302 [0.159, 0.476]	67.44 [60.26, 75.25]	0.527
		0.16	0.357 [0.195, 0.512]	65.95 [57.41, 73.80]	0.488
		0.25	0.377 [0.205, 0.525]	65.40 [57.73, 74.32]	0.492
RBE1.1	Anterior	0.05	0.367 [0.208, 0.513]	62.65 [54.57, 70.84]	0.477
		0.16	0.410 [0.214, 0.566]	61.66 [53.03, 70.76]	0.496
		0.25	0.435 [0.238, 0.583]	61.41 [52.73, 70.44]	0.504

4.6 Cluster Analysis

The DBSCAN of the $D50\%$ to the substructures medulla oblongata, midbrain and pons did not show signs of classifying cases and controls together in the same clusters and instead grouped patients within case-control groups together. The distributions of cases and controls in clusters are given in table 4.16, while the actual clusters are shown in figure 4.29. Most patients were included in one large cluster, while a high ratio of patients from case-control groups one and two were placed in a separate, smaller cluster. Additionally, two patients were labelled noise. Most patients from groups one and two received much lower dose to the medulla oblongata than the rest of the patients in the cohort. Thus, the $D50\%$ to the medulla oblongata was the most central parameter in determining the resultant clustering. The McNamara and Rørvik clusterings both ended up

with silhouette scores of 0.69, while RBE1.1 had a silhouette score of 0.66. Both variable RBE and RBE1.1 gave silhouette scores describing clusterings of decent quality, with the dissimilarity of the clusters being evident from figure 4.29. The distribution of cases and controls did not show clear patterns in the separation of cases and controls. The ratio of cases to controls in cluster 2 was 33%, which exceeds the ratio of 25% for the whole cohort but not by a significant margin, and the cluster is small.

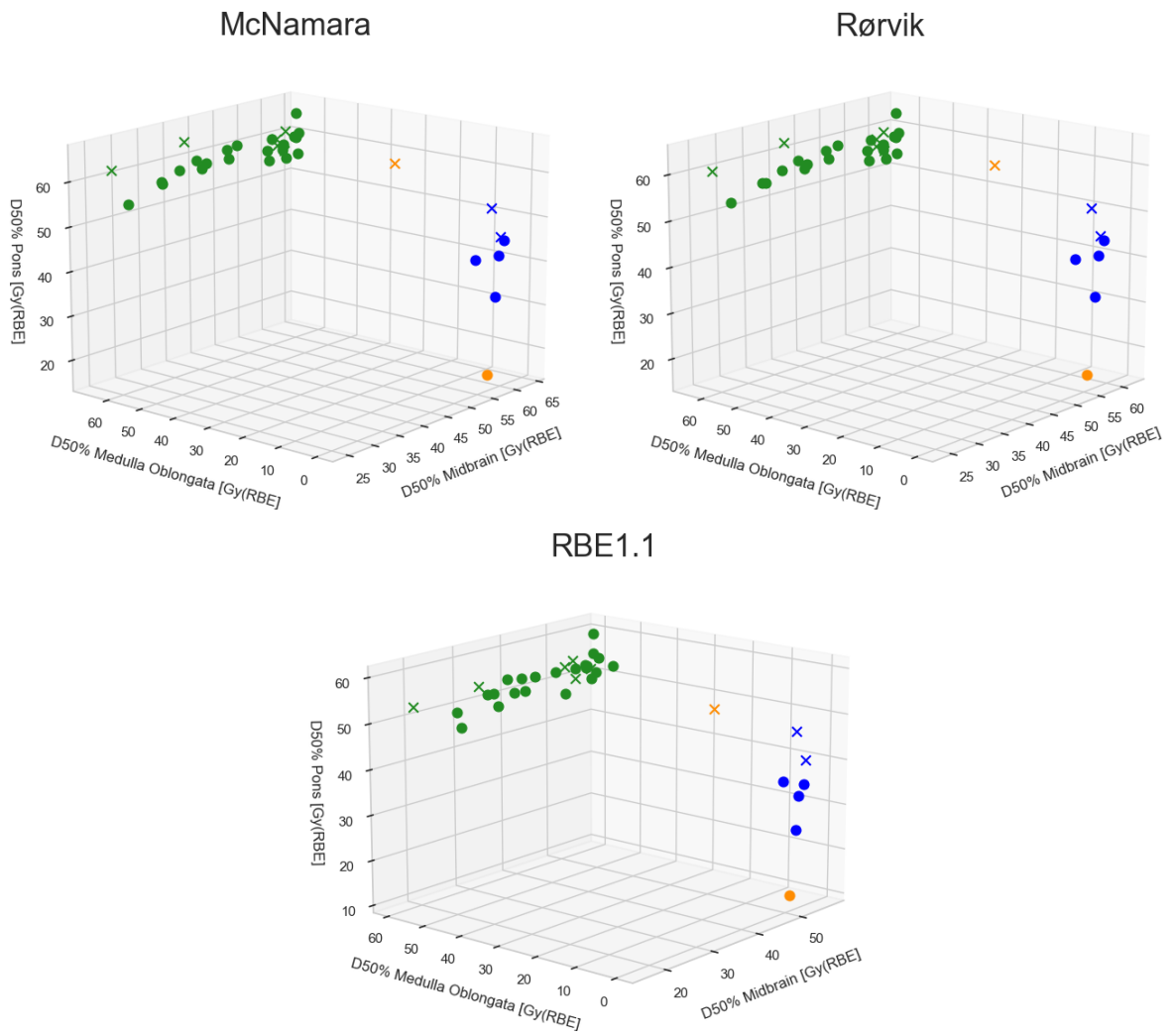

Figure 4.29: DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons with doses from RBE1.1, and the McNamara and Rørvik variable RBE models. Cases are given as crosses, while controls are given as dots.

Table 4.16: The distribution of cases and controls in the DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons. RBE-weighted doses are calculated with RBE1.1, and the McNamara and Rørvik RBE models.

Clusters	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	6	23	6	23	6	23
Cluster 2 (●)	2	4	2	4	2	4
Noise (●)	1	1	1	1	1	1

The hierarchical clustering of the $D50\%$ was similar to the DBSCAN. However, the clusters were slightly altered due to being obligated to incorporate the points deemed as outliers by the DBSCAN. The distribution of cases and controls is shown in table 4.17, while the clusters are shown in figure 4.30. Dendrograms of the clusters are given in figure 4.31 for the McNamara model and in figures B.1 and B.2 in appendix B for the Rørvik clustering and RBE1.1. The resultant clusters had improved silhouette scores compared to the DBSCAN, with silhouette scores of 0.77 for the two variable RBE models and 0.74 for RBE1.1. Similarly to the DBSCAN, the hierarchical clustering also showed case-control groups clustered together rather than cases and controls in separate clusters. The single smaller cluster was defined by case-control groups one and two in this analysis, as well, which is illustrated in figure 4.31, where the case-control groups of the patients are given in the *leaves*. The distribution of cases and controls, especially for RBE1.1, shows a higher ratio of cases to controls in the smaller cluster compared to the whole cohort.

Table 4.17: The distribution of cases and controls in the Hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons. RBE-weighted doses are calculated with RBE1.1, and the McNamara and Rørvik RBE models.

	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	7	22	7	22	6	22
Cluster 2 (●)	2	5	2	5	3	4

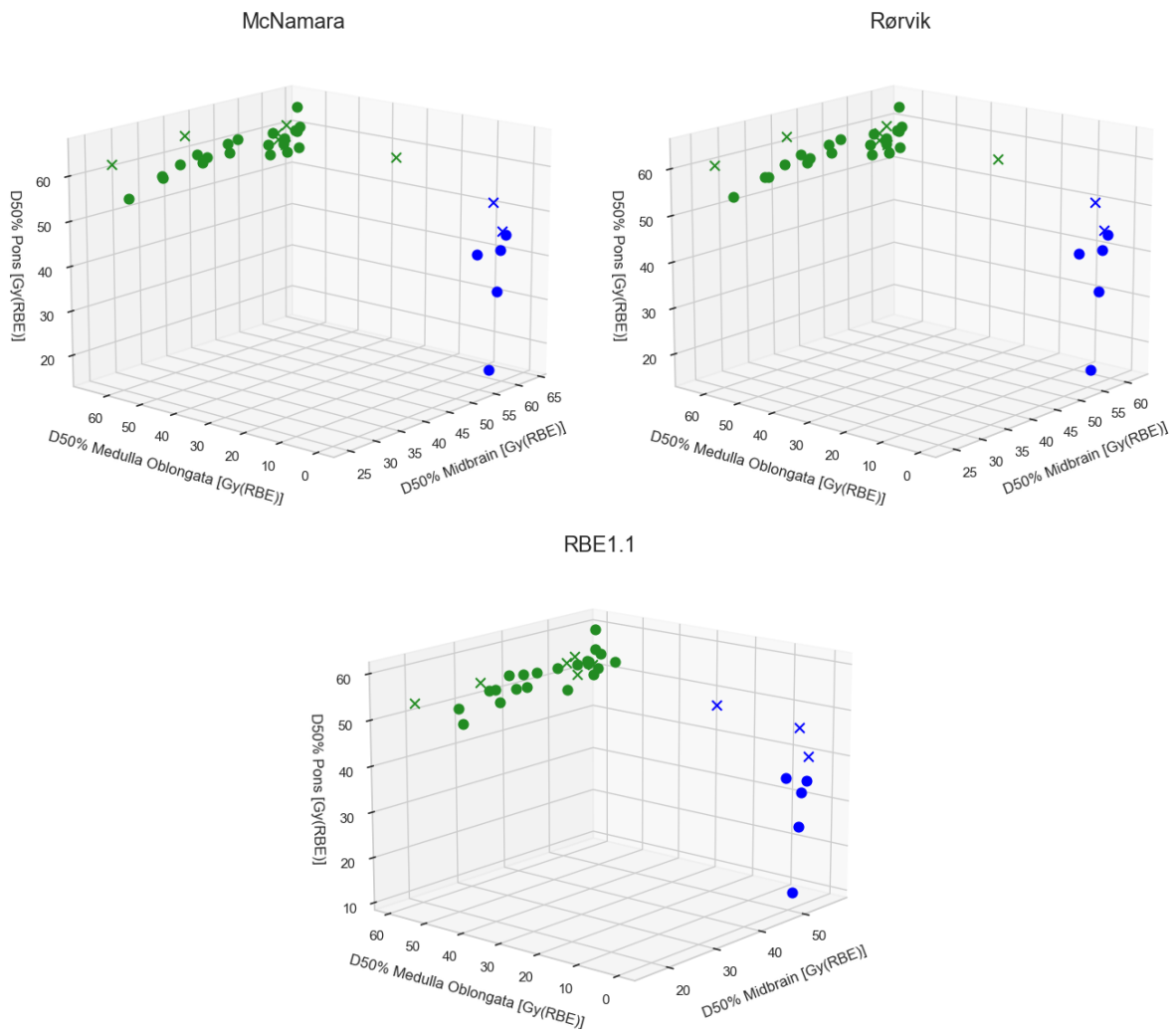

Figure 4.30: Hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with doses from RBE1.1, and the McNamara and Rørvik variable RBE models. Cases are given as crosses, while controls are given as dots.

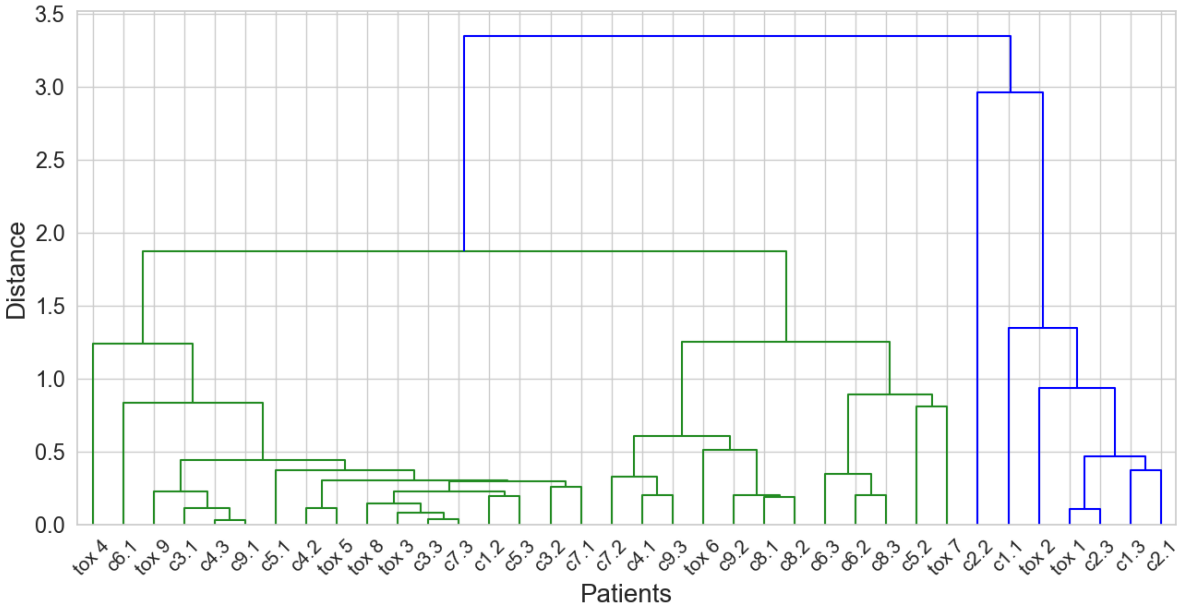

Figure 4.31: Dendrogram for the hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted doses from the McNamara RBE model. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

The clustering of the gEUD to the midbrain and the pons gave one large cluster, one smaller cluster and between two and five points labelled noise. The distribution of cases and controls are given in table 4.18, while the clusters are shown in figure 4.32. Most cases were part of the larger central cluster, thus setting a minimum threshold for the gEUD received by the cases to both the midbrain and the pons. However, the clusters were less distinct compared to the clustering of the $D50\%$, which is reflected in their silhouette scores. The variable RBE models had silhouette scores of 0.36, while RBE1.1 showed slight improvement at 0.40. Another difference from the $D50\%$ was that the gEUD values changed to a greater extent between RBE1.1 and the variable RBE models. This resulted in different clusters forming, particularly evident from the DBSCAN is the presence of more noisy points for RBE1.1.

Table 4.18: The distribution of cases and controls in the DBSCAN of the gEUD to the midbrain and the pons. RBE-weighted doses are calculated with RBE1.1, and the McNamara and Rørvik RBE models.

	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	8	21	8	21	8	21
Cluster 2 (●)	0	3	0	3	1	4
Noise (●)	1	4	1	4	0	2

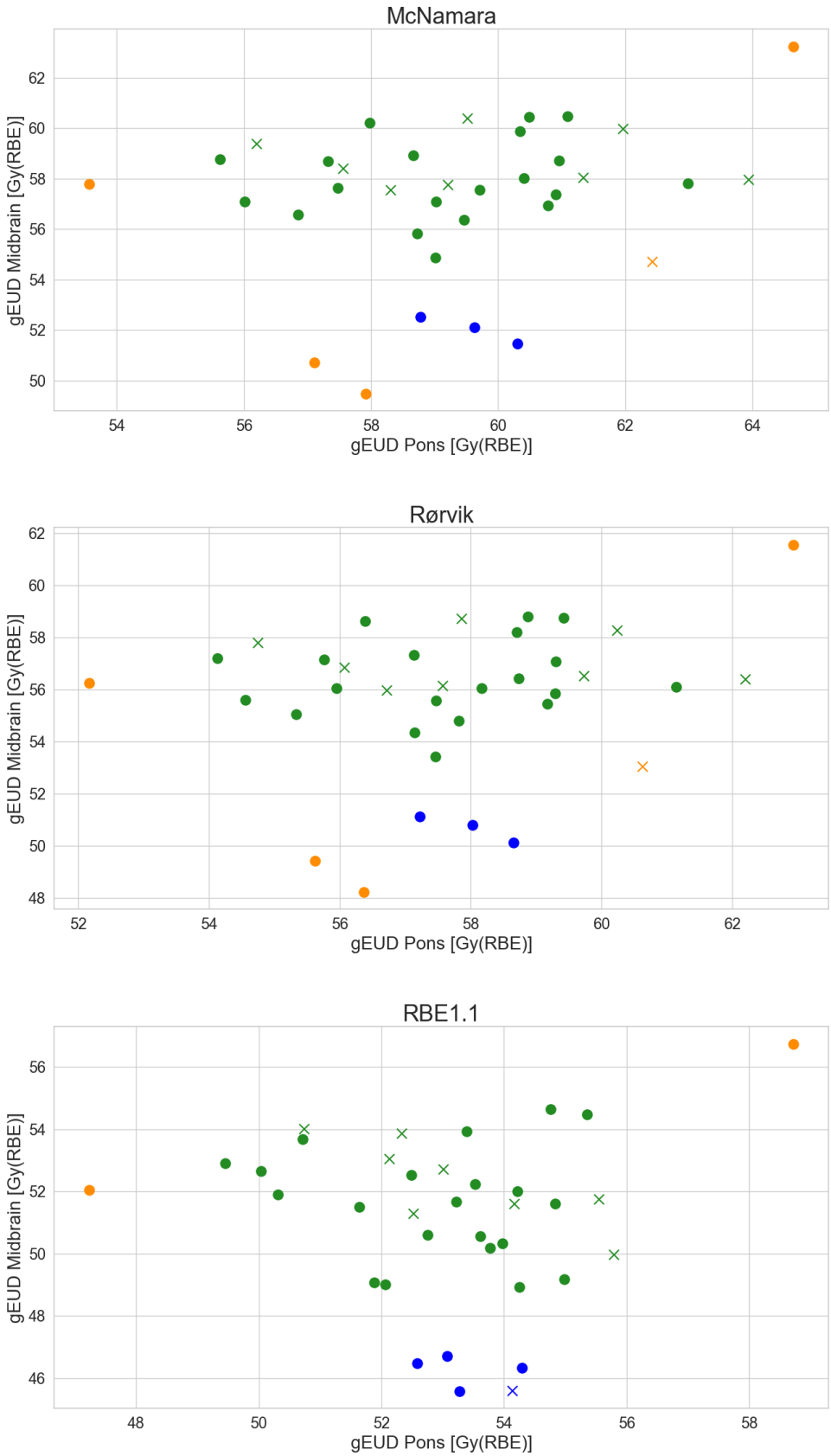

Figure 4.32: DBSCAN of the gEUD ($n = 0.05$) to the midbrain and the pons with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik variable RBE models. Cases are given as crosses, while controls are given as dots.

