Diss. ETH No. 12474

Range precision of therapeutic proton beams

A dissertation submitted to the

EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE ZÜRICH

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

for the degree of

Doktor der Naturwissenschaften

Doctor of Natural Sciences

presented by

Barbara Schaffner

Dipl. Phys. ETH

born July 28, 1968
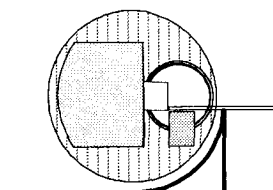
from Hausen b. Brugg, AG

accepted on the recommendation of

Prof. Dr. Peter Niederer examiner

Prof. Dr. Peter Rüegsegger co-examiner

Dr. Eros Pedroni co-examiner

1997

This work was performed at PSI in the Division of Radiation Medicine.

PAUL SCHERRER INSTITUT

Contents

Abstract

Radiotherapy is one of the most important therapeutic methods used against cancer. Proton beams - compared to the conventionally used photon and electron beams - offer a favourable depth dose curve with a pronounced peak at the end of the proton range, the so-called Bragg peak. High dose can be delivered to the target volume with relatively low dose being deposited proximal to the target and no dose distally. Due to the high dose gradients at the end of the protons range, a considerable sparing of healthy tissue can be achieved. This allows the delivery of higher doses to the target and less dose to the surrounding tissue than it is possible with conventional techniques. Thus better tumor control rates and/or less side effects are expected. The full potential of the Bragg peak can however only be used if the range of the protons and therefore the position of the peak is known very accurately. Additionally, density differences along the beam path can cause a spreading of the range spectrum, which results in a degradation of the shape of the Bragg peak. Therefore this thesis deals with the range precision of therapeutic proton beams.

The first part deals with methods and precision of determining the relative stopping power of the tissues in the patients body, which define the range of a proton beam of known energy. A theoretical method - based on literature data of the elemental composition of biological tissue - is presented to calibrate conventional X-ray CT images of patients into relative proton stopping power values. The validity of this calibration curve is confirmed with measurements on animal tissue samples. The uncertainty in the range of protons due to inaccurate calibration of CT images and beam hardening artefacts is estimated to be $\pm 1.1\%$ of the total range in soft tissue and $\pm 1.8\%$ in cortical bone. In practice one will account for this uncertainty with a safety margin added to the target volume.

In the second part the effects of multiple Coulomb scattering on the proton ranges and angular components of the phase space of the beam in the presence of inhomogeneities in the patient are discussed. These effects cause a dilution of the ranges in an initially monoenergetic beam. A simple and quick analytical calculation of the range dilution in a heterogeneous medium has been developed. Different analytical dose calculation algorithms have been written and their accuracy has been evaluated. In most cases an elemental pencil beam dose calculation has been found to be most accurate. The PSI spot-scanning system requires however a fast spot by spot dose calculation to be used for the iterative optimisation procedure of the spot intensities. Such an algorithm has been developed and found to be accurate enough in most situations.
By a comparison to Monte Carlo simulated dose distributions it is shown that the accuracy of all analytical algorithms is lowest in the areas where a high range dilution is obtained. Similarly, high dose errors due to random set-up errors of the patient during a fractionated treatment appear in the same areas. Thus it is crucial for a precise treatment to choose beam incidence angles, which avoid critical areas. The calculation of the range dilution and their display in the beams-eye-view is an important tool which helps identifying optimal beam angle(s).

Concluding, a strategy is presented to deal with uncertainties in the range of proton beams for therapeutic applications. The precise knowledge of stopping power properties of the tissue and of the effects of multiple Coulomb scattering on the dose distribution in

the presence of tissue heterogeneities and set-up errors allows the planner to take suitable actions such as the appropriate choice of safety margin and gantry angle.

If the irradiation can be performed such that no high range dilution is found in the target, the total uncertainty is dominated by the uncertainty in the CT-calibration distally of the target and laterally by set-up uncertainties. These errors can be easily dealt with by a safety margin.

In the case where a high range dilution in the target cannot be avoided, Monte Carlo simulations of the situation with different calibration curves and including set-up errors can be performed. Appropriate consequences can then be taken if necessary.

Following this procedure, an adequate coverage of the target and a maximal sparing of neighbouring tissue can be achieved. This is crucial for a successful treatment of individual patients and for proton therapy as a treatment modality.

Zusammenfassung

Radiotherapie ist eine der wichtigsten Therapien im Kampf gegen den Krebs. Protonen - im Gegensatz zu den konventionell gebräuchlichen Photonen und Elektronen - bieten eine vorteilhafte Tiefendosiskurve mit einem hohen Dosispeak am Ende der Reichweite, dem sogenannten Bragg Peak. Dadurch kann eine hohe Dosis im Target appliziert werden, während die Dosis proximal des Targets tief gehalten und die Dosis distal ganz vermieden werden kann. Durch diesen starken Dosisgradienten am Ende der Protonenreichweite kann das gesunde Gewebe beträchtlich geschont werden. Dadurch können höhere Dosen auf das Target und tiefere Dosen in der Umgebung appliziert werden als dies mit konventionellen Techniken möglich ist. Höhere Tumorkontrollraten und/oder weniger Nebenwirkungen werden erwartet. Um die ganzen Möglichkeiten der Bragg Kurve ausnützen zu können, muss die Reichweite der Protonen - also die Position der Spitze der Bragg Kurve - genau bekannt sein. Ein zusätzliches Problem entsteht durch Dichteunterschiede entlang des Strahls. Dadurch kann das Reichweitenspektrum der Protonen verbreitert werden, was zu einer Degradation der Bragg Kurve führt. Aus diesen Gründen handelt diese Dissertation von der Reichweitenpräzision von therapeutischer Protonenstrahlung.

Der erste Teil behandelt Methoden und Präzision der Bestimmung von relativen Stopping Power Werten im Patienten. Diese definieren die Reichweite eines Protonenstrahls bekannter Energie. Eine theoretische Methode - basierend auf Literaturwerten für die Zusammensetzung von biologischem Gewebe - zur Konvertierung von konventionellen Röntgen-CT Bildern von Patienten in relative Stopping Power Werte wird vorgestellt. Die Richtigkeit dieser Kalibrationskurve wurde durch Messungen an tierischen Gewebeproben bestätigt. Die Unsicherheit in der Protonenreichweite wegen Ungenauigkeiten der Kalibrationskurve und Strahlhärtungsartefakten in den CT Bildern wird auf $\pm 1.1\%$ der Reichweite in Weichteilen und $\pm 1.8\%$ in kompakten Knochen geschätzt. Für diese Unsicherheit wird in der Praxis eine Sicherheitsmarge zum Tumorvolumen addiert.

Im zweiten Teil werden die Auswirkungen der Coulomb Vielfachstreuung und der Winkelunsicherheit des Phasenraumes des Protonenstrahls auf die Dosisverteilung in der Umgebung von Inhomogenitäten diskutiert. Diese Einflüsse bewirken eine Unsicherheit in der Reichweite von Protonen mit derselben Anfangsenergie. Eine einfache und schnelle Berechnung der Verschmierung von Protonenreichweiten im heterogenen Medium wurde entwickelt. Verschiedene analytische Dosisberechnungsmethoden wurden entworfen und auf ihre Genauigkeit untersucht. Meistens erwies sich ein pencil-beam Algorithmus als der genauste. Für das PSI spot-scanning System wird jedoch ein schneller Algorithmus benötigt, der auf einer individuelle Berechnung aller applizierten Dosispunkte beruht. Damit wird iterativ die Gewichtung der einzelnen Spots in einer Optimierungsroutine berechnet. Ein solcher Algorithmus wurde gefunden und ist in den meisten Fällen genau genug.

Durch Vergleiche von Dosisverteilungen berechnet mit Monte Carlo Techniken zu den Resultaten der analytischen Rechnung wurde gezeigt, dass die Genauigkeit aller analytischen Algorithmen in denjenigen Bereichen am tiefsten ist, für die auch eine hohe Reichweitenverschmierung errechnet wird. Ebenso ergeben sich in denselben Gebieten grosse Dosisfehler aufgrund von zufälligen Positionierungsfehlern des Patienten während einer fraktionierten Behandlung. Für eine präzise Behandlung ist es deshalb äusserst

wichtig, den Eintrittswinkel des Strahls so zu wählen, dass kritische Gebiete vermieden werden. Die Berechnung der Reichweitenverschmierung und deren Darstellung in der beams-eye-view ist ein wichtiges Instrument für die Identifikation von optimalen Einstrahlrichtungen.

Zusammenfassend wird eine Strategie präsentiert, wie mit dem Problem der Reichweitenunsicherheit umgegangen werden soll. Der Planer benötigt ein genaues Wissen über das Bremsvermögen des Gewebes und die Effekte, die die Vielfachstreuung im inhomogenen Gewebe auf die Dosisverteilung hat. Dieses Wissen erlaubt ihm angepasste Massnahmen zu treffen, wie die Wahl eines geeigneten Sicherheitssaumes oder Gantry Winkels.

Wenn die Bestrahlungswinkel so gewählt werden können, dass keine hohen Reichweitenverschmierungen im Targetbereich auftreten, dann wird die totale Unsicherheit im distalen Bereich des Targets dominiert durch die Unsicherheit in der CT-Kalibration. Diese Fehler können einfach durch einen Sicherheitssaum kompensiert werden.

Im Fall, wo eine grosse Reichweitenverschmierung nicht vermieden werden kann, können Monte Carlo Simulationen durchgeführt werden, die auch unterschiedliche Kalibrationskurven und Positionierungsfehler berücksichtigen. Falls nötig können angemessene Massnahmen ergriffen werden.

Unter Beachtung dieser Abläufe, kann eine gute Abdeckung des Tumorvolumens bei gleichzeitiger maximalen Schonung des umliegenden Gewebes erreicht werden. Dies ist entscheidend für die erfolgreiche Behandlung von individuellen Patienten, aber auch für die Protonentherapie als Behandlungsmethode.

1 Introduction

1.1 Cancer and radiation therapy

Cancer is a major cause of death in industrialised societies and the incidence of cancer is expected to rise further with longer life expectancies. One out of three people in Europe will suffer from a malignant tumor during his/her lifetime. Only 45 % of these can be cured at present, where cure is defined as survival for five years without further symptoms. Therefore 18 % of the whole population will die with a malignant tumor. This high percentage emphasises the importance of improving treatment methods.

Surgery is the oldest and up to now most successful treatment of localised tumors. 22 % of all patients are cured by surgery alone.

The discovery of X-rays in 1895 by Röntgen and radioactivity in 1896 by Bequerel ('uranic rays') and in 1987 by the Curies ('radium rays') started intensive investigations in the properties and possible applications of the new rays. The effect of radiation on the human tissue was quickly discovered and applied to treat superficial tumors. Technical developments in the 1940s made it possible to produce new radioactive nuclides emitting high energy photons with sufficient intensities to be used for the treatment of deep seated tumors. ^{60}Co became available, which is the most commonly used isotope for external beam radiotherapy and is still used in many places today. Another major development during that time was the linear accelerator for electrons, which has become the most important radiotherapy source today.

At present, 12 % of all cancer patients are cured by radiotherapy alone and 6 % by it in combination with surgery. Latest developments in the field - multi-leaf-collimators, computer controlled intensity modulated beam delivery systems and the use of heavy particle beams - show a good potential for improvement of the cure rates.

Systemic therapies, like chemotherapy and radio-immunotherapy, are also expected to improve further. Their application as unique therapy is however limited to systemic diseases or very early stages. Alone or in combination with other methods they contribute the remaining 5 % of cured patients.

1.2 The biological effect of radiation

The biological targets of a radiation treatment are the tumor cells and especially their DNA. Ionising radiation may produce single or double strand breaks. This can happen either through direct ionisation of the DNA or indirectly through the production of highly aggressive OH° - radicals. Repair mechanisms of the cell are usually capable of repairing single strand breaks. It is therefore necessary to induce double strand breaks to cause a lethal chromosomal damage.

The amount of radiation deposited in the tissue is measured as absorbed energy per mass, or dose. Through the necessity of causing double strand breaks, the biological effect of the radiation on the tissue depends not only on the dose, but also on the quality of the radiation and also on specific properties of the tissue. The closer together ionisation events occur along a single photon or particle track, the higher is the probability of hitting the same DNA twice. An immediate double strand break can almost never be repaired correctly. Densely ionising radiation like neutrons and heavy ions is therefore very efficient. The relative biological effectiveness (RBE) of those radiation types with

so-called high LET (linear energy transfer) is therefore a few times higher than that caused by the same dose deposited by low LET radiation.

Damage caused by low LET irradiation like the commonly used photons and electrons, are usually single strand breaks. Multiple hits of the same DNA helix are therefore necessary to induce a lethal damage. If doses are low, the probability is high that only single strand breaks occur, which can be repaired by the cell.

The relation between dose and cell damage is subject to extensive experiments with cell cultures. The results are usually displayed on a semi-logarithmic scale as in Fig. 1 (fraction of cells surviving the irradiation vs. dose) and mathematically described by the linear-quadratic model or α–β-model. α is the parameter used to refer to the probability that the interacting lesions originate from a single track. α is responsible for the linear component of the cell survival curve (single hit, irreparable damage). β refers to the probability of interacting lesions arising from two independent tracks. β is responsible for the increasing curvature of the shoulder (multi-hit, reparable damage).
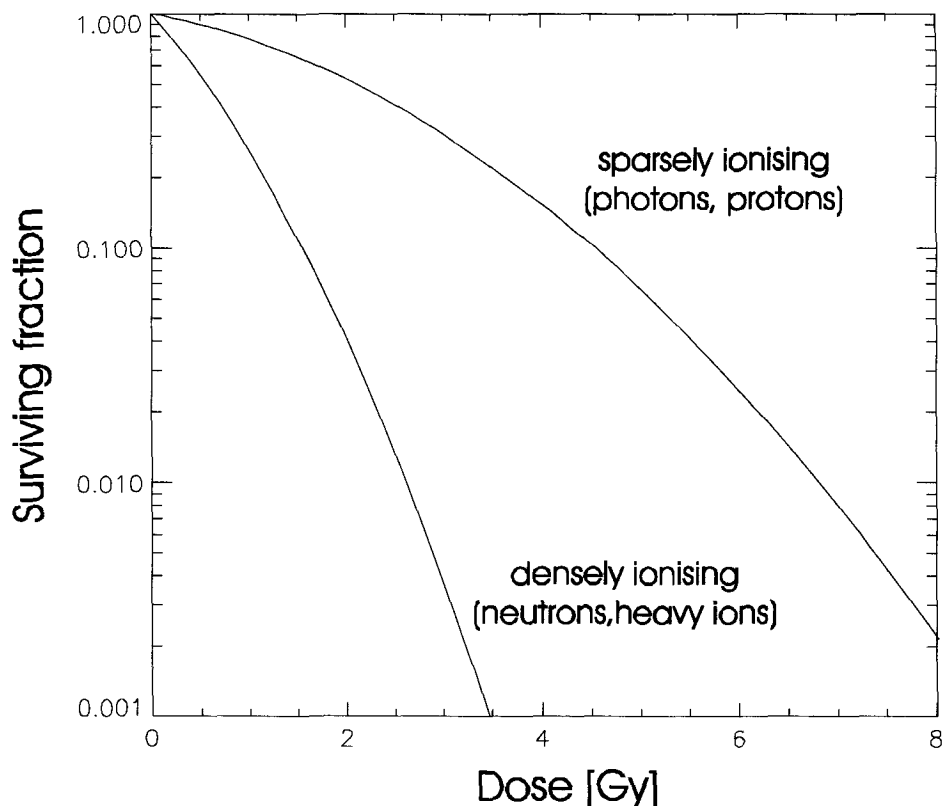
Fig. 1: Typical cell survival curves for cells irradiated with sparsely and densely ionising radiation respectively. Densely ionising particles cause a more severe damage to the cells, which can only be repaired to a very limited degree.

Fig. 1 shows two examples of cell survival curves. One for low LET irradiation (sparsely ionising) and one for high LET radiation (densely ionising) of the same cell line. The large shoulder of the photon/proton curve indicates that a lot of repair occurs in the cells after the irradiation. Almost no repair is possible in the cells irradiated with the same dose delivered by neutrons.

Repair is certainly not desired in tumor cells, but it is crucial to avoid relevant damage to normal tissue as much as possible. A fractionation of the whole treatment dose into small daily fractions allows all cells to recover. Small differences in the dose given to the tumor

and healthy tissue respectively or small differences in the radiosensitivity of different tissues are amplified by this procedure.

The fractionation of the dose is one way to increase the therapeutic ratio defined as: 'the ratio of the probability of eradicating the tumor within the irradiated volume to the probability of causing severe late damage to normal tissue'. Except for a few specialised treatments radiation therapy is always delivered as a fractionated treatment.

The shape of the cell survival curve is therefore a very important factor when comparing different radiation qualities. Protons are low LET radiation. Therefore the biological effect of protons is similar to photon and electron beams. Although recent experiments suggest a higher RBE of very low energy protons, a factor of 1.1 is commonly used for all therapeutic proton energies (Slater et al 1988a, Austin-Seymour et al 1985, Munzenrider et al 1985 and 1989, Habrand et al 1989).

1.3 Physical properties of photon and proton beams

Biological effects of photon and proton beams are similar, their physical properties however are very different. Soon after the production and acceleration of isolated protons became possible, R. Wilson (1946) realised the potential of their characteristic depth-dose curve (Bragg curve) for therapeutic applications.

The energetic proton loses its energy progressively along its track, with most of the energy deposited near to where the particle comes to rest. This high energy deposited at the stopping point forms the Bragg peak in the depth dose curve. The position of the Bragg peak is well defined by the initial energy of the protons. With a superposition of different energy protons a homogeneous dose distribution - the so-called spread-out Bragg peak (SOBP) - can be achieved in any depth of the body. At the same time healthy tissue proximal and distal of the target region can be spared partly and completely respectively. The depth dose curves of a single - and a spread-out Bragg peak are shown in Fig. 2.
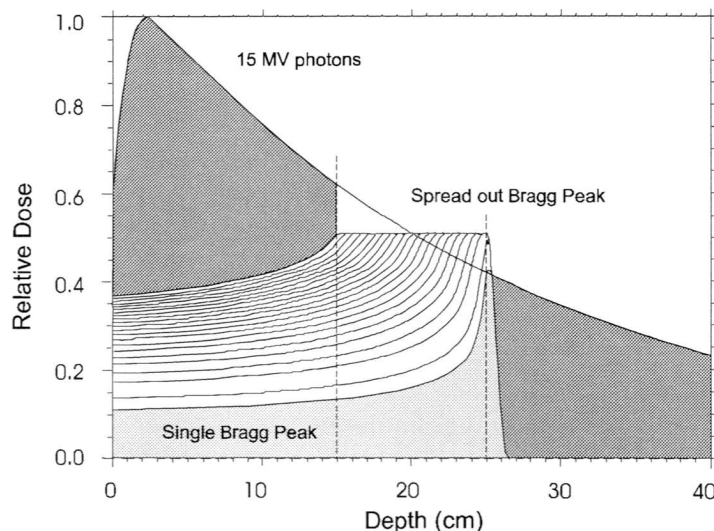

Fig. 2: Bragg peak and SOBP of protons compared to the depth dose curve of a 15MV photon beam. The dark grey area indicates the amount of healthy tissue, which can be spared by using protons instead of photons for a tumor lying between the two dashed lines.

In the same graph the depth dose curve of a 15 MV photon beam is shown. The dose deposited by photons basically follows an exponential decrease with depth. The dominant reaction of therapeutic photon beams with body tissue is the Compton effect. Dose is deposited by recoil electrons, which are ejected in the more forward direction and with increasing kinetic energy with higher energy photons. This causes a small low dose area - the build-up area - directly below the surface of the patient, where no equilibrium between ejected and absorbed electrons is reached yet. The resulting sparing of the skin is an advantage of photons, which cannot be obtained with protons. By using multiple beam directions with overlapping dose depositions in the target, the photon dose outside the target can be lowered below the target dose. Yet, the integral dose to healthy tissue is always higher when using photons than protons.

One expects therefore a gain in the therapeutic ratio by using protons. This can either be a decrease in the probability of severe side reactions when the tumor dose is kept constant, or an increase in the tumor control probability because higher doses can be delivered to the target.

