Dosimetric accuracy of synthetic-CT for pediatric brain proton radiotherapy

Bachelor Thesis

Nour Khairi
6208975

Supervisor:
Matteo Maspero
Department of Radiotherapy
University Medical Center Utrecht
May 2020

Abstract

Radiotherapy (RT) is the treatment of cancer with ionizing radiation. The delivery of radiation to a patient requires accurate planning by medical physicists, in order to target the tumour and avoid healthy tissue. Cancer patients generally undergo imaging scans in preparation for RT, typically involving magnetic resonance imaging (MRI) and computed tomography (CT) scans. CT images are a requirement for RT planning as they provide the tissue attenuation information needed for dose planning. The CT and MR workflow may result in uncertainties due to image registration. It also adds the burden of additional radiation exposure, more imaging sessions and possibly more uses of anaesthesia. With developments in computational methods, the generation of synthetic CT (sCT) for treatment planning is possible. These are fake CT images which contain the attenuation information required to perform dose calculations. In this study, we evaluated the feasibility of dose planning on deep-learning generated sCT for brain tumour pediatric patients undergoing proton RT. These sCT were generated in a previous study by a trained conditional generative adversarial network (cGAN) for a sample of patients [Bentvelzen et al., 2018]. We chose 10 of these patients to perform the evaluation. sCT data is evaluated based on image-similarity and dose comparison metrics. From the metric results, we found that the sCT are generally dosimetrically accurate and could be integrated in an MR-only workflow for proton radiotherapy.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Radiotherapy

Radiation therapy or radiotherapy (RT) aims at irradiating tumours with high energy radiation to destroy cancerous cells. The first successful cases of the therapeutic uses of radiation emerged in the late 19th century, shortly after the discovery of X-rays in 1895 [Baskar et al., 2014]. Ionizing radiation induces DNA damage, the main type being double stranded breaks (DSB) which can cause cell death if not repaired by the cell [Willers et al., 2004]. Cancer cells reproducing at a rapid rate are more sensitive to DNA damage by radiation than normal cells, making RT an effective method of treatment [Baskar et al., 2014]. Since the radiation affects both normal and cancerous cells, RT plans typically reflect an effort to spare healthy cells from harmful radiation. The goal is that substantial, curative doses are delivered to the tumour while decreasing the dose to healthy tissue. Tumour localisation is therefore extremely important [Jaffray, 2012]. The absorbed radiation dose is measured in gray (Gy) using the international system of units (SI), representing one joule per kilogram [van Assendelft et al., 1973].

Photons and electrons are the most commonly used therapeutic particles for RT in hospitals, but heavier particles such as protons and neutrons can also be employed, these have more significant biological effects due to their high density of ionisation [Joiner and van der. Kogel, 2009]. Protons in particular have shown some dosimetric advantages as opposed to photons [Baskar et al., 2014], as discussed further in section 1.3.

There are three main ways of administering radiation to the patient, the use of which is determined by the clinical task:

- Teletherapy: a beam of radiation is aimed at the patient, passing through the skin to arrive at the tumour. This method is also known as external beam radiotherapy (EBRT).
- Radioisotope therapy: radioactive material is given to the patient orally or through an injection.
- Brachytherapy: a radioactive source is placed close to or within a tumour.

The planning of EBRT will be the focus of this study.

1.2 RT for pediatric brain tumours

The most common form of solid tumours in children are brain tumours [Pollack et al., 2019]. They are only second to leukemia in overall malignancies in pediatric patients [Subramanian and Ahmad, 2019]. After 2016, central nervous system (CNS) tumours, arising from the brain or spinal chord, became the leading cause of childhood cancer-related death. This was also explained by

the decrease in mortality in leukemia patients [Subramanian and Ahmad, 2019]. Approximately 550 children are diagnosed with cancer per year in the Netherlands, and a third of them end up receiving radiotherapy. Within this group, 44% suffer from brain tumours [Guerreiro, 2019]. Little is known about the definite etiology of brain tumours, i.e, no direct cause is found for the pathology. However, various environmental and genetic risk factors have been found. Genetic predisposition syndromes and exposure to ionizing radiation (such as X-rays from CT and radiotherapy) represent the causes for a small percentage of total cases [Withrow et al., 2019]. Ionizing radiation, in particular, is associated with nonmalignant brain tumours such as meningiomas. The surveillance of trends in pediatric brain tumours is important as possible risk factors are constantly discovered [Withrow et al., 2019]. For example, the International Agency for Research on Cancer (IARC) has categorized radiofrequency (RF) electromagnetic fields as a potential carcinogenic. Children are particularly sensitive to irradiation of the cranium. Therefore, despite the non-ionizing nature of RF fields, there is some concern that increased exposure to RF fields will lead to increased cases of pediatric brain cancers [Withrow et al., 2019].
RT plays a major role in treating pediatric brain tumours. Prior to the development of RT, treatment of pediatric tumours depended on surgical resection. In the last 75 years, RT has become incorporated into the treatment of various brain tumours, either in the form of an adjuvant therapy (applied after the surgical removal), or as a definitive treatment [Knab and Connell, 2007]. However, long term side effects of RT on the CNS are significant impediments for childhood survivors of brain tumours. Nowadays, children affected by brain tumours live long enough for late effects of RT to manifest. Data from pediatric patients who have undergone whole-brain irradiation of a range of doses for CNS prophylaxis (a tumour preventative measure) highlights long term effects of CNS irradiation such as growth abnormalities, neuro-cognitive deficits, behavioural disturbances as well as a higher risk of second malignancies [Knab and Connell, 2007]. However, a wide range of CNS malignancies which were entirely fatal are now curable with the multimodality treatment approach combining surgery, chemotherapy and radiotherapy [Knab and Connell, 2007].

1.3 Photon or proton radiotherapy?

1.3.1 Particle Production

Photon RT

X-ray photons are typically used for EBRT [Kufe et al., 2003]. Therapeutic X-rays are produced by exploiting brehmsstrahlung or 'braking radiation', where charged particles are decelerated due to an electrostatic interaction with another charged particle, usually electrons interacting with the nuclei of atoms. Brehmsstrahlung is the mechanism behind linear accelerators (linacs). Linacs produce high energy (>0.33 MeV) X-ray photons, also known as megavoltage X-rays [Pereira et al., 2014]. In the linac, electrons are accelerated through structures called waveguides, following a linear path through a potential difference. High RF fields that are generated for the acceleration of electrons in the waveguides are amplified by klystrons [Podgorsak, 2005]. The path of the electron is steered by magnets, including the bend magnet which directs the electrons towards a high-density target material, typically tungsten. A schematic of the linac can be seen in Figure 1.1. The Coulomb interaction between the incident electron and the tungsten nuclei produces bremsstrahlung radiation. The electron decelerates and loses some of its kinetic energy as bremsstrahlung photons, which are emitted at a range of energies in between zero and the kinetic energy of the electron. Bremsstrahlung radiation is therefore a continuous spectrum. [Podgorsak, 2005].

Characteristic radiation is also produced from the interaction between the incident electron and the electrons of the tungsten. If the incident electron has sufficient energy, it can release an inner shell orbital electron from the atom. An electron from a higher energy level then drops down to fills the vacancy, emitting characteristic radiation in the process. The photons emitted have discrete energies that are characteristic of the target material, hence the name. [Podgorsak, 2005]. The clinical X-ray beam spectrum is a superimposition of the continuous bremsstrahlung

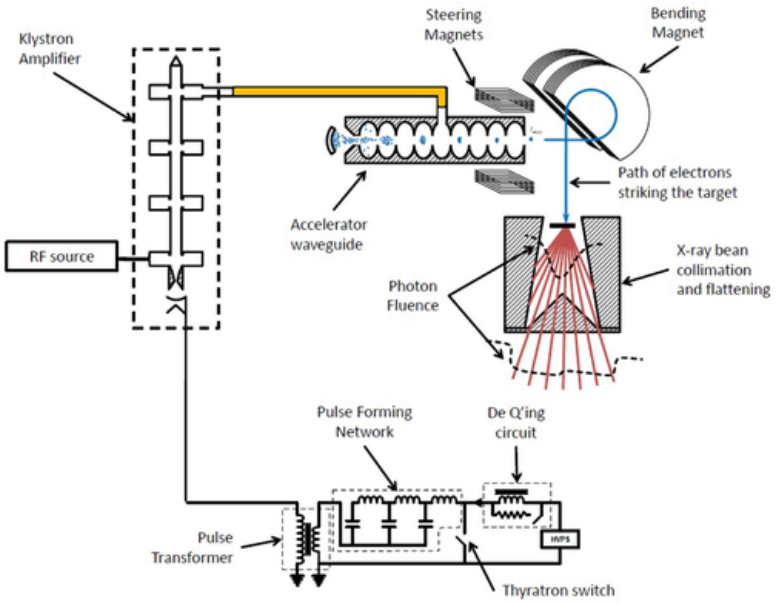

Figure 1.1: Schematic of a medical linac
[Anderson et al., 2015]

spectrum and the discrete line spectra of the characteristic radiation. The contribution of characteristic radiation in the megavoltage X-ray beams is negligible and brehmsstrahlung radiation dominates [Podgorsak, 2005]. Photon RT requires high intensity beams due to attenuation laws of X-rays and gamma rays. The change in intensity of the rays as they pass through the medium follows an exponential decay:

$$I = I_0 e^{-\mu x}$$

where I_0 is the initial intensity of the photons and I is the intensity across a distance x. μ is the linear attenuation coefficient, which is the fraction of photons attenuated per unit thickness of the medium [Cerry and Duxbury, 1998], expressed in units of cm^{-1}. A higher intensity beam therefore ensures that enough ionizing radiation reaches the tumour before being attenuated.

Proton RT

Circular particle accelerators, such as cyclotrons, are used to accelerate protons for proton beam RT. In a cyclotron, the proton is released from a source at the centre of a large electromagnet, and is accelerated along a spiral path by a uniform, static B field [Tsuboi et al., 2020]. An alternating RF voltage is applied between two D-shaped electrodes called 'dees', with a vacuum gap between them in which the particle moves from electrode to electrode (Figure 1.2). The frequency of the RF matches the rotating frequency (or gyrofrequency) of the proton, given by the following equation:

$$f = \frac{qB}{2\pi m}$$

where f is the gyrofrequency, q is the charge and m is the mass of the charged particle [Alonso and Finn, 1992]. This resonance causes an oscillating E field which accelerates the proton each time it passes through the gap, causing the radius of its orbit to gradually increase. The kinetic energy of the proton increases with every crossing of the gap, resulting in a high energy particle that can be used for therapeutic purposes [Podgorsak, 2005]. Cyclotrons and their variations are more

expensive than standard photon RT equipment [Podgorsak, 2005]. However, there are dosimetric advantages to using heavy particles such as protons instead of photons in cancer treatment with RT (section 1.3.2).

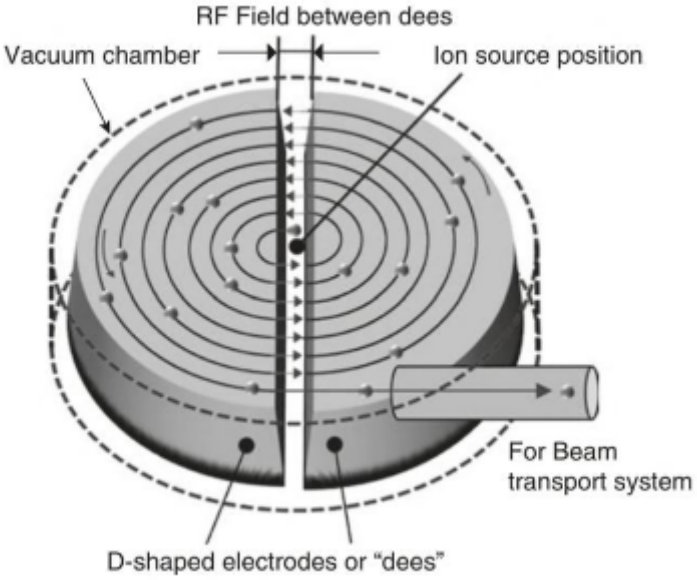

Figure 1.2: Schematic of a cyclotron
[Tsuboi et al., 2020]

