A physics-oriented review of Proton Beam Therapy and its uncertainties

1. Background and theory of proton beam therapy

1.1 Introduction to proton beam therapy

Proton beam therapy (PBT) is a highly debated form of cancer treatment due to its high costs and uncertain success. The UK Government has committed £250 million for the two NHS PBT centres. [1] The clinical evidence supporting its use, however, remains mixed. [2]

Protons lose energy by electromagnetically interacting with electrons in the orbitals of atoms in their paths. As the energy deposited by protons is inversely proportional to their velocity, they deposit more energy at a specific depth in the tissue (where the tumour would be, if the depth is calculated correctly). This deposition of high energy happens at the "Bragg-Peak". [10 – p. 2]

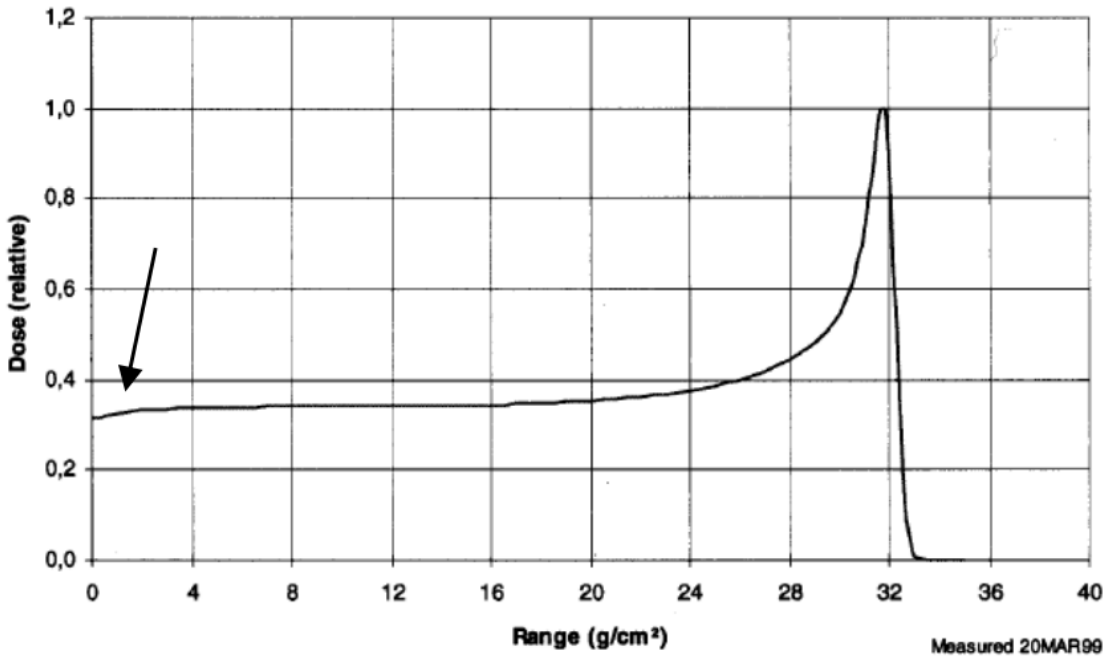

Figure 1 Bragg Peak (Harvard) used from [12]

Theoretically, PBT should benefit from the Bragg Peak characteristics of protons, enabling them to deliver high energy at the end of the beam while minimising the integral dose in the body. Normal tissue damage can, therefore, be reduced by using PBT, making it suitable for several types of cancer. Examples include in children, where PBT has the potential to reduce adverse effects like secondary cancers, cardiac disease and cognitive dysfunction. For skull based and sinonansal malignancies, PBT has helped minimise damage to critical tissues such as brain stem and optical structures when providing the high dose required to kill the tumour. [2]

However, treatment of lung cancer (the most common cancer worldwide [8]) has proven to be challenging due to inhomogeneities* in anatomy in the proton beam path and organ (respiratory) motion leading to uncertainties in the dose distribution within the target volume. Protons are especially sensitive to these variations as they show Bragg Peak characteristics. [5][2][10]