1.4 The history and role of protons in radiotherapy

Only a few years after R. Wilson published his paper "Radiological use of fast protons" (1946), the first patient was treated with protons in 1954 at the Lawrence Berkely Laboratory in California. The high technical effort necessary for such a treatment and the success of the cobalt unit at around the same time were reasons for not investing more in research on proton therapy. Only a limited number of treatments were carried out in physics research institutes, where facilities could be added to existing accelerators.

Tumors at the back of the eye were one of the first indications for proton therapy. Eye treatments require only proton energies up to 70 MeV and can now be delivered at different centres throughout the world. Today it is an established treatment and external photon beam therapy is usually not considered to be an alternative due to the vicinity of critical structures like the optic nerve, the macula and the lens. A large percentage of the total number of proton therapy patients up to now suffered from are eye diseases. At the OPTIS facility at PSI more than 2000 patients have been treated up to now.

The second site, where a large increase of the therapeutic ratio could be documented by using protons were chordomas and chondrosarcomas in the head (Suit et al 1982 and 1990). The high spatial precision in applying the radiation close to many radiosensitive organs was the main reason for this success. These results were obtained by a collaboration between the Harvard Cyclotron Laboratory and the Massachusetts General Hospital. Since 1961 they treated a large number of patients with proton energies up to 160 MeV.

Prostate and gynaecological tumors are other sites where the sharp delineation achievable with protons is expected to improve treatment outcome considerably (Slater et al 1992a, Smit 1992). Lying in the centre of the pelvis, they too are surrounded by critical structures like the bladder, the rectum and the femoral heads. Energies around 200 MeV are necessary to treat them. One of the first centres having these energies available was the Loma Linda University Medical Centre in California. It was also the first commercial proton therapy centre and the first facility to use a gantry - a rotating isocentric beam delivery system - and a dedicated accelerator. Operation was started in 1991 (Slater et al 1988b and 1992b). Preliminary results from a study on combined

photon-proton treatments of prostates showed a good tumor response and low side effects despite the use of high target doses (Yonemoto et al 1997).

The documented success of protons in these cases caused an increasing interest in proton therapy, which is demonstrated by an increasing number of new facilities built in the last two decades and planned for the near future (Fig. 3).
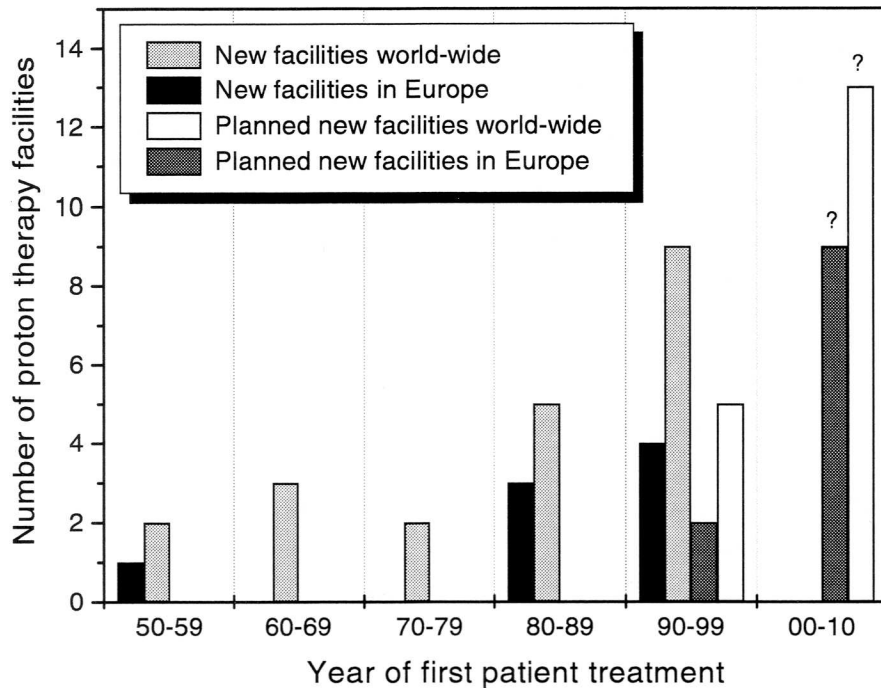

Fig. 3: Overview of the number of proton treatment facilities starting operation in the last 50 years and being planned for the future (Sisterson 1997). There is a continuos increase in the interest in the therapeutic application of protons.

It is now the goal and the task of people working in proton therapy to find the tumor sites which are expected to profit most from protons. It must be shown that the theoretical advantage of the proton depth dose curve leads to clinically significant differences in treatment outcome to justify the higher costs of proton treatments.

State-of-the-art photon treatments using inverse planning techniques and intensity modulated beam delivery systems can achieve a very good conformation of the dose to the target. They can, however, not overcome the limitations given by the physical properties of the photon beam, which do not allow to reduce the dose outside the target any further. In proton therapy on the other hand there are still large technical improvements possible until the intrinsic limitations of the beam are reached.

1.5 Proton therapy at PSI

A major step forward has been realised with the PSI spot-scanning system for deep-seated tumors (Pedroni et al 1995). The spot-scanning technique uses the superposition of some hundred to ten thousands of individual thin pencil beams of protons to impose a

uniform radiation dose within the tumor. A high flexibility is achieved with this system in combination with the gantry and an additional degree of freedom through a rotating patient support. This will allow an optimal conformation of the dose to the shape of the tumor volume. However, before the full flexibility of the system can be used with enough confidence to treat patients, many computer simulations and dosimetric measurements have to be done.

1.5.1 The PSI spot-scanning system and the gantry

PSI is the second place in the world to have a proton gantry. With its 4 m diameter it is far smaller than the gantries at Loma Linda (12 m diameter). The compactness of the gantry is believed to be one factor which will increase chances of transfering proton therapy from research institutes to a hospital environment, which is one goal of the PSI facility. The compact gantry became possible through some technical novelties (Pedroni et al 1995):

- The gantry does not rotate around a static patient couch in the centre, which is the conventional gantry design. The patient support of the PSI gantry is mounted eccentrically on the front end of the rotating structure and performs a counter rotation to keep the patient horizontal all the times. This design allows the reduction of the radius of the gantry to the minimal distance need to deflect the 200 MeV proton beam by 90° without using super-conducting magnets.

- Contrary to all existing proton facilities world-wide, an active beam steering system was developed. The use of the spot-scanning technique allows the application of the beam to the patient almost directly behind the 90° bending magnet, while scattered beam techniques require a lot of space for scattering elements, collimator, bolus and drift space to reduce the angular spread.

The main elements used in the spot application are schematically shown in Fig. 4.

The elements active during a treatment are the kicker magnet (not shown), which switches the beam on and off for each spot, and the sweeper, which displaces the beam in front of the 90° bending magnet in the gantry head. The displacement of the beam by the sweeper causes the beam to shift parallel to itself after leaving the gantry nozzle through the vacuum window. Outside the vacuum window is the range shifter, which scans the energy of the beam in discrete steps by inserting polyethylene plates in the beam path. Thus, these two elements scan the beam in two dimensions. The scanning in the third dimension could in principle also be done by magnets. This would, however, require a larger gantry than the one built at PSI. For the compact gantry it was chosen to shift the patient table in front of the beam. The three scanning axis form a Cartesian, right handed system S, T, U, which is rotated with the gantry angle α. S is along the beam axis, U along the sweeper direction and T is the direction of the table movement. Another degree of freedom for the choice of the beam incidence direction is obtained by a rotation of the patient table.

The order of the spot application is chosen in a way to minimise the treatment time. I. e. first a line of spots is applied in U by changing the sweeper setting (3 ms), then the next range shifter plate is set (30 ms) for the next line of spots until a full plane is finished. The last movement is done by the table (1-2 s) and then the next plane of spots is applied.

The spot-scanning is a highly flexible system, which allows the conformation of the dose to almost any shape of target by using only one or a few beam ports. It also allows the

delivery of inhomogeneous dose distributions, which introduces another degree of freedom for the intensity modulation. Additionally, the system is very efficient, since it does not need any patient specific hardware like a bolus or collimator to shape the beam.
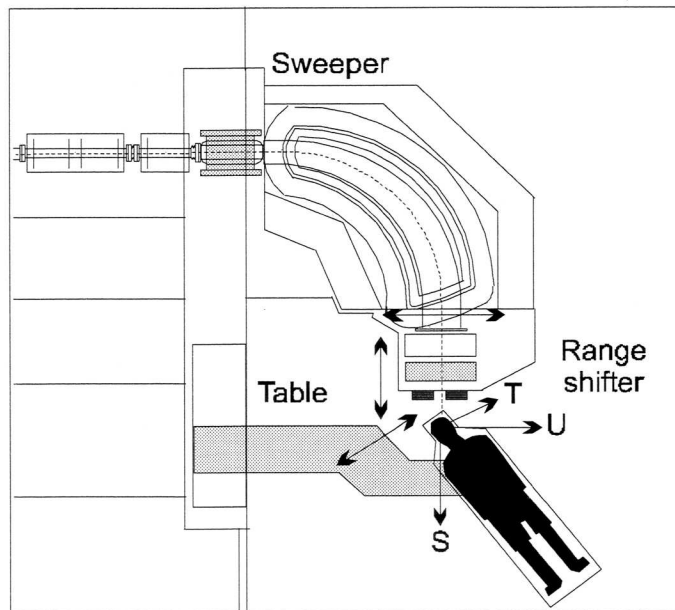

Fig. 4: The concept of the spot scanning system used at PSI. The beam is shifted with the sweeper magnet parallel to itself in the U-direction. The range shifter scans the energy of the beam in discrete steps by inserting polyethylene plates in the beam path. The displacement in the second lateral direction is done by a shift of the patient table. Additionally there is the option of rotating the patient support.

1.5.2 The treatment planning software

The fact that no patient specific beam modifying elements are used is a facilitation in treatment planning. Treatment planning is based on look-up-tables (LUTs) for the depth-dose and depth-width relation of a fixed energy proton pencil beam in water. Assuming a Gaussian shape of the beam, these values are sufficient to calculate the whole 3D shape of the pencil beam.

The depth dose LUT is obtained by fitting the data measured on the gantry to the theory as given by the Bethe-Bloch equation for energy loss through ionisation and flux reduction and energy loss through nuclear interaction (Scheib 1993).

The beam width calculation is based on the multiple Coulomb scattering theory in the appendix. Contributions from the initial phase space and corrections for the effect of the air gap between range shifter and patient are added in quadrature.

Typical values for the beam parameters are 1 cm and 1.1% fwhm for beam width and momentum band respectively.

In materials other than water, the calculation of the dose distribution is done using the concept of the water-equivalent-range (WER). The WER is given by the integration of the stopping power values relative to water along the beam path. The dose distribution, which would be obtained in water, is then scaled in depth by the WER. The necessary information about stopping power values is obtained from CT images of the patient using a calibration curve described in detail later on.

Once the target volume is drawn into the CT images only a few inputs from the planner are necessary to create a plan for a single field. The main one is the choice of the beam energy and incidence direction. Then the planning is done automatically by the software. A grid of possible Bragg peak positions in the patient is defined and the positions inside a certain margin around the target contour are selected. As a first approximation, beam weights which have been found to produce the flat SOBP shown in Fig. 2 are assigned to each row of selected spots. In the following optimisation procedure, the beam intensities are varied iteratively to achieve an optimal dose distribution as evaluated by using a least square algorithm. Usually this optimisation results in increased intensities of the border spots to enhance the steepness of the dose fall-off.

The set of spot parameters obtained from the optimisation procedure can be used for testing different dose calculation algorithms for the same situation, to perform a Monte Carlo simulation or a recalculation of the dose on a finer grid or in a phantom material. Details on different algorithms (section 3.3) and the Monte Carlo code (appendix B) are given later on.

1.6 Range uncertainties of the protons in the patient

The Bragg peak is the great advantage of protons, but can also cause major problems if applied improperly. The small high dose area and the steep dose gradients of a deposited proton pencil beam require a high precision in the control of the range of the protons. An accurate range control requires a high precision of the application technique (steering of the beam, patient setup and repositioning) and an accurate modelling of the beam.

Therefore, this thesis deals with all the problems in treatment planning related to an uncertainty in the range of the protons.

There are two main points, which have to be addressed:

A. Position of the Bragg peak in depth[1]

The depth of the peak in a homogeneous medium is defined by the initial energy of the beam and the stopping power of the medium. The energy of the beam can be measured rather accurately. Therefore, the precision of modelling the range of the beam depends mainly on the precision of the knowledge of the stopping power properties of the human tissues. A direct measurement of proton stopping powers in the human body is theoretically possible using energy loss proton tomography. Proton tomography images have already been taken in 1981 (Hanson et al). Due to the success of computer tomographic (CT) methods using photon attenuation measurements, this technique has not been developed any further.

In a first approximation proton stopping power and photon attenuation are both a function of the electron density. In high density tissue and at low photon energies (used in CT imaging), however, a considerable portion of the photon attenuation is due to the photoelectric effect, which is a function of Z^3. Hence, a calibration curved is needed to convert CT images of the patient into stopping power images. The definition of this calibration curve using literature data for real tissue compositions and reference measurements from the CT scanner is described in section 2.2. Measurements on animal tissue samples were performed to confirm the validity of the calibration curve (section 2.3).

---

[1] submitted for publication to Physics in Medicine and Biology

B. Shape of the Bragg peak[2]

The precise water equivalent range of the protons in the human body could be obtained by integrating the relative stopping power values from the calibrated CT images along proton trajectories, if their paths would be completely known in advance. Coulomb scattering however causes the path of the protons to deviate randomly from straight lines. A proton detected first at a point A (Fig. 5) and then at a point B can have taken a multitude of paths between the two points. The energy loss suffered by a proton following the different trajectories depends on the stopping power of the traversed materials. Thus the energy and therefore the residual range of a proton following trajectory t_1 is different from one following t_2.
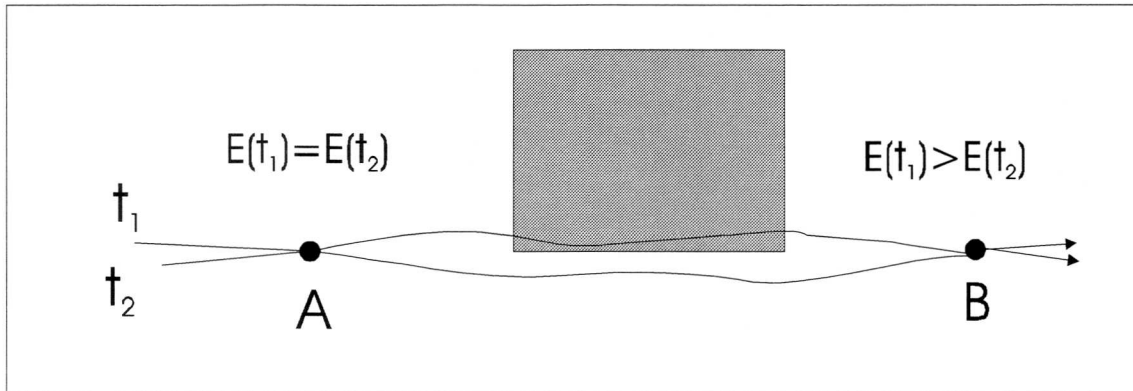

Fig. 5: Two possible proton trajectories close to an inhomogeneity interface. At point A the energy and thus the residual range of a proton is well known. If the same proton is detected at point B its energy depends on the path (t_1 or t_2) which has been taken between the two points.

Looking at a whole beam passing through inhomogeneous regions of a patient, one obtains a wide spread in the energy distribution of the protons. This effect causes a degradation of the whole Bragg peak (Urie et al 1986, Schneider and Pedroni 1995) and presents difficulties in the modelling of the dose distribution.

It is shown in section 3.3 that the use of the original dose calculation as described by S. Scheib (1993) does not model well enough the effects of multiple Coulomb scattering (MCS) in highly heterogeneous regions. The first goal was therefore to develop a way of finding the anatomical regions and beam angles which are likely to produce problems in the dose calculation. A simple way of estimating the variation in the energy spectrum at any point in the patient, which is equivalent to a dilution of the range of the protons, is described. A beams-eye-view display of the relative range dilution expected at the distal border of the target can be used to alert the planner of critical situations.

Since it is not always possible to avoid critical heterogeneous structures, different dose calculation models have been developed, which can deal better with inhomogeneities in the patient. Their advantages and disadvantages, which strongly depend on the context in which they shall be used, are discussed.

A comparison of the results of analytical models to results of a Monte Carlo simulation shows that the use of one of the new analytical models in the optimisation procedure helps to avoid unexpected hot and cold spots in the dose distribution. Thus it leads to better treatment plans and hopefully to better tumor control rates.

---

[2] shortened version is in preparation for publication

2 The precision of proton range calculations in proton radiotherapy treatment planning: Experimental verification of the relation between CT-HU and proton stopping power

2.1 Introduction

The use of protons in radiotherapy offers a good potential to spare healthy tissue due to their characteristic depth dose curve. In the so called Bragg curve relatively low dose is deposited in the entrance region, while a high dose peak can be positioned in the tumor. The fall off of the dose at the distal end of the peak is very steep. With the beam used at PSI - which has a relatively broad momentum band of 1.1 % - the fall off from 90 % to 10 % of the dose is within 7mm for an initial beam energy of 177 MeV.

The range of the protons - and therefore the position of the peak - is controlled by the initial energy of the protons and by the integrated stopping power in the tissues along the beam path. Ideally one would measure a 3D stopping power (SP) distribution directly with proton tomography using energy loss measurements of protons passing through the patient. Though some experimental work has been done (Hanson et al 1981) proton tomography is not available. Instead conventional computed tomography (CT) is used for acquisition of photon attenuation images. This information has to be converted into SP using a calibration curve. We use the method presented by Schneider et al (1996) to create the calibration curve. This curve - as well as any other calibration curve - is only valid for a well defined photon spectrum. The spectrum is given by the scanner properties (tube voltage, filter) and the material traversed by the photon beam up to the point of interest (beam hardening). Various authors have already reported on variations in CT measurements from scanner to scanner (Prasad et al 1979, Battista and Bronskill 1981, McCullough and Holmes 1985) and between different phantom sizes (Minohara et al 1993).

One could correct for beam hardening by using different calibration curves depending on the size of the scanned object. The validity of this approach for different sites in the human body has still to be verified. On the PSI gantry a proton radiography equipment will be installed soon (Pemler 1997). With this tool the residual energy of single protons can be measured after passing through the body. It will allow us to gain more detailed information on the SP distribution in the body. Thus we will be able to verify the validity of the calibration in each patient individually.

In this paper we will present measurements to assess the accuracy of the calibration curve for a well defined photon spectrum. We will also give an upper limit of the expected error of the CT data caused by beam hardening. This error margin can vary from scanner to scanner and is therefore only valid for our scanner, a Siemens Somatom CR. A better precision could probably be achieved using a dual energy CT scanner, which is however not available to us. Maximal proton range errors will be calculated for typical cases taking both sources of error - CT data acquisition and calibration curve - into account.

2.2 The calibration of the CT images into proton stopping power

The conventional method to obtain a relationship between CT numbers and relative electron density is to perform a calibration measurement with tissue substitute materials of known elemental composition. The curve is then obtained by interpolating between these data points. With this method the calibration curve (for relative electron density as well as for relative proton stopping power) strongly depends on the choice of the tissue substitutes used, as pointed out by Schneider et al. The reason for this is that photon attenuation for low energy X-rays is not only caused by Compton scattering - which is proportional to the volumetric electron density - but also by the photoelectric effect and coherent scattering, which depend also on the effective atomic number. Therefore Schneider et al suggest to use a stoichiometric calibration. They use a parametrization based on literature data (Jackson and Hawkes 1981, Rutherford et al 1976)

$$(1) \qquad \mu = \rho_e \left\{ K^{ph} \tilde{Z}^{3.62} + K^{coh} \hat{Z}^{1.86} + K^{KN} \right\}$$

$$(2) \qquad \tilde{Z} = \left[\sum \lambda_i Z_i^{3.62} \right]^{1/3.62}$$

$$(3) \qquad \hat{Z} = \left[\sum \lambda_i Z_i^{1.86} \right]^{1/1.86}$$

to obtain the expected scanner dependent Hounsfield units of given biological tissue compositions. ρ_e is the volumetric electron density, λ_i and Z_i the relative weights and the atomic numbers of the elements in the material. K^{ph}, K^{coh} and K^{KN} are scanner dependent constants to describe the photoelectric effect, coherent scattering (including contributions for binding corrections in Compton scattering) and Compton scattering on a free electron respectively. Throughout this paper we will use the definition of scaled Hounsfield units for CT numbers :

$$(4) \qquad HU_{sc} = 1000 \frac{\mu}{\mu_w}$$

where μ and μ_w are total photon absorption coefficients for the material and water respectively.