1.3.2 Clinical Relevance

Both photons and protons are exploited for EBRT. The use of protons instead of photons has physical advantages which can decrease the dose to normal tissue, resulting in potentially improved long-term side effects [Merchant et al., 2008]. When X-ray photons travel through a medium, they do not deposit all their energy in a direct manner [Marcu et al., 2012]. Energy is first transferred to the electrons in the medium, then it is absorbed through electron interactions and electron stopping in the medium [Marcu et al., 2012]. This is reflected by the dose deposition curve for an X-ray beam seen in Figure 1.3. The build-up region is where dose deposition increases with depth as energy is transferred to the medium. A maximum dose deposition is reached after which the energy loss is an exponential decay of dose with increasing depth [Marcu et al., 2012]. The shape of the photon dose deposition curve indicates that therapeutic X-rays deposit most of the dose near the entrance into the tissue, meaning that more dose is deposited on the way to the tumour rather than to the tumour itself. In addition organs beyond the tumour would also be irradiated. Proton therapy has the advantage of limiting the deposition of the maximum dose to the target volume. This is explained by the Bragg curve of the proton, which is the plot of the Linear Transfer Energy (LET) as a function of distance through a medium. The Bethe formula, given below, describes the energy loss of charged particles

$$-\frac{dE}{dx} = \frac{4\pi n z^2}{m_e c^2 \beta^2}\left(\frac{e^2}{4\pi\varepsilon_o}\right)^2 \left(ln\left(\frac{2m_e c^2\beta^2}{I(1-\beta^2)}\right) - \beta^2\right)$$

where E is energy, x is the distance travelled, n is the electron density of the medium, e and $\mathrm{m_e}$ are the charge and mass of an electron respectively, $\beta = \frac{v}{c}$ where v is the velocity of the particle, c is the speed of light in a vacuum, z is the multiple of the electron charge and I is the mean excitation

energy [Grimes et al., 2017]. This is reflected in the Bragg curve's characteristic (Figure 1.3). The Bragg peak is the point at which the maximum energy deposition occurs, and is determined by the initial energy of the proton or photon. Because the energy loss of protons is inversely proportional to the square of their velocity, the Bragg peak occurs just prior to the proton coming to a stop [Newhauser and Zhang, 2015]. Figure 1.3 shows two curves for the proton dose deposition. The pristine Bragg peak occurs when the medium is irradiated with a monoenergetic proton beam, i.e., no device is used for modulation of the beam. The spread-out peak occurs when the beam has been modulated, such that the axial dimension for the peak region is increased [Newhauser and Zhang, 2015]. For therapeutic protons, the Bragg peak would be conformed towards the depth of the tumour such that the proton slows down inside the tumour. Proton beams are often referred to as having no exit dose since after the Bragg peak, the dose deposition is reduced to zero as the proton comes to a stop [Merchant et al., 2008]. In other words, there is little to no radiation dose left to deposit to organs beyond the tumour. The deposited dose prior to reaching the Bragg peak is around 30% of the maximum dose [Mitin and Zietman, 2014]. This allows therapeutic proton beams to deposit most of their energy to the tumour itself, avoiding radiation exposure to healthy tissue. Overall, these differences in energy loss patterns between photons and protons result in nearly a 60% reduction in integral dose when using protons instead of photons in RT [Mitin and Zietman, 2014].

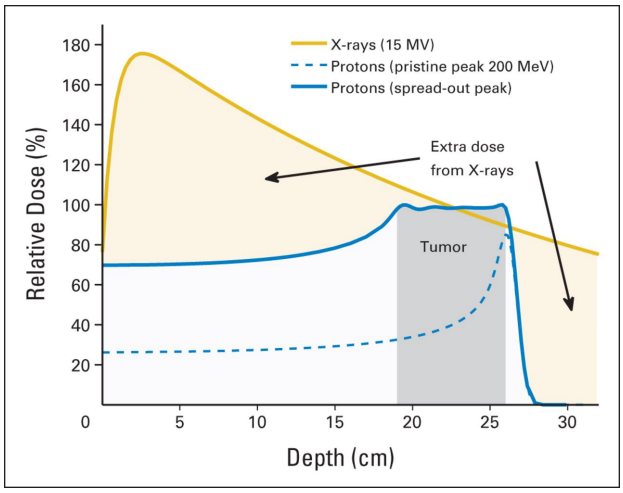

Figure 1.3: Dose depth profiles of X-ray photons in comparison to protons [Mitin and Zietman, 2014]

Proton therapy for pediatric brain tumour

There have been several studies comparing proton and photon treatments for pediatric patients suffering from solid tumours. They all point to a reduction in dose to normal tissues when protons are the treatment of choice [Lee et al., 2005, Mu et al., 2005, Weber et al., 2004]. A dosimetric comparisons of protons and photons for pediatric patients with brain tumours revealed that protons consistently lower the distribution of low and intermediate doses (0-40 Gy) to parts of the cerebrum [Merchant et al., 2008]. For patients with Medulloblastoma, for instance, critical healthy tissues (chochlae, hypothalamus, chiasm and pituitary) had substantially reduced doses when using protons in comparison to photons [Merchant et al., 2008]. Furthermore, models relating radiation dose to cognitive outcomes predict that reduced doses from proton treatments may have positive effects on cognitive function, measured through spelling scores and IQ tests [Merchant et al., 2008]. However, these are predictive models and it would require years of following up with patients to confirm cognitive advantages. Since RT can significantly compromise the health-related quality of life of pediatric patients due to their increased vulnerability to radiation, a study

investigated the effects of proton versus photon RT on the quality of life of children (2-18 years old) with brain tumours, the results revealed that patients treated with proton RT compared favourably to those who had undergone photon RT [Yock et al., 2014]. However, the true success of proton RT over photon RT in terms of sparing healthy tissue can only be determined once the survivorship of this new treatment is achieved [Merchant et al., 2008]. Therefore, there is no current consensus on whether photon or proton treatment is more favourable. A better assessment of the clinical advantages of protons will be possible once further data becomes available, and the late-effects of the treatment can be studied [Merchant, 2013]. There are still many unresolved practical issues regarding proton RT which would require further research if the treatment is to be widely introduced [Grimes et al., 2017, Newhauser and Zhang, 2015]. For instance, the finite range of proton beams adds a degree of freedom when it comes to planning the treatment [Paganetti, 2012]. The range refers to the depth of the point at which the proton energy finally reaches zero. The proton range in the tissue comes with significant uncertainties arising from various factors such as patient and imaging set-up [Paganetti, 2012]. Accurate range quantification and patient positioning is therefore required in order to fully reap the benefits of a proton treatment [Paganetti, 2012, Newhauser and Zhang, 2015].

Cost effectiveness

The main disadvantage to proton RT is the high cost of particle RT in general as opposed to photon RT. These costs have caused some to speculate whether proton RT is simply too expensive to be widely incorporated [Lievens and Van Den Bogaert, 2005]. A decision to employ particle treatment as a standard treatment is still pending in many countries, most likely due to a lack of definitive consensus on whether the many clinical benefits of proton therapy, for example, would be worth the costs [Peeters et al., 2010]. A cost analysis with input from various sources such as Belgium and the Netherlands resulted in a total capital cost of € 94.9 million for proton facilities, with one cyclotron, and € 23.4 million for photon facilities with more than one linac. [Peeters et al., 2010]. The ability to provide multiple linacs for a lower cost means more patients are able to be treated at a time in the photon facility. Furthermore, adult tissue is less vulnerable to the radiation effects of RT which result in secondary malignancies in younger patients [Mitin and Zietman, 2014]. Adult patients are furthermore less exposed to the developmental issues faced by pediatric patients. Without the cost factor, physicians would use the modality that would minimize the dose to normal tissue. However, due to the expense differences between photon and proton therapy, the dosimetric advantages may not be sufficient to justify the choice [Mitin and Zietman, 2014]. Patients and insurance companies must be convinced of a considerable clinical advantage before tolerating the high cost of proton RT, and other particle therapies [Mitin and Zietman, 2014].

1.4 The Clinical Pathway

Cancer patients follow a clinical pathway in which the cancer is diagnosed and staged then the appropriate choice of treatment is decided. After the diagnosis is made, the radiotherapy team enters a planning phase. The treatment of choice is often discussed in a multidisciplinary meeting. The radiation oncologist takes over when it is decided that RT is the appropriate treatment [Massoptier and Viard, 2014]. A schematic of the workflow that the patient undergoes is shown in Figure 1.4. During treatment, the full radiation dose is not delivered to the patient in one sitting. In fact, in the infancy of RT, it was realized that biological effects of the dose are dependent on the way it is administered over time [Bentzen, 2010]. A general consensus in RT is that the total radiation dose should be delivered in multiple smaller doses called "fractions" rather than a single large dose. This temporal programming of radiation delivery is known as fractionation [Bentzen, 2010]. Multiple sessions are therefore required to deliver the full dose to the patient. Other treatments (not shown Figure 1.4) are often used in conjunction with RT, such as chemotherapy and surgery. Once RT is prescribed as the treatment of choice, various steps are undertaken in

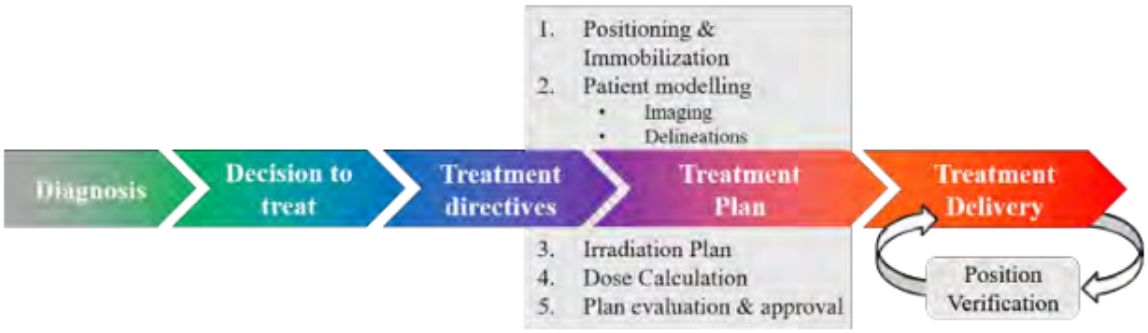

Figure 1.4: Clinical workflow undergone by a cancer patient [Maspero, 2018]

the planning stages. These procedures are explained in the following subsections.

1.4.1 Patient simulation

Positioning and Imaging

Since the dose is delivered in multiple fractions, the patient is required to be in the same position for each of their RT sessions in order to ensure full delivery of the dose to the tumour, while sparing healthy tissue. To ensure this, the patient is positioned in the treatment set-up, patient movement is minimized by immobilization devices such as head masks for those undergoing brain radiotherapy. Marks are typically made on the mask to allow for consistent patient positioning with respect to the treatment machine [Verellen et al., 2007]. After positioning, imaging is carried out to localize the tumour and to identify the surrounding radio-sensitive organs (OARs). Various modalities can be employed for this such as computed tomography (CT), magnetic resonance imaging (MRI) or positron emission tomography (PET). Since CT involves X-rays, It is often incorporated in the planning stage due to the attenuation information reflected in the images. Once the planning images are produced, the radiation oncologist delineates (outlines) the target volume; the tumour [Podgorsak, 2005]. This is known as the gross target volume (GTV). To account for the microscopic spreading of the tumour around the main volume, another delineation is made known as the clinical target volume (CTV) surrounding the GTV. Variations occurring due to organ movements, such as breathing and rectal or bladder content, are taken into account by the ITV which encompasses the CTV and GTV [ICRU, 2010, Podgorsak, 2005]. Set-up variations and machine tolerances are accounted for in the planning target volume (PTV) surrounding the ITV, CTV and GTV [ICRU, 2010, Podgorsak, 2005]. A schematic of these volumes is given in Figure 1.5.

CT in RT

As mentioned above, planning CT images are widely used in treatment planning. The use of CT planning has been crucial in the development of modern RT. CT images provide attenuation information such that dose distributions can be optimized and exposure to radio-sensitive organs can be minimized [Meyer et al., 2017]. Studies have shown the capacity of planning CT in detecting unseen clinical findings such as unexpected metastases which were undetected in the diagnostic stage, as well as the possibility of finding concurrent malignancies [Meyer et al., 2017]. CT images provide information by measuring x-ray attenuation coefficients of a volume element or 'voxel' within the tissue. These attenuation coefficients are converted to CT values (gray-scale values)

which correspond to the 2D image elements or pixels. CT values are then normalized according to the attenuation properties of water and represented as Hounsfield units (HU) as shown below

$$HU = 1000\frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}}$$

where μ_{water} and μ_{air} are the linear attenuation coefficients for water and air respectively. The Hounsfield scale was named after British physicist Godfrey N. Hounsfield who was one of the Nobel Prize-winning inventors of CT. It is a density scale where water is arbitrarily assigned a density of 0 HU and air a density of -1000 HU. Bone, which has a high density, has a value close to 1000 HU. HU values of tissues reflect their X-ray attenuation and are therefore related to their physical density [Broder, 2011]. CT images are displayed in grey-scale which is representative of the HU values of tissues. Tissues with lower densities appear black, while high-density tissues appear white [Broder, 2011]. All intermediate tissue densities are represented by the various shades of grey in between. The human eye cannot perceive extremely subtle variations in grey tones, therefore, a specific range or 'window' of HU values is chosen to be displayed in the CT image, depending on the HU value of the tissue of interest. This is known as windowing or grey-level mapping [Broder, 2011]. Calibrating from CT numbers to relative electron densities and stopping power in the TPS is an essential part of RT planning [Koniarova, 2019].