While large scale trials take place to confirm the safety and effectiveness of proton beam therapy, it is important that we tackle these issues of uncertainties so that irrespective of the results of those trials, we have the tools to improve this therapy with high theoretical potential. [2] Tackling proton range uncertainties would not only lead to reduced toxicities, by avoiding higher dose to normal cells due to reduced margins, but also increase effectiveness of the treatment by increasing the dose in the target volume. [2][18][10 p.559] The limit for uncertainty of the dose is defined by the ICRU to be 5%. [23]

There are several sources of uncertainty in PBT, including variations in anatomy affecting the dose distribution in the target volume and RBE values being assumed to be constant (1.1) rather than considering their dependence on proton beam energy and the types of tissues the beam is passing. [2] The uncertainty in conversion of the CT numbers (obtained from CT scans) to Stopping power ratios (SPRs) is one of the major uncertainties to be considered for treatment planning and dose calculation. [20] A major source of uncertainty, especially in analytical algorithms (used for treatment planning), is the range degradation by multiple coulomb scattering. In Monte Carlo algorithms, the mean excitation energies (I-values) are even more significant (according to Paganetti et al., 2012 [18]). Many of these sources are discussed in detail in this paper.

Many solutions have also been proposed and implemented to tackle these uncertainties. For example, proton beams are passed through the directions with least atomic variations, while avoiding high RBE (relative biological effectiveness) to critical tissues such as the spinal cord. One should also avoid beam directions skimming the heart or going through diaphragm (to avoid effects from organ motion) while using more fields which cancel each other's uncertainties. [10 p.558] Alongside, other strategies such as using Monte Carlo algorithms for dose calculation and robust optimisation techniques used are also in Intensity Modulated Proton Therapy to optimise treatment plans in face of all these uncertainties (especially organ motion). [10]

1.2 The physics behind the production of the proton beams

Protons are produced by stripping H^- ions of two electrons by passing it through a thin carbon foil in most proton accelerators. [4] Cyclotrons and synchrotrons are the two most common circular accelerators used to accelerate protons. [5]

*Inhomogeneities refer to changes in density, atomic number etc in the medium. They affect the electromagnetic interactions of the protons with the medium, leading to scattering.

1.2.1 Basic physics principles of cyclotrons and synchrotrons:

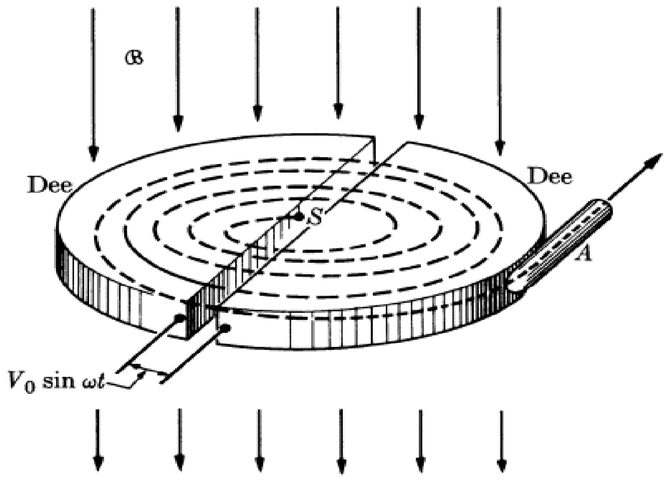

Figure 2: cyclotron basic principle. A constant magnetic flux density B, is applied on the two Dees to keep the particles orbiting in a circle while an alternating potential difference of V_0 sin(ωt) (with ω matching the angular frequency of the particle to maintain resonance) is applied, acting to accelerate the particles between the Dees. [3]

By equating centripetal acceleration and Lorentz force,

$\frac{mv^2}{r} = qvB$ (1)

where m is the mass, v is the velocity and q is charge of the particle while r is the radius of circle traced by it.

Radius of circle described by the particle can be obtained:

$r = \frac{mv}{qB}$ (2)

(r is proportional to v. Hence the radius traced by particle increases after each time the particle passes through the electric field causing it to accelerate and v to increase).