The hierarchical clustering of the gEUD showed deviations from the DBSCAN, which can be mainly attributed to the obligation of incorporating all points in clusters. The distribution of cases and controls in the hierarchical clustering of the gEUD is given in table 4.19, while the clustering itself is shown in figure 4.33. Dendrograms are provided for the McNamara model in figure 4.34, while appendix B includes the Rørvik and RBE1.1 dendrograms in figures B.3 and B.4, respectively.

The hierarchical clustering of the gEUD yielded better silhouette scores for fewer clusters, with McNamara having a silhouette score of 0.42 for two clusters but a local maximum for five clusters at 0.32. Similarly, Rørvik also had a maximum silhouette score of 0.44 for two clusters but with a local maximum of 0.32 for five clusters. However, the silhouette score of RBE1.1 was instead decreasing for increasing number of clusters, from 0.50 at two clusters to 0.30 at seven clusters. The five cluster models were chosen in all cases since more information is yielded by a higher number of clusters, and an equal number of clusters maintained grounds for comparison between the models. However, the high silhouette scores of the models with fewer clusters suggest that the data is too similar to yield a solid number of clearly separated clusters. All clusterings ended up with two reasonably similar central clusters and three outer clusters ranging from one to five points per cluster. Variable RBE had all cases incorporated in the two central clusters, while RBE1.1 had all but one case in the centre.

Table 4.19: The distribution of cases and controls in the hierarchical clustering of the gEUD to the midbrain and the pons. RBE-weighted doses are calculated with RBE1.1, and the McNamara and Rørvik RBE models.

	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	4	12	4	11	7	17
Cluster 2 (●)	5	8	5	9	1	4
Cluster 3 (●)	0	5	0	5	1	4
Cluster 4 (●)	0	1	0	1	0	1
Cluster 5 (●)	0	1	0	1	0	1

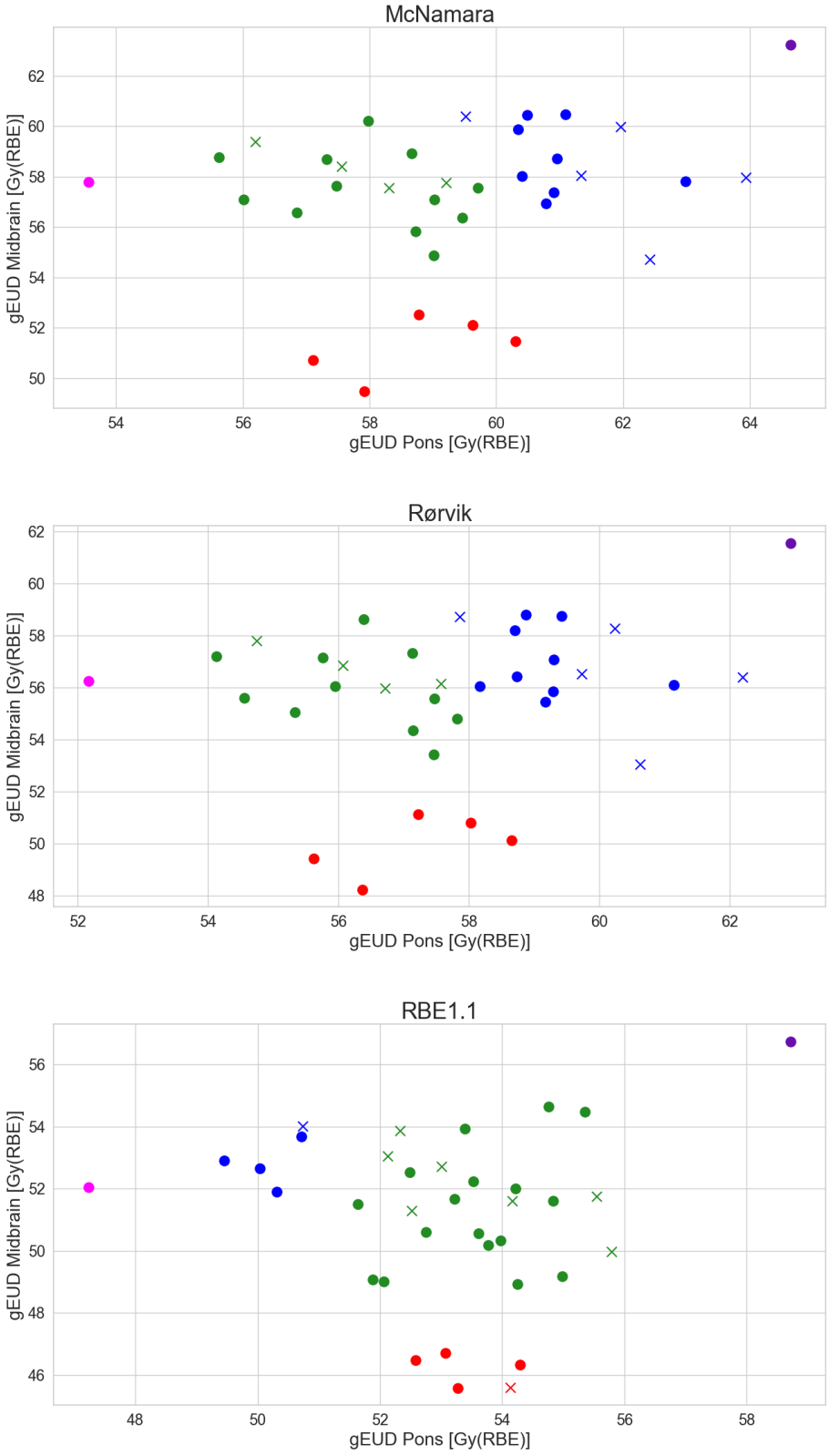

Figure 4.33: Hierarchical clustering of the gEUD ($n = 0.05$) to the midbrain and the pons with RBE-weighted doses from RBE1.1, and the McNamara and Rørvik RBE models. Cases are given as crosses, while controls are given as dots.

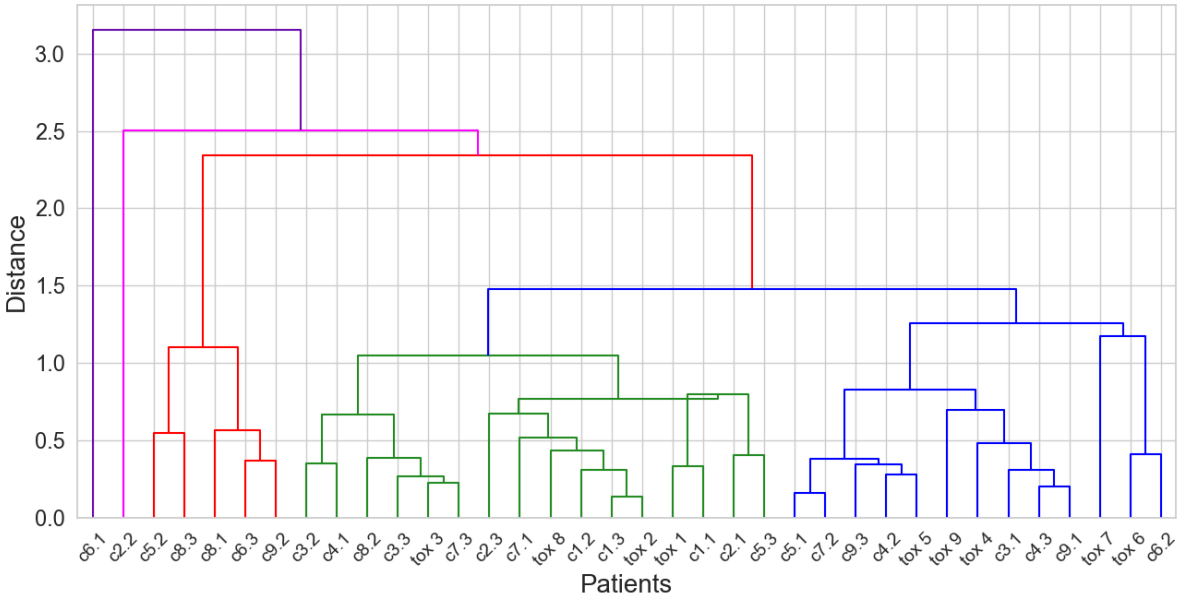

Figure 4.34: Dendrogram for the hierarchical clustering of the gEUD (n=0.05) to the midbrain and the pons with RBE-weighted doses from the McNamara RBE model. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

Further clustering was also performed to investigate the cohort, but little of note was discovered. The clustering of the $D50\%$ for the pons substructures is provided in appendices attached to this thesis, with appendix B showing the hierarchical clustering and appendix C showing the DBSCAN.

Chapter 5

Discussion

This thesis has investigated a cohort of paediatric proton therapy patients from a case-control study, where the cases experienced toxicity in the form of CTCAE grade 2+ brainstem necrosis following the treatment. RBE-weighted doses recalculated with variable RBE models from McNamara et al. [21] and Rørvik et al. [22], as well as $\mathrm{LET_d}$-weighted dose from Unkelbach et al. [70], showed that cases received higher RBE-weighted dose on average compared to controls. Similar trends were also observed for most substructures. Additionally, $\mathrm{LET_d}$ was found to be slightly higher for average cases compared to average controls for most substructures. The gEUD of the brainstem and its substructures showed similarly increased gEUD to cases compared to controls. At the same time, different volume parameter n was not observed to affect the differences to a large extent. Based on AUROC-values, the LKB models were shown to improve with lower n, and the curves were shown to be more shallow for substructures receiving less average RBE-weighted dose, such as the medulla oblongata. Additionally, this thesis has presented the first time LKB models for brainstem necrosis have been fitted to brainstem substructures. The LKB models for most substructures were also based on concentrated points around the TD50 which provided increased nuance in a dose range associated with uncertain risk of brainstem necrosis. Finally, the cluster analysis showed that data points were difficult to separate into solidly dissimilar clusters and that outliers and the matching of patients heavily impacted the clustering.

Average DVHs showed cases with a higher variable RBE-weighted dose than controls, particularly in the high dose range. The corresponding $D50\%$, $D10\%$ and $D_{\max}$ further showed that averages were systematically higher for cases compared to controls for variable RBE, while the same was not seen for RBE1.1. However, the 95% CIs were wide compared to the point estimate between the averages. Additionally, studying the difference between cases and the average of their matched controls showed similar results with consistently higher RBE-weighted doses to cases than controls for variable RBE and for most substructures. Still, few differences were shown to be statistically significant. Statistical significance was found for the $D10\%$ to the full brainstem and pons middle

posterior and $D_{\max}$ to the midbrain. The position of the statistically significant substructures varies and does not suggest prioritising any subvolumes over others. However, despite few differences being statistically significant, the systematic differences observed here for higher dose levels may warrant further study. This is strengthened by previous studies concluding with toxicity rates being linked with higher dose and smaller volume constraints [101, 102]. However, several other studies have also reported statistical significance or predictive ability for dose statistics incorporating higher volume fractions, and lower dose constraints [102, 114, 115].

The difference between variable RBE and RBE1.1 is likely influenced by LET changes since the variable RBE models predict increasing RBE with increasing LET. LET_d-volume histograms showed higher LET_d to cases compared to controls for most substructures and most LET_d intervals. The LET_d distributions of the case-control groups also showed that, on average, the cases received higher LET_d to the brainstem substructures than the controls, although the CIs were wide. Differences between cases and the average of their corresponding controls also showed a majority of cases with higher $L50\%$ and $L_{\max}$ than controls. However, no statistical significance was found, which, similarly to the variable RBE, suggests that further study of LET_d distributions and toxicity is needed.

Quantifying the effect of LET_d on toxicity is difficult since LET cannot be measured directly, so different simulations of the LET is the typical experimental approach, where MC methods are commonly utilised [25, 50, 116]. Additionally, many low dose areas are associated with moderate LET_d-values. The association between elevated LET and higher RBE is well-known [16] and one would expect this to translate to increased risk of toxicity compared to a constant RBE [117]. Low-grade toxicity in the form of MRI image changes following proton therapy of brain tumours has been described and associated with regional RBE and LET differences [26, 27, 28]. However, studies investigating MC simulated LET and brainstem necrosis as clinical endpoint have so far been inconclusive in finding any solid correlation [24, 25].

The $(\alpha/\beta)_\text{x}$ is, in addition to the LET_d, another central parameter for the variable RBE models studied for this thesis. While different brainstem substructures failed to show systematic differences between cases and controls specific to certain substructures, RBE-weighted dose was seen to increase with decreasing $(\alpha/\beta)_\text{x}$. However, the assumption that the full brainstem has a constant $(\alpha/\beta)_\text{x}$ may be overly simple. Parts of the motivation for this thesis are the regional differences in radiosensitivity of the brainstem

substructures [17, 34, 35], which suggests separate $(\alpha/\beta)_x$ for individual substructures. The assumption of constant $(\alpha/\beta)_x$ across the full brainstem means the recalculated variable RBE-weighted dose between substructures does not account for potential regional differences in $(\alpha/\beta)_x$-ratio. This would result in calculated RBE-weighted doses being more similar between substructures than the reality of the individual tissues of the brainstem would be described by. Higher $(\alpha/\beta)_x$ is also associated with higher NTCP [118], which emphasises the importance of accounting for $(\alpha/\beta)_x$ in treatment planning.

Differences in dose between patients, particularly across case-control groups and different substructures, were considerable. Notably, the medulla oblongata received a low dose for most patients in case-control groups one and two due to the position of the tumours treated. Excluding case-control groups one and two from the statistical analysis could possibly have led to statistical significance between case and controls for more substructures. In addition to significant dose differences, the LET_d of the radiation was also varied between patients. While the distal end of the Bragg peak was positioned to avoid the brainstem to the greatest extent possible, the LET_d-volume histograms still showed high average LET_d for several substructures. The inter-patient variations were also seen to decrease for higher dose statistics, which is partly explained by the $D10\%$ and $D_{0.1cc}$ being used as grounds for the matching of the case-control groups. Another possible explanation is the different volume coverage of the substructures, which would affect lower dose statistics to a greater extent than higher dose statistics.

The gEUD calculated with Lyman's proposed $n = 0.05$ [87] and Burman et al.'s $n = 0.16$ [88] showed higher average values for cases than controls, with differences increasing for variable RBE models and lower n. This is likely because a low n-value highly weights small volumes, which are more dependent on doses approaching maximum doses. This is consistent with the statistical significance seen for $D10\%$ and D_{max}. Additionally, the brainstem is generally considered a serial organ. Thus, damage to smaller volumes will significantly affect the function of the entire organ, leading to gEUD being calculated with low n-values.

The LKB models fitted in this thesis for both the full brainstem and brainstem substructures with RBE1.1 had reasonably similar TD50 values to the already established models for the full brainstem by Lyman [87] and Burman et al. [88], while variable RBE was slightly higher. In addition, the previously established model parameters tested on our data had among the higher AUROC-values observed in this thesis, particularly for the

variable RBE. However, the m-values for the models fitted in this thesis were consistently higher than parameters from previous LKB models [87, 88], and the curves established were subsequently more shallow. This was the result for both the full brainstem and its substructures. Stronger AUROC-values were also found for models with lower n-values, and when n was determined by fit to the data points, even lower n-values were preferred. This further demonstrates the brainstem being sensitive to dose changes in smaller volumes and strengthens the importance of high dose level constraints for the brainstem. However, no substructures were found to give solid ground for prioritising over the others. Further, the bootstrapped CIs were also fairly wide, even after the synthetic constraints were introduced. This is correlated with the substructures receiving the most varied dose between patients, but is also likely to be partly explained by the small dataset leading to high degrees of patients being represented several times within a bootstrapped sample and weighted very heavily for the resultant model, which will be partiularly evident if the repeated patients are outliers.

The AUROC-values were also not as high as one would prefer for a predictor, with a considerable number of the models having AUROC-values lower than 0.5. The highest values of around 0.65 were mostly seen for variable RBE and small n-values, and the midbrain and pons middle anterior were the substructures with the highest AUROC. However, the AUROC-values presented in this thesis compares poorly to previous LKB studies evaluating based on AUROC, with examples being AUROC-values of 0.82 [119], 0.78 [120], and 0.86–0.90 [121]. The poor comparison might be caused by the patients being so highly concentrated that they were hard to separate based on a single metric, as well as the fact that patients within a case-control group were associated with similar dosimetric values, despite differing in the clinical outcome. The validation of the generalisability of an NTCP model is an essential aspect of NTCP modelling and is discussed in three of the seven TRIPOD guidelines for clinical machine learning [37]. The AUROC gives an unbiased estimate of the performance and provides a direct point of comparing performance between otherwise different NTCP models. AUROC was particularly important for the LKB models presented in this thesis since the literature reports few LKB model fits for brainstem toxicity as clinical endpoint.

The dataset consists of patients within a critical dose range representing a significant, but uncertain risk of brainstem necrosis. The critical dose range is subsequently referring to the dose range representing clinically relevant doses to the brainstem with a significant risk of brainstem necrosis, in LKB terms, the adjacent area of the TD50. However, the

concentration of the patients around the TD50 means that the high and low dose regions are subjects of uncertainty. Similar LKB models from other studies represent wider dose spans and typically includes more patients to maximise accuracy. Thus, most LKB models studied for this thesis, while describing an assortment of different clinical endpoints, shared the inclusion of larger cohorts than ours at $N = 166$ [119], $N = 116$ [121], $N = 388$ [122], $N = 203$ [123], and N between 49 and 382 (data from eleven previous studies) [124]. The selection of patients for the case-control study meant that the fit to the critical dose range was prioritised but that the models did not accurately represent higher and lower dose ranges. The synthetic constraints were the chosen solution to the limited dose range. This was justified by the critical dose range remaining to a large degree defined by the cohort, while the low and high dose ranges were constrained to certain toxicity for high doses and no toxicity for low doses. Furthermore, the shallow curves due to large m-values overpredicted NTCP for lower doses, compared to observed incidence rates [25, 101, 102, 115, 125, 126]. Thus, the constraints were based on our knowledge of the low incidence of necrosis in lower dose ranges. This is supported by the fact that studies reporting brainstem toxicity typically concern tumours where excessive dose to the brainstem is unavoidable [25, 101, 102, 115], the observation of statistical significance for dose statistics to the full brainstem volume [102], as well as predictive ability for the percentage of the infratentorial brain volume irradiated by doses between 50 Gy(RBE) and 60 Gy(RBE) [114].