To obtain the K-constants we measured HU_{sc} data of cylindrical tissue substitute samples of 4 cm diameter. The reference samples (AP6, water, MS/SR 4, TSK/SR 1, HB/SR 4 (Constantinou 1978)) were scanned individually in the centre of a water phantom. Constantinou and Harrington (1992) report variations in CT numbers for high density material of more than 2 % by just exchanging positions of single samples in an ordinary multi-sample calibration phantom. To get good results it is therefore crucial to scan the samples individually in the centre of the phantom since this guarantees the same photon spectrum for all the samples.

From equations (1) and (4) we obtain

$$(5) \qquad \frac{HU_{sc}}{1000 \rho_e} = A \tilde{Z}^{3.62} + B \hat{Z}^{1.86} + C$$

where $A = \dfrac{K^{ph}(E)}{\mu_w(E)}$, $B = \dfrac{K^{coh}(E)}{\mu_w(E)}$ and $C = \dfrac{K^{KN}(E)}{\mu_w(E)}$ include all unknown parameters.

The left side of eq. (5) includes the measured HU_{sc} and the known electron density of the sample. To determine the coefficients A, B and C, a multiple linear regression fit was performed on the tissue substitute data with the left side of eq. (5) as dependent and the effective Z-numbers as independent variables.

For each tissue the corresponding relative stopping power value is calculated using the Bethe-Bloch formula, which can be approximated by:

$$(6) \quad SP = \frac{\rho_e^t}{\rho_e^w} \left\{ \log\left[2m_e c^2 \beta^2 / I^t \left(1 - \beta^2\right)\right] - \beta^2 \right\} / \left\{ \log\left[2m_e c^2 \beta^2 / I^w \left(1 - \beta^2\right)\right] - \beta^2 \right\}$$

where ρ_e^t and ρ_e^w are volumetric electron densities and I^t and I^w are the mean ionisation potentials of the tissue and water respectively. While the absolute SP clearly depends on the energy of the particle, it can be shown that the relative SP (6) is almost independent of β for the range of particle energies relevant to radiotherapy.

HU_{sc} values taken from eq. (5) and SP values from eq. (6) are calculated and plotted in Fig. 6 for human tissue compositions taken from literature (ICRU 44 1989, Woodard and White 1986).
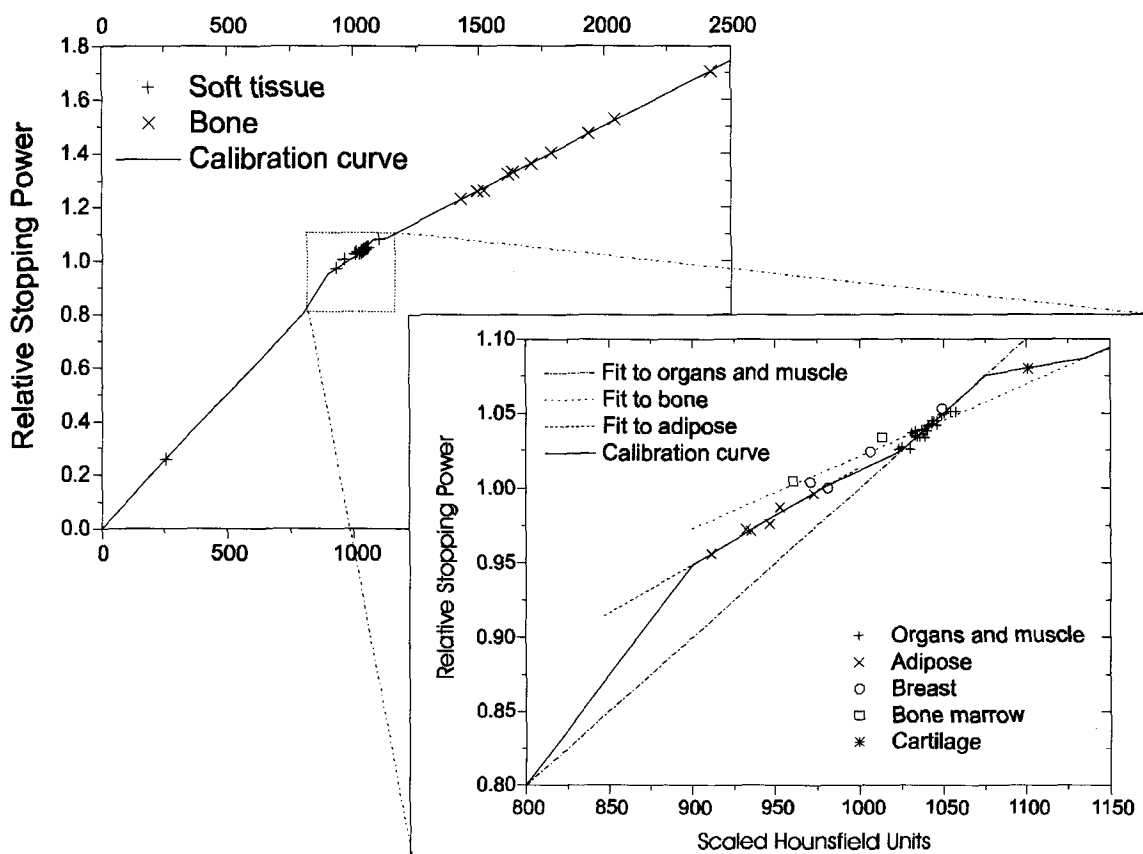

Fig. 6: Calculated data points (Scaled HU vs. relative SP) for biological tissues grouped into soft tissue and bone (behind). In the enlarged graph the soft tissue group is again split into five categories. The different lines show the three linear fits to the calculated soft tissue, adipose and bone data points and the chosen calibration curve.

While all the data points for real bone compositions lie on a straight line there is no unique definition of the calibration curve in the soft tissue region (Fig. 6).

For analysis we split the soft tissue data into five categories (organs and muscle, adipose tissue, breast, cartilage, bone marrow). This classification is not only motivated by anatomical criteria, but also reflects a difference in the elemental composition. $\tilde{Z}$ describes the increased contribution from the photoelectric effect to the total photon absorption coefficient eq. (1) for higher atomic numbers. The mean ionisation potential I also increases with the atomic number. These two factors are mainly responsible for the

deviation of the calibration curve from a straight line. The two most common elements by weight in soft tissue are oxygen and carbon. They add up to more than 85 % in all the soft tissues (except cartilage).The category adipose contains tissues (adult and children data from different sources) with a relatively high content of carbon (above 35%) and a relatively low content of oxygen (below 43 %) and therefor a relatively low effective atomic number. The category organs and muscle includes bladder, blood, brain, breast, GI tract, heart, kidney, liver, lung, lymph, muscle, ovary, pancreas, prostate, spleen, testis and thyroid. The carbon and oxygen content of the tissues in this category are below 17 % and above 69 % respectively. Their effective atomic numbers are clearly higher than those in adipose tissue.

Linear fits where done to the bone, the adipose and the organ and muscle categories separately. These fits (Fig. 6) are used for the definition of the calibration curve. The data points for breast, bone marrow and cartilage will be discussed later on.

To decide on how and where to connect the three linear fits, CT data sets of different sites in the human body (head, abdomen, pelvis) were analysed. Fig. 7 shows histograms of the frequency of HU_{sc} values. In the head CT a clear peak at 1030-1035 HU_{sc} from the brain can be seen, in the abdomen and pelvis CTs major peaks from muscle and organs appear between 1040 and 1060 HU_{sc}. We decided that the calibration curve should follow the soft tissue fit between 1025 and 1075 HU_{sc}.
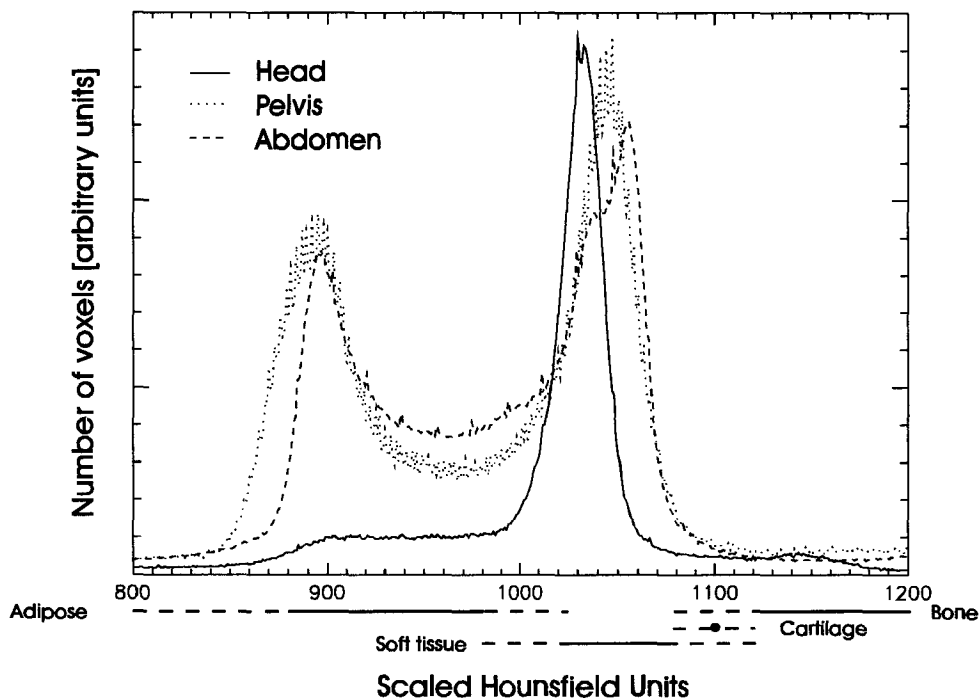

Fig. 7: Frequency of HU_{sc} values in CT data sets of different sites in the human body. The lines below indicate the regions described in the text, where the definition of the calibration curve Fig. 6 follows the data points for a particular category of tissues (straight line) or where a linear connection between the curves for two categories was chosen (dashed lines).

In the pelvis and abdominal CT a second peak appears between 880 and 910 HU_{sc}, which is caused by subcutaneous and retroperitoneal adipose tissue. This peak is clearly below the calculated data points based on the literature data. When looking at the CT images in more detail, it is obvious that the subcutaneous adipose tissue is compressed where the patient lies on it. We conclude that it is very difficult to define the density of adipose

tissue. This probably also affects the accuracy of the literature data. For this reason the definition of the calibration curve between 800 and 900 Hu_{sc} is somewhat arbitrary.

Based on this data the decision was taken to connect the soft tissue and the adipose fits from 800 to 900 HU_{sc} and 980 to 1025 HU_{sc} respectively and to connect the soft tissue and bone fits from 1075 to 1135 HU_{sc} through the data point for cartilage. Thus we obtain the calibration curve plotted as a solid line in Fig. 6.

Breast is considered to be a good model for a tissue with a high fat content which can vary between individuals. Since all the data points for breast lie within 1 % of the curve, we expect the curve to fit well also to other soft tissues with increased fat content.

The data points for bone marrow are the only ones to differ by more than 1 % from the calibration curve. Bone marrow however does usually not appear isolated in the human body but together with trabecular bone. In CT images used for proton therapy planning the pixel resolution is not small enough to separate trabecular bone and bone marrow. Pixels containing bone marrow will most likely also contain some bone and therefore display a CT value considerably higher than the one for marrow alone. Since the data points for both, red and yellow marrow, lie close to the fit to bone data, any mixture of bone and marrow is therefore expected to lie on the fit to the bone data.

With this method we cannot overcome the problem of beam hardening artefacts. From our measurements the variation in the measurement of HU_{sc} between a 15 cm and a 30 cm diameter phantom is around 5 % for a bone substitute material (HB/SR4). Variations for soft tissue were found to be below 1 %. We defined a small and a large body calibration curve from the fit parameters obtained with the tissue substitute samples scanned in the 15 cm and 30 cm phantom respectively. Our standard calibration curve is defined to be the average between the two curves. Fig. 8 shows a plot of the three curves.
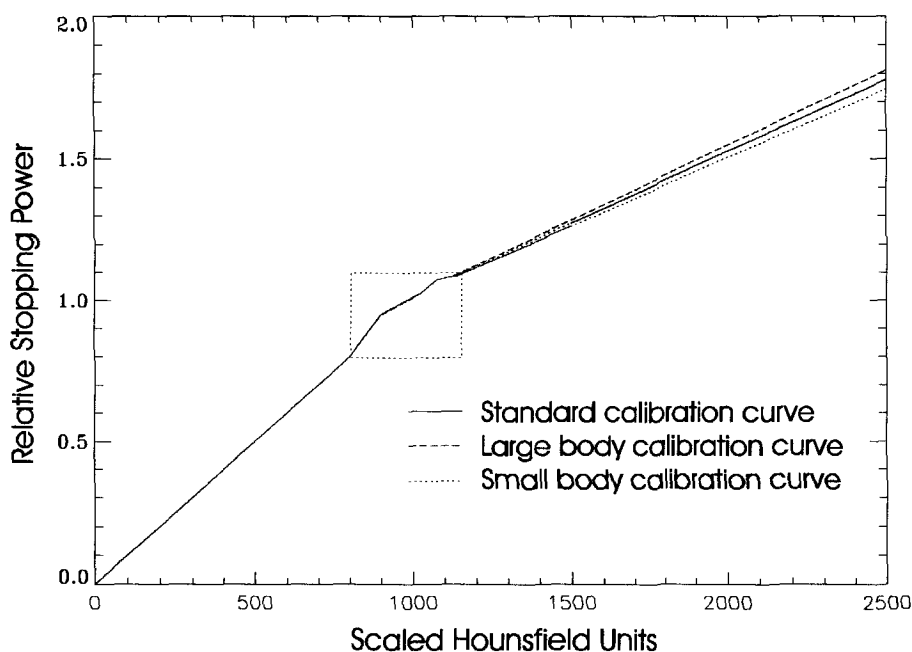

Fig. 8: The standard calibration curve obtained by averaging between the small and large body calibration curve. The dotted rectangle indicates the range plotted in Fig. 6.

There is practically no difference in the three curves up to about 1100 HU_{sc}. At 2000 HU_{sc} - a value higher than that is only found in cortical bone - the SP values obtained from the large and

small body curve respectively differ by $\pm 1.5\,\%$ from the value obtained from the standard curves.

2.3 Measurement of proton stopping powers for biological tissue

The theoretical calibration curve described in section 2.2 has been verified experimentally using animal tissue samples.

It is not the aim of this work to assign standard values for HU_{sc} or SP to any of these tissues, but verify that their HU_{sc} - SP combination matches the calibration curve created with the stoichiometric method. We do not expect the elemental composition of animal and human tissue to differ considerably. For these reasons a few samples from different animals and different organs were considered adequate. They are muscle (from a horse), adipose tissue, brain and blood (pig), bone , kidney, heart, spleen, liver and lung (cow).

2.3.1 Sample preparation

The samples were taken fresh from the butcher and all measurements done within 24h. With the exception of bone all samples were prepared in the following way. Four or five of them were positioned in a $10\,cm \times 10\,cm \times 30\,cm$ (inner width) perspex box, each in an individual thin plastic bag. They were cut in pieces (in the case of the brain three whole brains where used) and combined to fill the box completely and to produce an area of homogeneous tissue 10 cm across and at least 3 cm wide. Slight pressure was applied to the samples to avoid air gaps within the volume crossed by the beam during the SP measurement (section 2.3.3). From the blood the fibrin was taken out but no other treatment applied. The flat surface of the box assures a sharply defined Bragg peak and the wall (2 cm) simulates some surrounding tissue (see Fig. 9).
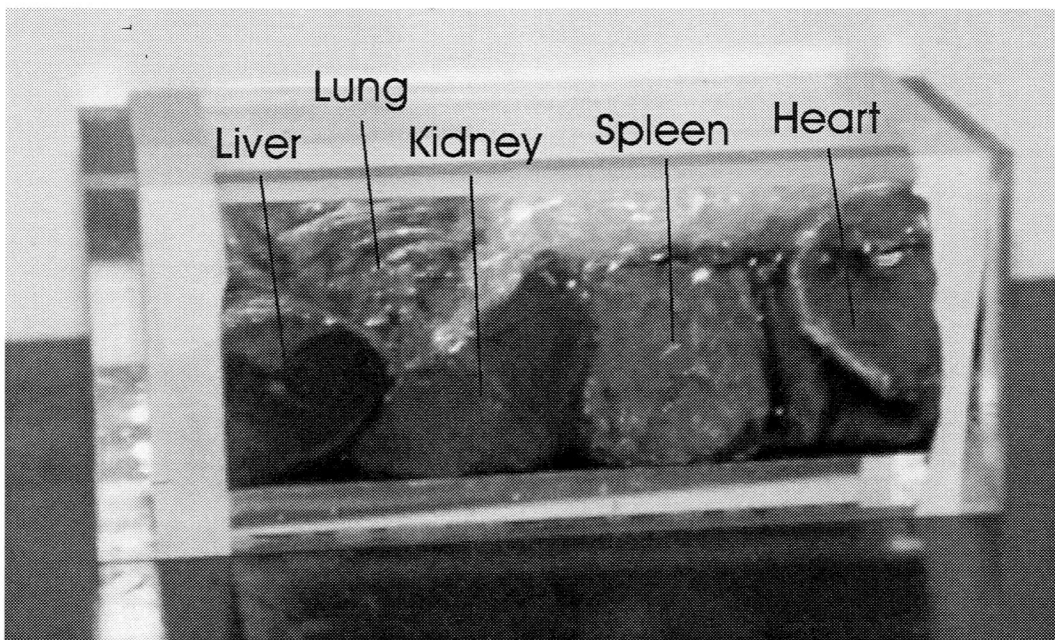

Fig. 9: Some of the animal tissue samples in the perspex box ready for measurement. The samples were positioned such that there was no air gap between the wall of the box and the samples in a well defined area. The stopping power measurements with the proton pencil beam was performed in that area.

2.3.2 Measurement of CT values

CT scans of the box with the samples were done in 0.4 cm steps and a three dimensional volume reconstructed. From the CT images the homogeneity of the samples was confirmed and the position defined where the range measurement with the proton beam should be performed. On some samples measurements on two different positions were taken (labelled tissue1 and tissue2). The HU_{sc} value for each tissue was defined as the mean value in a 0.8 cm × 0.8 cm × 10.0 cm volume of interest (VOI) centred around the above position. This definition requires a good homogeneity of the tissue throughout the VOI. For all the tissue samples except lung and bone - which will be treated separately - the one standard deviation for the pixel to pixel value was between 0.5 % and 2 %. These tissues will be referenced as homogeneous tissues

2.3.3 Measurement of the relative stopping power

The SP measurement was done using a proton pencil beam (ca. 1 cm fwhm) and range shifter plates which were inserted in the beam path to scan the energy of the beam in discrete steps from 177 to 70 MeV. After passing through the box with the samples, the integral dose over the whole spot was measured with a large parallel plate ionisation chamber (9 cm ∅). This setup is shown in Fig. 10.
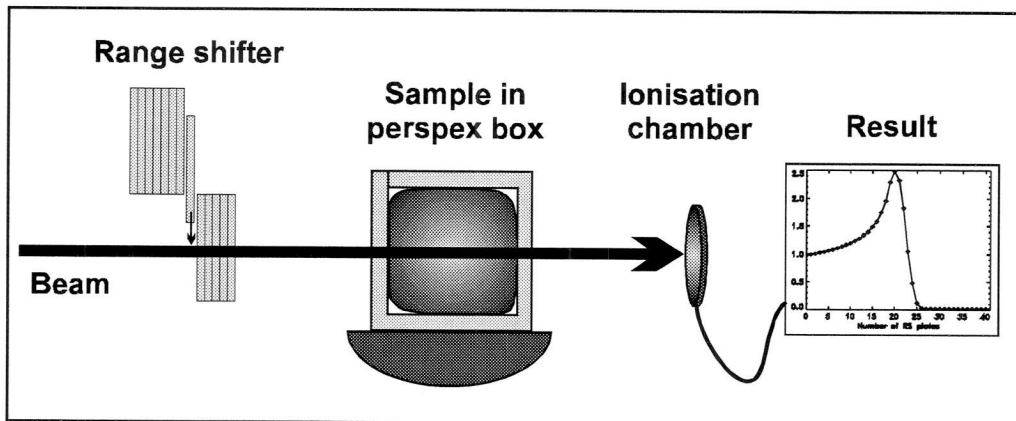

Fig. 10: The set-up of the proton range measurement. The beam energy is scanned in discrete steps using the range shifter. It passes through the box with the tissue samples and is then measured in an ionisation chamber. The signal of the chamber plotted versus the number of range shifter plates in the beam (i.e. the beam energy) shows the shape of a Bragg peak.