CT scans have become powerful diagnostic and planning tools in RT. However, they deliver significantly higher radiation doses than other diagnostic methods. This is of particular concern for pediatric patients as they are more vulnerable to malignancies related to radiation and have longer life spans to display late effects. Many epidemiological studies have been conducted to investigate the correlation between CT scans and the subsequent development of pediatric cancers, including brain tumours [Meulepas et al., 2019]. For example, studies revealed a positive correlation between radiation dose from CT scans and brain tumours in pediatric patients in the UK. Similar results were found amongst Australian and Taiwanese children. This is partially due to "con-founding by indication" [Meulepas et al., 2019]. In other words, for some of the patients, the need for CT scans is dependent upon symptoms that were leading to brain tumours in the first place. A recent epidemiological study in the Netherlands was conducted to investigate the possible correlation amongst children in the Netherlands. The results showed that high radiation doses delivered in head CT (20–50 mGy) may increase brain tumor risk within a decade of the scan [Meulepas et al., 2019].

MRI in RT

Conventionally, CT imaging was the main modality used for treatment planning. However, since the advancement of multimodality imaging, improvements have been made in the administration of EBRT in terms of facilitating accurate imaging of the target volume and the surrounding OARs. MRI became important in providing contrast information which can aid in segmentation of volumes of interest as well as providing motion information and post-treatment analysis to understand the biological effects of the therapeutic radiation [Chandarana et al., 2018]. MRI has become the primary modality used in RT treatment planning for brain tumours [Yu et al., 2019]. It plays this role due to its tissue delineation and definition of the tumour margins. There are, however, some disadvantages to MR imaging. These include the lack of electron density information and the frequency of geometrical inaccuracies due to distortions. CT images on the other hand have great geometrical accuracy and provide electron density information. Therefore, the use of CT complements the MRI images and thus the two modalities have been used in conjunction and CT and MR images are registered (mapped onto) one another [Beavis et al., 1998]. Using a combined MR-CT workflow at the planning stage provides all the information required to generate an effective RT plan. However, this method can result in residual errors as well as an added burden on the patient due to the multiple imaging sessions required. A possible solution to this is explored in section 1.5.

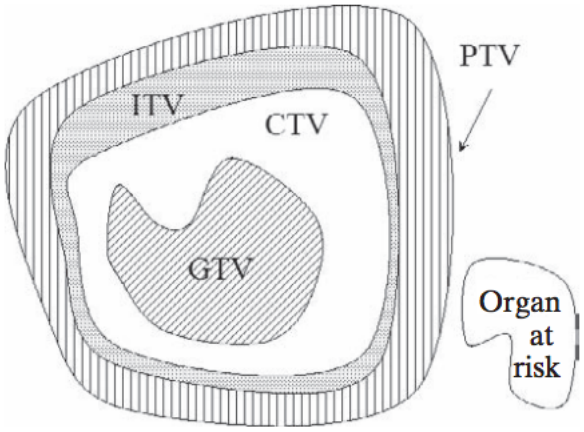

Figure 1.5: Schematic of the GTV, CTV, ITV, PTV and OAR
[Podgorsak, 2005]

1.4.2 Generating the RT plan

The possibility of producing and delivering an ideal treatment plan, in which 100% of the irradiation is administered to the target volume while 0% of the dose is received by OARs and healthy tissue is the "holy grail" of radiotherapy. This is practically impossible to generate due to the dose deposition of photons and protons as they travel through the tissue. Therefore, the trade-offs required in order to achieve the closest scenario to this are crucial in producing a clinically realistic yet optimum plan [McGowan et al., 2013]. An RT treatment plan is produced by optimising the position of the radiation fields to arrive at radiation doses that fit the defined dose constraints. Optimisation is carried out in a treatment planning system (TPS), which is a software that generates 3D dose distributions according to the mode of irradiation, the 3D patient model and the defined constraints and prescriptions for radiation dose. An example of a TPS is Raystation, shown in Figure 1.8. Usually, radiation doses are calculated using CT images registered to MR images. The registration process is not perfect and results in errors of the tumour volume [Nyholm et al., 2009]. For brain tumours, these errors tend to be small because of the rigidity of the skull [Balachandran et al., 2010]. However, they can be significant when the scans have differing setups.

Objective-based planning

An EBRT treatment plan includes various parameters such as the number of beams, their intensities as well as their geometric setup. Beam-shaping devices have facilitated intensity modulation for each beam, an example of this is the multi-leaf collimator (MLC) often used for photon beam RT [Voutilainen, 2016]. MLCs have movable leaves which function to block out fractions of the radiation, seen in Figure 1.6. Without the MLCs, conventional collimator jaws result in rectangular fields which do not correspond to the target volume shape and result in increased dose to surrounding tissue [Jeraj and Robar, 2004]. Intensity modulated radiation therapy (IMRT) is a technique where each beam is divided into 'beamlets', whose intensity can be chosen independently in order to attain better targeting and less side-effects. This method has increased the number of parameters in treatment planning [Voutilainen, 2016]. An analogous issue is seen in proton treatments. Proton RT involves the generation of a narrow mono-energetic beam (also known as a "pencil beam") which scans the target volume with the aid of magnetic deflection. [Paganetti and Bortfeld, 2006]. Pencil beams are scanned in a zig-zag motion in the x-y plane perpendicular to the direction of the beam (z). The depth of the beam in the z direction is modified by varying

its initial energy. A schematic of pencil beam scanning can be seen in Figure 1.7. Scanning starts at the deepest layer, which requires the highest energy pencil beam, the proton beam energy is then reduced for the next layer, and so on [Paganetti and Bortfeld, 2006]. Unlike IMRT, this method does not require the use of collimators such as the MLC [Paganetti and Bortfeld, 2006]. Intensity modulated proton therapy (IMPT) is therefore defined as the electromagnetic spatial control of pencil beams (or pencil beam scanning) with variable energies and intensities [Kooy and Grassberger, 2015]. Since both the energy and intensity of each pencil beam can be modified, IMPT significantly increases the degrees of freedom in treatment planning, making proton beam shaping highly accurate to the target volume, while also being highly computationally complex [Paganetti and Bortfeld, 2006].

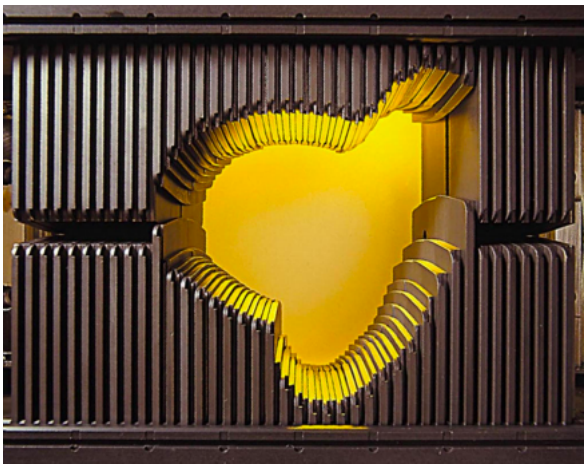

Figure 1.6: Varian's Millennium 80 MLC
[Voutilainen, 2016]

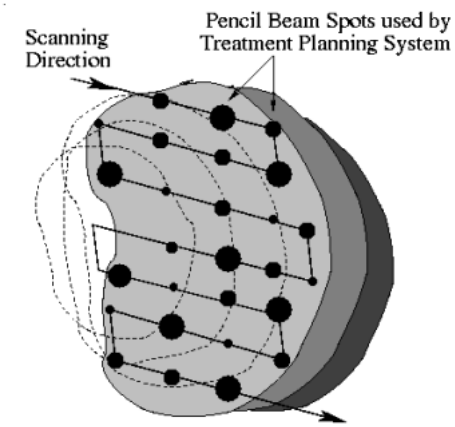

Figure 1.7: Pencil beam scanning method
[Trofimov and Bortfeld, 2003]

Inverse planning

As a result of increased parameters, a movement has been made into inverse planning, in which outcomes are specified by the planner and the program calculates which parameters would be closest in leading to the desired outcome [Voutilainen, 2016]. Inverse planning involves optimization methods which are used to arrive at an approximate solution rather than trying to arrive at solutions analytically. Optimization methods usually employ dose-volume "objectives" which are bounds specified for each structure. The target volume is assigned with an upper and lower dose objective while OARs are assigned a lower dose objective. Dose-volume objectives are usually sufficient in ensuring the prescribed dose to the tumour and the least radiation delivered to OARs [Voutilainen, 2016].The are several techniques for dose calculation in RT planning, including pencil beam algorithms and Monte Carlo methods explored below [Jabbari, 2011].

Dose calculation

The Monte Carlo method (MC) is the most accurate technique for dose calculation. It is widely used for photon RT planning. MC is a statistical simulation that relies on a repetition of random sampling to solve deterministic problems. In the context of radiative transport, MC yields a numerical solution to the Boltzmann transport equation which employs the fundamental laws of photon-atom and electron-atom interactions [Bielajew, 2007]. Physical interactions, such as compton scattering from atomic electrons, must be modelled by a Monte Carlo code written for RT applications. MC for radiotherapy simulates the path of individual particles with the use of sampling to a sufficient amount from the probability distributions which govern individual physical processes. It does this by making use of machine-generated random numbers [Jabbari, 2011]. Improvements have been made to the MC dose calculation algorithms used within the TPS in order to reproduce all beam geometries and modification devices accurately, as well as to take into account the heterogeneity effects in the patient geometry [Jabbari, 2011]. For proton RT, analytical algorithms are usually used to perform dose calculations (also called pencil beam algorithms) [Schuemann et al., 2015]. This approach assumes that any beam incident on the patient is a collection of narrower and smaller pencil beams. At the centre of each pencil beam is a central axis along which a dose is delivered. The pattern of dose deposition varies with intensity and spectrum of the beam [Carolan, 2010]. The total beam is modelled as a conglomeration of these infinitely small pencils after the collimator [Jabbari, 2011]. These assumptions compromise the accuracy of the analytical pencil beam method in comparison with the physically accurate MC approach [Jia et al., 2012]. MC simulations for proton therapy do exist but have not been widely incorporated into proton clinical settings due to their lengthy computational time [Schuemann et al., 2015].

1.4.3 Immobilisation and irradiation

Once the RT plan has been generated, treatment delivery can take place. The patient is positioned in such a way that the target volume with respect to the linac matches the location on the RT plan [Saenz et al., 2018]. Patient set-up and target localisation can be improved using image guided radiotherapy (IGRT), in which various imaging modalities such as cone-beam CT, PET or MRI are used to image the patient just before irradiation, in order to verify their set-up in relation to the radiation fields [Verellen et al., 2007]. While IGRT ensures the quality of the treatment by confirming the placement of the beams, it could increase fraction time as well as radiation exposure to the patient. It also assumes that errors in patient positioning can be corrected by rigid transformation (table shifts and rotations) which may be an oversimplification of the kind of revisions required. IGRT usually decreases positioning errors, thereby decreasing radiation delivered to the surrounding tissue [Dawson and Jaffray, 2007].

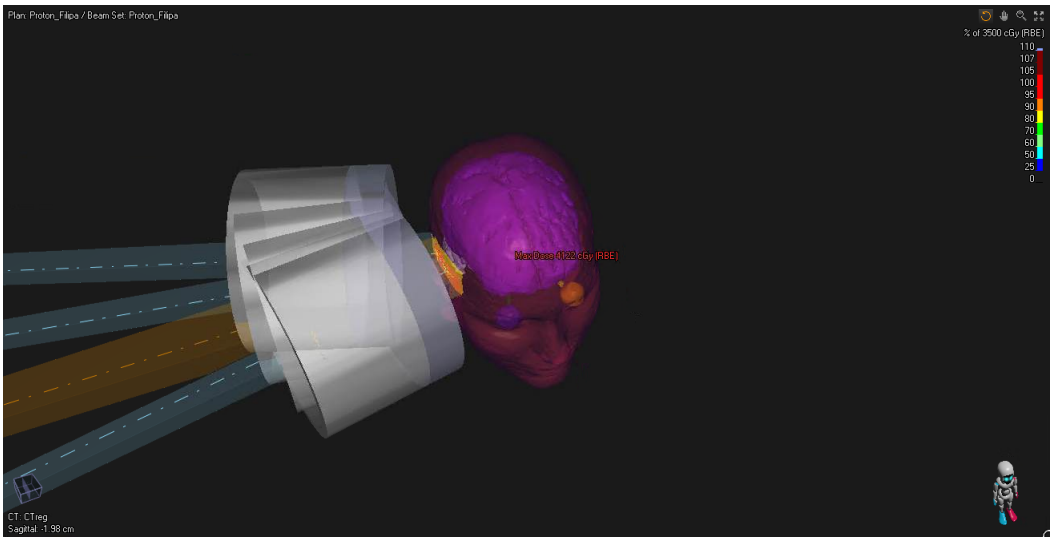

Figure 1.8: Example of beam configurations in the TPS Raystation.