Its angular velocity (v/r) can be obtained as:

$\omega = -\frac{q}{m}B$ (3)

With the negative sign indicating that the angular velocity opposes the B field for a positive charge. It is independent of the radius. [3]

Cyclotrons and synchrotrons

The maximum kinetic energy obtained by the particle in a cyclotron would be $\frac{mv^2}{2} = \frac{(mv)^2}{2m} = \frac{(qBR)^2}{2m}$ (from (1), mv=qBR). In a cyclotron, the energy is limited by relativistic effects as, if m is the relativistic mass of the particle (m= γm_0 where γ is the relativistic constant and m_0 is the mass of the particle in a stationary frame), angular velocity (equation (3)) of the particle would decrease as the relativistic mass increases. [3] Hence, increasing velocity would lead to the particle and oscillating potential no longer being in phase such that acceleration becomes difficult.

Synchrotrons solve this problem by changing the magnetic field or the voltage frequency as the momentum of the particles increases to maintain a constant radius to be traced by the particle. [5] They can accelerate protons to the required energy, which can be extracted and delivered to patients directly through gantries. [5][10]

Individual particles undergo betatron (breaking their circular orbits to wander in vertical and horizontal planes) as well as synchrotron oscillations (of individual out-of-step particles around the synchronous particles) yet the whole bunch remains stable. The full energy beam at the end is extracted using a fast kicker magnet or peeled off slowly using an energy-loss foil. [7]

For a video explanation of the basic physics of cyclotrons, see [9]

1.2.2 Beam spreading systems

We need to spread out the approximately 1 cm proton beam emerging from the accelerator so it can cover the whole tumour. The beam needs to be spread in the lateral direction to cover the cross-sectional area of the tumour as well as in the depth direction. There is a uniform dose distribution such that the overall beam forms a cylinder around the tumour. The lateral spreading is done with the help of passive scattering and magnetic beam scanning while the depth distribution is achieved with the help of the spread out Bragg Peak (SOBP). [10]

SOBP

A modulator wheel can be used to create beams of multiple energies forming a flat region of dose with an SOBP. [5][10]