The integral dose plotted against the number of range shifter plates shows a characteristic Bragg curve. Using the same setup but substituting the tissue samples by water, a reference curve was taken. With the shift Δx between the curve for water and for the tissue sample the relative SP of the tissue was calculated:

$$(7) \qquad\qquad SP = 1 + \frac{\Delta x}{10cm}$$

Fig. 11 shows the measured reference curve for water plotted together with the curve for muscle shifted by 2.6, 2.8 and 3.0 range shifter plates respectively. One range shifter plate is equivalent to 0.23 cm of water. The middle curve clearly fits the reference curve best. We conclude that this method of measuring the relative stopping power is very sensitive and expect the SP measurement to be accurate to 0.2 %, which corresponds to a shift of 0.1 range shifter plates.
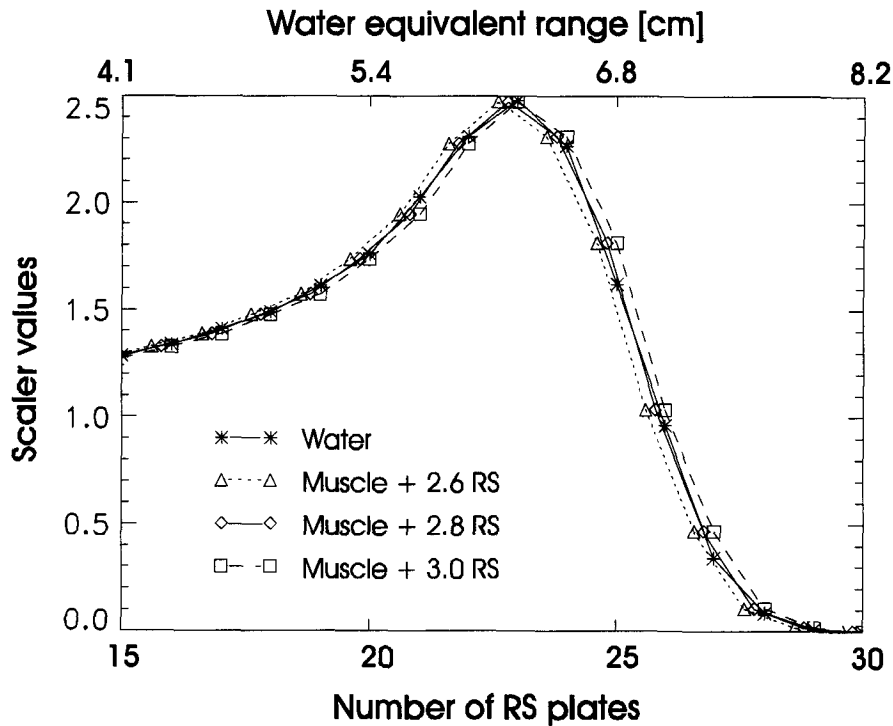
Fig. 11: An example of the measured curve for water plotted together with the curves for muscle, shifted by 2.6, 2.8 and 3.0 range shifter plates respectively. This plot shows that the measurement can be performed accurate to ± 0.1 range shifter plate. This corresponds to an accuracy of determining the relative stopping power of the tissue to ± 0.2 %.

2.3.4 Results for homogeneous material

For homogeneous tissue HU_{sc} and SP values were obtained by the method described above. All homogeneous tissue samples have densities between 0.9 and 1.1 g/cm3. They are considered representative for all human tissues except bone and lung.

In Fig. 12 results for the soft tissue area are plotted together with the small body calibration curve. The deviation between the measured SP for the biological tissue samples and the SP values obtained from the calibration curve for the measured HU_{sc} is below 1 %.
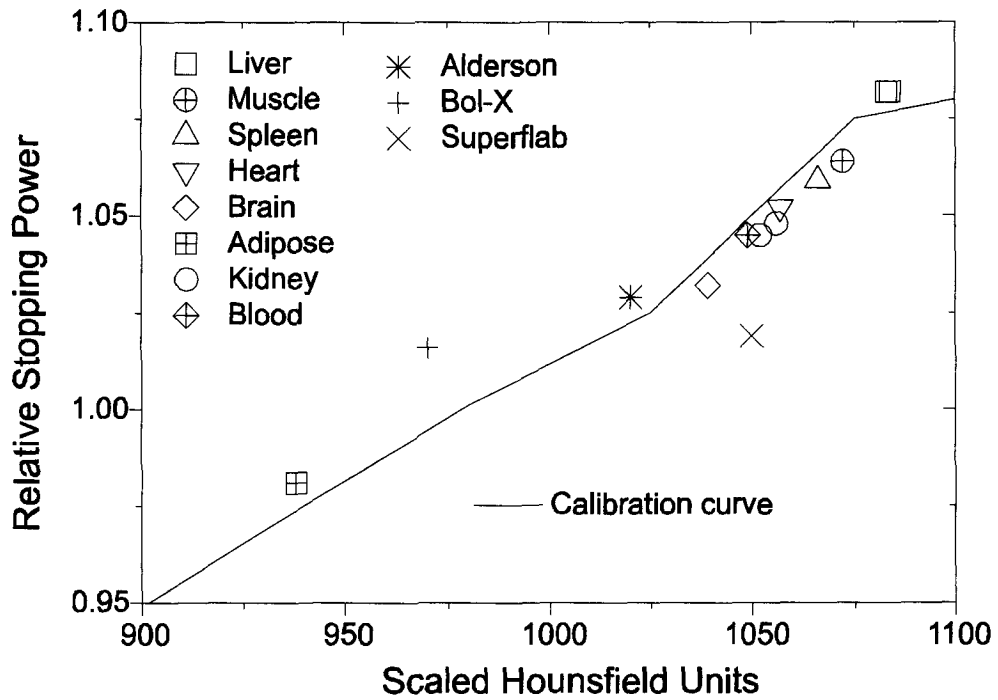
Fig. 12: Measured HU_{sc} - and SP data for biological tissues, two bolus materials and Alderson tissue plotted the small body calibration curve. The correlation between the calibration curve and the measurement is better than ± 1 % for all biological tissues. The correlation for the two bolus materials is not so good. This can however be corrected in the CT images which are used for treatment planning.

There is however a discrepancy between absolute calculated HU_{sc} and SP values (see section 2.2 for method and data) for certain tissues and their corresponding measured values. This can be seen from Tab. 1 which shows the measured data and the corresponding calculated data using equations (6) and (5).

Tissue	HU_{sc} meas.	SP meas.	$HU_{sc}m/SPm$	HU_{sc} calc.	SP calc	$HU_{sc}c/SPc$
Adipose	938	0.981	956.2	932	0.972	958.8
Brain	1039	1.032	1006.8	1040	1.038	1001.9
Kidney 1	1052	1.045	1006.7	1041	1.041	1000.0
Kidney 2	1056	1.048	1007.6	1041	1.041	1000.0
Heart 1	1057	1.052	1004.8	1054	1.051	1002.8
Heart 2	1057	1.052	1004.8	1054	1.051	1002.8
Spleen 1	1066	1.059	1006.6	1050	1.050	1000.0
Spleen 2	1066	1.059	1006.6	1050	1.050	1000.0
Muscle	1072	1.064	1007.5	1037	1.039	998.1
Liver 1	1084	1.082	1001.8	1049	1.049	1000.0
Liver 2	1083	1.082	1000.9	1049	1.049	1000.0
Blood	1049	1.045	1003.4	1057	1.051	1005.7

Tab. 1 : Measured and calculated data (eq. (6) and (5)) for the biological tissues. Although the absolute values for the relative SP and scaled HU for the individual tissues do not coincide very well, the ratios are accurate to better than ± 1 %.

This discrepancy can be explained by the density variation which can occur in biological tissues of the same elemental composition, especially after isolating tissue samples from

the body. It might also be due partially to the slight pressure applied to the samples (section 2.3.1). We are however not interested in absolute HU_{sc} and SP values but in the behaviour of the calibration curve between the two. Therefore we calculated the ratio between HU_{sc} and SP for measured and calculated values which is independent of the mass density. The results are listed in Tab. 1 and plotted in Fig. 13. The agreement between calculated and measured value is better than 1 %.
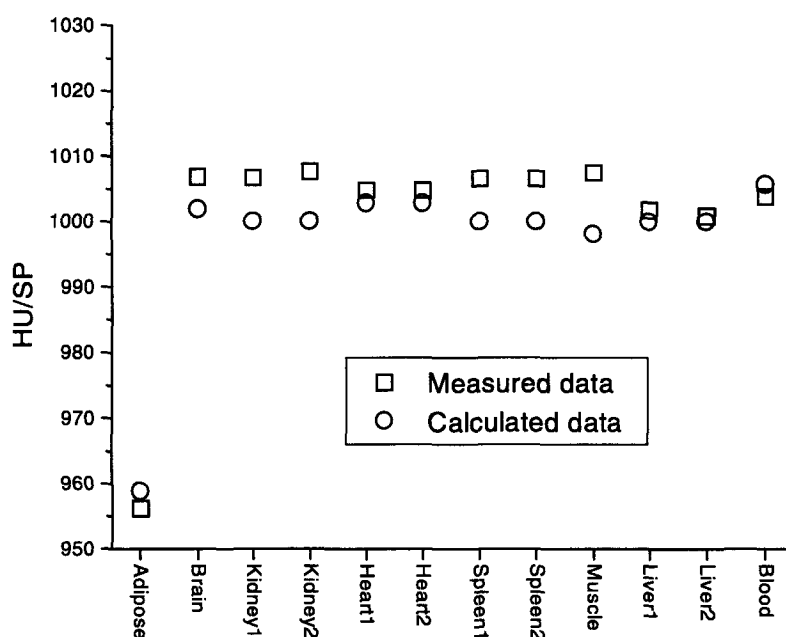
HU/SP

1030
1020
1010
1000
990
980
970
960
950

□ Measured data
O Calculated data

Adipose Brain Kidney1 Kidney2 Heart1 Heart2 Spleen1 Spleen2 Muscle Liver1 Liver2 Blood

Fig. 13: Ratios HU_{sc}/SP value of biological tissue for measurement and calculation. The ratio is almost exactly 1000 for all tissues in the organs and muscle category (section 2.2) while it is considerably lower for adipose tissue. This confirms our approach of using a linear calibration curve for soft tissue except adipose.

The fact that the ratio HU_{sc}/SP is almost 1000 for all soft tissues except adipose implies that HU_{sc} and SP are directly proportional to each other for all these tissues. It also implies that any combination or microscopic mixture of these tissues (esp. blood) and water will lie on that curve. At the same time it is clear from Fig. 13 that one has to allow for adipose tissue not lying on that curve, especially when considering the high amount of adipose tissue found even in a normal size person (see Fig. 7).

With the same method we measured HU_{sc} and SP of two bolus materials (Bol-X, Superflab). The results are also plotted in Fig. 12. These measurements demonstrate the importance of knowing exactly the characteristics of the materials used in the beam path. The two materials have approximately the same SP when exposed to a proton beam, but the photon attenuation properties are very different, which is expressed in a difference of about 80 HU between the two. Using bolus materials without corrections can introduce range errors of up to 3 % of the bolus thickness.

A homogeneous piece of Alderson material was measured as well. The result is in good agreement with the curve. It is an interesting fact, that both HU_{sc} and SP values of Alderson tissue are considerably lower than the values for all measured tissues in general and muscle in particular.

2.3.5 Analysis and results for inhomogeneous tissues

The analysis of measurements for inhomogeneous cases presents some problems, since an unavoidable degradation of the Bragg peak occurs (Urie et al 1986, Schneider and Pedroni 1995). It is also not possible to define a single value for the HU_{sc}. In the lung sample for example the mean value and one standard deviation in the VOIs (defined in section 2.3.2) for two different positions of the proton beam were $401 \pm 134\,HU_{sc}$ and $545 \pm 206\,HU_{sc}$. The second region includes a branch of the bronchus.

We therefore used an indirect method to check the validity of the calibration curve for lung and bone. The same method was used to evaluate measurements of different positions in the kidney through different amounts of adipose tissue.

Fig. 14 shows the concept of this indirect measurements:

Stopping power measurements using the range shifter scan were performed as described in section 2.3 with the exception that the bone was not put in the box but was measured in air. The measured integral dose after the sample is plotted in the graph on the right hand side of the picture (x) as a function of the residual water equivalent range of the beam used to measure the sample. The residual water equivalent range is defined to be the nominal range of the primary proton beam minus the water equivalent range of all absorbing material except the sample (range shifters, chamber wall and the wall of the perspex box in case of the lung measurement). In the same plot a non-degraded Bragg peak is plotted to illustrate the effect of degradation of the Bragg peak in the bone (dashed line).

Based on the calibrated CT images a Monte Carlo simulation (Tourovsky et al. 1993) of the range experiment was performed. The MC code normally used for dose calculations in the patient (Appendix B) was slightly modified to adapt to the setup of this measurement. One million protons per range shifter setting were traced through the three dimensionally reconstructed CT images of the bone or the box containing the tissue samples respectively. The energy lost by the individual protons in the chamber upon exit from the sample was added up for each range shifter setting. This sum was assumed to be proportional to the signal obtained from the ionisation chamber measurement. Thus the Bragg curve - obtained by plotting the total energy versus the number of range shifter plates - should match the curve obtained from the ionisation chamber measurements. This method was checked by simulating the measurements of the homogeneous tissues. The agreement in range was equal or better than 0.4 % for all the samples.

The MC code is based on calibrated CT images. For lung and kidney the small body curve was used to calibrate the CT data while a special calibration curve was created for the bone using measurements of the reference samples scanned in air. With this method we can neglect beam hardening artefacts in the case of these isolated tissue samples. The results of the MC simulation are also plotted as + in the same graph (Fig. 14).

Since the accuracy of positioning the beam is expected to be about 1 mm another four MC simulations were done assuming a shift of the beam of 1 mm in all four directions. While the shape of the curve was changed only slightly, the position of the maximum dose (obtained from the spline interpolation) was the same within + 0.09 cm and - 0.13 cm.
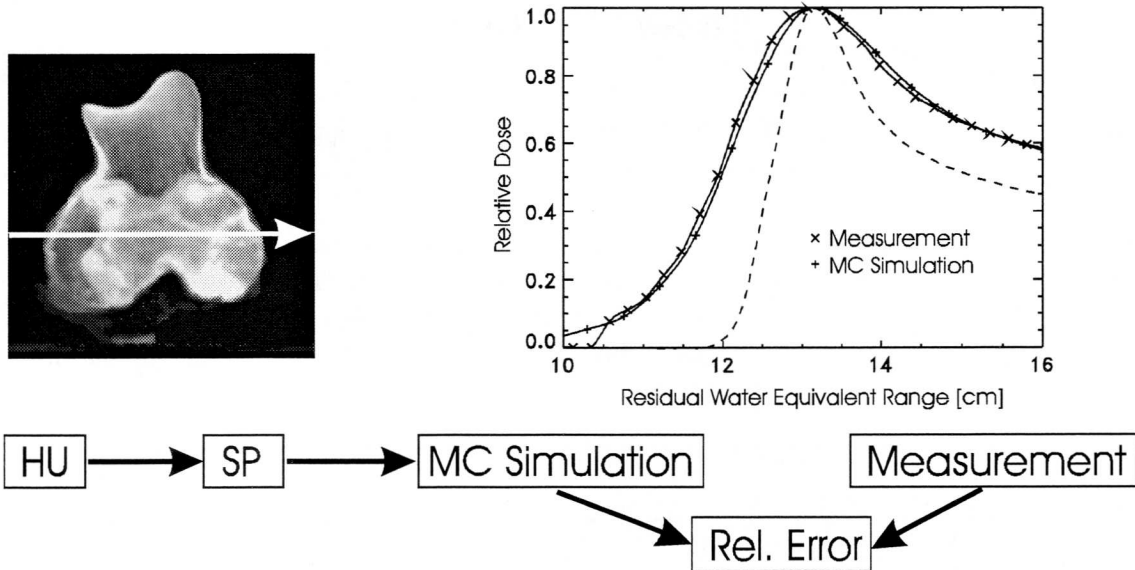
Fig. 14: CT images of the bone with measured and simulated Bragg curves for a beam path along the arrow. The correlation between the two curves is very good. Thus the calibration which has been used for the Monte Carlo simulation is believed to be correct. The dashed line shows a non-degraded Bragg peak to illustrate the amount of degradation in the bone.

From the good agreement between the curves predicted by the Monte Carlo simulation and the measured Bragg curves we conclude that the range control is better than 1% of the residual water equivalent range of the beam. The results from the comparison between simulated and measured data for lung and through the central part of the kidney with its high adipose content also showed a range control within 1 %.

We conclude from these results that the calibration curve in the high density (bone) as well as in the low density (lung) regions is accurate to 1 %.

The observed degradation of the Bragg peak is considerable and should not be neglected in treatment planning. This experiment also shows that the Monte Carlo code is a good option to verify plans and test analytical dose calculation algorithms.

2.4 Conclusion - An overall error estimate for proton ranges

One has to distinguish between errors in the calibration curve and variations in the acquisition of CT data due to beam hardening.

Errors in the calibration curve can arise from an inaccurate model for the HU_{sc} and SP calculation, from incorrect input data for the real tissue composition or for the elemental ionisation potential or from tissues which do not fit the average calibration curve because of a particular elemental composition. We measured HU_{sc} and SP of biological tissue samples. The experimental verification of the calibration curve yields results better than 1 % for all measured samples from lung to bone. From these results we conclude that it is possible to define a single calibration curve to fit all human tissues. Thus the use of an organ specific calibration, as proposed by Schneider and Pedroni (1995), does probably not bring big further improvements. If bolus material is used, its properties should however be checked and possibly a correction to the calibration curve be applied.

Without beam hardening artefacts one could expect to achieve an accuracy of ± 1% in the determination of the SP from CT measurements. However, variations in the HU_{sc} measurements due to beam hardening effects can cause considerable errors, especially in high density materials. The result of a CT measurement depends an the position of the

sample within the body and the size of the body. When an average calibration curve is used, maximal errors of ± 1.5 % in the determination of the stopping powers for bone have to be expected. For soft tissue the maximal error is ± 0.5%.

Adding the two contributions quadratically we get an overall approximation of the range error caused by CT and calibration curve of 1.8 % and 1.1 % for bone and soft tissue respectively.

Tab. 2 shows an estimate for absolute range errors in two typical cases based on these relative errors. The amount of material in the beam path as well as the water equivalent range (wer) changes of course from patient to patient. For bone a relatively high stopping power value of 1.8 was chosen to make sure to get an upper limit approximation for the absolute range error.

	Soft Tissue			Bone			Total
	amount	wer	abs. error	amount	wer	abs. error	abs. error
Brain	10 cm	10.3 cm	0.11 cm	1 cm	1.8 cm	0.03 cm	0.14 cm
Prostate (lateral beam)	15 cm	15.5 cm	0.17 cm	5 cm	9 cm	0.16 cm	0.33 cm

Tab. 2 Two typical proton treatment cases and expected range errors. The expected error in the position of the distal fall-off of the dose distribution is expected to be a few millimetres.

In practice we will use minimal and maximal calibration curves to calculate maximal range errors for different positions in each patient individually. Using the three curves the full dose calculation is done three times to indicate maximal errors to be expected in the final dose distribution due to uncertainties in the determination of the stopping power. Further investigations using range calibrated proton radiography are expected to allow us to certify these limits in practice and eventually bring the error margins further down.

3 Range dilution of the proton beam after heterogeneous structures and its influence on treatment planning

3.1 Introduction

When protons pass through a medium they are scattered in the Coulomb field of the nuclei. Their path is therefore not a straight line. Coulomb scattering happens randomly along the whole path of the proton, usually causing only small angular deviations from the initial trajectory. The probability distribution for the resultant angular spread at a certain depth is described by the Molière theory (Molière 1948). When occasional large angle scattering is neglected, the distribution can be approximated by a Gaussian (see appendix). In the following analysis of proton range dilution we will use that simplification.