1.5 MR-only radiotherapy

MR-only is a proposed workflow in which MRI without CT imaging is used at the treatment planning stage. The intensity of voxels in MR images do not correspond to tissue attenuation properties, contrasts in MR images rather reflect the proton densities and magnetic relaxation of the tissue. This makes dose calculations in an MR-only workflow are a challenge [Edmund and Nyholm, 2017]. In the usual MR-CT workflow, dose calculations are performed on MR registered onto CT images. Residual registration errors may persist after registration in this multi-modality approach [Ulin et al., 2010]. This means that systematic uncertainties are potentially induced, leading to a reduction in treatment effectiveness [Nyholm et al., 2009]. Proposing RT based on sole MRI eliminates these residual uncertainties [Edmund and Nyholm, 2017, Owrangi et al., 2018]. MR-only RT is furthermore beneficial considering the reduction in CT scans, with subsequent patient burden and overall cost reduction [Devic, 2012]. In addition, considering that the pediatric population is more sensitive to radiation effects compared to an adult population [Hall, 2006], MR-only RT is particularly attractive given the reduced exposure to the ionizing radiation delivered through CT at the planning stage. Patient immobilisation during both scanning and treatment is required for accurate tumour localisation and targeting. As this may be a challenge for pediatric patients, anaesthesia or sedation can be required for both CT and MRI scanning. This poses clinical risks such as oxygen de-saturation for pediatric patients, especially those with developmental delays. Furthermore, other underlying risks can be further increased due to patient co-morbidities in both pediatric and adult patients [Malviya et al., 2000]. An MR-only approach reduces the number of scanning sessions, and thus the need for anaesthetic use, which can be beneficial for patients of all ages.

1.5.1 sCT Generation with CNNs

Since tissue attenuation information, usually provided by the CT, is needed for dose calculations, the generation of synthetic CT (sCT) has been proposed. sCT can be generated by converting MRI into CT-like data. There are various methods of generating sCT, usually classified in voxel, atlas or hybrid approaches. Voxel-based methods are dominated by machine learning and use the voxel constrasts of MR images to assign them with corresponding HU values. Atlas-based methods mainly align the location of the MRI voxels to an atlas (template) by registration. Hybrid methods are a combination of the atlas and voxel approaches [Edmund and Nyholm, 2017]. In the sCT

used for this work, sCT for patients with primary brain tumors were generated using convolutional neural networks (CNNs).

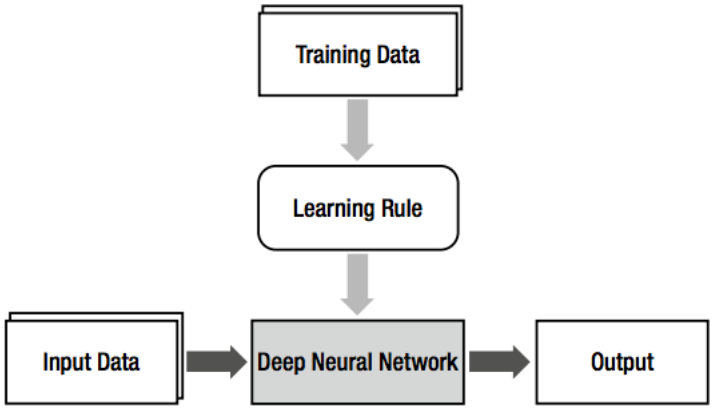

Figure 1.9: Concept of deep neural networks.
[Kim, 2017a]

CNNs are deep learning algorithms specialized for image data [Kim, 2017a]. CNNs were inspired by the visual cortexes of humans and other animals. Neurons in the visual cortex respond individually to regions in the visual field. This region is known as the receptive field of the neurons. Around the individual neuron, adjacent neurons have overlapping receptive fields such that a full map of the visual field is formed in the cortex [Hubel and Wiesel, 1968]. In a similar way, CNNs are fed with input images and they assign importance to different aspects of the image. [Kim, 2017a]. CNNs are a class of deep neural networks, which are multi-layered neural networks with one or more hidden layers [Kim, 2017b]. A schematic of how deep neural networks work, including CNNs, is shown in Figure 1.9, where the "Learning Rule" is the algorithm generating the model from the training data. Sufficient training is an important part of the success of a deep neural network and involves feeding the network with data such that it learns how to generate the desired output [Kim, 2017b]. sCT generation is here interpreted as an image-to-image translation problem, for which CNNs have already been employed for various anatomical sites [Wolterink et al., 2017, Dinkla et al., 2019, Maspero, 2018, Liu et al., 2019]. These approaches have shown promising results when trained with homogeneous datasets.

Aim of the study

In this study, we investigate a voxel-based method of sCT generation with a CNN carried out by Bentvelzen et al., 2018. We assess the HU assignment of the sCT compared to CT images, as well as their geometric and dosimetric accuracy. In evaluating this, we aim to assess the feasibility of dose planning on these sCT. We therefore also evaluate the success of trained CNNs in generating sCT that are accurate enough for proton RT planning in the case of pediatric patients with brain tumour.

Chapter 2

Methods

The success of trained CNNs in generating sCT is evaluated by quantifying the similarity of the sCT to real CT data through mathematical metrics. A subset of 10 patients of different demographics were selected out of the 52 patients used in a larger study [Bentvelzen et al., 2018]. These patients had undergone MR and CT scans as part of the clinical routine for RT planning. MR and CT scans for RT planning and had MR-based sCT generated was selected. The accuracy of sCT was assessed by image and dose evaluations. An overall workflow is seen in Figure 2.1, in which the image and dose evaluation step will be the focus of this study. This section will include the patient demographic, a brief look at the sCT generation method used by Bentvelzen et al., 2018 to create the sCT evaluated here, as well as the image-based similarity and dose comparison metrics used for evaluation.

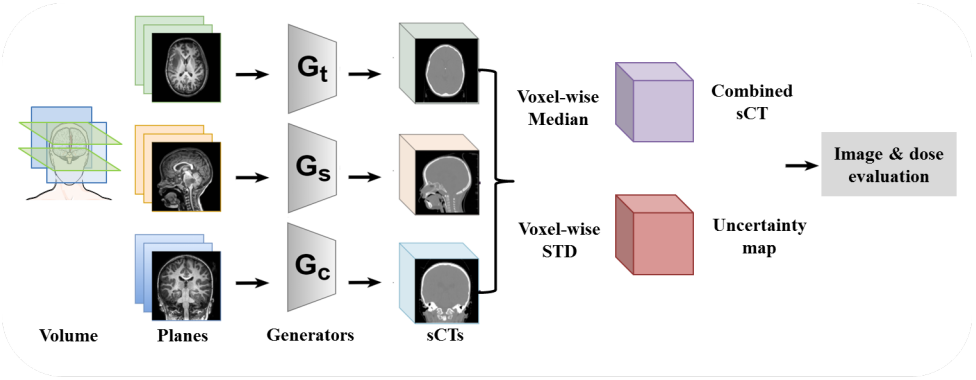

Figure 2.1: Workflow of the project
[Maspero et al., 2020]

2.1 Patient demographic

The test set in this evaluation included 10 patients with a male to female ratio of 3/7. 2 patients had embryonal tumours, 1 had an ependymal tumour. The rest of the patients had diffuse astrocytic and oligodendroglial gliomas which are the most common type of neuroepithelial tumour [Malzkorn, 2019]. Most patients were at Grade IV stage of cancer except for the patient with ependymal tumour who was at Grade III. The ratio of intact to post-surgery brains was 4/6 and the ages of the patients ranged from 2.6 - 17 years old. As for the treatment plans, these depend on various factors: clinical context, demographic of the patient and their tumour type [Bentvelzen et al., 2018]. Patients in the test set were chosen out of a larger study of 52 patients who had

undergone CT and MR scans to generate RT plans in preparation for treatment. Treatment plans varied from patient to patient. The total dose administered ranged from 20-54 Gy, and the number of fractions ranged from 10-30 fractions. The dose per fraction ranged from 1.8 - 3.5 Gy per fraction. The summary of patient information is given in Table 2.1, where the second column is the use of the contrast agent Gadolinium in MR scanning to enhance visibility of internal structures. The parameters for the MR scans undergone by all patients in the test set are given in Table 2.2, where the MR sequence type is a 3D T1-weighted gradient echo, the flip angle refers to the rotation of the net magnetization as a result of the applied radiofrequency pulse, the FOV is the field of view of the scan, and TR and TE are the echo and repetition times respectively [Bentvelzen et al., 2018].

Patient	Gado	Tumour Class	Stage	Gender	Age	Dose (Gy)	Fractions
9	N	Ependymal	Gr III	M	5.3	54.0	30
13	Y	Embryonal	-	F	2.6	54.0	30
22	N	Diffuse astrocytic and oligodendioglial	Gr IV	F	9.6	20.0	10
35	N	Diffuse astrocytic and oligodendioglial	Gr IV	F	17.0	35.0	10
38	N	Diffuse astrocytic and oligodendioglial	Gr IV	F	5.1	39.0	13
42	N	Embryonal	Gr IV	F	7.1	36.0	20
60	N	Diffuse astrocytic and oligodendioglial	Gr IV	M	12.8	54.0	30
65	Y	Diffuse astrocytic and oligodendioglial	Gr IV	M	9.0	36.0	20
79	Y	Diffuse astrocytic and oligodendioglial	Gr IV	F	13.6	39.0	13
85	N	Diffuse astrocytic and oligodendioglial	Gr IV	F	7.6	54.0	30

Table 2.1: Demographics of patients included in the test set

Parameter	Value
Sequence type	3D T1-W GE
Flip Angle (°)	8-12
TR (ms)	4.8-10.5
TE (ms)	2.2-5.9
FOV (mm^3)	200-256, 200-256, 150-220
Matrix Size	240-512x240-512, 240-512x240-512, 150-350x150-350
Acquisition time (s)	31-360

Table 2.2: Range of MRI scan parameters
Values are presented as ranges in the anterior-posterior, right-left and cranial-caudal directions [Bentvelzen et al., 2018].

2.2 sCT generation

sCT generation was performed in the previous study by Bentvelzen et al., 2018. The training of the network included a subset of 40 patients out of the total 52. The CT of the subset were rigidly registered onto MRI using the Elastix registration toolbox. This registered CT will be abbreviated as CT_{reg}. Examples of the CT, CT_{reg} and MR images are shown in Figure 2.2. The sCT images used in this project were generated by a conditional generative adversarial network (cGAN). GAN networks consist of a generator and a discriminator network. These two networks are trained together to generate sCT that are as close to real CT images as possible, achieved by exploiting the ability of the discriminator network to differentiate between real and synthetic images [Maspero et al., 2018a]. As Figure 2.2 shows, the generator network (G) and the discriminator network (D) have two opposing goals. G produces CT (x) that have the minimum pixel-wise differences with the real CT images (CT_{reg}) (x'), while D tries to discriminate between the sCT generated by G and real CT_{reg} images. G produces images that look convincing enough such that D cannot discriminate

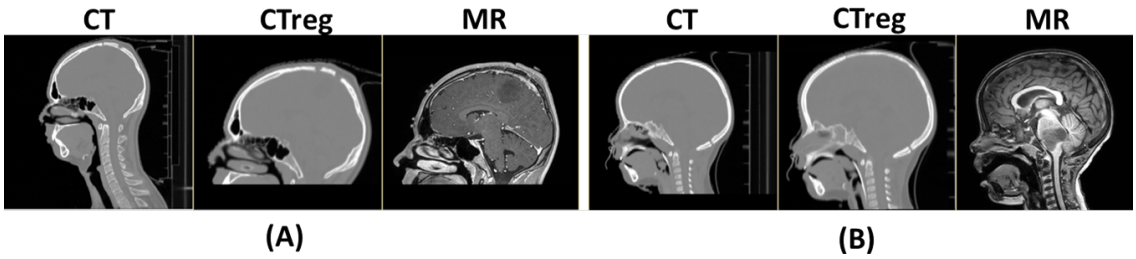

Figure 2.2: Two examples of subjects with CT, CT_{reg} and MRI. (A) 4 year old subject, no contrast agent, and a large FOV, (B) 17 year old subject, scan with gadolinium, and a small FOV.