While basing the cohort on a case-control design allows detailed study of a dose range associated with uncertain risk of brainstem necrosis, the risk of brainstem necrosis is difficult to translate to a general population from a specifically matched case-control study. An important aspect of clinical ML and modelling is the generalisability of a model to a population. This is quantified by minimising the bias of the model, which are trends specific to the dataset used to fit the model that do not translate over to a general population [84, 127]. Thus, the model would be most appropriately utilised on patients prescribed treatment resulting in similarly high dose to the brainstem, but these are generally the patients where the model would be useful regardless. However, by basing the study on concentrated points around the TD50 with three controls for every case, the LKB model might be partly biased towards a synthetic incidence rate of 25% brainstem necrosis in this dose range. Incidence rates of brainstem necrosis are generally found to be low but uncertain, and most studies consider lower doses than the ones included in this study which could potentially underpredict the true incidence rate. Typical rates of brainstem necrosis lie between 0% and 3.8% [25, 101, 102, 115, 125, 126]. However, the studies included base incidence rates on cohorts of between 70-313 patients, which are generally small samples to derive trends for entire populations from. The largest study by Indelicato et al. [102]

with 313 patients found a 3.8% 2-year incidence rate of brainstem necrosis with average $D50\% = 37.4$ Gy(RBE) and $D10\% = 51.3$ Gy(RBE) to the whole cohort. Our study had an average $D50\%$ of 49.1 Gy(RBE) and an average $D10\%$ of 55.4 Gy(RBE), both from the RBE1.1 factor, which one could assume to be associated with a higher toxicity rate.

A possible way of reducing bias is to fit the models based on all 954 patients the case-control groups were initially extracted from. This would achieve a stronger basis for the models in a wider dose interval including more patients. However, recalculation of all 954 patients would be time-consuming beyond the scope of this thesis, and many of the patients received doses lower than the ones associated with a significant risk of brainstem necrosis. Therefore, the synthetic constraint was considered a decent approximation, despite precedence for introducing synthetic constraints not being found in the literature studied for this thesis. The introduced constraints also result in parts of the modelling accounting for data points not directly associated with the patients in the cohort. However, constraints being applied to both variable RBE and RBE1.1 equally maintained grounds for comparison of the models for the different RBE-weighted doses and substructures. Additionally, the integrity of the critical dose range around the TD50 remained largely intact since the constraints were not represented in this dose range. The synthetic constraints were also not treated as actual patients for the AUROC scores, which means they were only part of the model calibration.

Another possible prospect for the future is to evaluate different models than LKB on the same patient cohort. Incorporating $\mathrm{LET_d}$ directly as a parameter could be an idea since this was also shown to be higher on average for cases compared to controls. One could also account for several degrees of toxicity, which is not wholly uncommon for LKB models either [119, 120, 122]. Additionally, while there is a lack of LKB models describing brainstem necrosis as clinical endpoint, other NTCP models fit to minor brainstem toxicity (i.e. MRI image changes) have been constructed. Peeler et al. [26] developed a probit model, which is very similar to a logit model, based on the $\mathrm{LET_t}$ of a cohort of 34 paediatric ependymoma (a type of CNS tumour). The model achieved an AUROC-value of 0.91. Bahn et al. [27] described a multivariate model incorporating absorbed dose, $\mathrm{LET_d}$ and distance to the ventricular system to establish a model for predicting radiation-induced brain lesions following brain tumours ($N = 110$). This achieved an AUROC-score of 0.94. Thus, there is a possibility that other models might have proved a better fit to our data set and by using a multivariate model, one could have evaluated the cohort based on both dosimetric and non-dosimetric parameters directly in the model.

The clustering generally showed the similarity of the patients within a case-control group instead of uncovering less apparent trends in the data. This was not wholly unexpected since they were matched based on similar parameters to the ones clustered. The clustering being based solely on dosimetric parameters is a weakness described by the TRIPOD guidelines [86]. However, non-dosimetric parameters were indirectly considered through the matching criteria of the case-control groups. Further, the clustering generally yielded solid silhouette scores for the simpler models only because the data was not significantly separated for most substructures. This also explains why the medulla oblongata was dominant in determining the clusters' shape. However, the data did include some considerable outliers, which currently formed individual smaller clusters. However, for larger datasets, the DBSCAN could prove a valuable asset in negating the effect of outliers.

Limitations of the cluster analysis performed for this thesis were primarily two-fold. The primary and fairly evident cause was the small size of the data set, which prohibited larger clusters from forming, and individual outliers were dominant in shaping the resultant clusters. The ability of clustering algorithms to investigate a large set of data quicker than could be achieved manually was also not utilised to its fullest potential in this thesis. Additionally, the parameters were purely dosimetric, meaning they could not be directly clustered against each other since all were co-dependent. This did not allow for the simultaneous investigation of several parameters, which is another advantage of cluster analysis as a data processing tool. Thus, this thesis has not found a strong cause for utilising cluster analysis on this cohort. However, while not widely utilised in clinical radiotherapy patient studies, cluster analysis is still valuable for clinical investigations. An important example is the exploration of gene expression, where cluster analysis is a valuable asset due to the large number of genes and complex genome systems to be studied [128, 129].

Chapter 6

Conclusion

In the assessment of the cases and controls, modest, but systematic increases in $\mathrm{LET_d}$ were observed for the average case compared to the average control. The differences were consistent across all $\mathrm{LET_d}$ ranges and most substructures. However, $L50\%$ and $L_{\max}$ showed significant CIs compared to the average differences between cases and controls, and no statistical significance was observed for studies of the cohort.

The average RBE-weighted doses were also higher for cases than controls for the variable RBE models, which was not seen for RBE1.1. In addition, the difference was found to increase for higher dose statistics, where statistical significance in $D10\%$ and $D_{\max}$ was found for certain studied substructures. This, along with increased $\mathrm{LET_d}$, supports RBE1.1 being an inaccurate assumption and that variable RBE may partly explain the incidence of toxicity.

The LKB models fitted for the brainstem and brainstem substructures showed similar fit parameters to previously established LKB models for the brainstem, and substructures receiving less consistent RBE-weighted dose across patients and case-control groups showed more shallow LKB curves. Based on AUROC-values, the LKB models improved with decreasing volume parameter (n) in the calculation of gEUD, which suggests prioritising dose constraints to smaller volumes. However, the LKB models were not associated with solid AUROC-values which could have been improved with a larger data set or exploring the fitting of alternative NTCP models to the dataset.

Cluster analysis showed few trends in the separation of cases and controls, but instead distinguished different case-control groups from each other. The dataset was likely too small for cluster analysis to be used to its full potential, but the DBSCAN's ability to filter out outliers was demonstrated.

Bibliography

[1] World Health Organization. *Cancer*. https://www.who.int/news-room/fact-sheets/detail/cancer. Accessed: 09.05.2021.

[2] Kreftregisteret. *Årsrapport for barnekreft 2019*. https://www.kreftregisteret.no/globalassets/publikasjoner-og-rapporter/arsrapporter/publisert-2020/arsrapport-2019-nasjonalt-kvalitetsregister-for-barnekreft.pdf. Accessed: 14.05.2021.

[3] World Health Organisation. *Global Intitiative for Childhood Cancer*. https://www.who.int/cancer/childhood-cancer/en/. Accessed: 09.05.2021.

[4] Leslie Robison and Melissa Hudson. "Survivors of childhood and adolescent cancer: Life-long risks and responsibilities". In: *Nature reviews. Cancer* 14 (Dec. 2013). DOI: 10.1038/nrc3634.

[5] Camilla Stokkevåg et al. "Normal tissue complication probability models in plan evaluation of children with brain tumors referred to proton therapy". In: *Acta Oncologica* 58 (July 2019), pp. 1–7. DOI: 10.1080/0284186X.2019.1643496.

[6] Paul B. Romesser et al. "Proton beam radiation therapy results in significantly reduced toxicity compared with intensity-modulated radiation therapy for head and neck tumors that require ipsilateral radiation". In: *Radiotherapy and Oncology* 118.2 (2016), pp. 286–292. ISSN: 0167-8140. DOI: https://doi.org/10.1016/j.radonc.2015.12.008.

[7] Tian Xiufang et al. "The evolution of proton beam therapy: Current and future status (Review)". In: *Molecular and Clinical Oncology* (Nov. 2017). DOI: 10.3892/mco.2017.1499.

[8] Tineke W.H. Meijer, Dan Scandurra, and Johannes A. Langendijk. "Reduced radiation-induced toxicity by using proton therapy for the treatment of oropharyngeal cancer". In: *The British Journal of Radiology* 93.1107 (2020), p. 20190955. DOI: 10.1259/bjr.20190955.

[9] Philip Mayles, Alan E. Nahum, and Jean-Claude Rosenwald. *Handbook of radiotherapy physics: theory and practice*. CRC Press, 2020.

[10] Serena Gianfaldoni et al. "An Overview on Radiotherapy: From Its History to Its Current Applications in Dermatology". In: *Macedonian Journal of Medical Sciences* 5 (July 2017). DOI: 10.3889/oamjms.2017.122.

[11] Philip P. Connell and Samuel Hellman. "Advances in Radiotherapy and Implications for the Next Century: A Historical Perspective". In: *Cancer Research* 69.2 (2009), pp. 383–392. ISSN: 0008-5472. DOI: 10.1158/0008-5472.CAN-07-6871.

[12] Harald Paganetti. *Proton Beam Therapy*. 2399-2891. IOP Publishing, 2017. ISBN: 978-0-7503-1370-4. DOI: 10.1088/978-0-7503-1370-4.

[13] International Commission on Radiation Units and Measurements. *ICRU Report 50 - Prescribing, Recording and Reporting Photon Beam Therapy.* 1999.

[14] Melody Xu, Abigail Berman, and Neha Vapiwala. "The Evolution of Proton Beam Therapy: Insights From Early Trials and Tribulations". In: *International Journal of Radiation Oncology*Biology*Physics* 90 (Nov. 2014), pp. 733–735. DOI: 10.1016/j.ijrobp.2014.07.035.

[15] Harald Paganetti et al. "Relative biological effectiveness (RBE) values for proton beam therapy". In: *International Journal of Radiation Oncology*Biology*Physics* 53.2 (2002), pp. 407–421. ISSN: 0360-3016. DOI: https://doi.org/10.1016/S0360-3016(02)02754-2.

[16] Harald Paganetti. "Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer". In: *Physics in Medicine and Biology* 59 (Oct. 2014), R419. DOI: 10.1088/0031-9155/59/22/R419.

[17] Sanford Meeks et al. "Calculation of cranial nerve complication probability for acoustic neuroma radiosurgery". In: *International journal of radiation oncology, biology, physics* 47 (July 2000), pp. 597–602. DOI: 10.1016/S0360-3016(00)00493-4.

[18] Charles Mayo, Ellen Yorke, and Thomas E. Merchant. "Radiation Associated Brainstem Injury". In: *International Journal of Radiation Oncology*Biology*Physics* 76.3 (2010), S36–S41. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2009.08.078.

[19] James Bates et al. "Visual decline in pediatric survivors of brain tumors following radiotherapy". In: *Acta Oncologica* 59 (Aug. 2020), pp. 1–6. DOI: 10.1080/0284186X.2020.1803500.

[20] Eivind Rørvik et al. "Exploration and application of phenomenological RBE models for proton therapy". In: *Physics in Medicine and Biology* 63 (Aug. 2018). DOI: 10.1088/1361-6560/aad9db.

[21] Aimee L. McNamara, Jan Schuemann, and Harald Paganetti. "A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published in vitro cell survival data". In: *Phys Med Biol* 60.21 (2015), pp. 8399–416. ISSN: 1361-6560 (Electronic) 0031-9155 (Linking). DOI: 10.1088/0031-9155/60/21/8399.

[22] Eivind Rørvik et al. "A phenomenological biological dose model for proton therapy based on linear energy transfer spectra". In: *Medical Physics* 44.6 (2017), pp. 2586–2594. ISSN: 0094-2405. DOI: 10.1002/mp.12216.

[23] Harald Paganetti. "Significance and Implementation of RBE Variations in Proton Beam Therapy". In: *Technology in Cancer Research & Treatment* 2.5 (2003), pp. 413–426. DOI: 10.1177/153303460300200506.

[24] Andrzej Niemierko et al. "Brain Necrosis in Adult Patients After Proton Therapy: Is There Evidence for Dependency on Linear Energy Transfer?" In: *International Journal of Radiation Oncology*Biology*Physics* 109.1 (2021), pp. 109–119. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2020.08.058.

[25] Drosoula Giantsoudi et al. "Incidence of CNS Injury for a Cohort of 111 Patients Treated With Proton Therapy for Medulloblastoma: LET and RBE Associations for Areas of Injury". In: *International Journal of Radiation Oncology*Biology*Physics* 95.1 (2016). Particle Therapy Special Edition, pp. 287–296. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2015.09.015.

[26] Christopher R. Peeler et al. "Clinical evidence of variable proton biological effectiveness in pediatric patients treated for ependymoma". In: *Radiotherapy and Oncology* 121.3 (2016), pp. 395–401. ISSN: 0167-8140. DOI: https://doi.org/10.1016/j.radonc.2016.11.001.

[27] Emanuel Bahn et al. "Late contrast enhancing brain lesions in proton treated low-grade glioma patients: clinical evidence for increased periventricular sensitivity and variable RBE". In: *International Journal of Radiation Oncology*Biology*Physics* 107 (Mar. 2020). DOI: 10.1016/j.ijrobp.2020.03.013.

[28] Alessandra Bolsi et al. "Pencil beam scanning proton therapy for the treatment of craniopharyngioma complicated with radiation-induced cerebral vasculopathies: A dosimetric and linear energy transfer (LET) evaluation". In: *Radiotherapy and Oncology* 149 (May 2020). DOI: 10.1016/j.radonc.2020.04.052.

[29] Jeffrey Smith, Ana Paula Abdala, and Julian Paton. "Structural and functional architecture of respiratory networks in the mammalian brainstem". In: *Philosoph-*

ical transactions of the Royal Society of London. Series B, Biological sciences 364 (Oct. 2009), pp. 2577–87. DOI: 10.1098/rstb.2009.0081.

[30] Anastasia Bohsali et al. "Imaging White Matter in Human Brainstem". In: *Frontiers in human neuroscience* 7 (July 2013), p. 400. DOI: 10.3389/fnhum.2013.00400.

[31] Brenda G. Clark et al. "The integral biologically effective dose to predict brain stem toxicity of hypofractionated stereotactic radiotherapy". In: *International Journal of Radiation Oncology*Biology*Physics* 40.3 (1998), pp. 667–675. ISSN: 0360-3016. DOI: https://doi.org/10.1016/S0360-3016(97)00734-7.

[32] Torunn I. Yock, Louis S. Constine, and Anita Mahajan. "Protons, the brainstem, and toxicity: Ingredients for an emerging dialectic". In: *Acta Oncologica* 53.10 (2014), pp. 1279–1282. DOI: 10.3109/0284186X.2014.957415.

[33] Daphne Haas-Kogan et al. "National Cancer Institute Workshop on Proton Therapy for Children: Considerations Regarding Brainstem Injury". In: *International Journal of Radiation Oncology*Biology*Physics* 101.1 (2018), pp. 152–168. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2018.01.013.

[34] Jinsoo Uh et al. "Differences in Brainstem Fiber Tract Response to Radiation: A Longitudinal Diffusion Tensor Imaging Study". In: *International journal of radiation oncology, biology, physics* 86 (Mar. 2013). DOI: 10.1016/j.ijrobp.2013.01.028.

[35] Chia-Ho Hua et al. "Brain Tumor Therapy-Induced Changes in Normal-Appearing Brainstem Measured With Longitudinal Diffusion Tensor Imaging". In: *International journal of radiation oncology, biology, physics* 82 (June 2011), pp. 2047–54. DOI: 10.1016/j.ijrobp.2011.03.057.

[36] Lawrence B. Marks et al. "Use of Normal Tissue Complication Probability Models in the Clinic". In: *International Journal of Radiation Oncology*Biology*Physics* 76.3, Supplement (2010). Quantitative Analyses of Normal Tissue Effects in the Clinic, S10–S19. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2009.07.1754.

[37] John Kang et al. "Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective". In: *International Journal of Radiation Oncology*Biology*Physics* 93.5 (2015), pp. 1127–1135. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2015.07.2286.

[38] Kreftregisteret. *Cancer in Norway 2019*. https://www.kreftregisteret.no/globalassets/cancer-in-norway/2019/cin_report.pdf. Accessed: 31.05.2021.

[39] Louis Constine et al. “Pediatric Normal Tissue Effects in the Clinic (PENTEC): An International Collaboration to Analyse Normal Tissue Radiation Dose–Volume Response Relationships for Paediatric Cancer Patients”. In: *Clinical Oncology* 31 (Mar. 2019). DOI: 10.1016/j.clon.2019.01.002.

[40] Lucio Cerrito. *Radiation and Detectors : Introduction to the Physics of Radiation and Detection Devices*. 1st ed. 2017. Cham: Springer International Publishing, 2017. DOI: 10.1007/978-3-319-53181-6.

[41] William R. Leo. *Techniques for nuclear and particle physics experiments: a how-to approach*. Springer, 1994.

[42] Radiologykey. *Radiation Interactions with Tissue*. https://radiologykey.com/4-radiation-interactions-with-tissue/. Accessed: 23.02.2021.

[43] Bernard Gottschalk. *Techniques of Proton Radiotherapy: Transport Theory*. 2012. arXiv: 1204.4470 [physics.med-ph].

[44] Thomas Bortfeld. “An analytical approximation of the Bragg curve for therapeutic proton beams”. In: *Medical Physics* 24.12 (1997), pp. 2024–2033. DOI: https://doi.org/10.1118/1.598116.

[45] International Commission on Radiation Units and Measurements. *ICRU Report 33 - Radiation Quantities and Units*. 1980.

[46] Herman Cember. *Introduction to health physics*. eng. New York, 2009.

[47] Eric J. Hall. *Radiobiology for the radiologist*. Wolters Kluwer, 2019.