Multiple small angle Coulomb scattering (MCS) causes a certain amount of confusion - a 'wiggling' - of the trajectory of the protons. The paths of protons of the same initial energy, which reach the same position in the patients body can therefore be quite different. If there are differences in the density of the material along these paths, the energy loss suffered by the protons up to that position is different. Consequently their expected residual range also differs.

It is furthermore practically impossible to produce a pencil beam with exactly parallel trajectories of the protons. There will always be a certain angular spread of the incident beam, which also contributes to the uncertainty in the path of the protons detected at a certain point in the patient.

In a body area with high density variations one can thus obtain a widening of the energy spectrum of the proton beam, which results in variations of the residual range between single protons and a dilution of the ranges in an initially monoenergetic proton beam. This effect can also be described as a degradation of the Bragg peak (Urie et al 1986, Schneider and Pedroni 1995).

This energy degradation of the beam is very difficult to model with an analytical calculation due to the random nature of the process. Monte Carlo (MC) simulations are best suited to model effects of MCS, but they are too slow to be used for routine treatment planning. A quick analytical estimation of the range dilution was therefore developed in the frame of this work. The purpose of this estimation is twofold:

- It shall be used for displaying the expected relative range dilution in the target region depending on the selected beam direction. Thus it serves as an on-line warning tool to indicate critical regions in the patients and shall help inexperienced planners to choose optimal gantry angles.

- The estimated range dilution is also used for the calculation of a correction term in one of the analytical dose calculation algorithm (section 3.3.5).

Analytical dose calculation algorithms are usually based on look-up-tables of depth-dose curves which describe the shape of the beam in water. Their capabilities of modelling the range dilution correctly are therefore very limited.

The first dose calculation algorithms for the treatment planning of proton radiotherapy use a simple ray casting technique (Chen et al 1979). They are written for scattered beam techniques. Compensation for tissue heterogeneities is done by compensating boluses as

described by Urie et al (1984). The effects of MCS as well as the scattering in beam modifying devices are important for the correct modelling of the lateral penumbra (Urie et al 1986). They could, however, only be done empirically.

At PSI a spot-scanning system for a compact proton gantry has been developed (Pedroni et al 1995). The compensation of tissue heterogeneities is achieved through the energy modulation of the individual pencil beams by using a range shifter. Therefore, a different treatment planning system has been developed (Scheib 1993). The dose model is based on a three dimensional mapping of a physical pencil beam measured in water and in air. The dose distribution of the scanning beam is scaled in depth according to the density integrated along the central axis of the beam spot. Since the physical pencil beam is relatively large, it has been found that neglecting the influence of off-axis heterogeneities can lead to considerable errors in the dose distribution.

A solution for modelling better the dose distribution obtained by both, scattered and scanned beams, are elemental pencil beam models. They are conceptually the same as the scanning pencil beam model, but use infinitesimally small pencil beams. Broad scattered beams (Petti 1992 and 1996 and Hong et al 1996) and scanning beams are then modelled as a superposition of elemental pencil beams.

In the spot-scanning technique an improved conformation of the dose to the target can be achieved by an iterative optimisation procedure to optimise the contributions of individual beam spots. For this optimisation a spot by spot dose calculation is needed. An algorithm based on the calculation of elemental pencil beams would be too slow for this purpose. Therefore, two modified versions of the original scanning beam model were developed. They are presented here and their results evaluated. The evaluation is done by comparing the results of the different dose calculation algorithms with the results of the PSI Monte Carlo code (Tourovsky et al 1993) for both geometrical and anatomical situations.

It will be shown that body areas with high range dilution coincide with areas where largest errors in the analytical dose calculations and errors due to set-up uncertainties are expected. Thus it is highly recommend to choose carefully the beam incidence angles using the above mentioned tool.

3.2 Estimation of the range dilution after inhomogeneities .

3.2.1 Definition and calculation of the range dilution

The range dilution of a proton beam passing through a heterogeneous target is defined to be the root-mean-square (rms) of the individual proton ranges. The proton ranges are calculated as a function of the position P and the beam incident direction in the CT data set by using representative proton trajectories.

Some major simplifications are assumed for this calculation of the range dilution (Fig. 15):

- The beam is assumed to extend infinitely in the lateral directions.
- Only protons stopping at P are considered, since they will cause largest errors in the dose calculation at point P.
- The representative trajectories of all protons stopping in the point P are straight lines
- The trajectories lie in a cone centred around the nominal beam path and originating in P.

- The probability of each ray is given by a Gaussian distribution for its angular displacement from the nominal beam direction.

In the present version of the code, water-equivalent-ranges (WER) are calculated along 25 discrete rays by integrating relative stopping power values (SP) to obtain a spectrum of possible ranges. The choice of the representative rays and their weights is done in the following way:

In x and y direction the angular displacement of the rays is chosen to be $\pm$ 1, 2, 3 σ_ϕ^{TOT} and $\pm$ 1/2, 3/2, 5/2 σ_ϕ^{TOT} in the diagonal, as shown in Fig. 15. The width of the distribution of angular displacements is σ_ϕ^{TOT} - the angular uncertainty, which will be defined later on. The contribution from each ray is weighted according to the probability of the ray (w_i), assuming a Gaussian distribution.

$$(8) \qquad w_i = \frac{A_i}{2\pi\sigma_\phi^{TOT\,2}} e^{-\frac{\phi_i^2}{2\sigma_\phi^{TOT\,2}}} \qquad \text{where} \qquad A_i = \frac{\pi}{4}\phi_i\sigma_\varphi$$
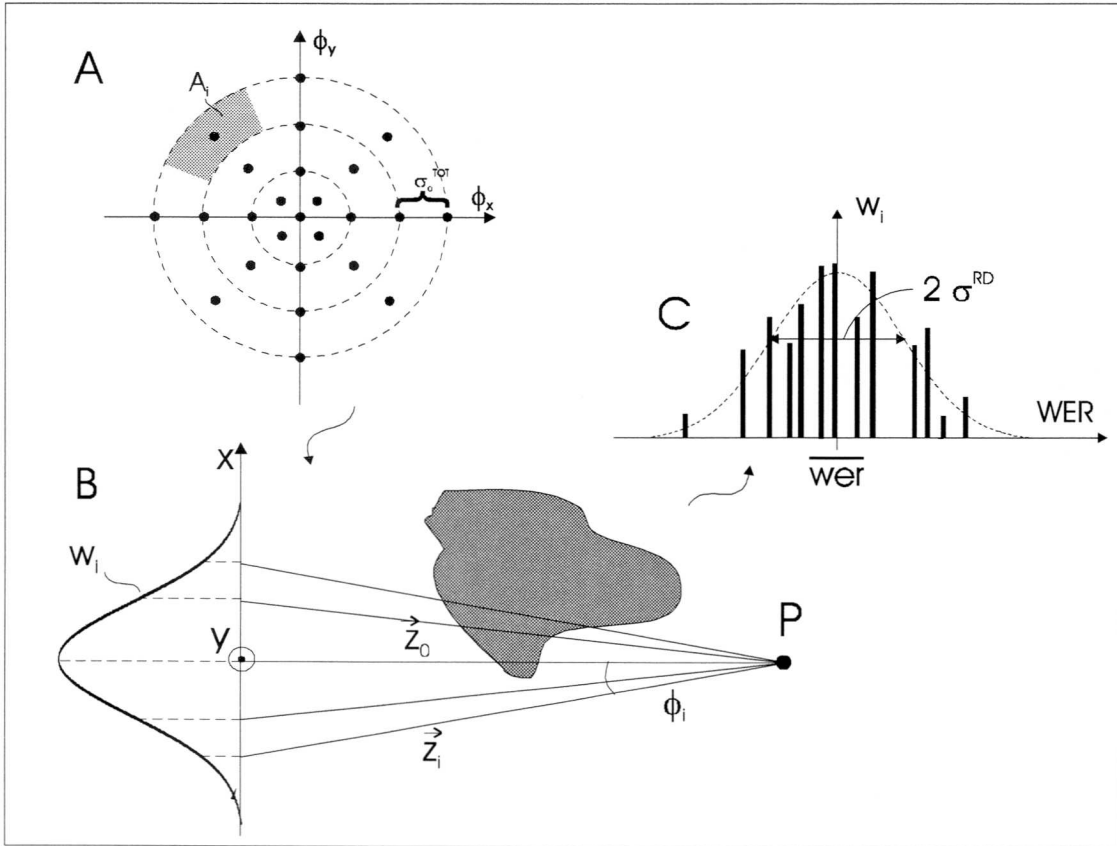
Fig. 15: Concept of the range dilution calculation based on the assumptions made in the text: The representative rays are arranged in a cone, each of them representative for a certain angular region (A). The integration of the density values is performed along those rays and weighted by the probability of the ray (B). The range dilution is then defined as the rms of the weighted contributions from the rays (C).

Thus, differences in the water equivalent range of two protons up to a certain point P are assumed to originate solely from the angular displacement between them. Although this is a rough approximation, the validity of this approach has been shown by Schneider et al

(1997) by comparing the obtained spectrum from 100 rays behind an interface with the results of a Monte Carlo simulation (Tourovsky et al 1993) of the same situation.

The range dilution (σ^{RD}) is then defined by the rms of the range spectrum. Using the assumptions made so far, the formulas follow straight forward:

(9)   Nominal range

$$wer(P) = \oint_{z_0} SP(x,y,z)d\bar{z}_0$$

(10)   Mean range

$$\overline{wer}(P) = \frac{\sum w_i \oint SP(x,y,z)\,d\bar{z}_i}{\sum w_i}$$

(11)   Absolute range dilution

$$\sigma^{RD}_{ABS}(P) = \sqrt{\frac{\sum w_i \left(\oint SP\,d\bar{z}_i - \overline{wer}\right)^2}{\sum w_i}}$$

(12)   Relative range dilution

$$\sigma^{RD}_{REL}(P) = \frac{100}{\overline{wer}(P)}\sigma^{RD}_{ABS}(P)\ [\%]$$

3.2.2 Definition and calculation of the angular uncertainty

For the definition of σ_ϕ^{TOT}, which describes the angular uncertainty at a point P in the patient, two effects have to be discussed . They are shown schematically in Fig. 16.

A) Angular uncertainty due to the properties of the incident protons, which are described by the phase space of the beam:

This can be modelled by a Gaussian distribution with width σ_ϕ^{EPS}, which contains contributions from the initial phase space of the beam and the scattering in the beam modulation device (modulator wheel or range shifter plates), but is constant at any depth of the patient. This component depends on the parameters of the beam delivery system.

B) Multiple Coulomb scattering:

MCS in the patient causes the path of the protons to deviate from a straight line. Towards the end of the protons range the angular uncertainty is increased drastically. This is an intrinsic property of protons and does not depend on the technique used. In Fig. 16 B the beam widths of initially infinitesimally small proton beams are plotted. Thus, protons being stopped in point P have a high probability of having started anywhere in the region between x_1 and x_2.

The two processes will first be discussed separately. The resulting uncertainties are added in quadrature.
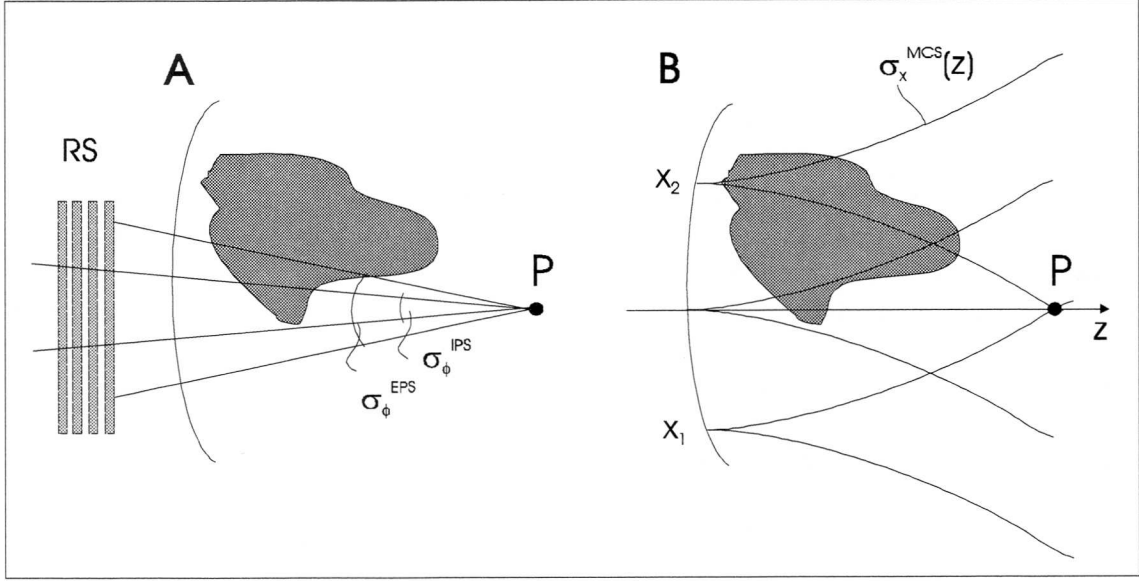
Fig. 16: Two effects contribute to the angular uncertainty in point P:
A. The effective phase space (initial phase space and MCS in the modulator device) causes an uncertainty of the angular distribution of the protons which is constant throughout the patient, but depends on the amount of scattering material in the beam.
B. Multiple Coulomb scattering in the patient causes an increasing angular uncertainty with depth in the patient.

3.2.2.1 The angular displacement of the incident beam

The width of the incident pencil beam used at PSI was measured in air at different distances from the nozzle using a small ionisation chamber (Scheib 1993) and has been parametrised by

(13) $$A_{IPS}(z) = A_0 + 2B_0 z + C_0 z^2$$

The spatial and angular distributions of the protons are assumed to follow a Gaussian distribution. It can be shown (Appendix A), that the standard deviations of those Gaussians are given by:

(14) $$\sigma_x^{IPS}(z) = \sqrt{\frac{A_{IPS}(z)}{2}} \qquad \sigma_\phi^{IPS}(z) = \sqrt{\frac{C_0}{2}}$$

When range shifter plates are introduced into the beam path, beam width and angular spread are increased through MCS in the plates. The same is true for other passive modulator devices, but we will discuss here only the results for the PSI system, which uses range shifter plates (Pedroni et al 1995). The parametrisation can still be written in the form of eq. (13), but with an additional contribution from the scattering in the plates:

(15) $$A_{EPS}(z) = (A_0 + A_{RS}) + 2(B_0 + B_{RS})z + (C_0 + C_{RS})z^2$$

The calculation of A_{RS}, B_{RS} and C_{RS} follows the theory in the appendix.
A_{EPS} parametrises the beam before it enters into the patient. It is called the effective phase space.

For the moment we are only interested in the angular uncertainty of the effective phase space. In analogy to eq. (14) this is given by:

$$(16) \qquad \sigma_\phi^{EPS}(z) = \sqrt{\frac{C_0 + C_{RS}}{2}}$$

Assuming a homogeneous proton beam flux produced by the scanning beams, the angular uncertainty of the protons along the beam axis is independent of the position and of the air gap between range shifter and patient. It depends however on the amount of scattering material in the beam, i.e. the number of range shifter plates (1 plate $\cong$ 0.454 cm water).

This is shown in Fig. 17, where the standard deviation of the angular distribution is plotted as a function of the number of range shifter plates in the beam path (diamonds and squares for the two orthogonal directions x and y). The contributions from MCS in the range shifter plates (+) and the initial phase space in the two directions (dashed and dashed-dotted lines) are plotted separately.

WER of range shifter plates [cm]
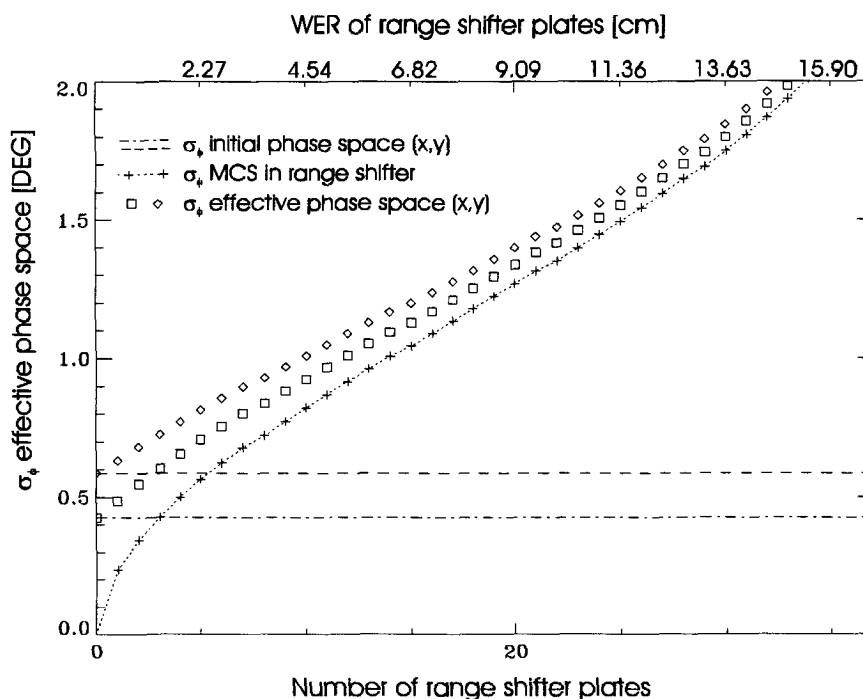
Fig. 17: The angular distribution of the effective phase space as a function of the number of range shifter plates in the beam for two orthogonal directions (x,y) and a beam energy of 177 MeV. To keep the angular uncertainty as low as possible, one should try to use only a few range shifter plates.

3.2.2.2 The angular uncertainty due to MCS in the patient

When a proton enters into the patient, it quickly loses its well defined direction through MCS. The angle θ between nominal and actual direction strongly depends on the depth in the patient and the initial energy of the proton. The probability distribution for the total angular deflections due to MCS can be calculated using the theory in the appendix.
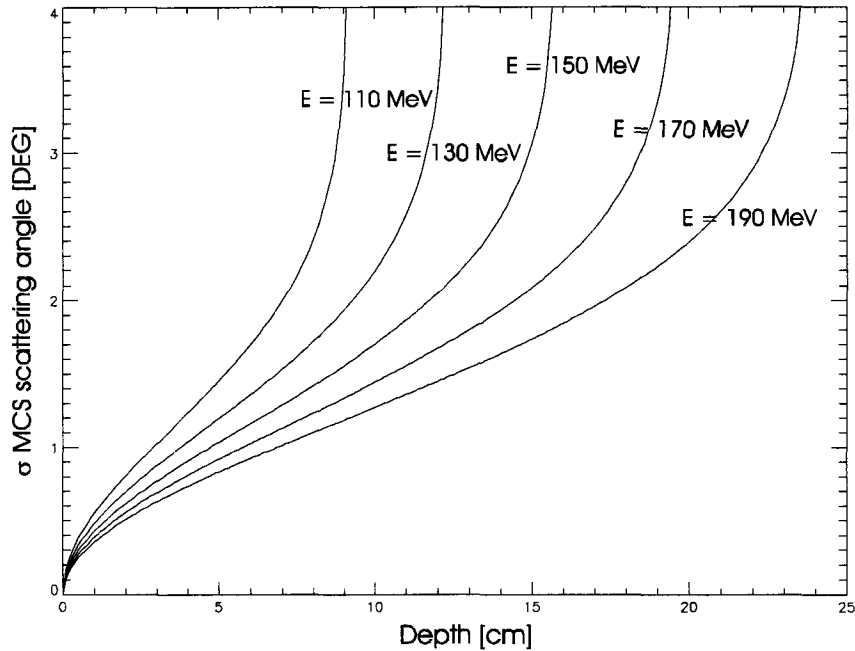
Fig. 18: Width of the distribution of scattering angle θ vs. the depth in water for different energies. The angular uncertainty caused by MCS in the patient increases drastically towards the end of the protons range.