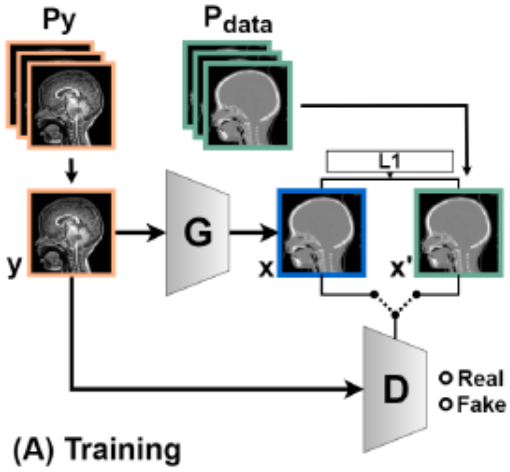

Figure 2.3: sCT generation method using the cGAN network

between them and the real CT_{reg}. It does this by utilizing the MR data (y). [Bentvelzen et al., 2018]. The success of this sCT generation approach for brain tumours would be evaluated by the affinity of generated sCT to CT_{reg}.

2.3 Evaluation metrics

Image-based similarities and dose comparisons were made. The metrics used for this evaluation are explained in the following subsections. All the following evaluations were performed in Matlab2019a.

2.3.1 Image-based similarity

The following metrics deal with the image quality assessment of the generated sCT to secure an image similarity to CT_{reg} sufficient enough for RT planning. The mean absolute error (MAE) and dice similarity coefficient (DSC) will assess the HU assignment of the sCT and their geometrical accuracy respectively.

Mean Absolute Error (MAE)

The mean error (ME) and mean absolute error (MAE) summarise the voxelwise differences in HU between $\mathrm{CT}_{\mathrm{reg}}$ and sCT. This is calculated for the intersection of the body contours of the sCT and $\mathrm{CT}_{\mathrm{reg}}$. The MAE is often calculated for the full FOV of the patient [Dinkla et al., 2019, Emami et al., 2018]. Considering the heterogeneous nature of the data and its FOV variation, the FOV size was homogenized in the analysis and a smaller FOV was considered which includes the superior part of the head to the base of the skull. Both the ME and MAE are measures of errors in HU between the two sets of data. The MAE, however, gives a magnitude of the average error. The mathematical form of the MAE and the ME is the following:

$$ME = \frac{1}{N}\sum_{i=1}^{n} sCT_i - CT_{\mathrm{reg}_i}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{n} |sCT_i - CT_{\mathrm{reg}_i}|$$

where N is the number of voxels in the head volume of the patient and CT_{reg_i} and sCT_i are the i^{th} reference voxel of the $\mathrm{CT}_{\mathrm{reg}}$ and comparison voxel of the sCT respectively [Edmund and Nyholm, 2017]. MAE values in the body contours are calculated for each patient in the small FOV. An average MAE across all patients is computed to investigate if the magnitude of error is comparable to that of similar studies.

Dice Similarity Coefficient (DSC)

The dice similarity coefficient (DSC) is a geometrical metric that assesses the spatial overlap between sCT and CT [Zou et al., 2004], where the agreement between sCT volume contours with those of the CT is quantified. The DSC is given by:

$$DSC = \frac{2|sCT_v \cap CT_v|}{|sCT_v| + |CT_{\mathrm{reg}_v}|}$$

where sCT_v and CT_v are the volumes of the structure in the sCT and $\mathrm{CT}_{\mathrm{reg}}$ respectively. A score between 0 and 1 is given, where 0 means no overlap and 1 means a perfect overlap [Yang et al., 2018]. High scores (≈ 1) would indicate high geometric consistency between the CT and sCT. The DSC is calculated for body and bone segmentations for each patient. A mean DSC is presented for comparison with other studies.

2.3.2 Dose Comparison

Rigid registration of sCT to planning CT was performed and delineations from the clinical plan such as target volumes and OARs were transferred to the registered sCT. Proton plans were generated in the TPS Raystation (v8.99.30.40, RaySearch Laboratories AB, Sweden) on the planning CT with the aim of a keeping the CTV median dose comparable to that of photon plans. The plans were optimised considering a 2% range and 2mm set-up uncertainties, and pencil beam scanning (PBS) was simulated for all the treatment plans. Three beams were configured at 160°, 180 and 2000°. The beam configuration was modified if the dose constraints were not met for a particular patient. For every patient, the PBS plans were then recalculated on the sCT [Bentvelzen et al., 2018]. The following metrics compare the CT-based dose calculation data with the $\mathrm{CT}_{\mathrm{reg}}$ dose recalculation.

Dose Volume Histogram (DVH)

The cumulative dose volume histogram (DVH) is a graphical representation of the percentage of the prescribed dose received by a percentage of a volume of interest, resulting from a particular RT plan. The percentage of the volume of a structure is on the y-axis, this volume receives a

dose greater than or equal to the corresponding dose on the x-axis. The DVHs are calculated for each patient and represented in graphs with a CT and sCT curve for each structure to highlight dosimetric differences. Dose differences to OARs (cerebrum, chiasm, lenses, brainstem, pituitary gland and cochleas) [Bentvelzen et al., 2018]) are retrieved and analyzed.

Voxel-wise Dose Differences (DD)

Relative voxel-wise dose differences (DD) evaluate the local dose differences between the CT and sCT data based on voxels, rather than the organ-based approach seen in the DVH metric. DDs are calculated by subtracting planning CT doses from sCT doses and representing this as a relative dose to the original planning CT dose. DDs are calculated for the body, the high dose regions receiving 90% of the dose, and the low dose regions receiving 10% of the dose. These are abbreviated by D>90% D>10% respectively.

$$DD\% = \frac{sCT - CT}{CT} 100$$

Gamma Index (γ)

The clinical CT plan represents the reference dose distribution, while recalculations done on the registered sCT represent the evaluated dose distribution. The Gamma index (γ) introduced by [Low et al., 1998] provides a dimensionless metric that combines the dose difference with the distance difference for each point on the evaluated dose distribution. It adds criteria (limits) on the differences in dose and distance to evaluate similarities between dose distributions. To find the gamma index, the following is calculated for each reference point:

- the minimum Euclidean difference between the reference and evaluated point (Δr)
- the dose difference between the reference and evaluated point (ΔD)

The following is computed for each reference point r_R:

$$\Gamma = \sqrt{\frac{\Delta r^2}{\delta r^2} + \frac{\Delta D^2}{\delta D^2}}$$

where δr is the distance difference criteria and δD is the dose difference criteria.

The gamma index is given by the minimum value calculated for all evaluated points r_E

$$\gamma = min(\Gamma) \forall (r_E)$$

Gamma criteria are denoted in the form $\delta D(\%)/\delta r(mm)$. A percentage of the evaluated points which pass the distance difference and dose difference criteria is therefore given. For this evaluation, the 2%/2mm and 3%/3mm criteria are used to assess dose differences. Gamma calculations are performed on the body, high dose D>90% and low dose D>10% regions.

Chapter 3

Results

3.1 Image-based similarity

Image-based similarity results are displayed as an overall mean or average over all patients in the test set, with a standard deviation (± SD). The range of results is also given. (min - max). This is summarised in Table 3.1.

MAE and DSC

The average MAE result over all patients in the test set was 65.7 HU Average DSC results are both close to 1, particularly for DSC for body segmentations. DSC for the body is slightly lower, with a higher SD in comparison to that of bone segmentations.

Metric	Average Result	Range
MAE (HU)	65.7 ± 13.5	41.9 —82.5
DSC Body	0.98 ± 0.004	0.98—0.99
DSC Bone	0.82 ± 0.08	0.66—0.87

Table 3.1: Average and range of MAE (HU) and DSC results

3.2 Dose Comparison

Averages and ranges for DD and gamma criteria are presented in Table 3.2 along with their SD for the body, low dose (D>10%) and high dose regions (D>10%). The table shows that overall dose differences (DD) are within 1% and gamma pass rates for the $\gamma 3\%/3mm$ and $\gamma 3\%/3mm$ are $\approx$ 100%. Tables 3.3 and 3.4 show DD and gamma results for an exemplary patient (pt 9) and an anomalous patient (pt 35). Graphs of the DVH, along with MR images, real CT (rCT), sCT and image differences, are displayed for both patients in Figures 3.1 and 3.2. In addition, the DVH and images of a patient with a high dose difference in the left cochlea are shown in Figure 3.3.

Voxel-wise differences (DD)

Average DD for the body and low dose regions (D>10%) had the same value. High dose regions (D>90%) had less dose differences on average. Overall, mean DDs were well within 1%, as seen in Table 3.2. For the exemplary patient (pt. 9), DD results were 0 for the body and (D>10%). Therefore, no voxel-wise differences between the sCT and CT for those regions, as seen in Table 3.3. For the anomalous patient (pt. 35), the body and low dose region had a significant voxel-wise difference of -4.32 (Table 3.4)

	Body	D>10%	D>90%
DD (%)			
Average	-0.45 ± 1.4	-0.45 ± 1.4	-0.01 ± 0.2
Range	-4.32—0.32	-4.32—0.32	-0.33—0.45
$\gamma 2\%/2mm$			
Average	99.0 ± 1.6	99.0 ± 1.6	97.5 ± 3.7
Range	94.8—100.0	94.8—100.0	89.2—100.0
$\gamma 3\%/3mm$			
Average	99.7 ± 0.4	99.7 ± 0.4	99.5 ± 0.9
Range	98.7—100.0	98.7—100.0	97.8—100.0

Table 3.2: Dose comparison averages and range of results
Metrics: relative voxel-wise differences (DD), gamma pass rates 2%/2mm and 3%/3mm for the body, low dose regions D>10% and high dose regions D>90%

Pt. 9	Body	D>10%	D>90%
DD (%)	0.00	0.00	0.09
$\gamma 2\%/2mm$	100.0	100.0	100.0
$\gamma 3\%/3mm$	100.0	100.0	100.0

Table 3.3: DD and gamma pass rates of an exemplary patient

Pt. 35	Body	D>10%	D>90%
DD (%)	-4.32	-4.32	0.00
$\gamma 2\%/2mm$	94.8	94.8	89.2
$\gamma 3\%/3mm$	98.7	98.7	98.0

Table 3.4: DD and gamma pass rates of the anomalous patient

DVH

Differences in the DVH can be read from the x-axis of the graph. As Figure 3.2 shows, DVH differences were barely noticeable in the exemplary patient (pt. 9), where the dashed lines represent the sCT and the continuous lines represent the CT. However, in the anomalous patient, significant differences can be seen, particularly for the chiasm. In Figure 3.3, the DVH of pt. 38 displays negligible differences, except in the left chochlea (CochleaL), where the difference between the dashed sCT curve and the continuous CT curve is significant.

Gamma criteria

The average pass rate for $\gamma 2\%/2mm$ and $\gamma 3\%/3mm$ was close to 100%. There was a slightly lower $\gamma 2\%/2mm$ pass rate in the high dose region (D>90%) on average, as well as a higher SD. For the exemplary patient (pt.9), pass rates for both gamma criteria were at 100%, whereas the anomalous patient (pt. 35) displayed significantly lower pass rates for the $\gamma 2\%/2mm$ criterion, the lowest being 89.2% in the high dose region. The graphical representation of the gamma pass rate for pt. 9 in Figure 3.4 shows the pass rate with increasing dose. The pass rate curves for $\gamma 2\%/2mm$ and $\gamma 3\%/3mm$ (shown in the graph as $Gam2_{sCT}$ and $Gam3_{sCT}$) are superimposed on each other as the pass rate was 100% for both criteria. In Figure 3.5, however, we see that pt. 35 displays a lower $\gamma 2\%/2mm$ pass rate, which remains almost consistent with increasing dose.