[48] Alexander Vaiserman et al. “Health Impacts of Low-Dose Ionizing Radiation: Current Scientific Debates and Regulatory Issues”. In: *Dose-Response* 16 (July 2018), p. 155932581879633. DOI: 10.1177/1559325818796331.

[49] Raymond E. Zirkle, Dorothy F. Marchbank, and Kathryn D. Kuck. “Exponential and sigmoid survival curves resulting from alpha and X irradiation of aspergillus spores”. In: *Journal of Cellular and Comparative Physiology* 39.S1 (1952), pp. 75–85. DOI: https://doi.org/10.1002/jcp.1030390408.

[50] Fada Guan et al. “Analysis of the track-and dose-averaged LET and LET spectra in proton therapy using the geant4 Monte Carlo code”. In: *Medical Physics* 42 (Oct. 2015), pp. 6234–6247. DOI: 10.1118/1.4932217.

[51] Stephen J. McMahon. “The linear quadratic model: usage, interpretation and challenges”. In: *Physics in Medicine & Biology* 64.1 (2018), 01TR01. DOI: 10.1088/1361-6560/aaf26a.

[52] Riccardo Ridolfi. "Study of the track reconstruction in the FOOT experiment for Hadrontherapy". PhD thesis. Mar. 2018. DOI: 10.13140/RG.2.2.15768.29440.

[53] Tod W. Speer et al. "Encyclopedia of Radiation Oncology". In: (Jan. 2013), pp. 155–155. DOI: 10.1007/978-3-540-85516-3_2000.

[54] National Cancer Institute. *Common Terminology Criteria for Adverse Events v5.0*. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_50. Accessed: 24.05.2021.

[55] Kristian S. Ytre-Hauge. "Measurements and Monte Carlo Simulations of Neutron Doses from Radiation Therapy with Photons, Protons and Carbon Ions". PhD thesis. University of Bergen, 2013.

[56] Wilhelm C. Röntgen. In: *Nature* 53.1369 (1896), pp. 274–276. ISSN: 1476-4687. DOI: 10.1038/053274b0.

[57] Paulo Martins. "A brief history about radiotherapy". In: 4 (Feb. 2018), pp. 8–11. DOI: 10.5281/zenodo.3824294.

[58] David Thwaites and John Tuohy. "Back to the future: The history and development of the clinical linear accelerator". In: *Physics in medicine and biology* 51 (Aug. 2006), R343–62. DOI: 10.1088/0031-9155/51/13/R20.

[59] Cepheiden. *Depth Dose Curves*. https://commons.wikimedia.org/wiki/File:Dose_Depth_Curves.svg. Accessed: 19.05.2021.

[60] Faiz M. Khan and John P. Gibbons. *Khan's The Physics of Radiation Therapy*. Lippincott, Williams & Wilkins, 2014.

[61] M Teoh et al. "Volumetric modulated arc therapy: a review of current literature and clinical use in practice". In: *The British Journal of Radiology* 84.1007 (2011), pp. 967–996. ISSN: 0007-1285. DOI: 10.1259/bjr/22373346.

[62] Edward C Halperin. "Particle therapy and treatment of cancer". In: *The Lancet Oncology* 7.8 (2006), pp. 676–685. ISSN: 1470-2045. DOI: https://doi.org/10.1016/S1470-2045(06)70795-1.

[63] R.R. Wilson. "Radiological use of fast protons". In: 47 (Jan. 1946), pp. 489–491.

[64] Particle Therapy Co-operative Group. *Particle therapy facilities in clinical operation (last update: April 2021)*. https://www.ptcog.ch/index.php/facilities-in-operation. Accessed: 16.04.2021.

[65] Sara St. James, Clemens Grassberger, and Hsiao-Ming Lu. "Considerations when treating lung cancer with passive scatter or active scanning proton therapy". In: *Translational Lung Cancer Research* 7.2 (2018), pp. 210–215. ISSN: 2218-6751. DOI: 10.21037/tlcr.2018.04.01.

[66] Hanne Kooy and Clemens Grassberger. "Intensity Modulated Proton Therapy". In: *The British journal of radiology* 88 (May 2015), p. 20150195. DOI: 10.1259/bjr.20150195.

[67] Laura Toussaint et al. "Towards proton arc therapy: physical and biologically equivalent doses with increasing number of beams in pediatric brain irradiation". In: *Acta Oncologica* 58 (July 2019), pp. 1–6. DOI: 10.1080/0284186X.2019.1639823.

[68] International Commission on Radiation Units and Measurements. *ICRU Report 78 - Prescribing, Recording, and Reporting Proton-Beam Therapy*. 2007.

[69] Christian P. Karger et al. "The RBE in ion beam radiotherapy: In vivo studies and clinical application". In: *Zeitschrift für Medizinische Physik* (2021). ISSN: 0939-3889. DOI: https://doi.org/10.1016/j.zemedi.2020.12.001.

[70] Jan Unkelbach et al. "Reoptimization of Intensity Modulated Proton Therapy Plans Based on Linear Energy Transfer". In: *International Journal of Radiation Oncology*Biology*Physics* 96.5 (2016), pp. 1097–1106. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2016.08.038.

[71] Stephen McMahon, Harald Paganetti, and Kevin Prise. "LET-weighted doses effectively reduce biological variability in proton radiotherapy planning". In: *Physics in Medicine and Biology* 63 (Oct. 2018). DOI: 10.1088/1361-6560/aae8a5.

[72] Nadine Barrie Smith and Andrew Webb. *Introduction to Medical Imaging: Physics, Engineering and Clinical Applications*. Cambridge Texts in Biomedical Engineering. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511760976.005.

[73] Kim Nalee et al. "Atlas-based auto-segmentation for postoperative radiotherapy planning in endometrial and cervical cancers". In: *Radiation Oncology* 15 (May 2020). DOI: 10.1186/s13014-020-01562-y.

[74] Marco Schwarz. "Treatment planning in proton therapy". In: *The European Physical Journal Plus* 126 (July 2011), pp. 1–10. DOI: 10.1140/epjp/i2011-11067-y.

[75] Martin Soukup, Matthias Fippel, and Markus Alber. "A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations". In: *Physics in medicine and biology* 50 (Dec. 2005), pp. 5089–104. DOI: 10.1088/0031-9155/50/21/010.

[76] Linda Hong et al. "A pencil beam algorithm for proton dose calculations". In: *Physics in Medicine and Biology* 41.8 (1996), pp. 1305–1330. DOI: 10.1088/0031-9155/41/8/005.

[77] Pedro Andreo. "Monte Carlo simulations in radiotherapy dosimetry". In: *Radiation Oncology* 13.1 (2018). ISSN: 1748-717X. DOI: 10.1186/s13014-018-1065-3.

[78] Jay L. Devore and Kenneth Berk. *Modern mathematical statistics with applications*. Springer, 2011.

[79] Bradley Efron. "Bootstrap Methods: Another Look at the Jackknife". In: *The Annals of Statistics* 7.1 (1979), pp. 1–26. DOI: 10.1214/aos/1176344552.

[80] Martin Carolan et al. "An MLE method for finding LKB NTCP model parameters using Monte Carlo uncertainty estimates". In: *Journal of Physics Conference Series* 489 (2014), p. 012087. ISSN: 1742-6588. DOI: 10.1088/1742-6596/489/1/012087.

[81] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability 57. Boca Raton, Florida, USA: Chapman & Hall/CRC, 1993.

[82] Trevor Hastie et al. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol. 27. Nov. 2004, pp. 83–85. DOI: 10.1007/BF02985802.

[83] Kelly Zou, James O'Malley, and Laura Mauri. "Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models". In: *Circulation* 115 (Mar. 2007), pp. 654–7. DOI: 10.1161/CIRCULATIONAHA.105.594929.

[84] Ethem Alpaydin and Francis Bach. *Introduction to Machine Learning (3rd edition)*. MIT Press, 2014.

[85] L. F. Shackell et al. "The Relation of Dosage to Effect". In: *Journal of Pharmacology and Experimental Therapeutics* 24.1 (1924), pp. 53–65. ISSN: 0022-3565.

[86] Gary S Collins et al. "Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement". In: *BMC Medicine* 13.1 (2015), p. 1. ISSN: 1741-7015.

[87] John T. Lyman. "Complication Probability as Assessed from Dose-Volume Histograms". In: *Radiation Research* 104.2 (1985), S13. ISSN: 0033-7587. DOI: 10.2307/3576626.

[88] Chandra M. Burman et al. “Fitting of normal tissue tolerance data to an analytic function”. In: *International Journal of Radiation Oncology*Biology*Physics* 21.1 (1991). Three-Dimensional Photon Treatment Planning Report of the Collaborative Working Group on the Evaluation of Treatment Planning for External Photon Beam Radiotherapy, pp. 123–135. ISSN: 0360-3016. DOI: https://doi.org/10.1016/0360-3016(91)90172-Z.

[89] Bahman Emami et al. “Tolerance of normal tissue to therapeutic irradiation”. In: *International Journal of Radiation Oncology*Biology*Physics* 21.1 (1991). Three-Dimensional Photon Treatment Planning Report of the Collaborative Working Group on the Evaluation of Treatment Planning for External Photon Beam Radiotherapy, pp. 109–122. ISSN: 0360-3016. DOI: https://doi.org/10.1016/0360-3016(91)90171-Y.

[90] Andrzej Niemierko and Michael Goitein. “Calculation of normal tissue complication probability and dose-volume histogram reduction schemes for tissues with a critical element architecture”. In: *Radiotherapy and Oncology* 20.3 (1991), pp. 166–176. ISSN: 0167-8140. DOI: https://doi.org/10.1016/0167-8140(91)90093-V.

[91] Andrzej Niemierko and Michael Goitein. “Modeling of normal tissue response to radiation: The critical volume model”. In: *International Journal of Radiation Oncology*Biology*Physics* 25.1 (1993), pp. 135–145. ISSN: 0360-3016. DOI: https://doi.org/10.1016/0360-3016(93)90156-P.

[92] Anders Brahme, Panayiotis Mavroidis, and Bengt Lind. “Dose-Response Relations for Tumors and Normal Tissues”. In: vol. 9. Sept. 2014, pp. 167–189. ISBN: 978-0-444-53633-4. DOI: 10.1016/B978-0-444-53632-7.00902-3.

[93] Laura Cella et al. “Multivariate Normal Tissue Complication Probability Modeling of Gastrointestinal Toxicity After External Beam Radiotherapy for Localized Prostate Cancer”. In: *Radiation oncology (London, England)* 8 (Sept. 2013), p. 221. DOI: 10.1186/1748-717X-8-221.

[94] Issam El Naqa et al. “Multivariable modeling of radiotherapy outcomes, including dose–volume and clinical factors”. In: *International Journal of Radiation Oncology*Biology*Physics* 64.4 (2006), pp. 1275–1286. ISSN: 0360-3016. DOI: https://doi.org/10.1016/j.ijrobp.2005.11.022.

[95] Mary Feng et al. “Machine Learning in Radiation Oncology: Opportunities, Requirements, and Needs”. In: *Frontiers in Oncology* 8 (Apr. 2018). DOI: 10.3389/fonc.2018.00110.

[96] Martin Ester et al. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. KDD'96. Portland, Oregon: AAAI Press, 1996, pp. 226–231. DOI: 10.5120/739-1038.

[97] Nadia Rahmah and Imas Sitanggang. "Determination of Optimal Epsilon (Eps) Value on DBSCAN Algorithm to Clustering Data on Peatland Hotspots in Sumatra". In: *IOP Conference Series: Earth and Environmental Science* 31 (Jan. 2016), p. 012012. DOI: 10.1088/1755-1315/31/1/012012.

[98] Jörg Sander et al. "Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications". In: *Data Mining and Knowledge Discovery* 2 (2004), pp. 169–194.

[99] Peter Rousseeuw. "Rousseeuw, P.J.: Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. Comput. Appl. Math. 20, 53-65". In: *Journal of Computational and Applied Mathematics* 20 (Nov. 1987), pp. 53–65. DOI: 10.1016/0377-0427(87)90125-7.

[100] Susan Fink and Brad Cookson. "Apoptosis, Pyroptosis, and Necrosis: Mechanistic Description of Dead and Dying Eukaryotic Cells". In: *Infection and immunity* 73 (May 2005), pp. 1907–16. DOI: 10.1128/IAI.73.4.1907-1916.2005.

[101] Michelle Gentile et al. "Brainstem Injury in Pediatric Patients With Posterior Fossa Tumors Treated With Proton Beam Therapy and Associated Dosimetric Factors". In: *International Journal of Radiation Oncology*Biology*Physics* 100 (Nov. 2017). DOI: 10.1016/j.ijrobp.2017.11.026.

[102] Daniel J. Indelicato et al. "Incidence and dosimetric parameters of pediatric brainstem toxicity following proton therapy". In: *Acta Oncologica* 53.10 (2014), pp. 1298–1304. ISSN: 0284-186X. DOI: 10.3109/0284186x.2014.957414.

[103] Lars F. Fjæra et al. "Implementation of a double scattering nozzle for Monte Carlo recalculation of proton plans with variable relative biological effectiveness". In: *Physics in Medicine & Biology* 65 (Oct. 2020). DOI: 10.1088/1361-6560/abc12d.

[104] Colin G. Orton and Lionel Cohen. "A unified approach to dose-effect relationships in radiotherapy. I: Modified TDF and linear quadratic equations". In: *International Journal of Radiation Oncology*Biology*Physics* 14.3 (1988), pp. 549–556. ISSN: 0360-3016. DOI: https://doi.org/10.1016/0360-3016(88)90273-8.

[105] Lars F. Fjæra et al. "Spatial Agreement of Brainstem Dose Distributions Depending on Biological Model in Proton Therapy for Pediatric Brain Tumors". In: *Advances in Radiation Oncology* 6 (Aug. 2020). DOI: 10.1016/j.adro.2020.08.008.

[106] Camilla Stokkevåg et al. "OC-0612 A case-control study of brainstem substructures and morbidity following pediatric proton therapy". In: *Radiotherapy and Oncology* 133 (2019). ESTRO 38, 26-30 April 2019, Milan, Italy, S323–S324. ISSN: 0167-8140. DOI: https://doi.org/10.1016/S0167-8140(19)31032-1.

[107] Helge E. S. Pettersen. *DVHToolkit Version 1.7*. https://github.com/BergenParticleTherapy/DVHToolkit. Accessed: 05.05.2021.

[108] John Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9 (June 2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[109] Wes McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

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

[111] L. Grippo, F. Lampariello, and S. Lucidi. "A truncated Newton method with nonmonotone line search for unconstrained optimization". In: *Journal of Optimization Theory and Applications* 60 (1989), pp. 401–419. DOI: https://doi.org/10.1007/BF00940345.

[112] David Wales and Jonathan Doye. "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms". In: *The Journal of Physical Chemistry A* 101 (Apr. 1998). DOI: 10.1021/jp970984n.

[113] Fabian Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (Jan. 2012).

[114] Erin Murphy et al. "Necrosis After Craniospinal Irradiation: Results From a Prospective Series of Children With Central Nervous System Embryonal Tumors". In: *International journal of radiation oncology, biology, physics* 83 (Aug. 2012), e655–60. DOI: 10.1016/j.ijrobp.2012.01.061.

[115] Jillian Gunther et al. "Imaging Changes in Pediatric Intracranial Ependymoma Patients Treated With Proton Beam Radiation Therapy Compared to Intensity Modulated Radiation Therapy". In: *International Journal of Radiation Oncology*Biology*Physics* 93 (May 2015). DOI: 10.1016/j.ijrobp.2015.05.018.

[116] Jakob Ödén et al. "Spatial correlation of linear energy transfer and relative biological effectiveness with suspected treatment related toxicities following proton therapy for intracranial tumors". In: *Medical Physics* 47 (Nov. 2019). DOI: 10.1002/mp.13911.

[117] Minna Wedenberg and Iuliana Toma-Dasu. "Disregarding RBE variation in treatment plan comparisonmay lead to bias in favor of proton therapy". In: *Medical physics* 41 (Sept. 2014), p. 091706. DOI: 10.1118/1.4892930.

[118] Harald Paganetti. "Relating the proton relative biological effectiveness to tumor control and normal tissue complication probabilities assuming interpatient variability in α/β". In: *Acta Oncologica* 56.11 (2017), pp. 1379–1386. DOI: 10.1080/0284186X.2017.1371325.

[119] Giuseppe Palma et al. "NTCP Models for Severe Radiation Induced Dermatitis After IMRT or Proton Therapy for Thoracic Cancer Patients". In: *Frontiers in Oncology* 10 (Mar. 2020). DOI: 10.3389/fonc.2020.00344.

[120] Zeming Wang et al. "Lyman-Kutcher-Burman normal tissue complication probability modeling for radiation-induced esophagitis in non-small cell lung cancer patients receiving proton radiotherapy". In: *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 146 (Mar. 2020), pp. 200–204. DOI: 10.1016/j.radonc.2020.03.003.

[121] Giuseppe Palma et al. "Modelling the risk of radiation induced alopecia in brain tumor patients treated with scanned proton beams". In: *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 144 (Dec. 2019), pp. 127–134. DOI: 10.1016/j.radonc.2019.11.013.

[122] Sarah Gulliford et al. "Parameters for the Lyman Kutcher Burman (LKB) model of Normal Tissue Complication Probability (NTCP) for specific rectal complications observed in clinical practise". In: *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 102 (Nov. 2011), pp. 347–51. DOI: 10.1016/j.radonc.2011.10.022.

[123] Laura Dawson et al. "Analysis of radiation-induced liver disease using the Lyman NTCP model". In: *International journal of radiation oncology, biology, physics* 53 (July 2002), pp. 810–21. DOI: 10.1016/S0360-3016(02)02846-8.

[124] V.A. Semenenko and Xiaogao Li. "Lyman-Kutcher-Burman NTCP Model Parameters for Radiation Pneumonitis and Xerostomia Based on Combined Analysis of Published Clinical Data". In: *Physics in medicine and biology* 53 (Mar. 2008), pp. 737–55. DOI: 10.1088/0031-9155/53/3/014.

[125] Shannon M. MacDonald et al. "Proton radiotherapy for pediatric central nervous system ependymoma: clinical outcomes for 70 patients". In: *Neuro-Oncology* 15.11 (2013), pp. 1552–1559. ISSN: 1522-8517. DOI: 10.1093/neuonc/not121.