The width of the angular distribution for different energies is plotted in Fig. 18 vs. the depth in water. The quick increase of the angular uncertainty of the protons up to a non defined state at the end of their range makes it difficult to define a mean scattering angle from this graph.
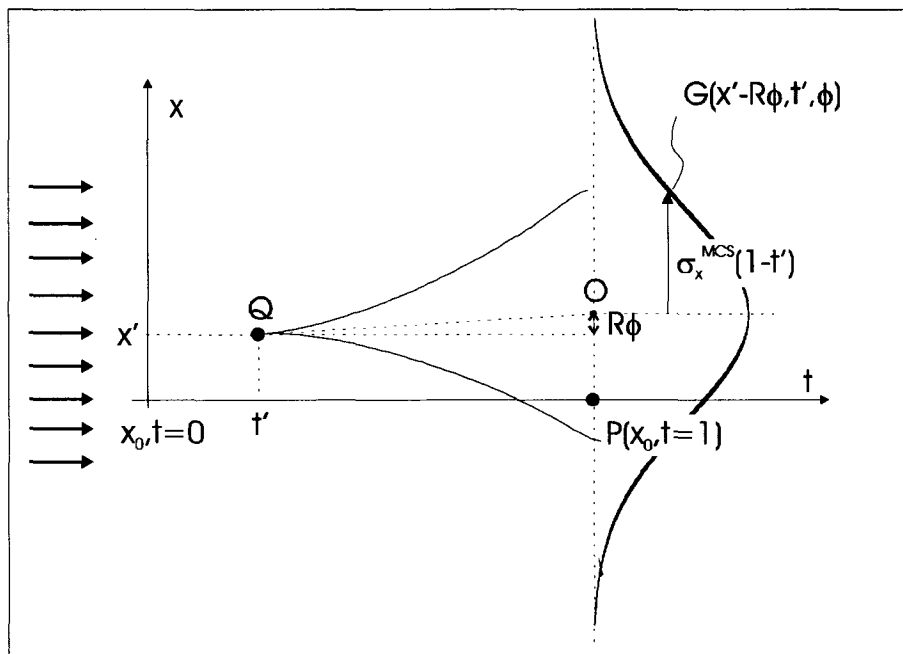

Fig. 19: Propagation of a proton beam from Q at (x', t') with known angular deviation φ to the stopping position at t=1. The probability for a proton from that beam to be stopped in x_0 is given by the Gaussian distribution $G(x'-R\phi,t',\phi)$.

We therefore decided to use a different approach and calculated the probability distribution of the origins of protons stopping at a point P. For simplification we assume

a homogeneous incident proton flux. The probability distribution of the stopping positions of protons passing through a point Q at (x', t') with an angular deviation ϕ from the nominal beam direction is given by a Gaussian with a standard deviation of $\sigma_x^{MCS}(1-t')$ (see MCS theory in the appendix and Fig. 19) centred at point O at (x'+Rϕ, t=1). To obtain the probability distribution independent of the angle ϕ, we multiply this distribution with the probability for the angle ϕ and integrate over ϕ.

Thus:

$$P(x',t') = \frac{1}{2\pi\sigma_x(1-t')\sigma_\phi(t')} \int_{-\infty}^{\infty} e^{-\frac{([x'-x_0]-R_0[1-t']\phi)^2}{2\sigma_x}} e^{-\frac{\phi^2}{2\sigma_\phi}} d\phi$$

(17)
$$= \frac{1}{\sqrt{2\pi}\Sigma} e^{-\frac{[x'-x_0]^2}{2\Sigma^2}}$$

$$where \quad \Sigma^2 = R^2\sigma_\phi^2(t') + \sigma_x^2(1-t')$$

R_0 : nominal range of the protons

$t = z/R_0$: depth relative to the nominal range
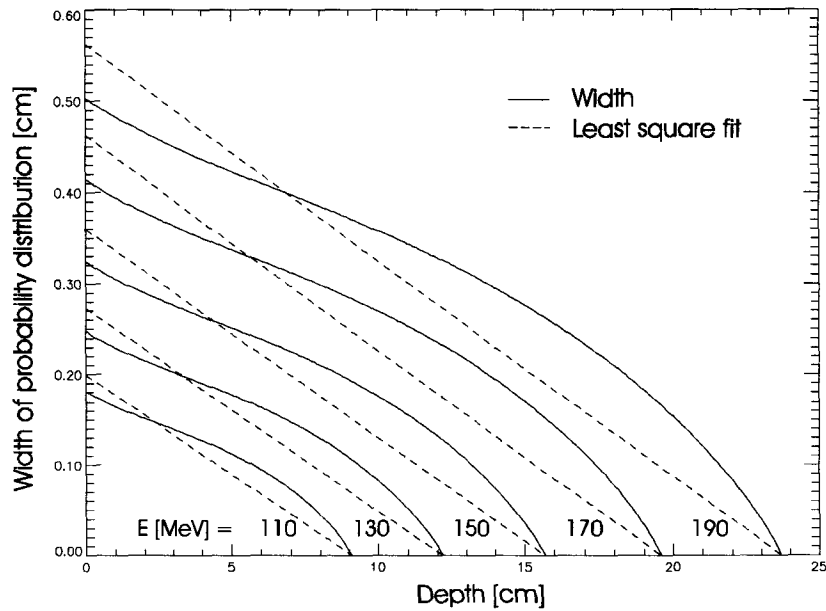
$R = R_0(1-t')$: residual range

Fig. 20: Width of the probability distribution of the origin of protons for different initial beam energies and least squares fits through the stopping point. The gradient of the fits is almost a constant.

The width $\Sigma(t)$ of the probability distribution in water is plotted in Fig. 20 for different energies.

For the estimation of range dilution we want to integrate the density values in the patients body along straight lines lying in a cone. We therefore use a least squares fit with the restriction to go through the stopping point P to approximate the function $\Sigma(t)$ with a straight line. We define the effective scattering angle due to MCS (σ_ϕ^{MCS}) to be the opening angle of the cone given by the linear fit. σ_ϕ^{MCS} has been found to be $1.31 \pm 0.04°$ for all initial energies from 110 -190 MeV. Considering the errors introduced by the

linear fit, the variations from energy to energy can be neglected and the effective scattering angle can be regarded as a constant for all beam energies used in proton therapy.

3.2.2.3 The effective total angular uncertainty

The effective total angular uncertainty (σ_ϕ^{TOT}) is the squared sum of contributions from the initial phase space, MCS in the range shifter plates and in the patient. The total effective angular uncertainty, as plotted in Fig. 21 for two different energies, is therefore changing with the number of range shifter plates. An important consequence from these considerations is that one can reduce the angular uncertainty by choosing the most appropriate initial energy. This is the energy which allows the deposition of the most distal spot in the target without using range shifter plates.
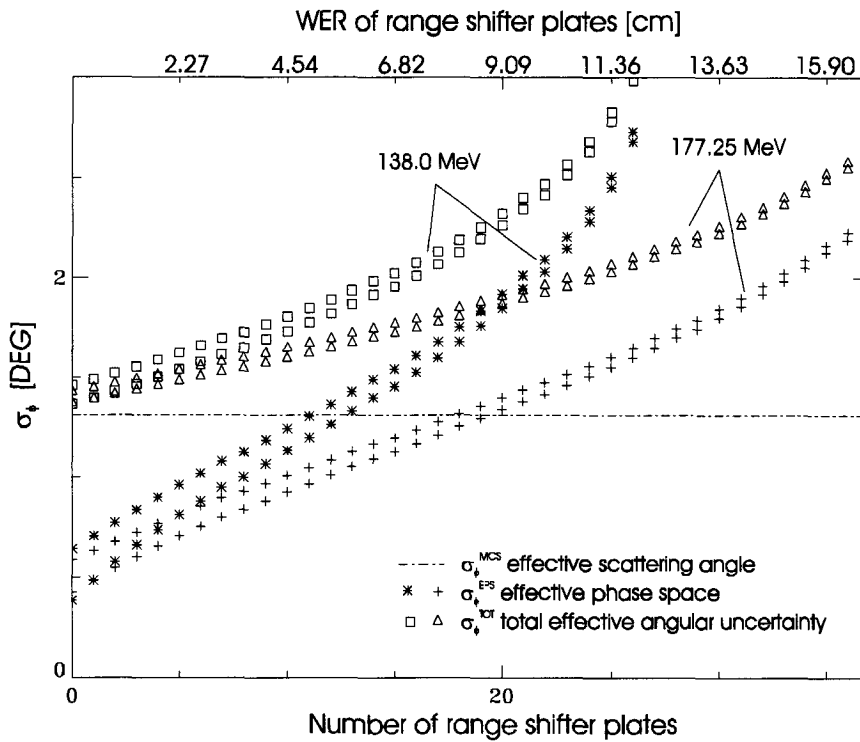
Fig. 21: The effective total angular uncertainty as a function of the number of range shifter plates for two different beam energies. It consists of contributions from the effective phase space and MCS in the patient. The contributions from MCS are dominant for a large number of range shifter settings.

3.2.3 Comparison of the results to proton radiography measurements

The range dilution of protons passing through a heterogeneous medium can be calculated using eq. (11) with the weighting function eq. (8) and the effective total scattering angle σ_ϕ^{TOT}. An experimental check for the correlation between calculated range dilution and actual widening of the proton energy spectrum can be performed by using proton radiography.

The proton radiography system developed by Schneider and Pedroni (1995) allows one to measure directly the range spectrum of protons after passing through a heterogeneous

medium. The rms of the range spectra measured in each pixel of the detector matrix are expected to correlate with the range dilution.

The calculation of the range dilution and proton radiography measurements of the energy spectrum were done for the head of an Alderson phantom. The results are displayed as images with black representing a high degradation and white no degradation. The two images in Fig. 22 - calculation on the left, measurement on the right - show a qualitative agreement between measurement and calculation. Areas with a broad measured energy spectrum can be clearly identified also in the calculation as areas with a high range dilution. They are especially the areas where the beam is tangential to the skull, the oral and nasal cavities and the auditory canal.

Thus the degradation of the proton beam is found to be high in regions with high density gradients parallel to the beam. We will classify heterogeneous structures parallel to the beam as critical heterogeneities. Critical heterogeneities can be identified by using the range dilution calculation.
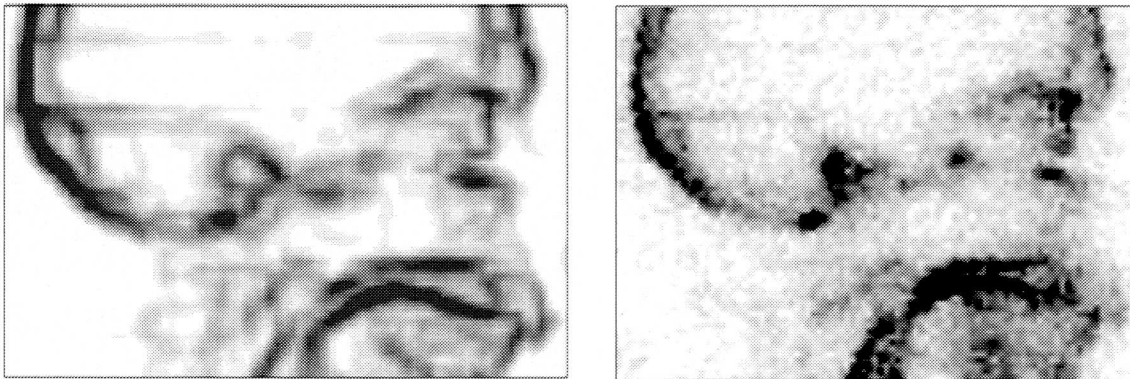

Fig. 22: Comparison of the range dilution calculated for the Alderson head (left) and width of the energy spectrum measured with proton radiography (right). Dark areas indicate in both pictures a high degradation of the proton beam.

3.2.4 Changes in the depth dose curve of a proton beam due to range dilution

The PSI treatment planning software uses a look-up-table (LUT) of the depth dose curve of protons in water, which has been described by S. Scheib (1993) in his thesis. It is characterised by the nominal beam energy and the width of the initial energy spectrum (momentum band). The dose deposited by a single proton of known initial energy is given by the Bethe-Bloch equation (Bichsel 1972). A discrete sampling of the energy spectrum is used to obtain the depth dose deposited by an average proton. An extra component is added to the initial energy spectrum to account for range straggling in the patient. Nuclear interactions are modelled as a flux reduction of primary protons and an empirical model for the deposited dose is used.

The range dilution information can be included in the LUT for the depth dose curve as an additional broadening of the initial energy spectrum. The resulting depth dose curves for a relative range dilution from 0 % to 9 % standard deviation are shown in Fig. 23. The influence of range dilution on the shape of the Bragg curve is very high in the peak region, while there are almost no changes in the plateau. This underlines the importance of the peak region and justifies our approach of calculating the range dilution only for protons stopping in the point of interest.

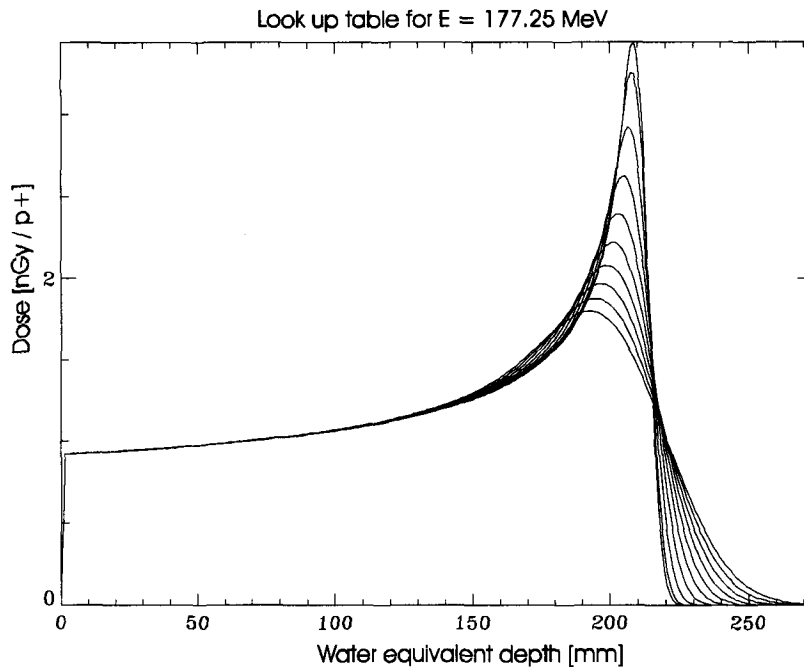Look up table for E = 177.25 MeV

Fig. 23: Depth dose curve for a 177 MeV beam with a momentum band of 1.1% and with a relative range dilution of 0 % - 9 %. The range dilution causes drastic effects in the Bragg peak region, but has almost no effect on the plateau dose.

3.3 Analytical dose calculation models

3.3.1 The need for analytical dose models

The confusion created by MCS on the trajectories of protons traversing complex density distributions is probably the most important effect limiting the precision of analytical dose calculation algorithms in proton radiotherapy. The best solution is to model these effects with Monte Carlo simulations, which become more and more available for the specific needs of radiotherapy treatment planning. Many papers published by various authors (Carlson et al 1997, Garelli et al 1996, Petti 1992 and 1996, Ragona et al 1996, Tourovsky et al 1993) represent the interest in this field. With new powerful machines and codes optimised for speed, they can be as fast as simulating 1 million proton histories in 1.5 - 2 min. on a DEC AXP 500/333 workstation (PSI code written by A. Tourovsky). This is sufficient to simulate a one field plan for head treatments with a good statistical resolution - as those shown later on - in 20-30 min.

Still there are strong reasons to continue to use analytical calculations and improve them also in the future:

- Simple analytical calculations are still considerably faster than MC simulations. This is important during the planning process to get a quick overview on the dose distributions obtained at different gantry angles.

- In body regions without critical heterogeneities a good analytical dose calculation is usually accurate enough.

- The PSI spot-scanning technique requires an inverse optimisation algorithm to compute iteratively the intensities of individual beam spots. For this purpose only an

analytical, spot by spot dose calculation can be used. The quality of the treatment will depend strongly on the precision of this algorithm.

New codes being written for proton dose calculations nowadays use elementary pencil beam models (Petti 1992 and 1996, Hong et al 1996). Their basic feature is the convolution of the proton flux Φ with an elemental pencil beam dose distribution or dose kernel K, which describe the dose as deposited in water by an infinitesimally small proton pencil beam. The dose kernels must include the effects of MCS in the patient, flux reduction and energy deposition through nuclear interactions and ionisation losses of primary protons. Density heterogeneities are taken into account along the central axis of the elemental pencil beam. In scattered beam dose calculations the proton flux Φ is usually taken to be a constant value inside the aperture of the collimators and zero outside. Ideally however, Φ is a function of the position in the patient and the angular distribution of the protons as schematically depicted in Fig. 24.
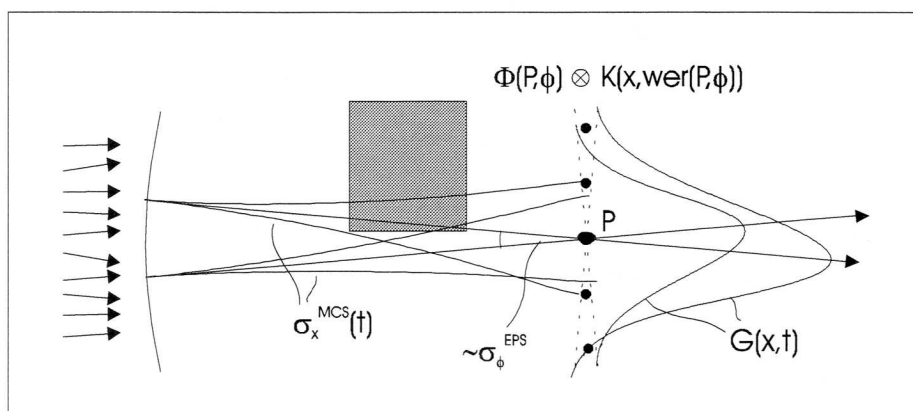

Fig. 24: Sketch of a pencil beam model of the phase space. This model uses look-up-tables of Gaussian shaped elemental pencil beams in water. A scaling of the depth-dose is done along the central axis of the pencil beam to account for differences in the density of the material. Ideally the model would also include the effects of angular uncertainties of the effective phase space.

Analytical pencil beam models have the disadvantage to neglect two important effects:
- Differences in the MCS effects due to different positions in depth of heterogeneous structures cannot be modelled in the dose kernel.
- Differences in the radial spread of the pencil beam caused by high Z materials compared to water are usually not modelled. We will neglect this effect in all our models. A solution to account for them has recently been proposed by J. Deasy (1997).

There is also a practical limitation to the use of pencil beam models for treatment planning for the PSI spot-scanning technique. In order for the flux-dose model to be used as a spot by spot calculation suitable for the optimisation, the proton flux and the convolution with the dose kernel must be calculated separately for each beam spot. This will however require a lot of computing power. Considering that an optimisation typically consists of 100-150 loops, each with a full dose calculation, it is clear that this approach is not practical.

3.3.2 Model 1 - a pencil beam model based on the total proton flux

Despite neglecting effects of the position of heterogeneities and of high Z materials, pencil beam models produce good results in most clinical situations. Therefore a pencil

beam algorithm, which is capable of simulating the PSI spot scanning dose delivery, has been implemented. It cannot be used in the optimisation procedure itself, but for the final dose calculation after the optimisation.

It is easy to imagine that the full implementation of the above mentioned ideal pencil beam model in three dimensions and performed on a spot by spot basis will require too much computing power. We therefore chose to neglect the angular disturbance of the effective phase space, i.e. to integrate the flux spectrum over all angles ϕ. A comparison to Fig. 21 shows that for the PSI gantry the contribution of the effective phase space to the total angular uncertainty is small compared to the effects of MCS, if only a few range shifter plates are used. We aim to use the minimal number of range shifter plates possible to deposit the most distal spots, which contribute the majority of the total dose. Thus, this effect is not expected to introduce large dose errors. Neglecting it allows however a major simplification of the algorithms.

If an iterative spot by spot calculation is not needed, the computing time can be further reduced by summing up the contributions of all beam spots to the total proton flux prior to the convolution with the dose kernel. The characteristics of the kernel depend on the energy of the beam entering into the patient, which is a function of the number of range shifter plates used. Therefore only flux contributions from beam spots applied with the same number of range shifter plates can be convoluted together. The total dose is calculated at the end by a summation of all dose components.

This will be the only model presented here, where flux and dose calculation are performed separately. It is therefore called the flux-dose calculation.

The details of the algorithm are as follows:

Flux calculation

In section 3.2 the concept of the effective phase space has been introduced, where the beam width is parameterised by eq. (15). For the calculation of the proton flux the beam width is given by:

$$(18) \qquad \sigma_i^{EPS}\left(z, N_{RS}\right) = \sqrt{\frac{A_{EPS}\left(z, N_{RS}\right)}{2}} \qquad i = x, y$$

Since the effective phase space includes initial phase space and contributions from the range shifter plates, it is a function of the number of plates (N_{RS}) in the beam path.