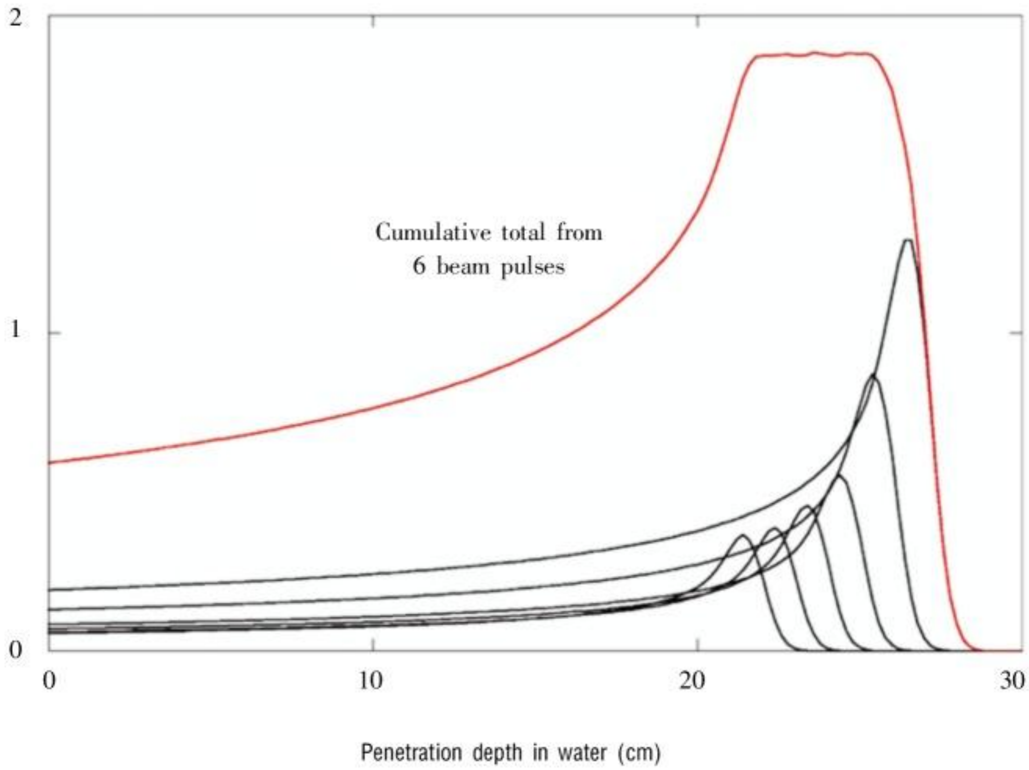

Figure 3 - the Spread-Out Bragg Peak (SOBP). Formed from beam pulses of 6 different energies.

A range compensator is used to conform the beam to the distal end of the tumour (i.e. cover it depth wise). [10 p. 116] The distal peaks have a higher contribution the SOBP but lower skin dose while the proximal peaks have the opposite properties. [10 p.143]

2. Uncertainties in Proton Beam therapy

2.1 Mean Excitation Energy

The mean excitation energy, or I-value describes how easily a molecule or an atom can absorb energy from a projectile (the protons in the beam in this case). This is converted into electronic and vibrational excitation. [6] It is used in the Stoichiometric calibration for converting CT numbers to stopping power ratios as well as in Monte Carlo algorithms when applying the Bethe-Block equation to determine the range of protons (both of which are discussed in the following sections). [13][18] The discrepancies in the I-values of water from about 67 to 80 eV can affect the stopping power ratios (1%) and hence the predicted range of proton beams (of the order ±1.5-2%). [14][18] While in tissue, uncertainty of 10-15% can introduce an uncertainty of 1.4% in SPRs of human tissues relative to water and about 1.5% in the proton range. [14][18]

2.2 Converting CT numbers to stopping power ratios

2.2.1 Stoichiometric Calibration

In PBT planning, CT numbers represent the ability of the tissue to decrease the intensity of proton beams compared to water. SPRs are used to determine the range of protons for treatment planning. [10][14] The method of stoichiometric calibration suggested by Schneider et al is the most common method used for converting CT numbers to Stopping power ratios. [10][13] This is because it found a straight-line calibration curve between these two values (for high CT numbers) by using chemical composition of real tissues instead of tissue substitutes (which had a more varying data).

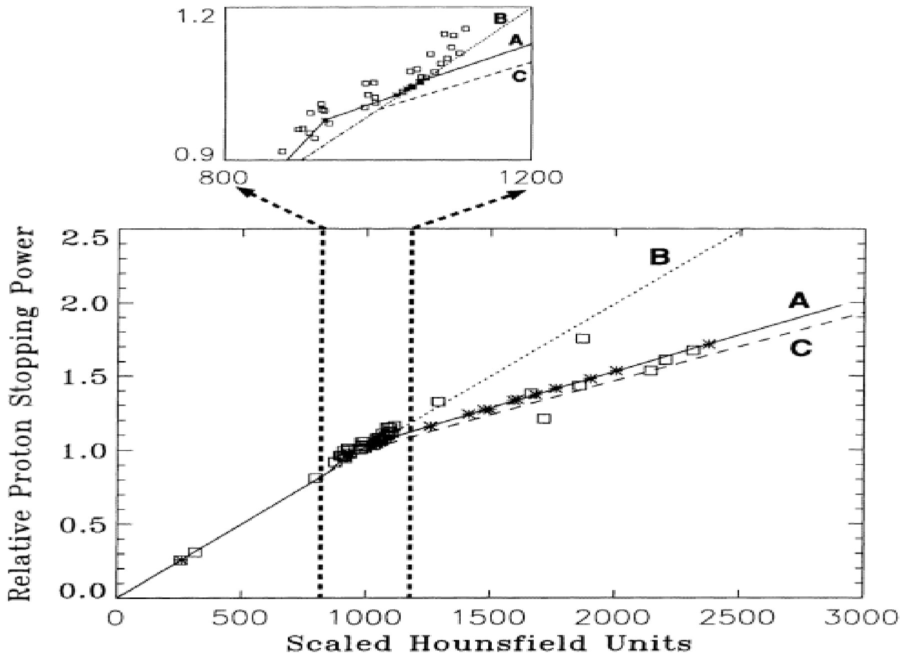

Figure 2. Calibration curves for the transformation of Hounsfield values into relative proton stopping power (ρ_s). The solid line shows the stoichiometric calibration (A) for biological tissues, the dotted line the tissue substitute calibration for Mylar/Melinex/PTFE (B) and the dashed line the tissue substitute calibration for B110/SB5 (C). The squares represent calculations for tissue substitutes and the stars are calculations based on the chemical composition of real tissues. The small plot shows in detail the Hounsfield number range corresponding to soft tissue.