[126] J. Vogel et al. “Risk of brainstem necrosis in pediatric patients with central nervous system malignancies after pencil beam scanning proton therapy”. In: *Acta Oncologica* 58 (Sept. 2019), pp. 1–5. DOI: 10.1080/0284186X.2019.1659996.

[127] Gregory Dexter et al. “Generalization of Machine Learning Approaches to Identify Notifiable Conditions from a Statewide Health Information Exchange”. In: *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science* 2020 (May 2020), pp. 152–161.

[128] Daxin Jiang, Chun Tang, and Aidong Zhang. “Cluster analysis for gene expression data: a survey”. In: *IEEE Transactions on Knowledge and Data Engineering* 16.11 (2004), pp. 1370–1386. DOI: 10.1109/TKDE.2004.68.

[129] Feng Nan et al. “Application of Improved SOM Network in Gene Data Cluster Analysis”. In: *Measurement* 145 (June 2019). DOI: 10.1016/j.measurement.2019.01.013.

Appendix A

Scatter Plots

Figure A.1 shows a scatter plot of $D50\%$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem.

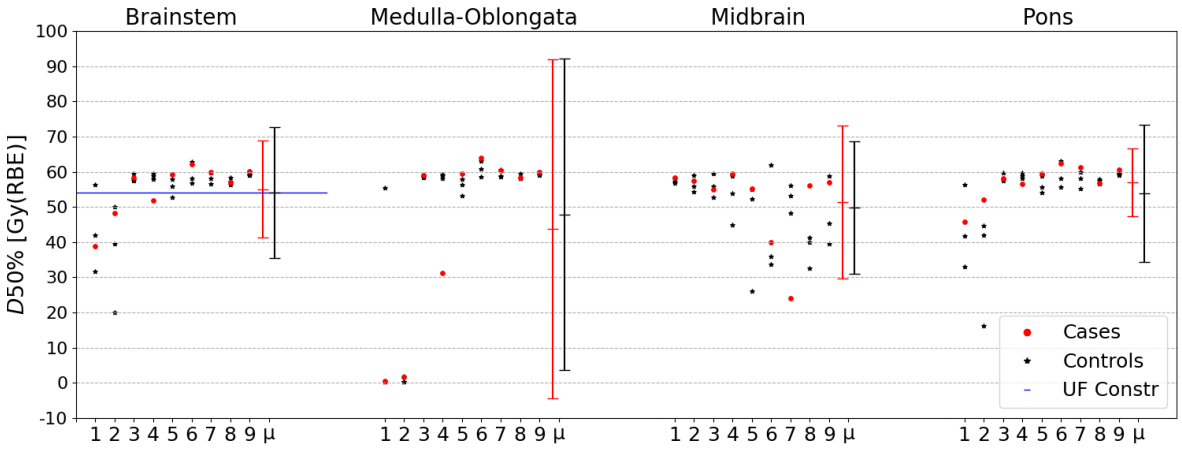

Figure A.1: $D50\%$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume. Constraint for $D50\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.2 shows a scatter plot of $D50\%$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

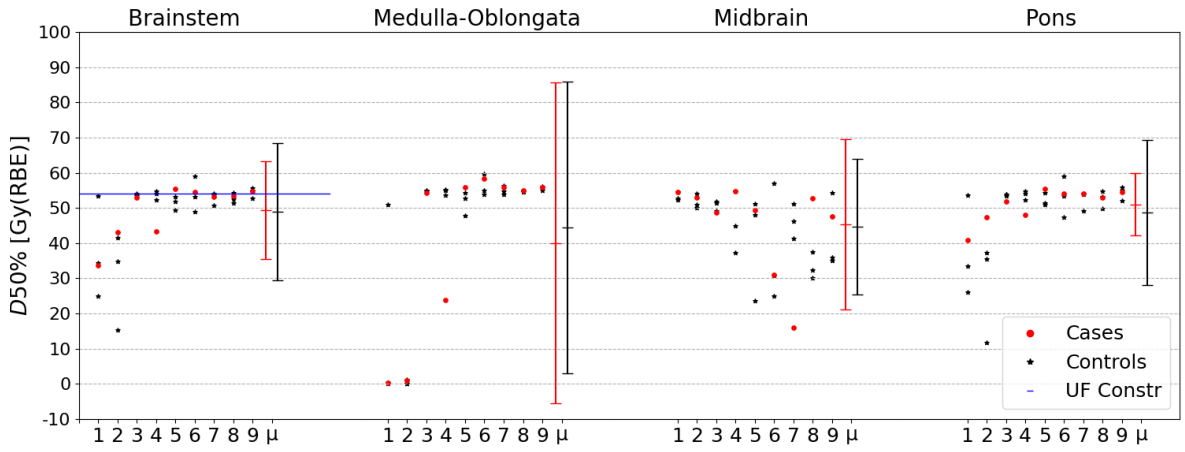

Figure A.2: D50% to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1. Constraint for $D50\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.3 shows a scatter plot of $D50\%$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem.

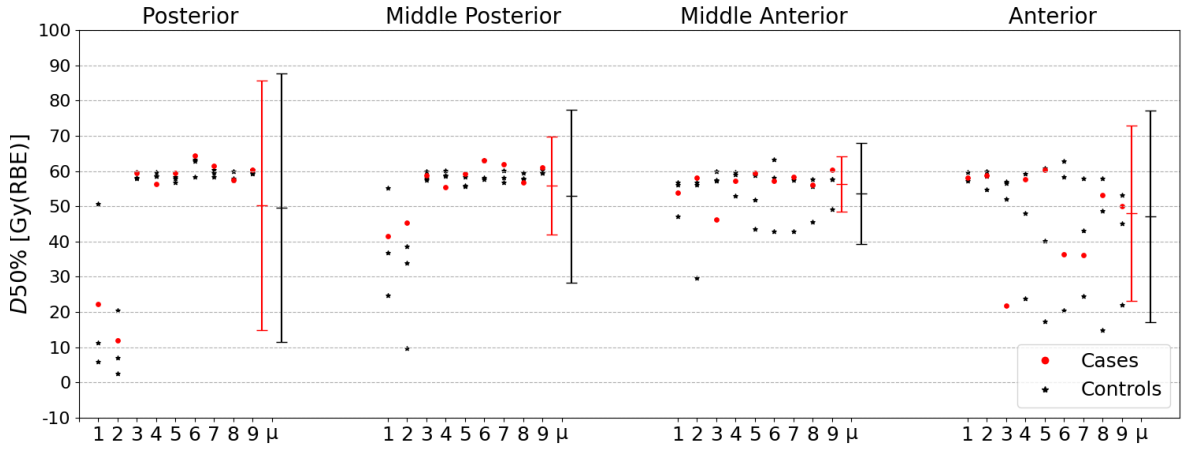

Figure A.3: D50% to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

Figure A.4 shows a scatter plot of $D50\%$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

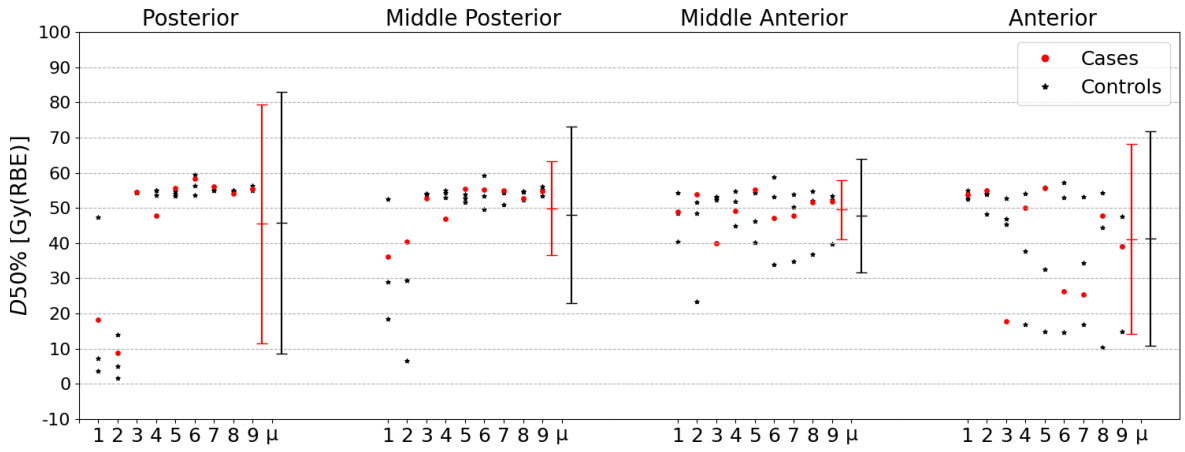

Figure A.4: $D50\%$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.5 shows a scatter plot of $D10\%$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

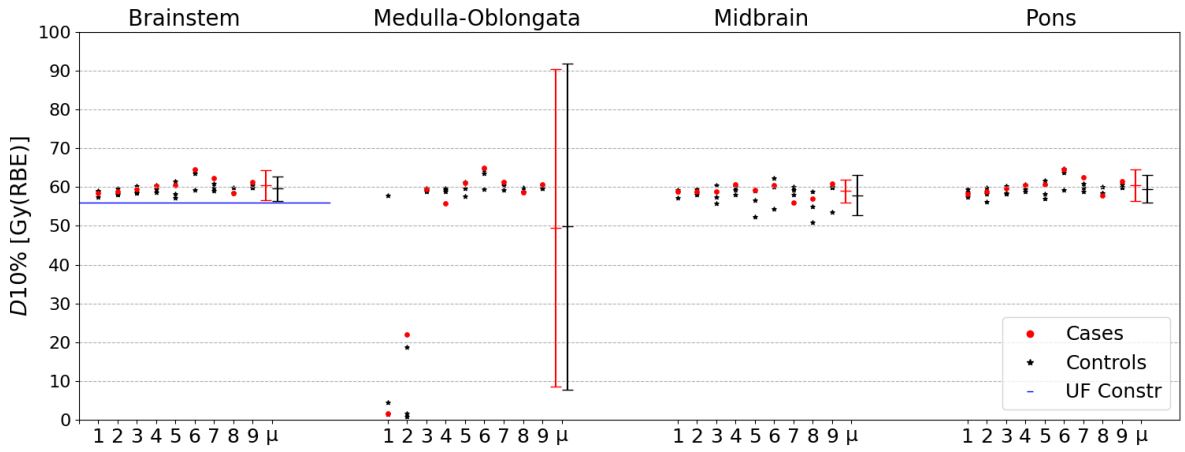

Figure A.5: $D10\%$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume. Constraint for $D10\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.6 shows a scatter plot of $D10\%$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% confidence intervals. Dose distribution is recalculated with the RBE1.1 factor.

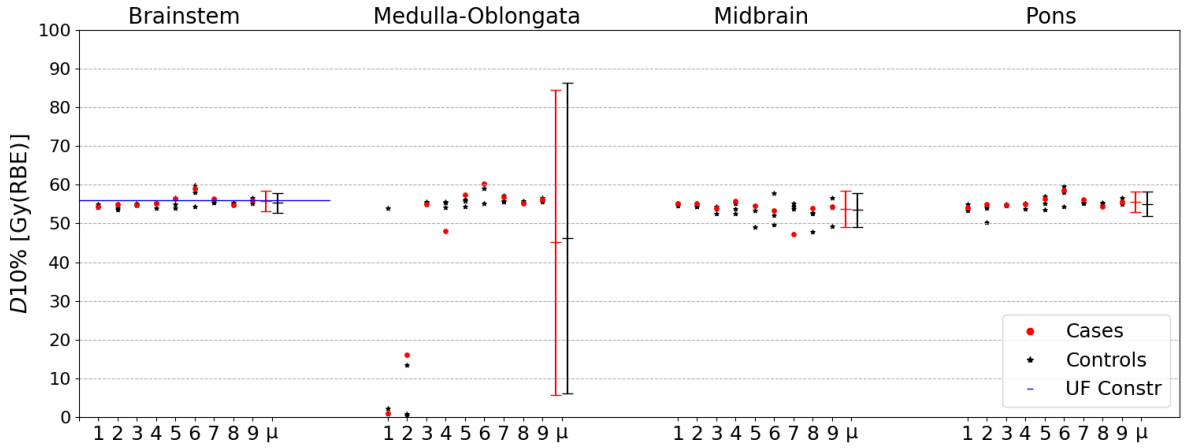

Figure A.6: $D10\%$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, with average case and control with 95% CI. RBE-weighted doses are found with RBE1.1. Constraint for $D10\%$ utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.7 shows a scatter plot of $D10\%$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem.

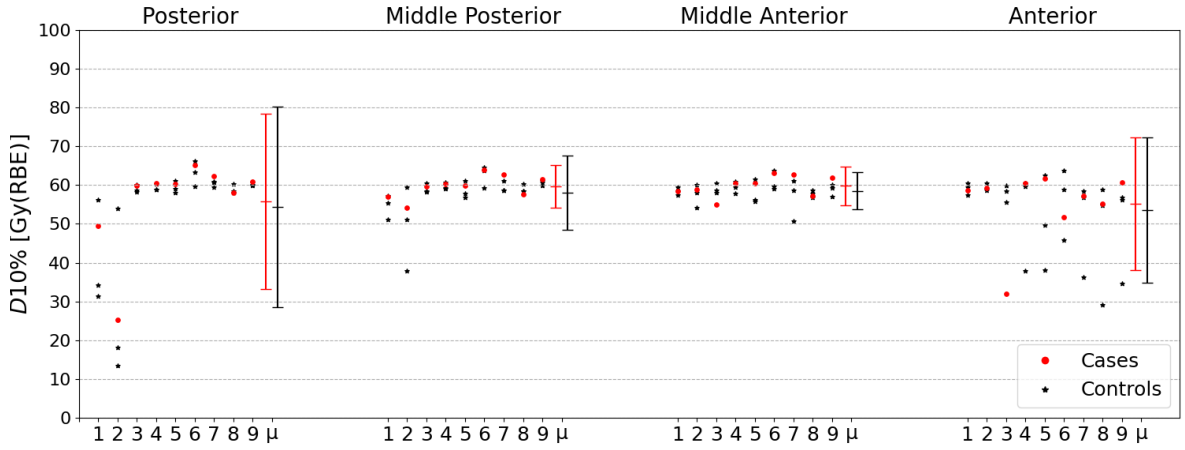

Figure A.7: $D10\%$ to the substructures of the pons for all case-control groups, including average case and control and averages with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem volume.

Figure A.8 shows a scatter plot of $D10\%$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

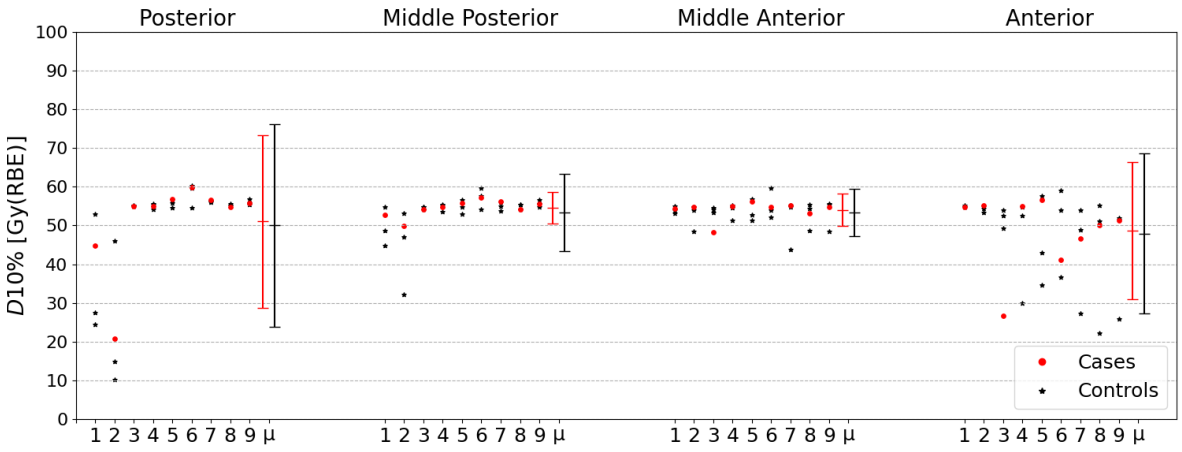

Figure A.8: $D10\%$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.9 shows a scatter plot of $D_{\max}$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

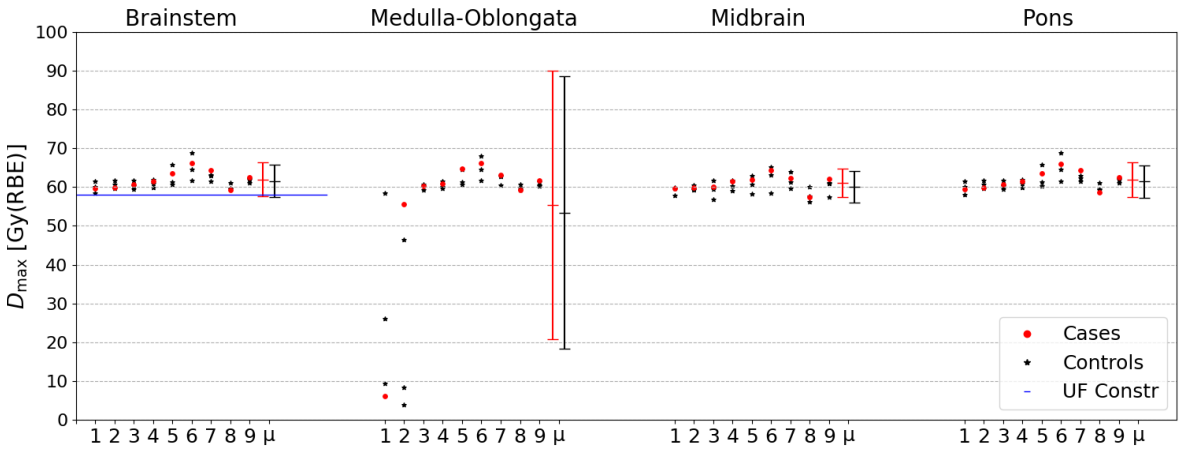

Figure A.9: $D_{\max}$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, incuding average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the whole brainstem volume. Constraint for D_{max} utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.10 shows a scatter plot of $D_{\max}$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

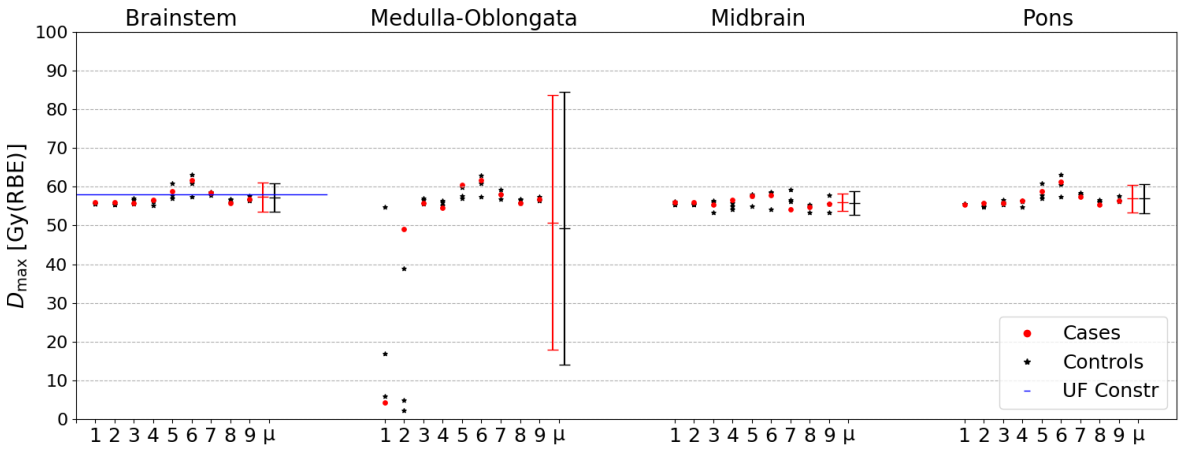

Figure A.10: $D_{\max}$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1. Constraint for D_{max} utilised clinically at the University of Florida is also given as a horizontal blue line [33].