The flux at any position in space for a single Gaussian shaped spot follows straight forward:

$$(19) \qquad \Phi\left(x, y, z; N_{RS}\right) = N_{p^+} \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x_0-x)^2}{2\sigma_x^2}} e^{-\frac{(y_0-y)^2}{2\sigma_y^2}}$$

N_p^+ : number of protons deposited by the spot, also called beam intensity or weight

x_0, y_0 : position of the central axis of the pencil beam

σ_x, σ_y : beam width in x and y direction due to the effective phase space

Dose kernel

A dose kernel as deposited by an infinitesimally small proton beam, which hits the patients surface is calculated using a Gaussian approximation. Thus:

$$K(x, y, wer; N_{RS}) = ID(wer)\frac{1}{N}e^{-\frac{(x_0-x)^2}{2\sigma^2}}\, e^{-\frac{(y_0-y)^2}{2\sigma^2}}$$

where $\quad wer = wer(x_0, y_0, z)$

(20)

$$\sigma = \sigma_{MCS}\left(wer\_pat(x_0, y_0, z); N_{RS}\right)$$

$$N = dx\, dy\sum_{x,y} e^{-\frac{(x_0-x)^2}{2\sigma^2}}\, e^{-\frac{(y_0-y)^2}{2\sigma^2}}$$

ID : integral dose, taken from the depth-dose LUT

wer : total water-equivalent-range along the central axis of the kernel

wer_pat : water-equivalent-range in the patient

dx, dy : size of the dose calculation grid in x and y

N   : normalisation factor

Contrary to the models presented later on, the beam width for the calculation of the kernel is given only by the width due to MCS in the patient.

Because the pencil beam is infinitesimally small at the surface, σ_{MCS} is zero and the usual normalisation factor $2\pi\sigma^2$ has to be substituted by N. This normalisation guarantees the numerical integral dose in each slice to be correct.

Total dose

From the undisturbed proton flux $\Phi(x,y,z;N_{RS})$ and the dose kernel $K(x,y,wer;N_{RS})$, the total dose for each range shifter setting is calculated by a convolution between the two. Since the calculation is done on a discrete grid, the integrals in the convolution become sums:

$$D(x, y, z; N_{RS}) = \sum_{\tilde{y}=-p_y}^{p_y} \sum_{\tilde{x}=-p_x}^{p_x} \Phi(x - \tilde{x}, y - \tilde{y}, z; N_{RS})\, K(\tilde{x}, \tilde{y}, wer(\tilde{x}, \tilde{y}, z); N_{RS})$$

(21)

where $\quad p_x = \dfrac{4\sigma_{MCS}}{dx}\quad, p_y = \dfrac{4\sigma_{MCS}}{dy}$

Proton flux, dose kernel and final dose distribution obtained from this model for a simple test case are displayed in Fig. 25. A 2 cm thick slab of a relative stopping power of 1.8 (~ cortical bone) is immersed in water. A single scanning pencil beam passes along the interface between the slab and water. For this and the following calculations an initial beam energy of 177 MeV, an air gap of 10 cm and 15 range shifter plates ($\cong$ 6.8 cm of water) were chosen.
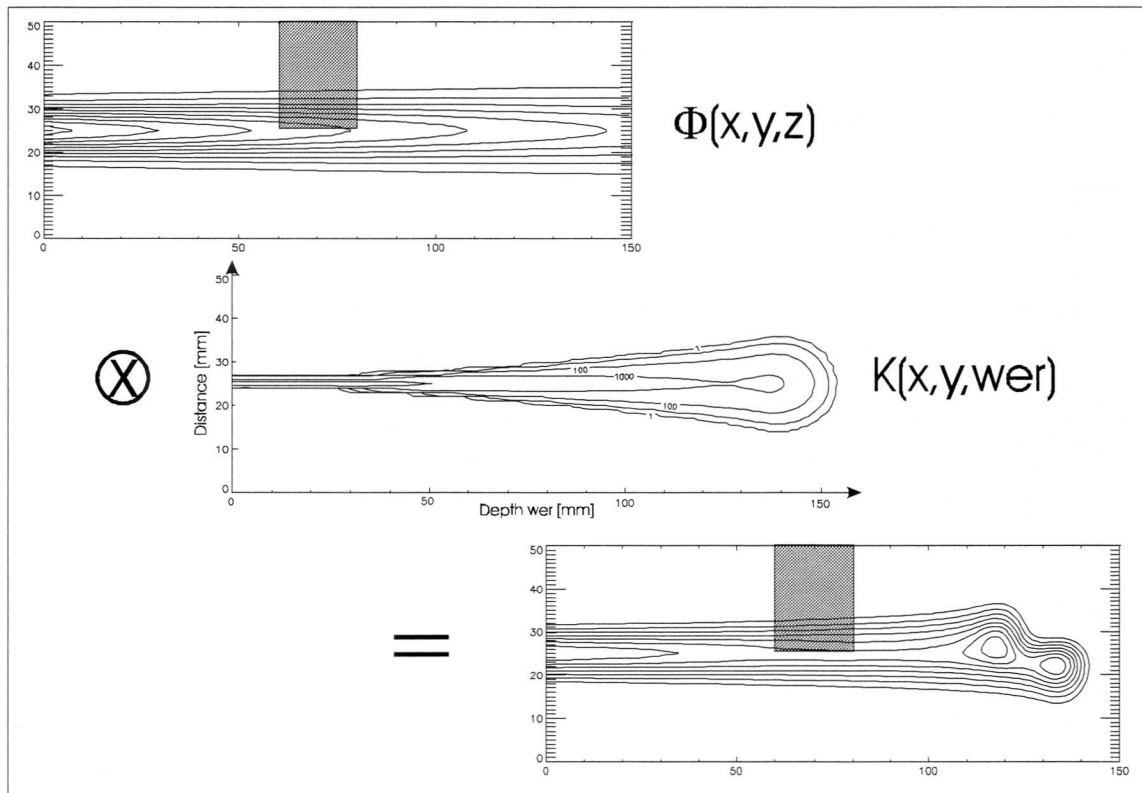
Fig. 25: Flux distribution, dose kernel and final dose as a result of the flux-dose calculation for a beam spot in a test phantom. The final dose distribution is calculated as the convolution between the undisturbed flux distribution and the dose kernel. Note that a logarithmic scale was used for the contour lines of the kernel. The kernel was calculated on a discrete grid of 1 mm^3 which causes the ripples in the contour.

Results of the MC calculation (Fig. 26) are taken as reference values. A comparison between the two results shows that the flux-dose calculation models the situation quite well. Yet the shadowing effect from the slab in the peak region was found to be underestimated. This is not obvious from the plots here, but can be clearly seen when plotting the differences between MC and analytical results (see also section 3.3.6).
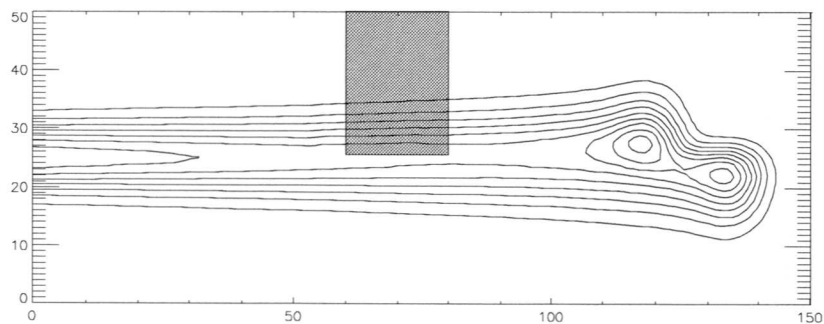
Fig. 26: A Monte Carlo simulation was performed for the same situation as in Fig. 25. The result looks quite similar to the one of the flux-dose calculation.

3.3.3 Model 2 - the scanning beam model

Due to the need for a spot by spot dose calculation, the original dose model implemented at PSI (Scheib 1993) consisted of a superposition of individual scanning pencil beams, which are based on measurements in water. The scanning pencil beam can be regarded as the convolution between the proton flux of a single physical pencil beam and the dose kernel. Only heterogeneities along the central axis of the scanning pencil beam are taken into account in the calculation of the dose for the whole spot application.

This model was developed with the assumption that the phase space is small compared to the broadening of the beam due to MCS. If the phase space was zero, it would be the same as the elemental pencil beam model. In practice, however, the beam width due to the effective phase space can be quite large, especially when a large number of range shifter plates and a large air gap have to be used (see also section 3.2.2). MC simulations therefore showed, that the resulting dose distributions from the scanning beam model were not accurate enough in the presence of critical heterogeneities.
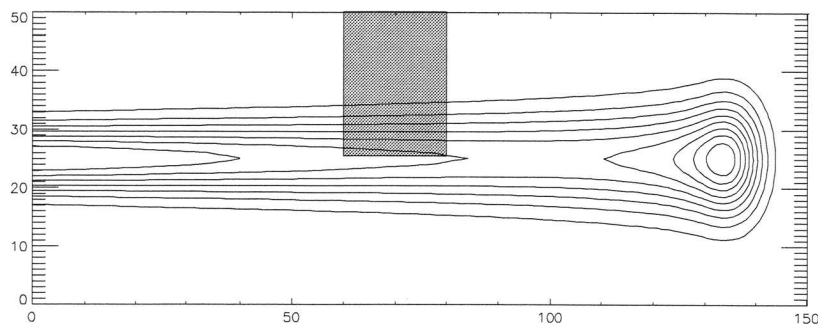

Fig. 27: Result of the original scanning beam dose calculation (model 2) for the test situation. The high density slab close to the beam axis has in this specific situation no influence on the dose distribution, since it is not 'seen' by the algorithm.

The problem with model 2 can be demonstrated clearly when comparing its results with the results of the MC simulation for the test situation. Since model 2 uses only the density information along the central axis of the beam, the effect of the off-axis slab is not modelled. A comparison to the MC result shows that there are large dose errors in the shadow of the slab. New algorithms - based on the scanning pencil beam model - were therefore evaluated to account better for heterogeneities (sections 3.3.4 and 3.3.5).

3.3.4 Model 3 - a simple ray casting model specialised for spot-scanning

For the spot-scanning treatment planning we need a fast spot by spot dose calculation, which produces results comparable in accuracy to the flux-dose model. We therefore modified model 2 to account better for off-axis heterogeneities. Instead of scaling the whole broad pencil beam in depth according to the density along the central axis, we scale the depth-dose distribution along parallel rays independently from each other. Thus the variation in the energy loss of the protons due to heterogeneous structures is modelled as a projection of the density. However, we neglect all disturbance of the dose distribution behind the interface due to angular uncertainties in the phase space and MCS.The dose in each dose grid point is calculated by:

$$D(x, y, z) = N_{p^+} ID(wer) \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x_0-x)^2}{2\sigma_x^2}} e^{-\frac{(y_0-y)^2}{2\sigma_y^2}}$$

(22) $\quad$ *where* $\quad$ $wer = wer(x, y, z)$

$$\sigma_i = \sigma_i^W(wer(x, y, z)) \quad with \quad i = x, y$$

σ_x, σ_y : beam width in x and y directions; it includes all contributions to the beam width in the homogeneous case, i.e. initial phase space, MCS in the patient and range shifter and the air gap between range shifter and patient.

As expected, this dose calculation leads to a very sharp edge in the shadow of the slab in the test case due to the absence of all disturbing effects (Fig. 28). This dose distribution is clearly unphysical. Yet the algorithm is a clear improvement compared to the scanning beam model (Fig. 27).
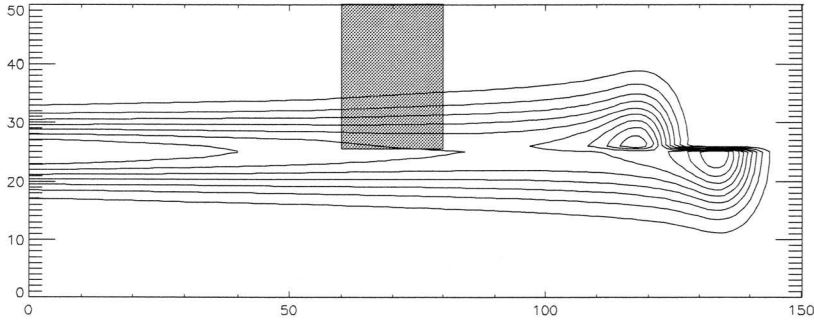
Fig. 28: The result for the simple ray casting dose calculation. Due to the absence of any disturbing effects, the shadow behind the slab is very sharp.

3.3.5 Model 4 - a ray casting model including a correction for range dilution

Behind an interface a broadening of the energy spectrum of the protons occurs, which can in a good approximation be modelled as a dilution in the residual range of the protons along different trajectories. These effects have been discussed in section 3.2.1. For a further improvement of the dose model, range dilution - i.e. the effects of MCS on the range and the angular spread in the effective phase space - has been included in the algorithm.

Relative range dilution eq. (12) can be simulated as a contribution to the width of the initial energy spectrum of the beam. The resulting depth dose curves have been shown in Fig. 23. The single LUT used so far for the depth-dose relationship is therefore substituted by ten LUTs with the ten curves shown in Fig. 23. Instead of the nominal WER eq. (9) along each ray, the mean WER eq. (10) has to be used. The change in beam width due to a broader initial energy spectrum is negligible.

With these two changes to the algorithm the formula for the dose deposited by one spot becomes:

$$D(x, y, z) = N_{p^+} ID\left(\overline{wer}, \sigma_{RU}\right) \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x_0 - x}{2\sigma_x}\right)^2} e^{-\left(\frac{y_0 - y}{2\sigma_y}\right)^2}$$

(**23**)   *while* $\overline{wer} = \overline{wer}(x, y, z)$

$$\sigma_i = \sigma_i^W\left(\overline{wer}(x, y, z), \sigma_{RV}\right) \quad with \quad i = x, y$$

This refined ray casting model was again used to calculate the dose distribution for the test situation (Fig. 29). The sharp edge observed in the dose distribution obtained from model 3 is smoothed due to the range dilution correction.
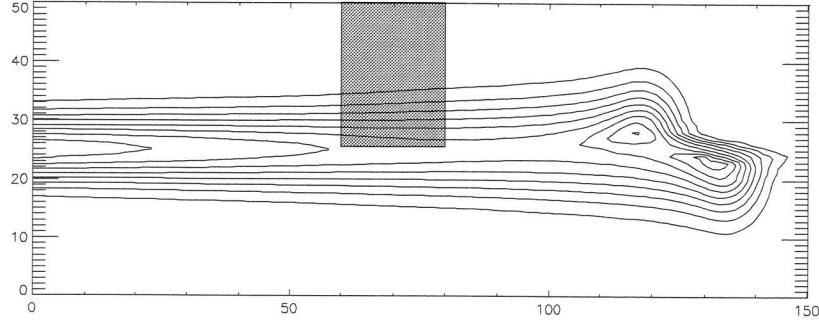
Fig. 29: Result of the refined ray casting model. The sharp edge obtained by the simple ray casting model (Fig. 28) is smoothed through the range dilution correction.

The cost in terms of computing time for the range dilution correction is relatively low. The calculation of the mean range and the range dilution correction takes considerably more time than the nominal range calculation, which is used in model 3. The most time consuming process of the planning is however the optimisation procedure, which takes exactly the same time in both models.

3.3.6 Comparison between the analytical algorithms for simple geometries

Comparison between the ray casting models with and without range dilution correction

The easiest way of comparing the results of different models is the comparison of the total squared differences (Δ^2_{sq}) between the results of the analytical calculations and of the MC simulation. When total squared differences between the MC simulation and the analytical dose calculations are calculated, a considerable improvement from the simple to the refined ray casting model is found. Maximal differences range from 40-60% of the peak dose for model 3 (M3) and 20-40% for model 4 (M4). The maximal differences seem to be very large, but they appear only in a few voxels (0.1 x 0.1 x 0.1 cm resolution), as will be shown in histograms later on. These errors were determined by studying different situations, such as different numbers of range shifter plates, different air gaps and, most interesting, variations in the position of the slab in depth.

The position in depth of the high density slab has a strong influence on the whole dose distribution. Yet we don't know of any other analytical model besides our refined ray casting model, which accounts for the difference in depth of heterogeneities.

Due to the sampling of angularly displaced rays, the area affected by the range dilution and integral density corrections increases with the distance from the slab. This effect is

shown in Fig. 30 where the dose contours for two extreme positions of the slab (right below the surface of the phantom and before the peak) are overlaid. The contours are shifted along the beam axis for matching the slab positions. Plot A is for model 3 , plot B for model 4. In the area between the two dashed lines the effect of the range dilution correction in model 4 can be clearly seen. The step-like dose distribution in the range dilution corrected calculation has its origin in the discrete sampling of the representative rays. This effect disappears when the slab is closer to the peak or if the density distribution is smoother.
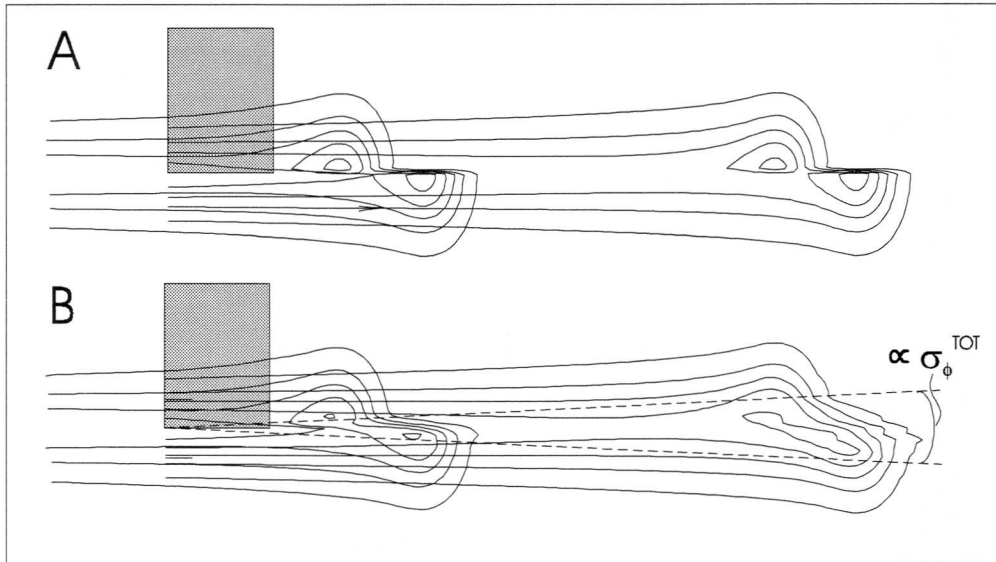
Fig. 30: The effect of the position of the slab in depth on the results of the two ray casting dose calculations.
A: Simple ray casting model; the result is independent of the position
B: Range dilution corrected ray casting: the result depends on the position

Comparison between the flux-dose model and the ray casting models

The main differences between the models are due to the different treatment of inhomogeneities in the patients body. These differences are evaluated for the following situations. In an otherwise homogeneous water phantom a 2 cm thick high density slab is placed on different positions in depth for the different simulations. The central axis of the beam passes along the interface. A beam energy of 177 MeV, no range shifter plates and no air gap have been chosen.