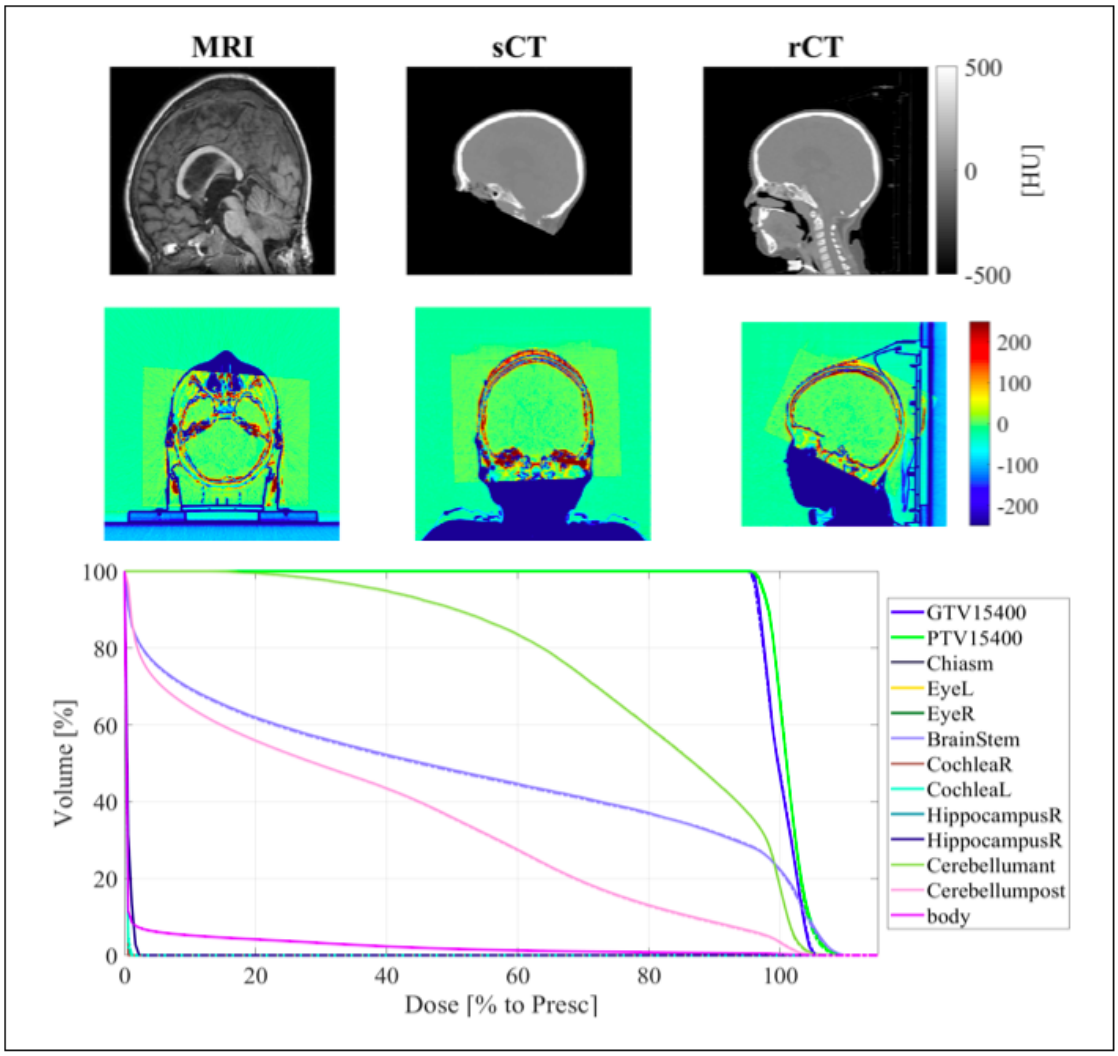

Figure 3.1: Pt. 9: Exemplary patient DVH and images

a) MRI, sCT and real CT (rCT) images. b) CT-sCT image differences c) DVH displaying low organ-based dose differences between CT (continuous line) and sCT (dotted line).

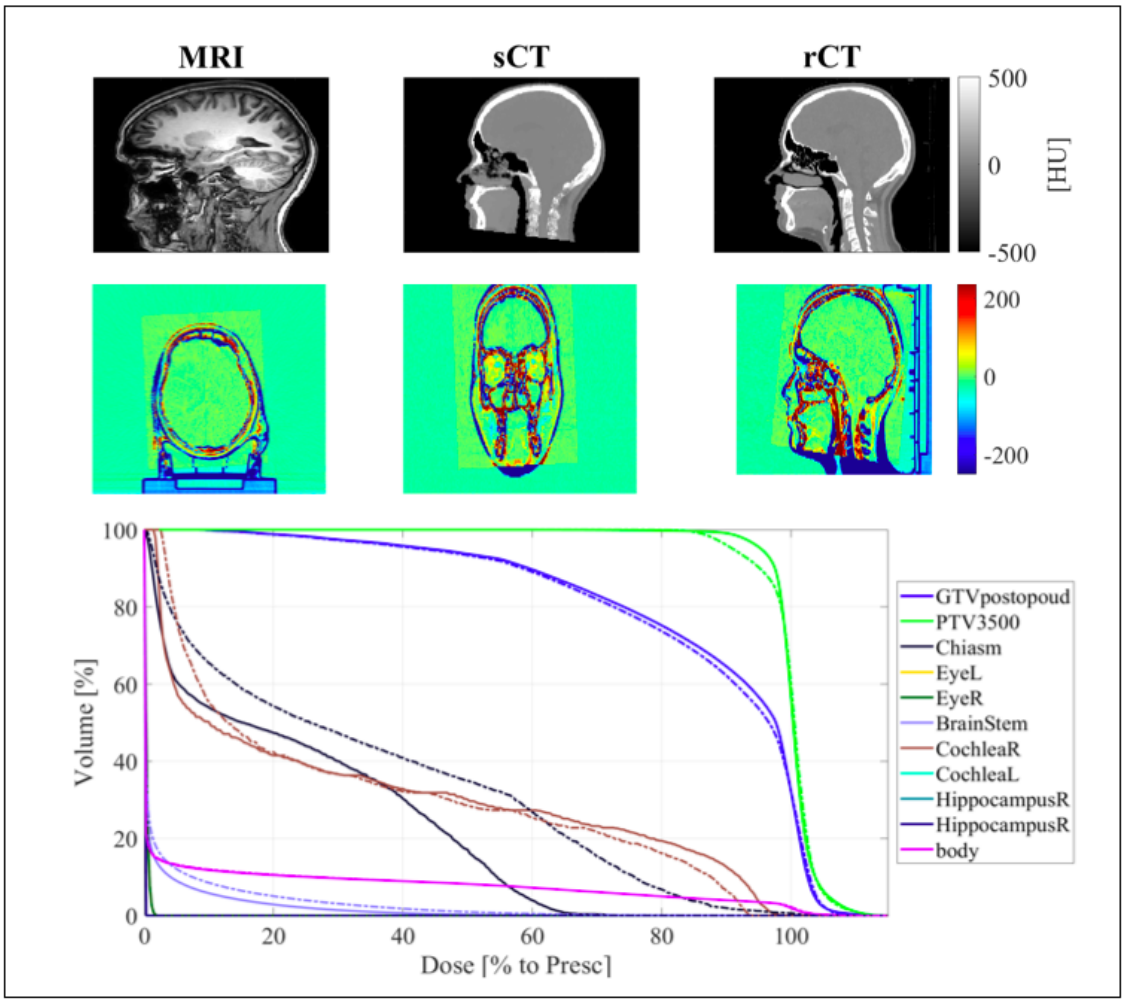

Figure 3.2: Pt 35: Anomalous patient DVH and images

a) MRI, sCT and rCT images. b) CT-sCT image differences c) DVH displays a relatively high dose difference in the Chiasm.

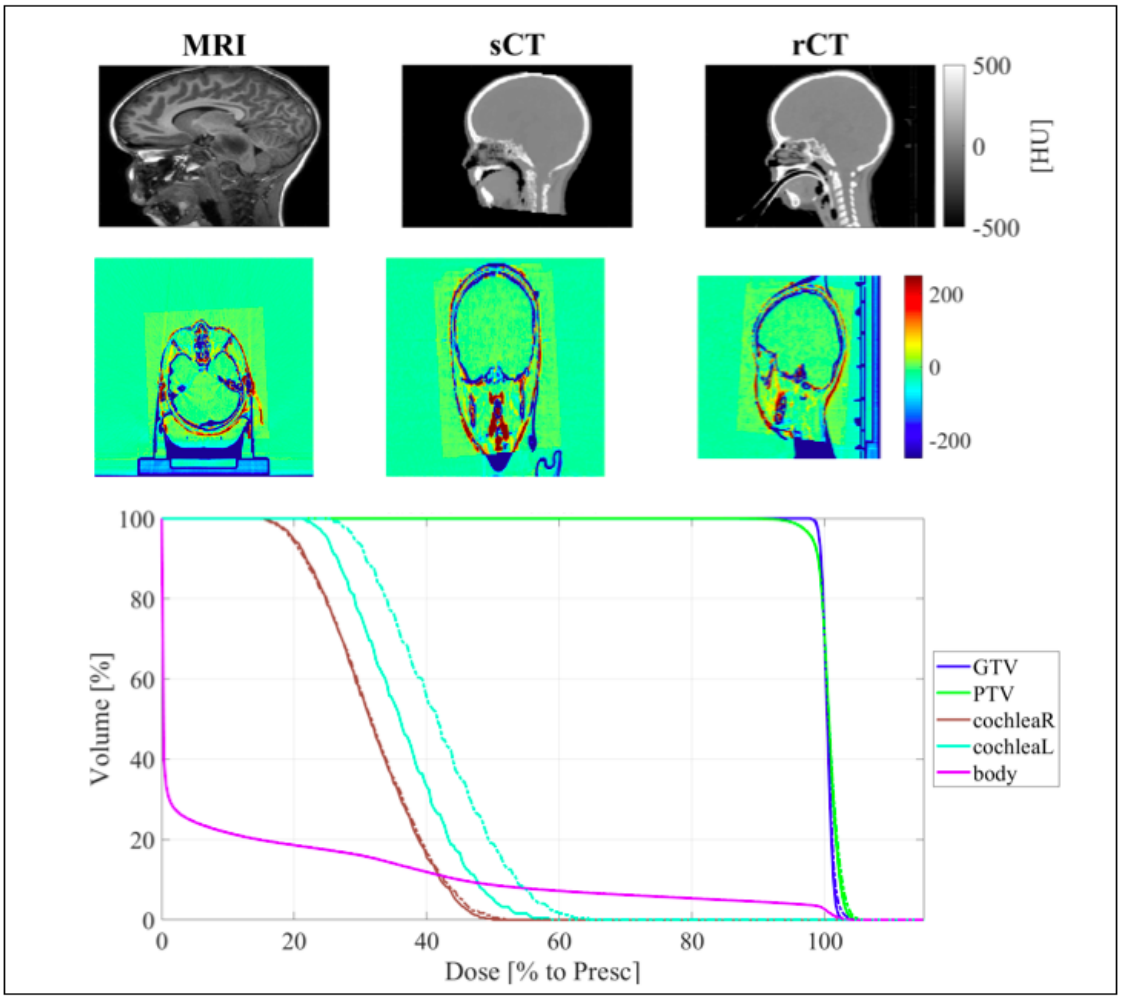

Figure 3.3: Pt 38 DVH and images

a) MRI, sCT and rCT images. b) CT-sCT image differences c) DVH displays a relatively high dose difference in the left chochlea (ChochleaL)

Chapter 4

Discussion

To assess the feasibility of proton planning on deep-learning based sCT, the sCT and CT of a test set of patients were compared, and the differences between them were measured through image-based metrics as well as dose-based metrics. The image based similarity evaluation resulted in an overall mean MAE of 65.7 ± 13.5 HU. This is comparable to similar studies, as seen in Table 4.1. For example, the study by Spadea et al., 2019 resulted in a mean MAE of 52 ± 7.0. It should be noted that differences in evaluation methods, such as the choice of a whole or small FOV, impacts the ability to make direct comparisons with the other studies. Nonetheless, the comparable MAE reveals that HU assignment of these sCT is generally appropriate. DSC on the body and bone are 0.98 ± 0.004 and 0.82 ± 0.08 respectively, these mean results are close to 1, indicating good overlap between the volumes in CT and sCT. DSC scores can also measure the geometric consistency of the body contours. Other studies show similar results (see Table 4.1 for DSC on the body).

The dosimetric evaluation resulted in mean voxel-wise dose differences (DD) of $-0.01 \pm 0.2\%$ for the high dose region ($D > 90\%$) and $-0.45 \pm 1.4\%$ for the body and low-dose region ($D > 10\%$). The latter result indicates a slightly larger voxel-wise dose difference in comparison to other studies in Table 4.1. Another proton study by Liu et al. had a mean DD of -0.07 ± 0.07 while Bentvelzen et al.'s photon study yielded a mean DD of $0.00 \pm 0.3\%$, both indicating very minimal voxel-wise differences. The dose differences in this study can be explained by a few anomalous results, the most prominent being those belonging to pt 35, where the DD was at -4.32% for the body and $D > 10\%$ (Table 3.4). This patient, however, had a 1 year gap between the MR and CT scan undergone in preparation for their treatment. Since the sCT is based on the MR data, the difference between the CT and sCT could be a result of growth, leading to a lack of correspondence. Another contribution is from pt 38, whose DD on the body and $D > 10\%$ and was -0.62%. In terms of organ-based doses, they displayed high dose differences in the left cochlea, as seen in their DVH in Figure 3.3. The cochlea has a small volume and is located in a low dose region, which generally leads to higher uncertainties in dose calculation, as noted by Bentvelzen et al. Furthermore, differences in the resolution of air regions can arise between CT and sCT, leading to differences in dose calculations. Overall, however, the mean DD of the test set was within 1%, which indicates only small local dosimetric differences between sCT and CT voxels.

In terms of organ-based differences, the DVH graphs reveal minimal dose differences to OARs as the sCT and CT curves are generally almost identical, as seen in the DVH of the exemplary patient 9. In addition to pt. 38 mentioned earlier, pt. 35 displays high dose differences in the chiasm, which can be explained by the time between their MR and CT scans. Keeping OARs dose at a minimum is essential to RT planning and the future quality of life of the patient, hence dose calculation on sCT must be accurate. One of the ways this can be improved is by ensuring that the MR data used to generate the sCT is recent enough, in order to match the current organ positions and body contours of the patient.