Figure A.11 shows a scatter plot of $D_{\max}$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

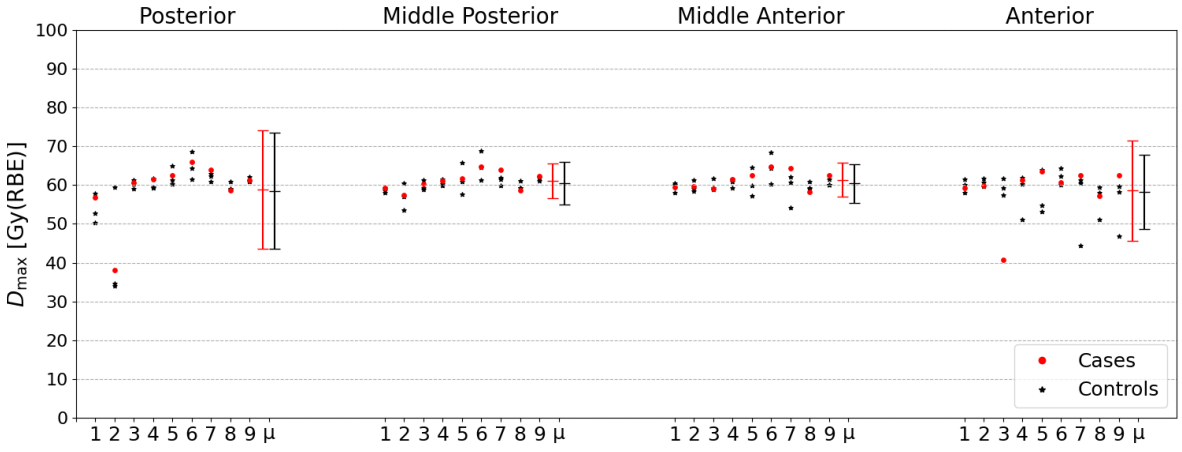

Figure A.11: $D_{\max}$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume.

Figure A.12 shows a scatter plot of $D_{\max}$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs and dose distribution from RBE1.1.

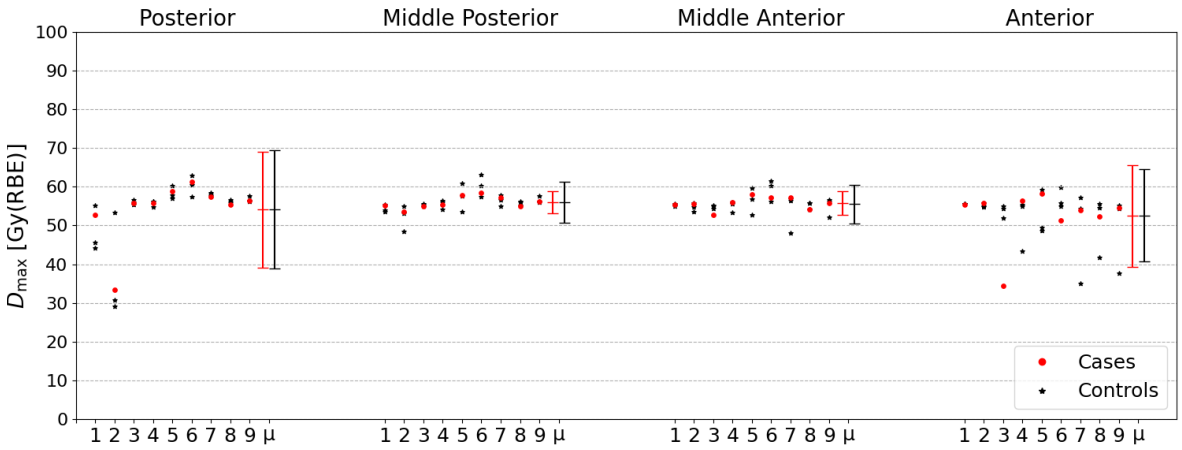

Figure A.12: $D_{\max}$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.13 shows a scatter plot of the gEUD with $n = 0.05$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

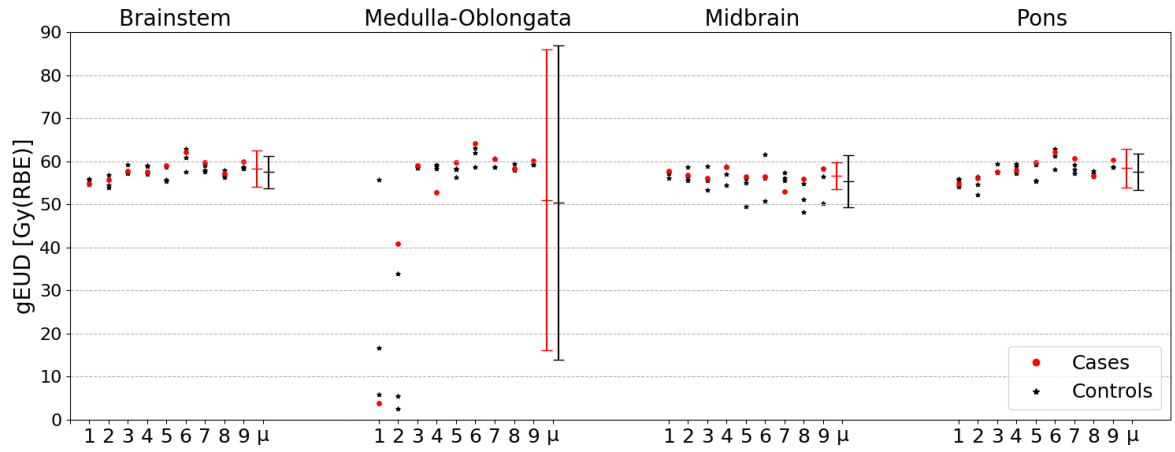

Figure A.13: gEUD with $n = 0.05$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control and averages with 95% CI. RBE-weighted doses are found with the Rørvik model assuming $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume.

Figure A.14 shows a scatter plot of the gEUD with $n = 0.05$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

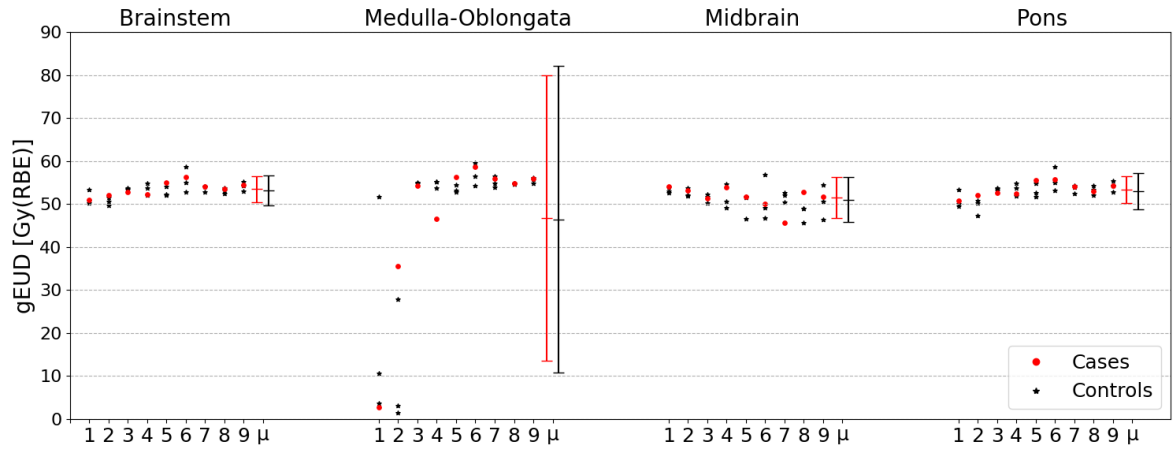

Figure A.14: gEUD with $n = 0.05$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.15 shows a scatter plot of the gEUD with $n = 0.16$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

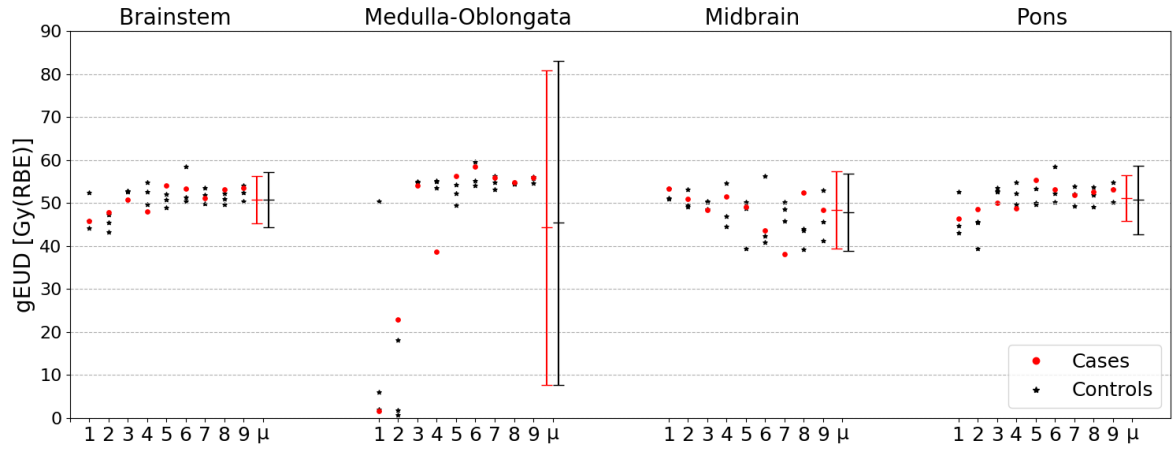

Figure A.15: gEUD with $n = 0.16$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem volume.

Figure A.16 shows a scatter plot of the gEUD with $n = 0.16$ to the full brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

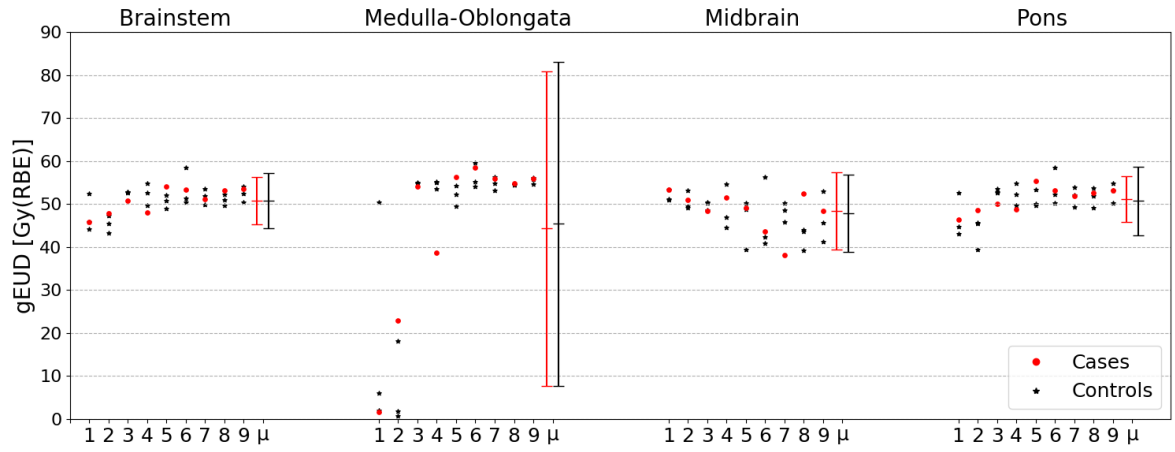

Figure A.16: gEUD with $n = 0.16$ to the brainstem, the medulla oblongata, the midbrain and the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.17 shows a scatter plot of the gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem.

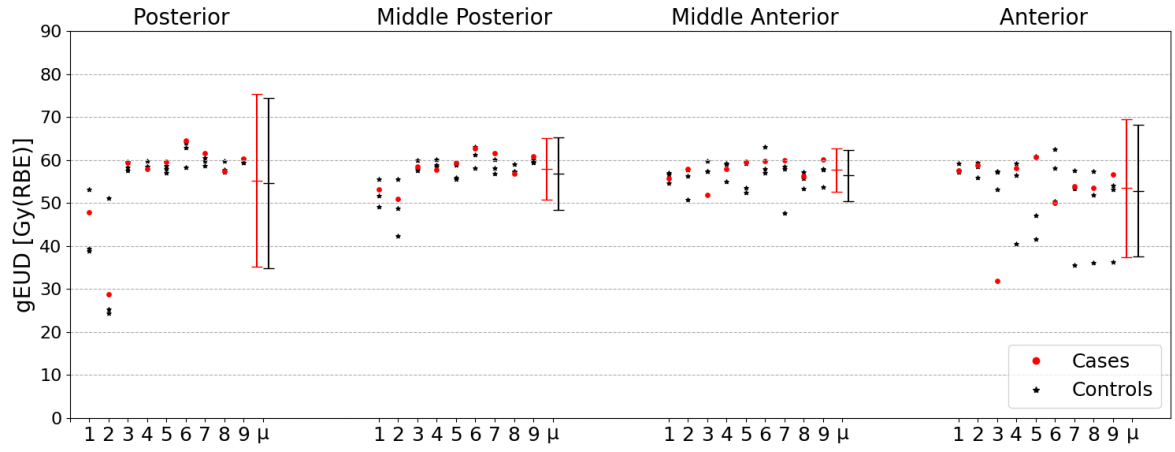

Figure A.17: gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming $(\alpha/\beta)_x = 2.1$ for the entire brainstem volume.

Figure A.18 shows a scatter plot of the gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

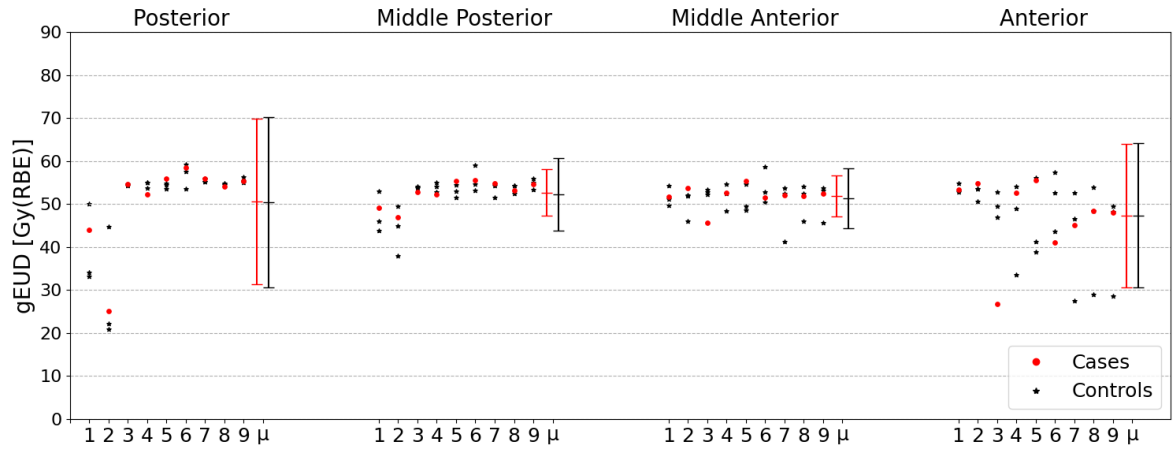

Figure A.18: gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Figure A.19 shows a scatter plot of gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the Rørvik variable RBE model assuming a constant $(\alpha/\beta)_x = 2.1$ for the full brainstem.

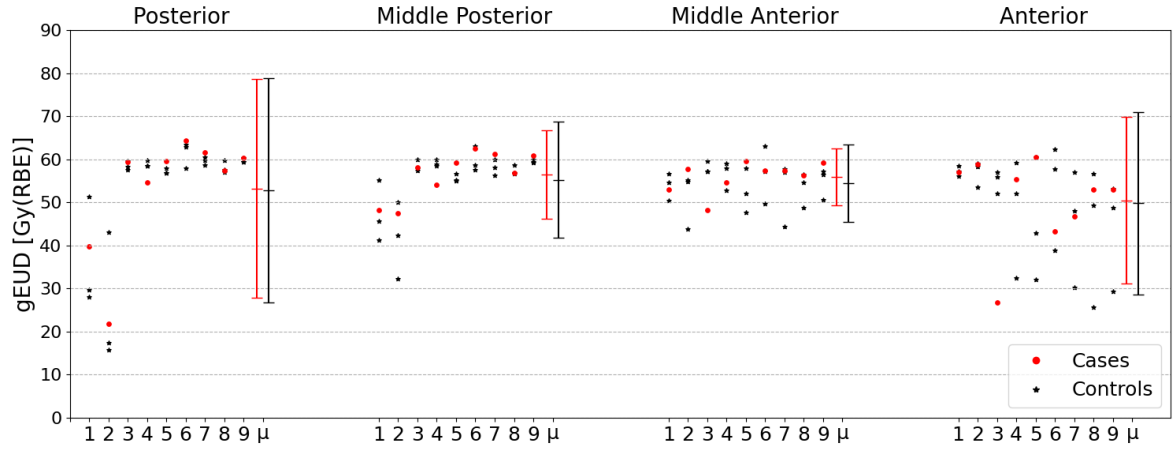

Figure A.19: gEUD with $n = 0.16$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with the Rørvik model assuming $(\alpha/\beta)_x = 2.1$ for the entire brainstem volume.