Figure 4: borrowed with caption from [13]

This is the method used:

Tissue substitutes (not necessarily tissue-like materials) of known chemical composition and density are used to find scanner specific parameters like K^{ph} (constant of photoelectric interaction), K^{coh} (constant for coherent scattering), K^{KN} (constant for Compton scattering). [14] Here is a table summarising the 3 phenomena:

Phenomenon	**Interaction of photons with**	**Energy transfer from photon to charged particle during interaction**	**Result of interaction**
Photoelectric effect	Tightly bound electrons	All energy transferred	Electron ejected with: kinetic energy = photon energy – binding energy of electron
Coherent (Rayleigh) scattering	Whole atom	Elastic scattering (no energy transferred)	Photon scattered at a small angle
Compton (incoherent) scattering	Loosely bound electrons	Some energy transferred (amount depends on angle of scatter)	Photon scattered with reduced energy and (Compton) electron backscattered

[21]

CT scans work by getting X-Ray (photon) images of the body from different angles and combining them to form a 3D image. [22] Hence these scanner specific parameters depend on photon interactions with matter. They are calculated by measuring their Hounsfield values (using the CT scanner that will be used for treatment) and then using the following equations:

$$H = 1000\frac{\mu}{\mu_w}$$

where H is the Hounsfield value, μ is the attenuation coefficient and μ_w is the same for water.

$$\mu = \rho\, N_g(Z, A)\{K^{ph}\, \tilde{Z}^{3.62} + K^{coh}\hat{Z}^{1.86} + K^{KN}\}\ (*)$$

where the Z values are the effective atomic numbers, ρ is the density and N_g is the number of electrons per unit volume. [14]

These are used to compute Hounsfield values for ICRP tissues (not substitutes).

Use this (and chemical composition of ICRU tissues) to calculate proton stopping powers:

$$S_{rel} = \rho_{e,rel}\ \times \frac{\ln\left(\frac{2m_e c^2\beta^2}{(1-\beta^2)I_{m,x}} - \beta^2\right)}{\ln\left(\frac{2m_e c^2\beta^2}{(1-\beta^2)I_{m,w}} - \beta^2\right)}$$

Where β is $\frac{speed\ of\ proton}{speed\ of\ light, c}$, $I_{m,w}$ is the mean excitation energy of water and $I_{m,x}$ the same for the medium while $\rho_{e,rel}$ is the relative electron density (between the medium and water) m_e is the mass of an electron.

Then form calibration curve for the tissues. [13] This method is less sensitive to the choice of tissue substitutes and is therefore the most widely used method in proton therapy centres. [14]

2.2.2 Sources of uncertainty

There are several sources of uncertainty in SPR calculation as discussed by Yang et al such as the determination of the scanner specific parameters and imaging uncertainty when scanning tissues for calibration. [14] A comparison of theoretical and measured CT numbers was found to have the most negative percentage difference for lung tissue. [14] Other sources of uncertainty in these calculations are from I-values and the deviation of human tissue composition in individual patients with different age, health etc. from ICRU standards. [14] A simulated population-based study resulted in the overall percentage uncertainty for calculating SPR in individualized bone (1.60%) and soft (1.20%) tissues being much higher than for the reference human tissues (0.29% and 0.43% respectively) showing stoichiometric calibration does not take such variations into account very well. [14] CT imaging (e.g. CT numbers in field of view getting affected by those outside it, such as the patient's couch, patient size and artefacts etc.) and the implementation of dose calculation algorithms (e.g. ignoring the dependence of SPR on proton energy) are further sources of uncertainty. [10][14] The difference between uncertainties in different tissue types showed that it may be more appropriate to use alternatives to the water-equivalent depth approach. [14] The total estimated uncertainty in the calculated SPRs is estimated to be 3.5%. [25]