Researcher	Site	Particle	Test set	MAE (HU)	DSC	DD(%)	Gamma (%)
This study	Brain	Proton	10	65.7±13.5	0.98±0.004	-0.45±1.4	99.0±1.6
[Bentvelzen et al., 2018]	Brain	Photon	12	75.2±15.6	0.98±0.01	0.00±0.3	99.7±0.4
[Emami et al., 2018]	Brain	Photon	15	89.3±10.3	-	-	-
[Dinkla et al., 2018]	Brain	Photon	26	67±11	0.98±0.01	0.00±0.2	95.8±2.1
[Maspero et al., 2018b]	Prostate	Photon	30	65±10	-	0.00±0.1	96.8±2.1
[Kazemifar et al., 2019]	Brain	Photon	14	47.2±11.0	0.96±0.02	-	99.2±0.8
[Spadea et al., 2019]	Brain	Proton	15	52±7.0	-	-	-
[Liu et al., 2019]	Prostate	Proton	17	51.3±16.9	-	-0.07±0.07	98±3.0

Table 4.1: Comparison with the literature

The gamma pass rates for both criteria ($\gamma 2\%/2mm$ and $\gamma 3\%/3mm$) were close to 100%. For $\gamma 2\%/2mm$, the pass rate was 99.0 ± 1.6 on average for the body and $D > 10\%$. This means that approximately 100% of the reference CT points were correspondent with the evaluated points of the sCT, based on distance and dose. The pass rates are comparable to other studies as seen in Table 4.1, where $\gamma 2\%/2mm$ results are reported.

The test set size in the literature in Table 4.1 ranged from 10-30, the smallest being the test set chosen for this evaluation, which is one limitation of this study. Future research could include a larger test sets for more accurate results, similar to that conducted by Maspero et al., 2018 which involved 30 patients. A larger test set would yield more precise mean values for the evalutation metrics, allowing for better identification of anomalies. The clinical applicability of the sCT needs to be further researched. For example, as shown in Figure 1.4, treatment delivery involves position verification, where imaging modalities are used to account for changes in the position of the patient at each fraction. The role of the sCT in the position verification process needs to be further explored. Furthermore, a future study could also involve the evaluation of proton range calculations, as uncertainties arising from this affect the dose distribution in the patient [Shafai-Erfani et al., 2019], although this is generally accounted for at the optimization stage. Another area to investigate would be the possibility of quantitative measures to assess the quality of the sCTs without the need for CT data [Bentvelzen et al., 2018, van Harten et al., 2020].

This study has positive implications on pediatric patients undergoing proton radiotherapy. CT scanning delivers a high radiation dose at the planning stage [Meulepas et al., 2019]. However, patients with brain tumour already receive high doses of radiation for treatment, making the added dose from CT negligible. However, the implementation of sCT could alleviate the discomfort of multiple scanning sessions. It removes the burden of increased doses of anaesthesia, with their accompanying negative side-effects on the health of the patients [Malviya et al., 2000]. Furthermore, uncertainties can arise as a result of the multi-modal approach due to the registration of CT and MR [Ulin et al., 2010]. If significant, this leads to the delivery of added radiation dose to surrounding tissue, reducing the quality of treatment. Dosimetrically accurate sCT can eliminate this by removing the need for image registration. Lastly, this approach could provide a more streamlined workflow, increasing the time and cost efficiency of the treatment. The results reported here move towards a successful MR-only method that would get rid of some of the disadvantages of the multi-modal approach. Overall, they indicate that the cGAN method of sCT generation could be used for an MR-only workflow with brain tumour pediatric patients.

4.1 Conclusion

This study aimed to assesss the dosimetric feasibility of deep-learning generated sCT, for the case of pediatric patients with brain tumour undergoing proton therapy. Its methodology involved the selection of a test set from a sample of patients who have undergone CT and MR scans in preparation for radiotherapy, and had sCTs produced as part of a previous study. To perform an evaluation of the sCT performance, the affinity of sCT to real CT data was measured. This was conducted with image-based similarity metrics; MAE and DSC as well as dosimetric comparisons; DD, DVHs and Gamma criteria. Averaged results revealed an accurate HU assignment for sCT,

comparable to similar studies. the mean DSC indicated overlap of segmentations of the sCT to the CT. The mean voxel-wise differences and DVHs yielded only small differences in local and OAR-based doses, and most points (approaching 100%) on the sCT passed the gamma criteria. This correspondence of sCT to CT indicates that dose calculations for the purpose of proton planning are feasible on these sCT. This proton-based study adds to the bulk of growing research on the MR-only approach to radiotherapy planning. MR-only provides a shorter, more efficient workflow relative to the multi-modal MR-CT approach. It further relieves the burden of added imaging sessions and exposure to anaesthesia, amongst other advantages. With further larger studies and investigations into clinical applicability, deep-learning based sCT could be integrated into clinical settings to reap the benefits of the MR-only approach.

Bibliography

[Alonso and Finn, 1992] Alonso, M. and Finn, E. J. (1992). *Physics*. Addison-Wesley. 3

[Anderson et al., 2015] Anderson, R., Lamey, M., MacPherson, M., and Carlone, M. (2015). Simulation of a medical linear accelerator for teaching purposes. *Journal of Applied Clinical Medical Physics*, 16(3):359–377. 3

[Balachandran et al., 2010] Balachandran, R., Welch, E. B., Dawant, B. M., and Fitzpatrick, J. M. (2010). Effect of MR distortion on targeting for deep-brain stimulation. *IEEE transactions on bio-medical engineering*, 57(7):1729–1735. 9

[Baskar et al., 2014] Baskar, R., Dai, J., Wenlong, N., Yeo, R., and Yeoh, K.-W. (2014). Biological response of cancer cells to radiation treatment. *Frontiers in molecular biosciences*, 1:24. 1

[Beavis et al., 1998] Beavis, A. W., Gibbs, P., Dealey, R. A., and Whitton, V. J. (1998). Radiotherapy treatment planning of brain tumours using MRI alone. *British Journal of Radiology*, 71(MAY):544–548. 8

[Bentvelzen et al., 2018] Bentvelzen, L. G., Maspero, M., Philippens, M., Breeuwer, M., Raaijmakers, A., Bosnacki, D., and Yeo, R. (2018). *Deep learning-based synthetic CT generation for brain MR-only radiotherapy A multi-view approach on heterogeneous MRI of pediatric patients*. PhD thesis. ii, 14, 15, 16, 17, 18, 25

[Bentzen, 2010] Bentzen, S. M. (2010). Fractionation Effects in Clinical Practice. *Leibel and Phillips Textbook of Radiation Oncology*, pages 40–54. 6

[Bielajew, 2007] Bielajew, A. (2007). The Monte Carlo Simulation of Radiation Transport. *Handbook of Radiotherapy Physics: Theory and Practice*, pages 75–87. 11

[Broder, 2011] Broder, J. (2011). *Imaging of Nontraumatic Abdominal Conditions*. Elsevier. 8

[Carolan, 2010] Carolan, M. G. (2010). Pencil Beam Dose Calculation Algorithm. 11

[Cerry and Duxbury, 1998] Cerry, P. and Duxbury, A. (1998). *Practical Radiotherapy: Physics and Equipment*. Greenwich Medical Media. Greenwich Medical Media. 3

[Chandarana et al., 2018] Chandarana, H., Wang, H., Tijssen, R. H., and Das, I. J. (2018). Emerging role of MRI in radiation therapy. *Journal of Magnetic Resonance Imaging*, 48(6):1468–1478. 8

[Dawson and Jaffray, 2007] Dawson, L. A. and Jaffray, D. A. (2007). Advances in Image-Guided Radiation Therapy. *Journal of Clinical Oncology*, 25(8):938–946. 11

[Devic, 2012] Devic, S. (2012). MRI simulation for radiotherapy treatment planning. *Medical Physics*, 39(11):6701–6711. 12

[Dinkla et al., 2019] Dinkla, A. M., Florkow, M. C., Maspero, M., Savenije, M. H., Zijlstra, F., Doornaert, P. A., van Stralen, M., Philippens, M. E., van den Berg, C. A., and Seevinck, P. R. (2019). Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network. *Medical Physics*. 13, 17

[Dinkla et al., 2018] Dinkla, A. M., Wolterink, J. M., Maspero, M., Savenije, M. H., Verhoeff, J. J., Seravalli, E., Išgum, I., Seevinck, P. R., and van den Berg, C. A. (2018). MR-Only Brain Radiation Therapy: Dosimetric Evaluation of Synthetic CTs Generated by a Dilated Convolutional Neural Network. *International Journal of Radiation Oncology Biology Physics*, 102(4):801–812. 25

[Edmund and Nyholm, 2017] Edmund, J. M. and Nyholm, T. (2017). A review of substitute CT generation for MRI-only radiation therapy. *Radiation Oncology*, 12(1):1–15. 12, 17

[Emami et al., 2018] Emami, H., Dong, M., Nejad-Davarani, S. P., and Glide-Hurst, C. K. (2018). Generating synthetic CTs from magnetic resonance images using generative adversarial networks. *Medical Physics*, 45(8):3627–3636. 17, 25

[Grimes et al., 2017] Grimes, D. R., Warren, D. R., and Partridge, M. (2017). An approximate analytical solution of the Bethe equation for charged particles in the radiotherapeutic energy range. *Scientific Reports*, 7(1):9781. 5, 6

[Guerreiro, 2019] Guerreiro, F. (2019). *Towards MRI-guided radiotherapy for pediatric patients with abdominal tumors*. PhD thesis, Utrecht University, The Netherlands. 2

[Hall, 2006] Hall, E. J. (2006). Intensity-Modulated Radiation Therapy, Protons, and the risk fo second cancers. Technical report. 12

[Hubel and Wiesel, 1968] Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1):215–243. 13

[ICRU, 2010] ICRU (2010). International Commission on Radiation Units and Measurements - Report 83. *Journal of the ICRU*, 10(1):1–106. 7

[Jabbari, 2011] Jabbari, K. (2011). Review of fast monte carlo codes for dose calculation in radiation therapy treatment planning. *Journal of medical signals and sensors*, 1(1):73–86. 11

[Jaffray, 2012] Jaffray, D. A. (2012). Image-guided radiotherapy: from current concept to future perspectives. *Nature Reviews Clinical Oncology*, 9(12):688–699. 1

[Jeraj and Robar, 2004] Jeraj, M. and Robar, V. (2004). Multileaf collimator in radiotherapy. *Radiology and Oncology*, 38(3):235–240. 9

[Jia et al., 2012] Jia, X., Pawlicki, T., Murphy, K. T., and Mundt, A. J. (2012). Proton therapy dose calculations on gpu: advances and challenges. *Translational Cancer Research*, 1(3). 11

[Joiner and van der. Kogel, 2009] Joiner, M. and van der. Kogel, A. (2009). *Basic clinical radiobiology*. Hodder Arnold, London. 1

[Kazemifar et al., 2019] Kazemifar, S., McGuire, S., Timmerman, R., Wardak, Z., Nguyen, D., Park, Y., Jiang, S., and Owrangi, A. (2019). MRI-only brain radiotherapy: Assessing the dosimetric accuracy of synthetic CT images generated using a deep learning approach. *Radiotherapy and Oncology*, 136:56–63. 25

[Kim, 2017a] Kim, P. (2017a). Convolutional Neural Network BT - MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence. pages 121–147. Apress, Berkeley, CA. 13

[Kim, 2017b] Kim, P. (2017b). Deep Learning BT - MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence. pages 103–120. Apress, Berkeley, CA. 13

[Knab and Connell, 2007] Knab, B. and Connell, P. P. (2007). Radiotherapy for pediatric brain tumors: When and how. *Expert Review of Anticancer Therapy*, 7(12 SUPPL.):69–77. 2

[Koniarova, 2019] Koniarova, I. (2019). Inter-comparison of phantoms for CT numbers to relative electron density (RED)/physical density calibration and influence to dose calculation in TPS. *Journal of Physics: Conference Series*, 1248(1). 8

[Kooy and Grassberger, 2015] Kooy, H. M. and Grassberger, C. (2015). Intensity modulated proton therapy. *The British Journal of Radiology*, 88(1051):20150195. 10

[Kufe et al., 2003] Kufe, D. W., Holland, J. F., Frei, E., and Society, A. C. (2003). *Cancer Medicine*. Cancer Medicine 6. BC Decker, 6th edition. 2

[Lee et al., 2005] Lee, C. T., Bilton, S. D., Famiglietti, R. M., Riley, B. A., Mahajan, A., Chang, E. L., Maor, M. H., Woo, S. Y., Cox, J. D., and Smith, A. R. (2005). Treatment planning with protons for pediatric retinoblastoma, medulloblastoma, and pelvic sarcoma: How do protons compare with other conformal techniques? *International Journal of Radiation Oncology*Biology*Physics*, 63(2):362–372. 5