Figure A.20 shows a scatter plot with the spread of gEUD with $n = 0.05$ to the substructures of the pons for all case-control groups and mean case and control with 95% CIs. Dose distribution is recalculated with the RBE1.1 factor.

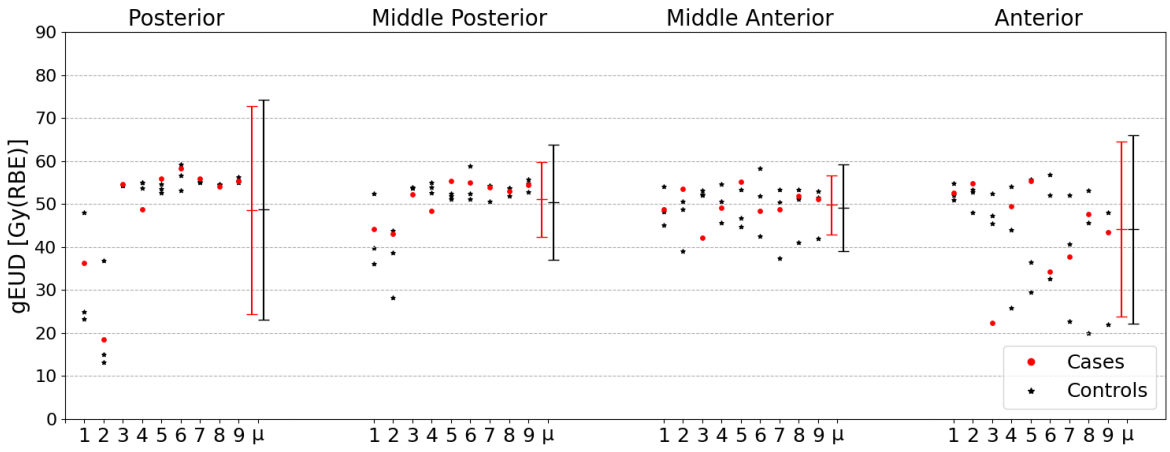

Figure A.20: gEUD with $n = 0.16$ to the substructures of the pons for all case-control groups, including average case and control with 95% CI. RBE-weighted doses are found with RBE1.1.

Appendix B

Hierarchical Clustering

Figure B.1 shows a dendrogram for the hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

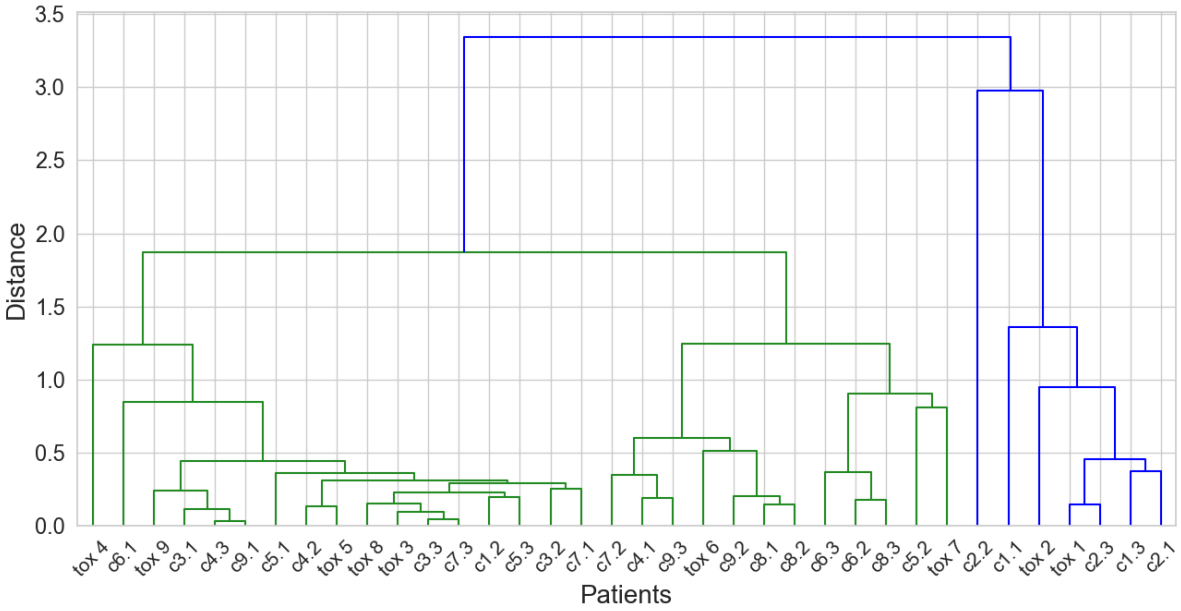

Figure B.1: Dendrogram showing the hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

Figure B.2 shows a dendrogram for the hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from RBE1.1.

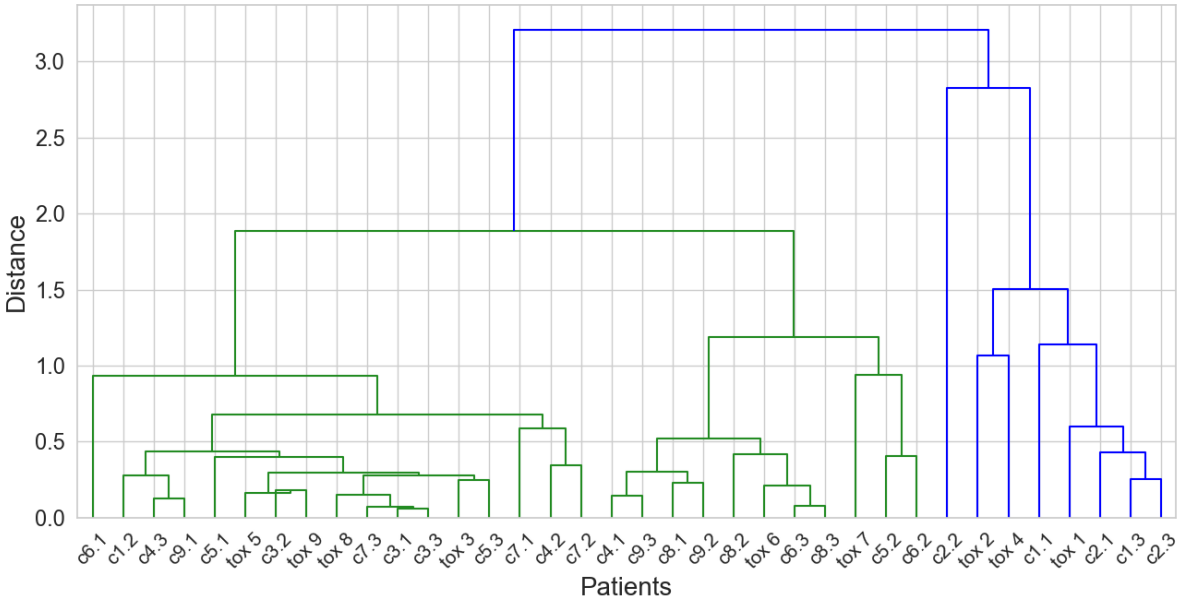

Figure B.2: Dendrogram showing the hierarchical clustering of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted doses from RBE1.1. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

Figure B.3 shows a dendrogram for the hierarchical clustering of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model assuming a constant $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem.

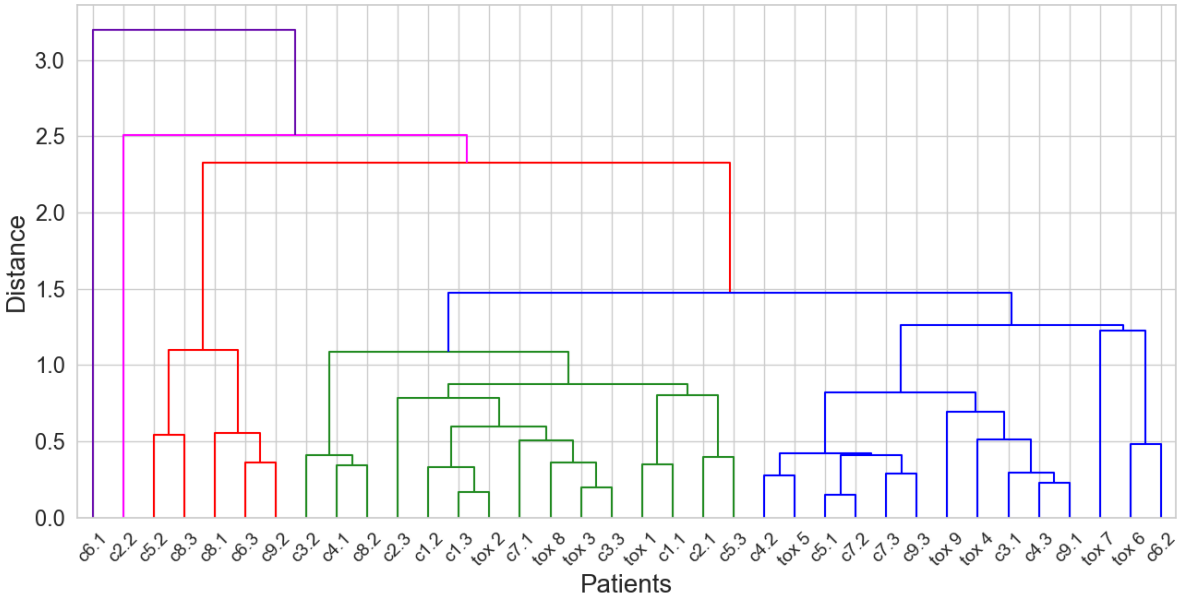

Figure B.3: Dendrogram showing the hierarchical clustering of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted doses from the Rørvik RBE model. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

Figure B.4 shows a dendrogram for the hierarchical clustering of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted dose from RBE1.1.

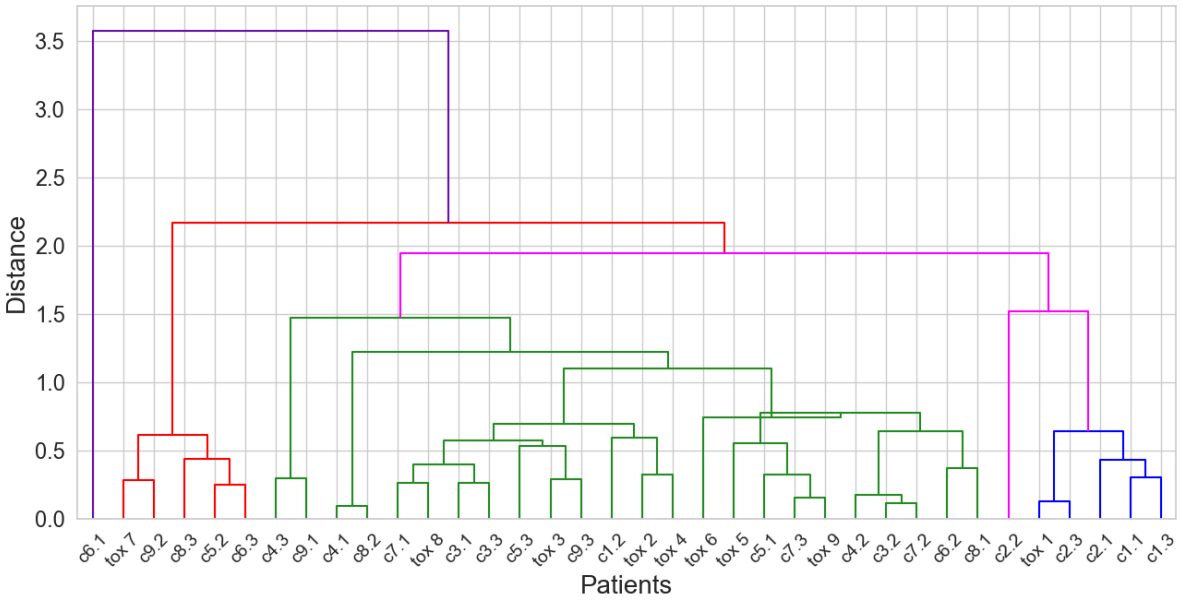

Figure B.4: Dendrogram showing the hierarchical clustering of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted doses from RBE1.1. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

The hierarchical clustering of the $D50\%$ to the pons substructures, posterior, middle posterior, middle anterior and anterior with RBE-weighted doses from the McNamara variable RBE model assuming $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem. The clustering yielded a silhouette score of 0.65. The dendrogram is shown in figure B.5.

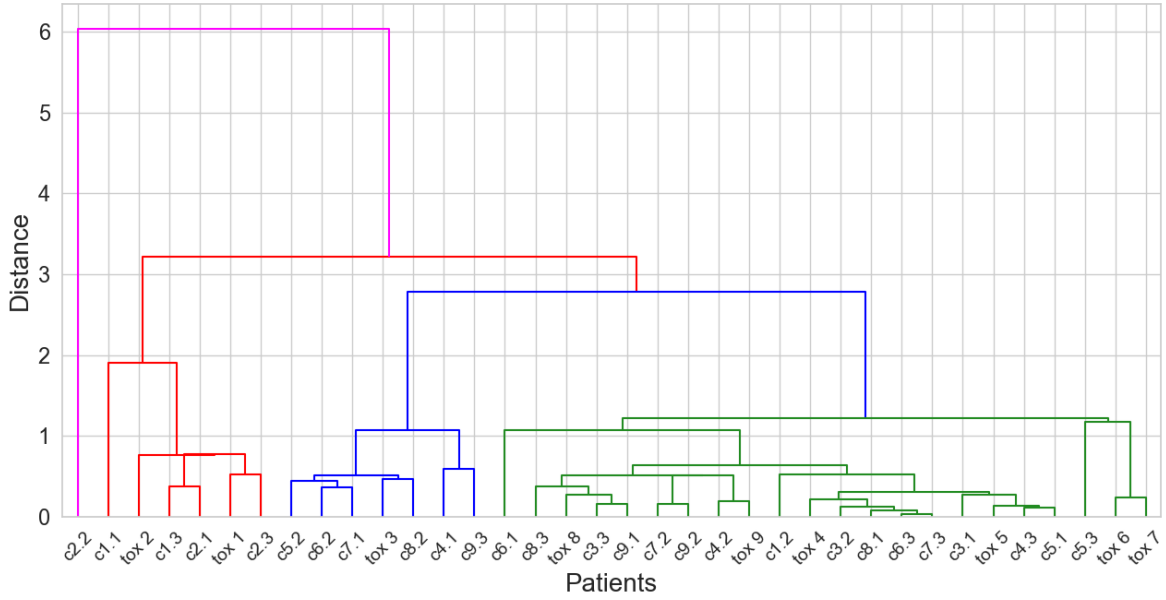

Figure B.5: Hierarchical clustering of $D50\%$ of pons substructures: posterior, middle posterior, middle anterior and anterior With RBE weighted doses from the McNamara variable RBE model, assuming $\alpha/\beta = 2.1$ for the entire brainstem pons. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

The hierarchical clustering of the $D50\%$ to the pons substructures, posterior, middle posterior, middle anterior and anterior with RBE-weighted doses from the Rørvik variable RBE model assuming $(\alpha/\beta)_{\mathrm{x}} = 2.1$ for the full brainstem. The clustering yielded a silhouette score of 0.62. The dendrogram is shown in figure B.6.

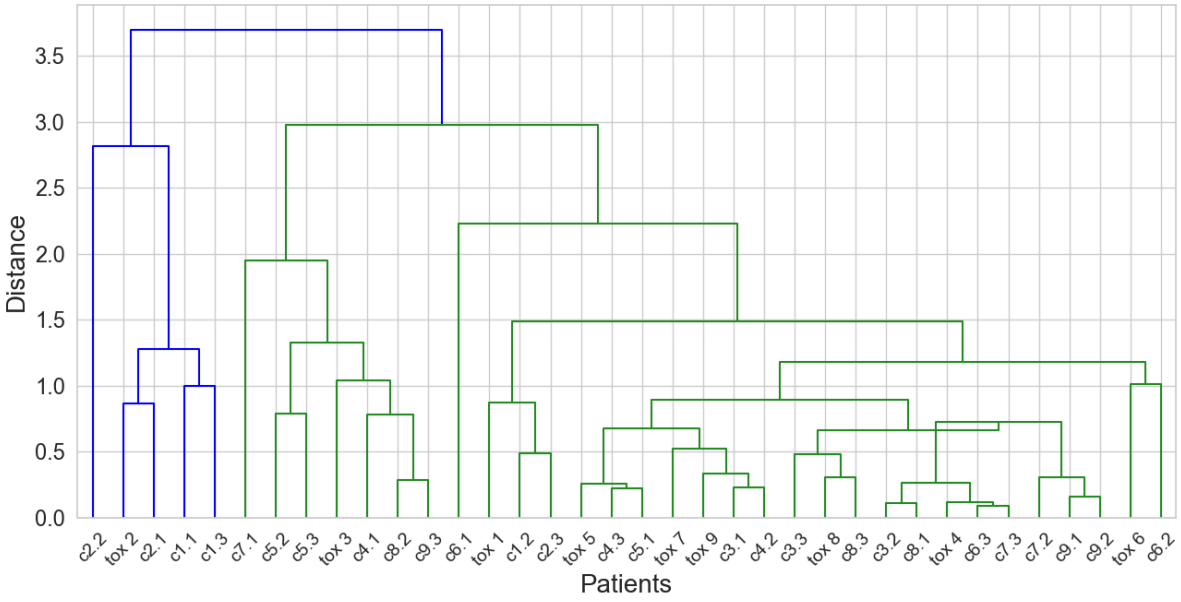

Figure B.6: Hierarchical clustering of $D50\%$ of pons substructures: posterior, middle posterior, middle anterior and anterior With RBE weighted doses from the Rørvik variable RBE model, assuming $\alpha/\beta = 2.1$ for the entire brainstem pons. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

The hierarchical clustering of the $D50\%$ to the pons substructures, posterior, middle posterior, middle anterior and anterior with RBE-weighted doses from RBE1.1. The clustering yielded a silhouette score of 0.61. The dendrogram is shown in figure B.7.