Dual Energy CT as solution

Dual Energy CT (DECT) allows us to determine the effective atomic number as well as the electron density ratio simultaneously. [25] Since Z values of human tissues were discovered to be directly related to the I-values, this has allowed for computation of SPRs that is affected much less by variations in the tissue composition. [25] This method has an uncertainty of 2.2%. [25]

2.3 Multiple coulomb scattering (MCS)

2.3.1 Moliere theory

Multiple coulomb scattering occurs by electromagnetic collisions of the protons with nuclei causing them to scatter by a few degrees. This is described by a well-tested theory called Moliere theory, and occurs simultaneously with stopping (electromagnetic collisions of the protons with electrons leading to the proton losing energy and eventually stopping, described by the well tested Bethe Block theory). [10] The simplest case of Moliere theory is of proton hitting a target made of a single element of known atomic weight and atomic number. It obtains a distribution function of the scattering angle, $f(\theta)$. This is a power series involving:

- characteristic single scattering angle, χ_c. On average a proton would undergo only one scatter above this angle.
- The characteristic multiple scattering angle, $\theta_M = \frac{1}{\sqrt{2}}(\chi_c\sqrt{B})$, where B is the reduced target thickness obtained from χ_c and χ_a (the small angle at which the scattering cross section** no longer follows the Rutherford law of $\frac{1}{\theta^4}$ due to electrons shielding the nucleus).
- The reduced angle $\theta' \equiv \frac{\theta}{(\chi_c\sqrt{B})}$ which normalises the scattering angle to form a dimensionless value.

**the scattering cross section can be interpreted as the probability of a particle being scattered

These combine to for the distribution function $f(\theta) = \frac{1}{2\pi\theta_M^2}\frac{1}{2}\left[f^{(0)}(\theta') + \frac{f^{(1)}(\theta')}{B} + \frac{f^{(2)}(\theta')}{B^2}\right]$ where $f^{(n)}(\theta') = \frac{1}{n!}\int_0^\infty y\, dy\, J_0(\theta' y) e^{\frac{y^2}{4}}\left(\frac{y^2}{4}\ln\left(\frac{y^2}{4}\right)\right)^n$

See [10] p. 38 to 40 for more details about the formulae. Moliere generalised these equations to arbitrary energy loss for protons (due to "stopping") interacting with compounds and mixtures. [10]

2.3.2 Effects of MCS in PBT

Study and implementation of MCS is necessary in treatment planning, not only to ensure accurate dose calculations (see section 2.3.1), but also to be able to understand other scattering effects some of which are discussed here. [23]

Cyclotrons are only able to produce proton beams of fixed energy. [5] The energy is then lowered using degraders which can cause beam spread. This is due to energy loss and multiple scattering as the beam passes through the degrader material, leading to increase in beam size. Hence magnetic bending systems called collimator system and energy selection system (ESS) are used to limit this energy spread. [10]

Neutrons created by interactions with materials in the collimator can cause background radiation in the patient, contribute to their whole-body dose and possibly increasing incidence of secondary tumours while also wasting up to 90% of proton intensity. [5][10] Only protons with the correct energy are passed into the beam transport system. The energy adjustment can also be done in a nozzle, just before it is delivered to the patient, using a range shifter (a stack of plates), a ridge filter (a plate with specially shaped ripples) or a range modulation wheel (a rotating wheel with azimuthally changing thickness). [10 p. 71]

2.4 Treatment planning methods in face of uncertainties

2.4.1 The use of Monte Carlo in PBT

Background

Analytic algorithms, like the pencil beam algorithm (with a faster broad beam version) developed by Hong et al in 1996 (see [16]) are unable to incorporate the effect of range degradation by MCS in their dose calculations. [18] This is because, in complex structures, modelling MCS in materials using their stopping power and water equivalent depth can cause a shift in the distal falloff of the proton beam. [18] The effect of this is more enhanced in small fields in head and neck and can affect majority of the treatment volume. [18] Hence, pioneering figures like Harald Paganetti have since helped in the development of Monte Carlo algorithms which have led to significant improvement in reducing uncertainty, leading to agreement between the measured and calculated dose which would lead to better treatment plans for the patients. [18]