[Lievens and Van Den Bogaert, 2005] Lievens, Y. and Van Den Bogaert, W. (2005). Proton beam therapy: Too expensive to become true? *Radiotherapy and Oncology*, 75(2):131–133. 6

[Liu et al., 2019] Liu, Y., Lei, Y., Wang, T., Kayode, O., Tian, S., Liu, T., Patel, P., Curran, W. J., Ren, L., and Yang, X. (2019). MRI-based treatment planning for liver stereotactic body radiotherapy: validation of a deep learning-based synthetic CT generation method. *The British Journal of Radiology*, 92(1100):20190067. 13, 25

[Low et al., 1998] Low, D. A., Harms, W. B., Mutic, S., and Purdy, J. A. (1998). A technique for the quantitative evaluation of dose distributions. *Medical Physics*, 25(5):656–661. 18

[Malviya et al., 2000] Malviya, S., Voepel-Lewis, T., Eldevik, O. P., Rockwell, D. T., Wong, J. H., and Tait, A. R. (2000). Sedation and general anaesthesia in children undergoing MRI and CT: adverse events and outcomes. *British Journal of Anaesthesia*, 84(6):743–748. 12, 25

[Malzkorn, 2019] Malzkorn, B. (2019). Integrated diagnostics of diffuse astrocytic and oligodendroglial tumors. *Pathologe*, 40(April):9–17. 14

[Marcu et al., 2012] Marcu, L., Bezak, E., and Allen, B. (2012). *Biomedical Physics in Radiotherapy for Cancer*. CSIRO Publishing. 4

[Maspero, 2018] Maspero, M. (2018). *MR-only Radiotherapy of prostate cancer*. Phd thesis, Utrecht University, The Netherlands. 7, 13

[Maspero et al., 2020] Maspero, M., Bentvelzen, L. G., Savenije, M. H. F., Serravalli, E., Janssens, G. O., van den Berg, C., and Phlippens, M. E. P. (2020). Deep learning-based synthetic CT generation for paediatric brain MR-only photon and proton radiotherapy. *Radiotherapy and Oncology 2020 (submitted)*. 14

[Maspero et al., 2018a] Maspero, M., Savenije, M. H., Dinkla, A. M., Seevinck, P. R., Intven, M. P., Jurgenliemk-Schulz, I. M., Kerkmeijer, L. G., and Van Den Berg, C. A. (2018a). Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Physics in Medicine and Biology*, 63(18). 15

[Maspero et al., 2018b] Maspero, M., Savenije, M. H. F., Dinkla, A. M., Seevinck, P. R., Intven, M. P. W., Jurgenliemk-Schulz, I. M., Kerkmeijer, L. G. W., and van den Berg, C. A. T. (2018b). Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. 25

[Massoptier and Viard, 2014] Massoptier, L. and Viard, R. (2014). *Innovative imaging to improve radiotherapy treatments*. SUMMER. 6

[McGowan et al., 2013] McGowan, S. E., Burnet, N. G., and Lomax, A. J. (2013). Treatment planning optimisation in proton therapy. *British Journal of Radiology*, 86(1021). 9

[Merchant, 2013] Merchant, T. E. (2013). Clinical controversies: Proton therapy for pediatric tumors. *Seminars in Radiation Oncology*, 23(2):97–108. 6

[Merchant et al., 2008] Merchant, T. E., Hua, C.-h., Shukla, H., Ying, X., Nill, S., and Oelfke, U. (2008). Proton versus photon radiotherapy for common pediatric brain tumors: Comparison of models of dose characteristics and their relationship to cognitive function. *Pediatric Blood & Cancer*, 51(1):110–117. 4, 5, 6

[Meulepas et al., 2019] Meulepas, J. M., Ronckers, C. M., Smets, A. M., Nievelstein, R. A., Gradowska, P., Lee, C., Jahnen, A., Van Straten, M., De Wit, M. C. Y., Zonnenberg, B., Klein, W. M., Merks, J. H., Visser, O., Van Leeuwen, F. E., and Hauptmann, M. (2019). Radiation exposure from pediatric CT scans and subsequent cancer risk in the Netherlands. *Journal of the National Cancer Institute*, 111(3):256–263. 8, 25

[Meyer et al., 2017] Meyer, H. J., Uhlig, J., Schramm, D., Bach, A. G., and Surov, A. (2017). Radiological findings on irradiation-CT in patients with non metastatic breast cancer. *British Journal of Radiology*, 90(1075). 7

[Mitin and Zietman, 2014] Mitin, T. and Zietman, A. L. (2014). Promise and pitfalls of heavy-particles therapy. *Journal of Clinical Oncology*, 32(26):2855–2863. 5, 6

[Mu et al., 2005] Mu, X., Björk-Eriksson, T., Nill, S., Oelfke, U., Johansson, K.-A., Gagliardi, G., Johansson, L., Karlsson, M., and Zackrisson, D. B. (2005). Does electron and proton therapy reduce the risk of radiation induced cancer after spinal irradiation for childhood medulloblastoma? A comparative treatment planning study. *Acta Oncologica*, 44(6):554–562. 5

[Newhauser and Zhang, 2015] Newhauser, W. D. and Zhang, R. (2015). The physics of proton therapy. *Physics in Medicine and Biology*, 60(8):R155–R209. 5, 6

[Nyholm et al., 2009] Nyholm, T., Nyberg, M., Karlsson, M. G., and Karlsson, M. (2009). Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments. *Radiation Oncology*, 4(1). 9, 12

[Owrangi et al., 2018] Owrangi, A. M., Greer, P. B., and Glide-Hurst, C. K. (2018). MRI-only treatment planning: Benefits and challenges. 12

[Paganetti, 2012] Paganetti, H. (2012). Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in Medicine and Biology*, 57(11). 6

[Paganetti and Bortfeld, 2006] Paganetti, H. and Bortfeld, T. (2006). Proton Therapy - New Technologies in Radiation Oncology. pages 345–363. Springer Berlin Heidelberg, Berlin, Heidelberg. 9, 10

[Peeters et al., 2010] Peeters, A., Grutters, J. P. C., Pijls-Johannesma, M., Reimoser, S., De Ruysscher, D., Severens, J. L., Joore, M. A., and Lambin, P. (2010). How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. *Radiotherapy and Oncology*, 95(1):45–53. 6

[Pereira et al., 2014] Pereira, G. C., Traughber, M., and Muzic, R. F. (2014). The role of imaging in radiation therapy planning: Past, present, and future. *BioMed Research International*, 2014(2). 2

[Podgorsak, 2005] Podgorsak, E. B. (2005). Radiation Oncology Physics: A Handbook for Teachers and Students. chapter Chapter 5:, pages 123–160. International Atomic Energy Agency, Montreal, Quebec, Canada. 2, 3, 4, 7, 9

[Pollack et al., 2019] Pollack, I. F., Agnihotri, S., and Broniscer, A. (2019). Childhood brain tumors: Current management, biological insights, and future directions. *Journal of Neurosurgery: Pediatrics*, 23(3):261–273. 1

[Saenz et al., 2018] Saenz, D. L., Astorga, N. R., Kirby, N., Fakhreddine, M., Rasmussen, K., Stathakis, S., and Papanikolaou, N. (2018). A method to predict patient-specific table coordinates for quality assurance in external beam radiation therapy. *Journal of applied clinical medical physics*, 19(5):625–631. 11

[Schuemann et al., 2015] Schuemann, J., Giantsoudi, D., Grassberger, C., Moteabbed, M., Min, C., Dowdell, S., and Paganetti, H. (2015). Dose calculation accuracy in proton therapy. *Radiotherapy and Oncology*, 115:S53–S54. 11

[Shafai-Erfani et al., 2019] Shafai-Erfani, G., Lei, Y., Liu, Y., Wang, Y., Wang, T., Zhong, J., Liu, T., McDonald, M., Curran, W. J., Zhou, J., Shu, H.-K., and Yang, X. (2019). MRI-Based Proton Treatment Planning for Base of Skull Tumors. *International journal of particle therapy*, 6(2):12–25. 25

[Spadea et al., 2019] Spadea, M. F., Pileggi, G., Zaffino, P., Salome, P., Catana, C., Izquierdo-Garcia, D., Amato, F., and Seco, J. (2019). Deep Convolution Neural Network (DCNN) multiplane approach to synthetic CT generation from MR images - application in brain proton therapy. *International Journal of Radiation Oncology*Biology*Physics*. 25

[Subramanian and Ahmad, 2019] Subramanian, S. and Ahmad, T. (2019). Cancer, childhood brain tumors. 1, 2

[Trofimov and Bortfeld, 2003] Trofimov, A. and Bortfeld, T. (2003). Optimization of Beam Parameters and Treatment Planning for Intensity Modulated Proton Therapy. *Technology in Cancer Research & Treatment*, 2(5):437–444. 10

[Tsuboi et al., 2020] Tsuboi, K., Sakae, T., and Gerelchuluun, A. (2020). *Proton Beam Radiotherapy: Physics and Biology*. Springer Singapore. 3, 4

[Ulin et al., 2010] Ulin, K., Urie, M. M., and Cherlow, J. M. (2010). Results of a multi-institutional benchmark test for cranial CT/MR image registration. *International Journal of Radiation Oncology Biology Physics*, 77(5):1584–1589. 12, 25

[van Assendelft et al., 1973] van Assendelft, O. W., Mook, G. A., and Zijlstra, W. G. (1973). International System of Units (SI) in physiology. *Pflügers Archiv European Journal of Physiology*, 339(4):265–272. 1

[van Harten et al., 2020] van Harten, L. D., Wolterink, J. M., Verhoeff, J. J. C., and Išgum, I. (2020). Automatic online quality control of synthetic CTs. In *Proc.SPIE*, volume 11313. 25

[Verellen et al., 2007] Verellen, D., Ridder, M. D., Linthout, N., Tournel, K., Soete, G., and Storme, G. (2007). Innovations in image-guided radiotherapy. *Nature Reviews Cancer*, 7(12):949–960. 7, 11

[Voutilainen, 2016] Voutilainen, A. (2016). Spatial Objectives in Radiation Therapy Treatment Planning. 9, 10, 11

[Weber et al., 2004] Weber, D. C., Trofimov, A. V., Delaney, T. F., and Bortfeld, T. (2004). A treatment planning comparison of intensity modulated photon and proton therapy for paraspinal sarcomas. *International Journal of Radiation Oncology*Biology*Physics*, 58(5):1596–1606. 5

[Willers et al., 2004] Willers, H., Dahm-Daphi, J., and Powell, S. N. (2004). Repair of radiation damage to DNA. *British Journal of Cancer*, 90(7):1297–1301. 1

[Withrow et al., 2019] Withrow, D. R., De Gonzalez, A. B., Lam, C. J., Warren, K. E., and Shiels, M. S. (2019). Trends in pediatric central nervous system tumor incidence in the United States, 1998–2013. *Cancer Epidemiology Biomarkers and Prevention*, 28(3):522–530. 2

[Wolterink et al., 2017] Wolterink, J. M., Dinkla, A. M., Savenije, M. H., Seevinck, P. R., van den Berg, C. A., and Išgum, I. (2017). Deep MR to CT synthesis using unpaired data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10557 LNCS:14–23. 13

[Yang et al., 2018] Yang, J., Veeraraghavan, H., Armato, S. G., Farahani, K., Kirby, J. S., Kalpathy-Kramer, J., van Elmpt, W., Dekker, A., Han, X., Feng, X., Aljabar, P., Oliveira, B., van der Heyden, B., Zamdborg, L., Lam, D., Gooding, M., and Sharp, G. C. (2018). Auto-segmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017. *Medical Physics*, 45(10):4568–4581. 17

[Yock et al., 2014] Yock, T. I., Bhat, S., Szymonifka, J., Yeap, B. Y., Delahaye, J., Donaldson, S. S., MacDonald, S. M., Pulsifer, M. B., Hill, K. S., DeLaney, T. F., Ebb, D., Huang, M., Tarbell, N. J., Fisher, P. G., and Kuhlthau, K. A. (2014). Quality of life outcomes in proton and photon treated pediatric brain tumor survivors. *Radiotherapy and Oncology*, 113(1):89–94. 6

[Yu et al., 2019] Yu, H., Oliver, M., Leszczynski, K., Lee, Y., Karam, I., and Sahgal, A. (2019). Tissue segmentation-based electron density mapping for MR-only radiotherapy treatment planning of brain using conventional T1-weighted MR images. *Journal of Applied Clinical Medical Physics*, 20(8):11–20. 8

[Zou et al., 2004] Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M. C., Kaus, M. R., Haker, S. J., Wells 3rd, W. M., Jolesz, F. A., and Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology*, 11(2):178–189. 17