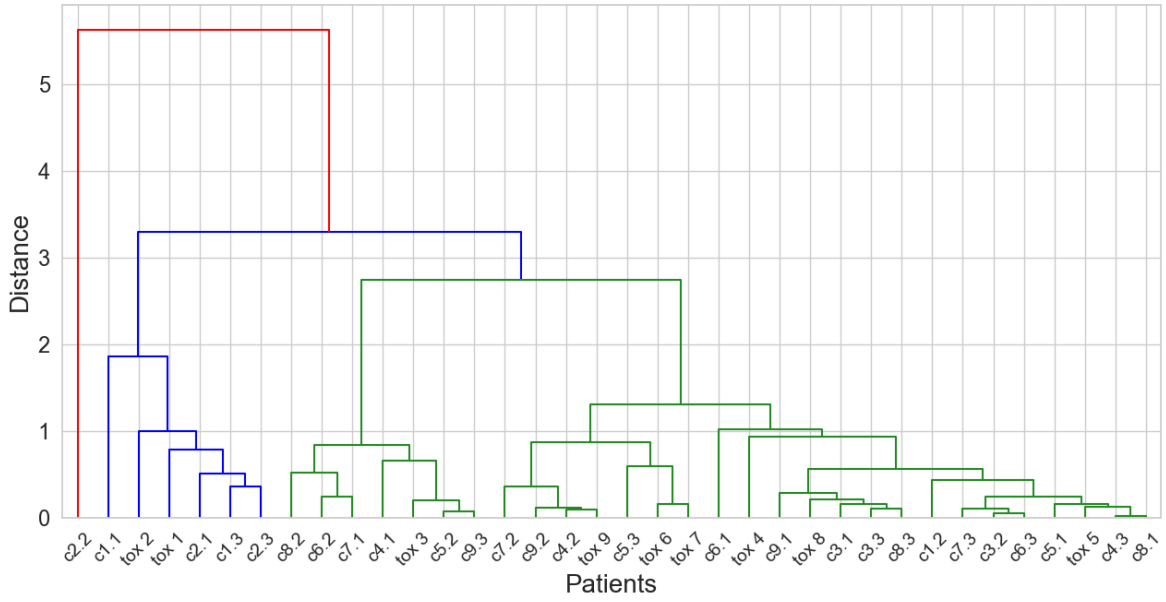

Figure B.7: Hierarchical clustering of $D50\%$ of pons substructures: posterior, middle posterior, middle anterior and anterior. The label tox n refers to the case of group n, while $cn.i$ refers to its corresponding matched controls.

The distribution of cases and controls for all three clusterings of D50% to the pons substructures are given in table B.1.

Table B.1: The distribution of the cases and controls for the hierarchical clustering of the $D50\%$ to the midbrain and the pons. RBE-weighted doses are calculated with RBE1.1, and the McNamara and Rørvik RBE models.

	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	6	16	8	23	7	22
Cluster 2 (●)	1	6	1	4	2	4
Cluster 3 (●)	2	4	-	-	0	1
Cluster 4 (●)	0	1	-	-	-	-

Appendix C

DBSCAN

The DBSCAN of D50% of the pons substructures, posterior, middle posterior, middle anterior and anterior with RBE weighted doses from the Rørvik RBE model and McNamara RBE model, both assuming $(\alpha/\beta)_x = 2.1$ for the full brainstem used a minimum points to form cluster of 5 and an epsilon of 1.0. The clusterings had silhouette scores of 0.65. RBE1.1 used the same parameters, but had a silhouette score of 0.60. The distribution of cases and controls are provided in table C.1

Table C.1: The distribution of cases and controls for the DBSCAN of the D50% for the pons posterior, pons middle-posterior, pons middle-anterior and pons anterior using different RBE weighted doses.

	McNamara Cases	McNamara Controls	Rørvik Cases	Rørvik Controls	RBE1.1 Cases	RBE1.1 Controls
Cluster 1 (●)	6	17	6	17	7	22
Cluster 2 (●)	2	3	2	3	2	3
Cluster 3 (●)	1	5	1	5	-	-
Noise (●)	0	2	0	2	0	2

Appendix D

k-distance Graphs

k-distance graph used to determine epsilon for the DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the McNamara RBE model, shown in figure 4.29, is given in figure D.1.

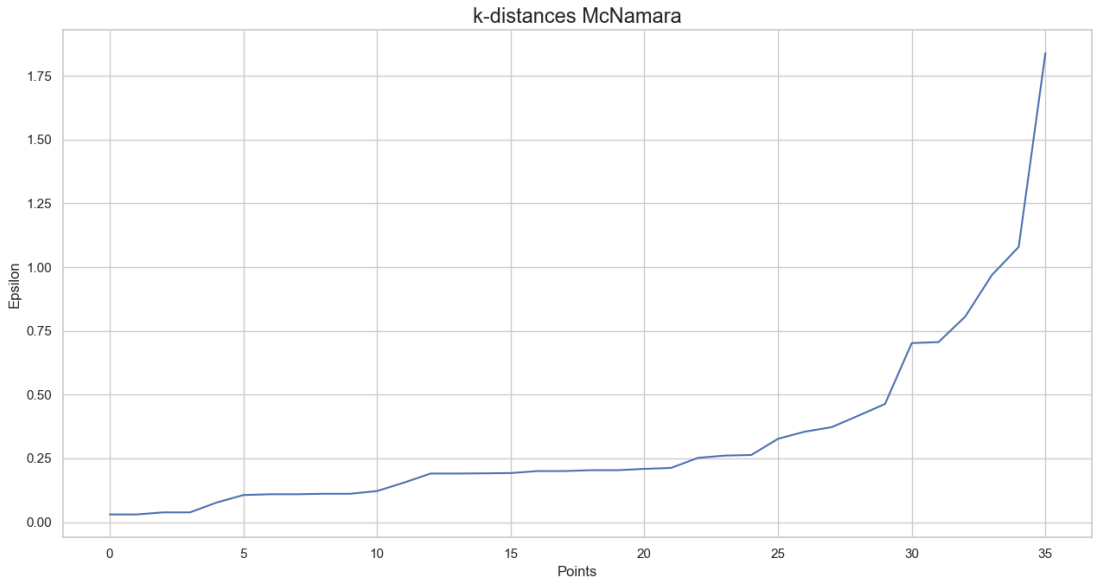

Figure D.1: k-distance graph of the D50% to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the McNamara RBE model. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.

k-distance graph used to determine epsilon for the DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model, shown in figure 4.29, is given in figure D.2.

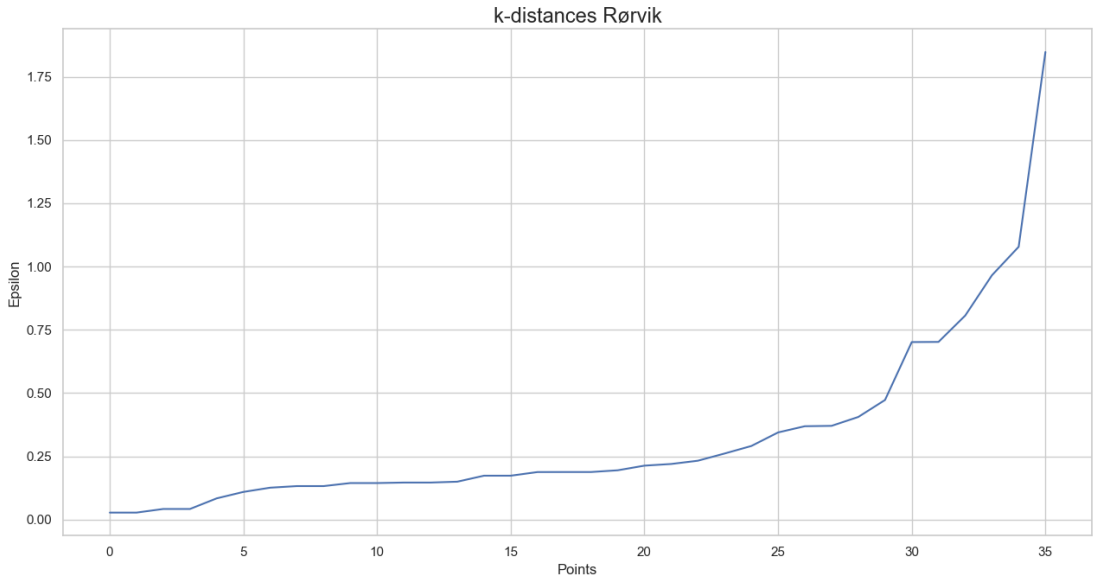

Figure D.2: k-distance graph of the D50% to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.

k-distance graph used to determine epsilon for the DBSCAN of the $D50\%$ to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from RBE1.1, shown in figure 4.29, is given in figure D.3.

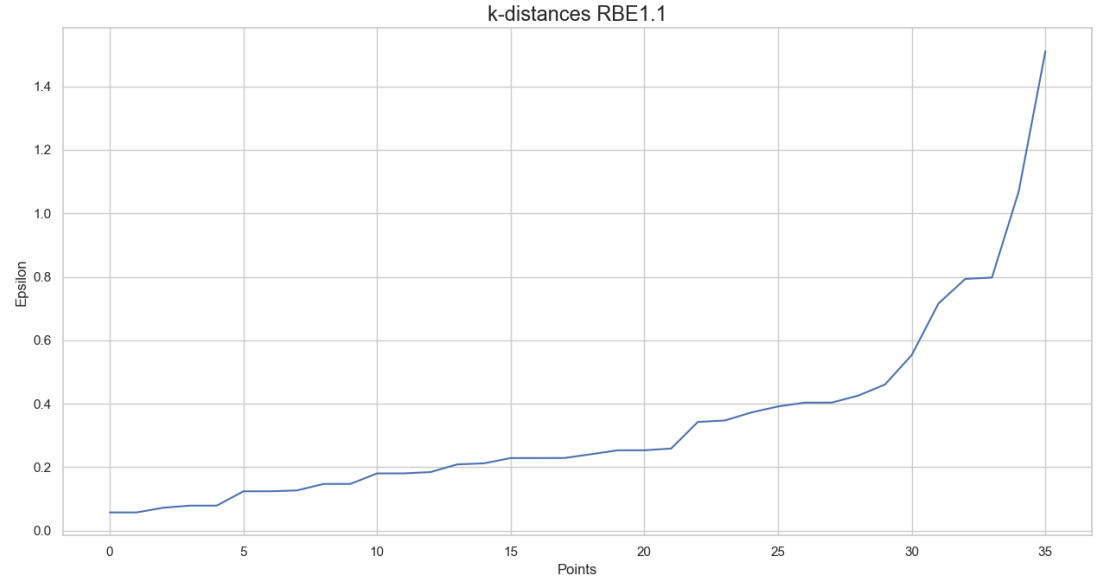

Figure D.3: k-distance graph of the D50% to the medulla oblongata, the midbrain and the pons with RBE-weighted dose from RBE1.1. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.

k-distance graph used to determine epsilon for the DBSCAN of the gEUD to the midbrain and the pons with RBE-weighted dose from the McNamara RBE model, shown in figure 4.32, is given in figure D.4.

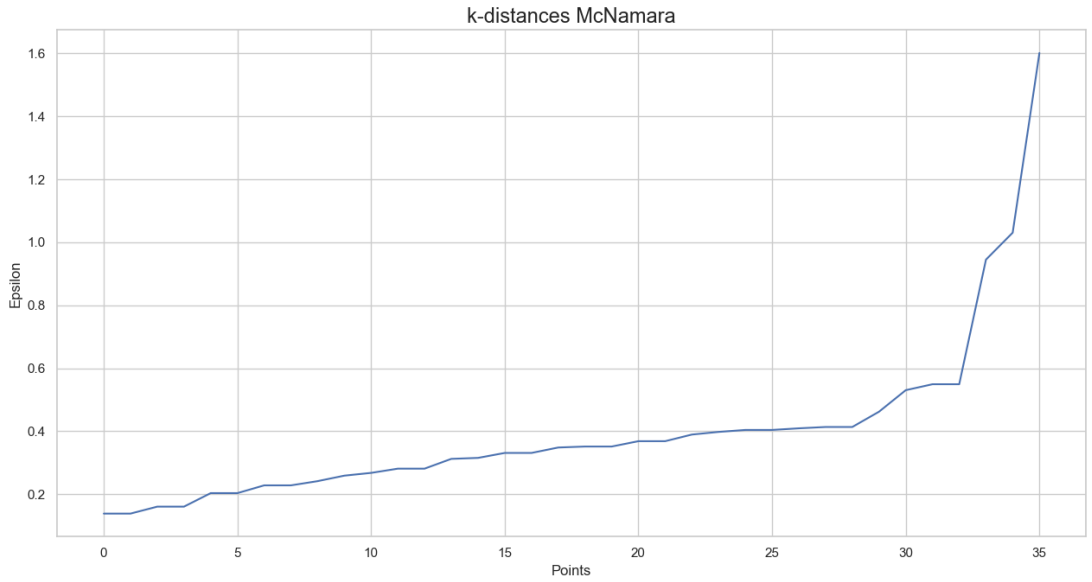

Figure D.4: k-distance graph of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted dose from the McNamara RBE model. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.

k-distance graph used to determine epsilon for the DBSCAN of the gEUD to the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model, shown in figure 4.32, is given in figure D.5.

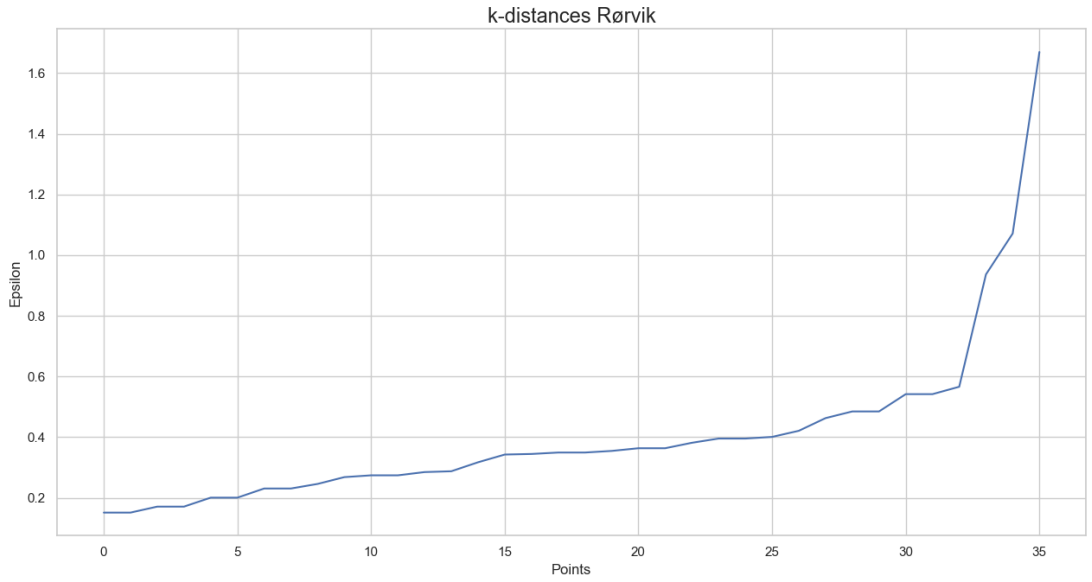

Figure D.5: k-distance graph of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted dose from the Rørvik RBE model. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.

k-distance graph used to determine epsilon for the DBSCAN of the gEUD to the midbrain and the pons with RBE-weighted dose from RBE1.1, shown in figure 4.32, is given in figure D.6.

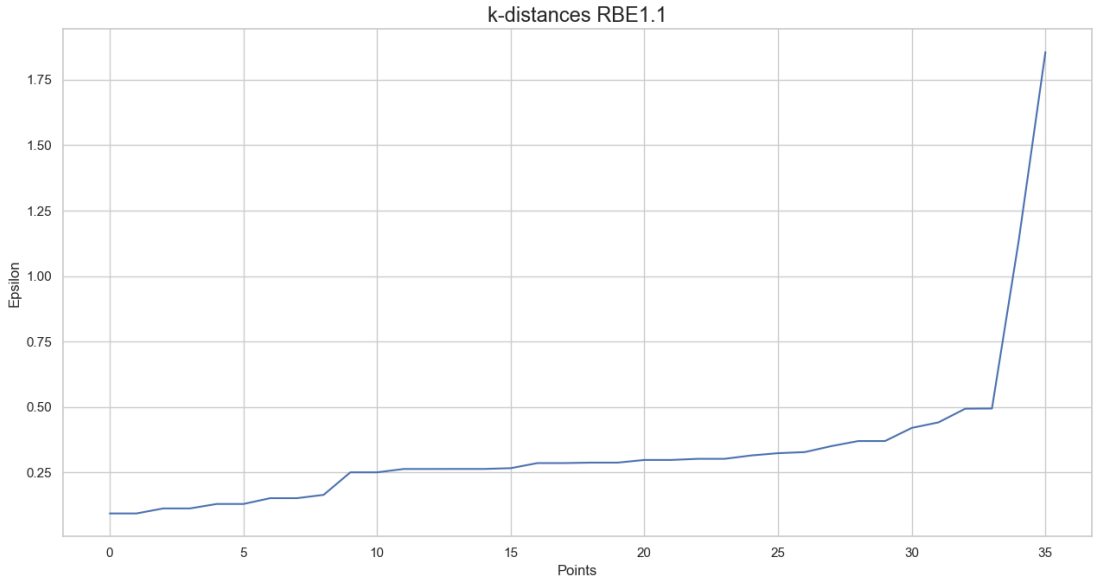

Figure D.6: k-distance graph of the gEUD with $n = 0.05$ to the midbrain and the pons with RBE-weighted dose from RBE1.1. The x-axis shows the number of points with $k = 2$ neighbours within a radial distance epsilon, while the y-axis shows the corresponding nearest neighbour distance, epsilon.