Development of Monte Carlo algorithm

GEANT4 incorporates MCS based on the Urban theory of scattering which is slightly different from the Moliere theory in that it predicts the net effect of multiple scatterings rather than simulating each individual event. [23][24] Designing the nozzle with a high accuracy is very helpful in achieving agreement between modelled and measured dose distributions. [23] Full list of theories considered in the GEANT4 code to model interactions of particles is available on their website. [24]

Evidence supporting use of Monte Carlo

Several studies highlight the superiority of Monte Carlo (MC) algorithms compared to analytical algorithms (AA). In an inter-institutional study using pencil beam scanning proton beam to irradiate lung phantoms, AAs were shown to dramatically overestimate dose to the target with 7.2% (AA) versus 1.6% (MC) absolute dose agreement on average. [17] The maximum dose difference (between measured and calculated) for AA (30%) was much higher than for MC (12%) in the internal gross target volume (used to minimise effects of organ motion). [17] The irradiation test used gamma analysis with relatively loose criteria. Yet only 1/5 institutions passed with analytical algorithm while 4/5 passed with MC. [17] Other studies such as the one by Scheuman et al. went as far as to consider MC calculations as the gold standard and any deviations of Analytical Dose Calculations (ADCs) from it as signs of failure. [19] However, MC has its own uncertainties and cannot be considered as gold standard. [18]

Uncertainties in the Monte Carlo algorithm

There are several sources of uncertainties in the Monte Carlo algorithm, namely the physics settings and CT to Hounsfield Unit conversion. [18] While there are discrepancies due to varied implementations of the multiple coulomb scattering models in MC, the uncertainties in the mean excitation energy are more significant. [18] It is common for the geometries provided by the manufacturer, including the geometrical position and physics constants of various materials, to be uncertain. [15] It is also difficult to measure beam parameters precisely. [15] Even if they are all correctly known and implemented, lack of knowledge of the arrangement of the treatment head materials as a function of time makes it challenging to model the Bragg Peak as a function of time. [15] If there are tissues with high density or atomic number in the beam path, they can significantly affect the accuracy of the calculations made by MC. [23] Moreover, change in density and atomic number leads to an increase in coulomb scattering. [19] Since the SPR uncertainty in the elements in human tissues (and stoichiometric calibration in hydrogen) are negligible in MC, their contributions are expected to be lower with MC than AA. [18]

Possible solutions and improvements

Manufacturers should make sure the geometrical specifications of the treatment head match the reality. Standardisation of treatment head geometries would be extremely difficult considering different manufacturers and manufacturing times but would be beneficial in reducing uncertainties. Understanding of the Monte Carlo code and its physics settings is necessary for its implementation.

3. Conclusion

In this report, a summary of the theory behind the production of proton beams was provided. The uncertainties associated with the conversion of CT Number to Hounsfield Units (done using stoichiometric calibration) and range degradation due to Multiple Coulomb scattering (explained via Moliere Theory) were discussed in detail. Comparisons between Analytical Algorithms and Monte Carlo (MC) algorithms when computing dose distribution in the target volume were made and evidence in favour of MC Algorithms provided. However, it was important to acknowledge the uncertainties in MC despite them being considered as the gold standard. The concerns regarding uncertainties in the introduction were shown to be true due to their impact on proton range and hence conforming with the target volume. Mean Excitation energies were discovered to have deep roots in the uncertainties in both CT to HU conversion, and MC. However, it was acknowledged that advancements in technology, such as DECT continue to be developed and help tackle these uncertainties. Even though Proton Beam Therapy is still in large scale trials, the work of academic figures such as Dr Harald Paganetti and Dr Ming Yang has lead to significant development in PBT in the last few decades. With continuing research and efforts to improve its cost and availability, it could be a very beneficial source of treatment for many cancer patients.

4. References

[1] NHS, accessed: July 2024, NHS commissioning » Proton beam therapy (england.nhs.uk)

[2] Mohan R., 2022, "A Review of Proton Therapy – Current Status and Future Directions - PMC (nih.gov)"

[3] Alonso M., Finn E.J., 1992, "Physics", 581-583, 499

[4] Berz M., Makino K., Wan W., 2021, "An introduction to Beam Physics",

[5] Liu, H. et al, 2011, "Proton therapy in clinical practice - PMC (nih.gov)"

[6] Sauer, S., 2015, "Chapter Three - The Mean Excitation Energy of Atomic Ions" https://www.sciencedirect.com/science/article/abs/pii/S0065327615000027#:~:text=The%20mean%20excitation%20energy%20describes,vibrational%20(including%20fragmentation)%20excitation.

[7] Perkins, D.H., 1972, "Introduction to High Energy Physics", 4th edition

[8] World Cancer Research Fund, accessed: July 2024, Lung cancer statistics | World Cancer Research Fund International (wcrf.org)

[9] Cyclotron & Synchrotron (Particle Accelerator) (youtube.com), accessed: July 2024

[10] Paganetti, H, 2019, "Proton Therapy Physics", second edition

[11] Cancer, C.J., 2011, "Cumulative total from 6 beam pulses - PMC"

[12] Harvard University, accessed: July 2024, Slide 1 (harvard.edu)

[13] Schneider, U et al., 1996, "The calibration of CT Hounsfield units for radiotherapy treatment planning" https://iopscience.iop.org/article/10.1088/0031-9155/41/1/009/pdf

[14] Yang, M. et al., 2012, "Comprehensive analysis of proton range uncertainties related to patient stopping-power ratio estimation using the stoichiometric calibration" https://iopscience.iop.org/article/10.1088/0031-9155/57/13/4095/pdf

[15] Bednarz B. et al, 2011, "Uncertainties and correction methods when modeling passive scattering proton therapy treatment heads with Monte Carlo"

[16] Hong, L et al, 1996, "A pencil beam algorithm for proton dose calculations", pdf (iop.org)

[17] Pencil Beam Algorithms Are Unsuitable for Proton Dose Calculations in Lung Paige A., 2017, Taylor, MS, Stephen F. Kry, PhD, and David S. Followill, PhD, https://www.sciencedirect.com/science/article/pii/S036030161731012X

[18] Paganetti H., 2012, Range uncertainties in proton therapy and the role of Monte Carlo simulations - PMC (nih.gov)

[19] Jan Scheumann et al, 2015, "Assessing the Clinical Impact of Approximations in Analytical Dose Calculations for Proton Therapy" https://www.sciencedirect.com/science/article/pii/S036030161500379X?casa_token=7Kfy_6r2CoQAAAAA:fisqlig1exVZILiwcvsu_v-jjgg1zqpmoFQXyTLflo8WJgskrUYzQ34ntGkUNUitpb98jVOt-_I

[20] Park H. et al, 2021, "Monte Carlo methods for device simulations in radiation therapy (iop.org)"

[21] E.B. Podgorsak, 2005, "Radiation Oncology Physics: A Handbook for Teachers and Students"

[22] National Institute of Biomedical Imaging and Bioengineering, https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct#:~:text=The%20term%20%E2%80%9Ccomputed%20tomography%2C%E2%80%9D,images%2C%20or%20%E2%80%9Cslices.%E2%80%9D

[23] Paganetti, H., 2004, "Accurate Monte Carlo simulations for nozzle design, commissioning and quality assurance for a proton radiation therapy facility", https://aapm.onlinelibrary.wiley.com/doi/epdf/10.1118/1.1762792

[24] Geant4 Collaboration, 1999-2024, accessed: December 2024, Multiple Scattering — Physics Reference Manual 11.2 documentation

[25] B Li *et al,* 2017, "Comprehensive analysis of proton range uncertainties related to stopping-power-ratio estimation using dual-energy CT imaging" https://iopscience.iop.org/article/10.1088/1361-6560/aa7dc9