In Fig. 31 the Δ^2_{sq} are plotted as a function of the position of the slab. This graph shows clearly the improvement from the simple ray casting model (M3) to the refined ray casting model using the range dilution correction (M4). It also shows how much the accuracy of the flux-dose calculation (M1) depends on the position of the slab.
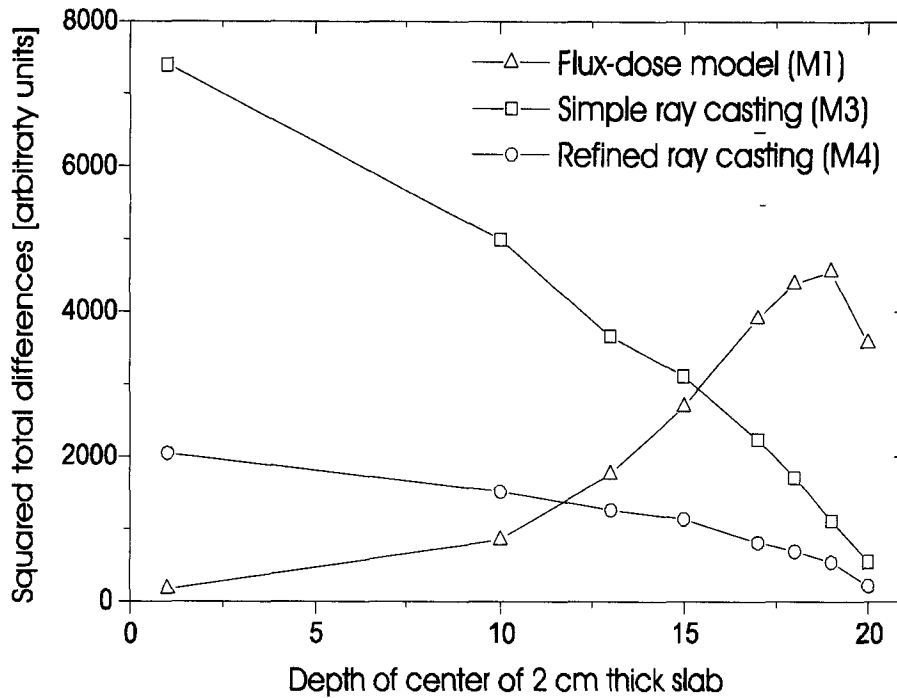
Fig. 31: Total squared differences between analytical models and MC simulation as a function of the position of a high density slab. The initial energy is 177 MeV, which corresponds to a nominal range of 21 cm. Clearly the accuracy of the ray casting models increases the closer the slab position is to the Bragg peak. The difference between the two ray casting models is due to the range dilution correction. The flux-dose model on the other hand gives best results when the slab is positioned far from the Bragg peak.

More detailed information about the quality of the results and the different properties of the dose models can be obtained by looking at histograms and images of differences between the results of analytical calculation and MC simulation.

Fig. 32 shows difference images for two extreme slab positions. When the slab is directly below the surface (position 1 cm in Fig. 31 and Fig. 32 case A), the flux-dose calculation (M1) leads to an almost perfect result. In the Bragg peak the differences between the results of the MC simulation and the flux-dose calculation (M1) do not exceed 3 % of the peak dose. Errors larger than 5 % occur only directly behind or along the interface. When the slab is very close to the position of the peak (position 18 cm, case B), the scattering is however overestimated by M1 because the shadowing effect of the slab cannot be modelled. This causes dose errors, which are almost as bad as those from the uncorrected ray casting technique (not shown) for a slab positioned at the surface. The refined ray casting (M4) models better the shadowing of the inhomogeneity close to the Bragg peak, but underestimates the effects of increased MCS towards the end of the proton range.
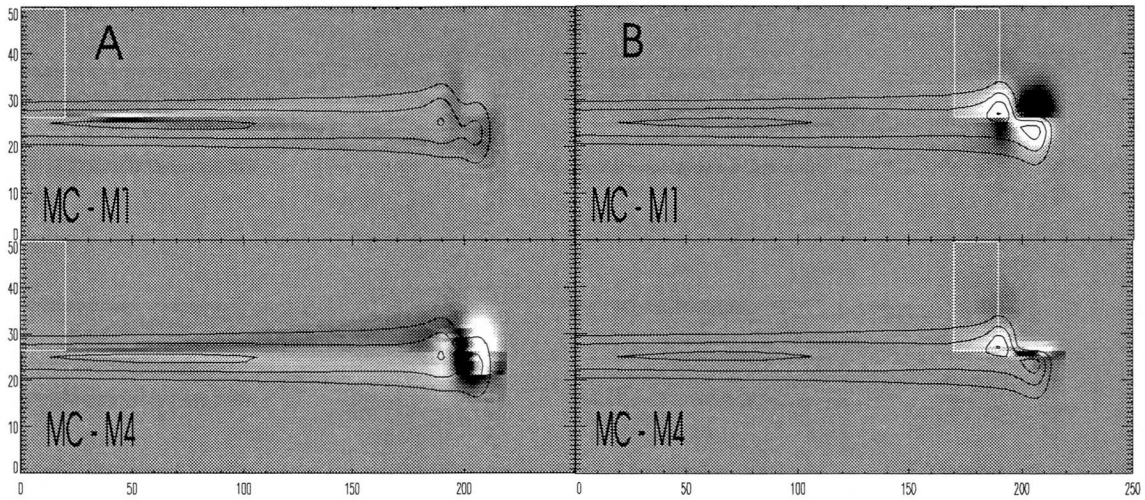
Fig. 32: Differences between MC results and analytical dose calculation for models 1 and 4. In the black areas the MC result is >10% lower than the analytical calculations, in the white areas it is >10% higher. The two images left (A) show the situation with the high density slab below the surface, in the images on the right side (B) the slab is positioned directly in front of the peak. The contours are taken from the MC dose distribution.
Note that the scaling is not the same in x and z!

Fig. 33 shows histograms for the two extreme cases A and B. The large deviations between simple ray casting algorithm (M3) and MC simulation in case A are caused by the sharp edge in the dose distribution. In that situation where the slab is directly below the surface of the phantom, MCS causes a large uncertainty in the average trajectory of the protons along the residual path. The range dilution correction (M4) can account for some of the effect, which reduces the amount of voxels with very high errors. The results for the flux-dose model (M1) confirm the statements made so far, i.e. the good result in case A and the large amount of voxels with dose errors above 10 % in case B.
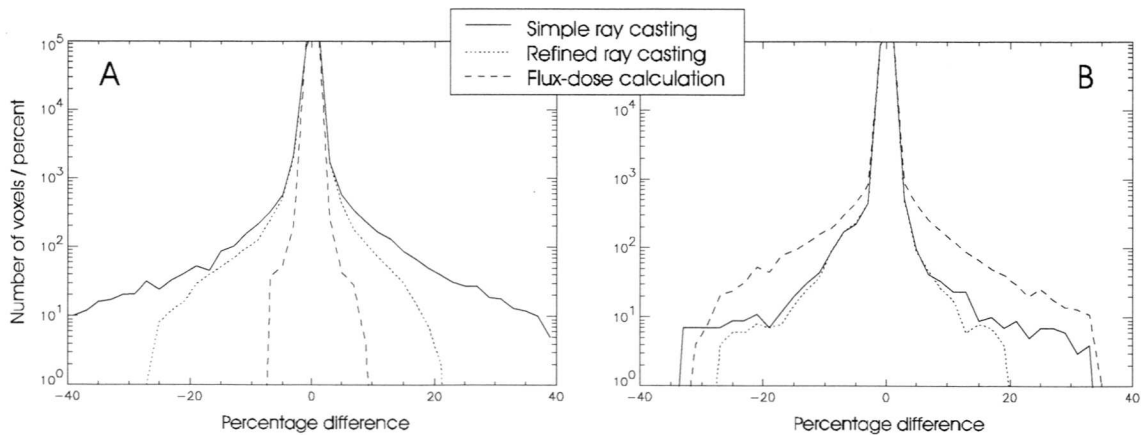

Fig. 33: Statistical analysis of differences between analytical calculation and MC simulation for two extreme positions of the slab. The two cases correspond to those shown in Fig. 32. This shows again that the flux-dose model is more accurate in case A, while the ray casting models are superior in case B.

These comparisons show two things clearly:

- The ray casting model is an improvement compared to the scanning beam model and can be further improved by applying a range dilution correction.

- The flux-dose calculation is superior to the ray casting the closer heterogeneous regions lie to the surface of the patient, but inferior the closer they are to the target. This indicates that it is important to used both models and compare their results.

The simulation of collimators presents a situation similar to one with the slab at the surface of the phantom. The modelling of collimators is therefore best done by using the flux-dose model. Collimators are needed in scattered beam techniques to conform the dose laterally to the target. At PSI there are plans to install optionally collimators (Pedroni et al 1995) in addition to the conformation by dynamic scanning to improve the lateral dose fall-off in the treatment of more superficial tumors. The need for a possibility to simulate the effects of a collimator was one additional reason to develop the flux-dose model.

3.3.7 Comparison of algorithms in anatomical cases

A comparison of the dose calculations for clinical cases can give more information on the preferred use of one or the other model in specific situations.

During the treatment planning process dose calculations are used in two different situations. In the optimisation loop only models based on a spot by spot calculation can be used. The models satisfying this requirement are the scanning beam model (M2) and the two ray casting models (M3 and M4). For the final dose calculation any model - including the flux-dose calculation (M1) and the MC simulation - can be used.

Improvements achieved by using the simple ray casting model instead of the scanning beam model in the optimisation loop

In a first comparison the results of the optimisation procedure using the scanning beam model are compared to those using the simple ray casting model.

The case selected for this comparison is a meningioma.

The chosen gantry angle of -40° is not ideal for a real treatment for three reasons:

- The left eye and optic nerve lie in the beam path.

- The alignment of the beam along the long thin extension of the tumor makes the case very sensitive to positioning errors.

- A relatively long bone-brain interface is aligned with the beam direction.

The last reason for not choosing -40° as a gantry angle make this a very good example to demonstrate the effects of beam parallel density heterogeneities on the dose.

The planning on this case was done using a beam energy of 138 MeV. The optimisation was done on a calculation grid with the same resolution as the grid given by the spot positions, i.e. the dose calculation points were aligned on the central axis of the scanning pencil beams. The separation between the spot positions along the table axis is usually chosen to be equal to the thickness of the CT slices. In this case a grid of 0.5 x 0.5 x 0.5 cm was chosen.

The two algorithms were used to create a beam data file. This contains the prescription for the lateral position, the number of range shifter plates and the weight of all spots, which were obtained from the automatic optimisation procedure. MC simulations of the dose delivered by the optimised beam data set were performed afterwards to evaluate the quality of the results of the analytical calculations.

The results of the optimisation using the two analytical dose calculations (M2 and M3) are shown in Fig. 34 along with the corresponding results of the MC simulations. In picture A2 one can recognise clearly the cold spot, which appears in the result of the MC simulation in the distal end of the upper part of the target. The dose calculated using the scanning beam model (A1) does not show this cold spot and the algorithm can therefore not compensate for it in the optimisation loop. The pictures below (B1 and B2) show the same situation using the ray casting model instead of the scanning beam model in the optimisation. Since the dose calculation using the ray casting model is more accurate, the algorithm can account for the effects of the inhomogeneities by weighting the pencil beams in that region differently. The final results of the analytical dose calculation and the MC simulation coincide much better in this case.

Fig. 34: see appendix (p. 80)

Intercomparison of the accuracy of the models 1, 3 and 4

The three improved analytical models were compared among each other using the example of the meningioma shown above, a chordoma and a prostate case.

As a first test the three models are used to calculate the final dose distribution for the meningioma case using the beam data obtained by M2 (Fig. 34 A). They all predict the cold spot correctly (Fig. 35), which has been missed by M2. As expected from the results in section 3.3.6, the range dilution correction (M4) smoothes the sharp edges in the dose distribution obtained by the simple ray casting (M3), but the effect is not very large. In the flux-dose calculation (M1) the scattering effect along and directly behind density interfaces is overestimated. This shows up along the bone-brain interface indicated by the arrow, where the flux-dose calculation predicts a narrow area, which obtains less than 90% of the nominal dose. This dip in the dose distribution also appears in the MC results, but is less pronounced.
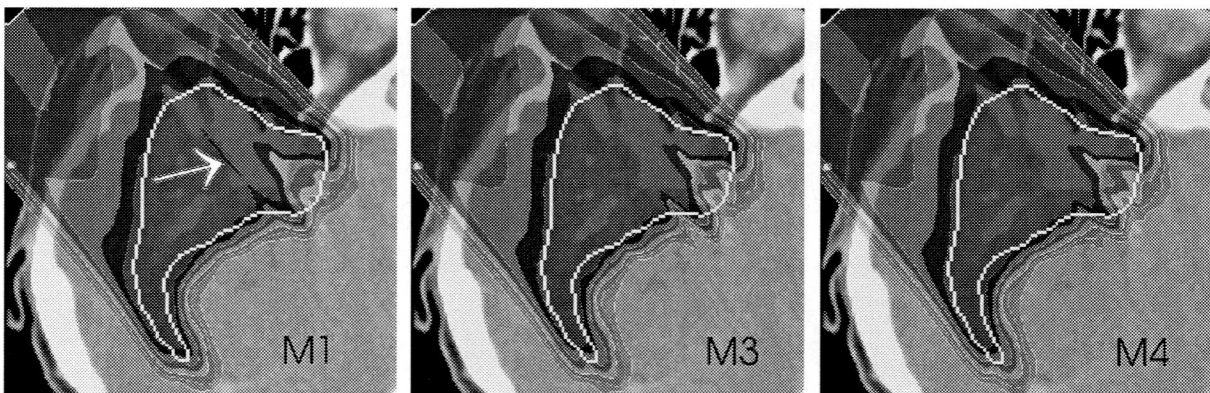

Fig. 35: Dose predicted by the improved analytical models for the situation simulated in Fig. 34 A. The models used are (from left to right) the flux-dose calculation, the simple ray casting and the refined ray casting. The resolution of the dose grid was doubled for all these calculations to 0.25x0.25x0.25 cm. The results obtained from the three models are very similar, except for the narrow low dose area which appears in M1 (arrow).

Differences between the results of all three models are however not very large. A more detailed analysis has been done looking at differences between the results of these analytical calculations and the MC simulation. The comparison in Fig. 36 shows images of these differences. All these calculations are based on the beam data obtained from the ray casting optimisation M3 (Fig. 34 B).
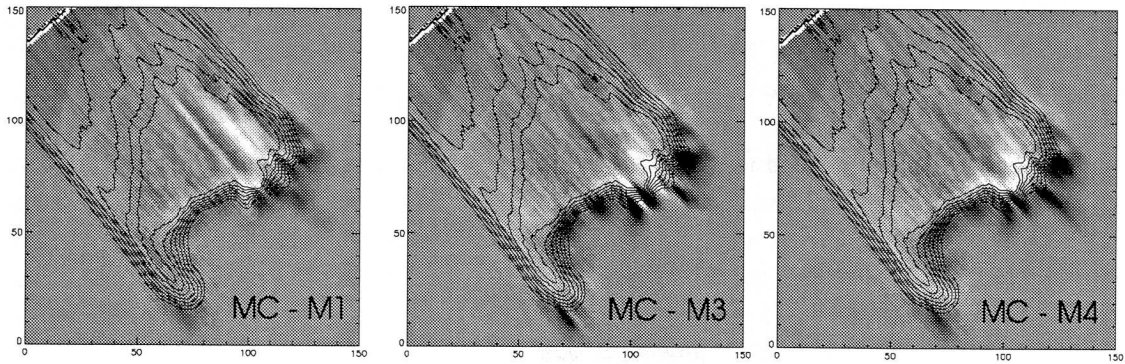
Fig. 36: Differences between results of MC simulation and analytical dose calculations. Black represents an area where the MC predicts a 10 % (or greater) lower dose than the analytical calculation, white a 10% higher dose (or greater). A linear scaling is applied to the grey scale between ± 10 %. Relatively high errors can be found for the flux-dose calculation along the bony interface in the target region. In the ray casting models highest errors are found along the distal border of the target.

These difference-images show again what has been observed in the simulation of the geometrical case. Largest errors occur for the ray casting models along the distal border of the target, i.e. in regions where most dose is delivered by the peak of the Bragg curve. However, that portion of the dose, which is delivered by the plateau of the Bragg curve, is modelled well even in heterogeneous density regions. The flux-dose calculation does better in regions past the inhomogeneity, even in the region of the Bragg peak, but models the dose badly close to an interface.

Overall, the flux-dose calculation models this situation best, which can be seen looking at the histogram of differences in Fig. 37. In the difference histograms of a prostate case with a beam planned to come lateral through the hip, overall a considerable improvement can be achieved using the flux-dose calculation, since the position of the inhomogeneity lies at some distance from the target. In a chordoma case however, where a beam was planned through the nasal cavities, the performance of all analytical algorithms is more or less the same.
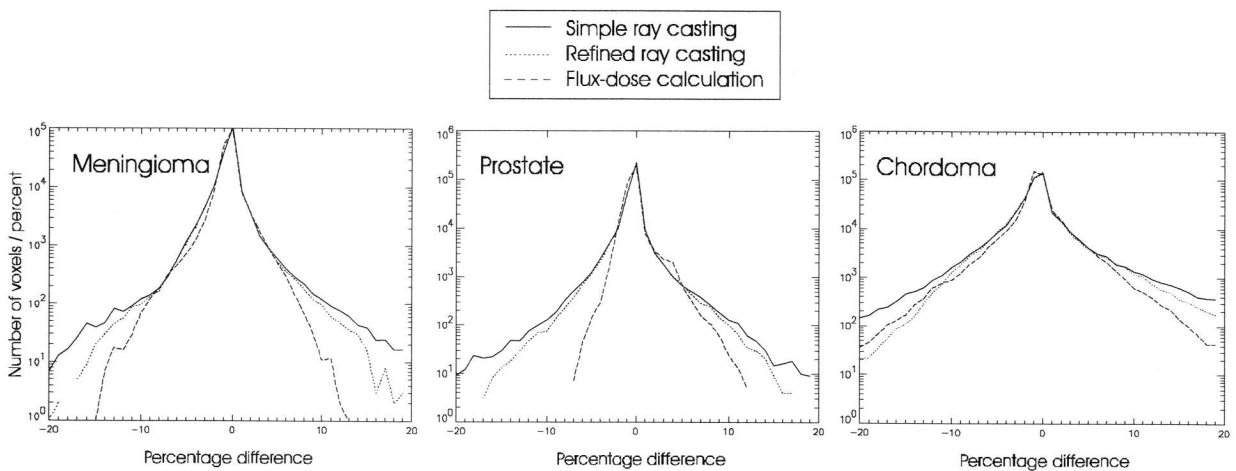

Fig. 37: Histogram of differences between MC simulation and the three analytical models for the meningioma case, a chordoma and a prostate case. In clinical cases the flux-dose calculation is usually superior to the ray casting models.

The images and histograms of differences represent differences in the dose calculations throughout the whole volume. They do not necessarily coincide with the situation in the target. Looking at the pictures in Fig. 36 one observes that the areas where the analytical calculations - especially both ray casting calculations - predict too low doses mainly lie inside the target, those where the predicted dose is too high past the distal edge. The reason for this is that the analytical calculations tend to underestimate the scattering of the protons in materials denser than water. Through the extra scattering, the projection of the paths of the protons on the beam direction in reality and in the MC simulation becomes slightly shorter than when it was calculated from the straight line approximation in the analytical calculations.

Tab. 3 shows a statistical analysis of the dose errors in the target volume. The percentage of the target volume, where the analytical dose calculations underestimate the dose is usually much larger than the volume which receives less dose than expected. The relatively large percentage of underestimation of the dose obtained from the flux-dose model is due to the inaccuracy of the algorithm along density interfaces.

	Meningioma		Prostate		Chordoma	
	underest.	overest.	underest.	overest.	underest.	overest.
flux-dose (M1)	4.7 %	4.9 %	1.0 %	0 %	29.2 %	10.6 %
ray casting (M3)	6.8 %	0.5 %	11.4 %	0.1 %	32.0 %	5.2 %
refined ray cast. (M4)	6.6 %	0.5 %	12.2 %	0 %	29.3 %	5.2 %

Tab. 3: Percentage of target volume, where the analytical dose calculation algorithms underestimate or overestimate the dose by more than 5 % evaluated for three selected cases. Especially the ray casting calculations tend to underestimate the dose in the target volume rather than to overestimate it.

For a good tumor control probability it is crucial that the whole target receives a certain minimal dose; higher doses at single positions in the target can be accepted more easily (Goitein and Niemierko 1996). In the same way it is better to overestimate the dose in the normal tissue distally of the target. The errors in the results obtained by the ray casting dose calculations therefore belong to the less serious errors. Care must however be taken to ensure that the distal border of the target is covered adequately.

3.4 Range dilution and dose errors

3.4.1 Accuracy of the dose calculation in the presence of inhomogeneities

The analytical dose calculation models presented in the previous section have their weaknesses and strengths in modelling the dose distribution in different anatomical situations. None of them is the algorithm of choice for all situations, but the ray casting and flux-dose models are all reasonably accurate in most cases. The areas where largest dose errors between analytical dose calculation and MC simulation are found, are the same in all these models. They coincide with areas of large range dilution. Fig. 38 shows the relative range dilution for the meningioma case discussed earlier. They appear in the same areas as the highest dose errors of Fig. 36.

If one can select gantry angles in such a way that the predicted range dilution is minimal in the target, one can expect to get a good description of the dose distribution using an analytical dose calculation. We thus conclude that the best method to avoid inaccuracies

in the analytical prediction of the dose distribution is to avoid critical heterogeneities with interfaces parallel to the beam.
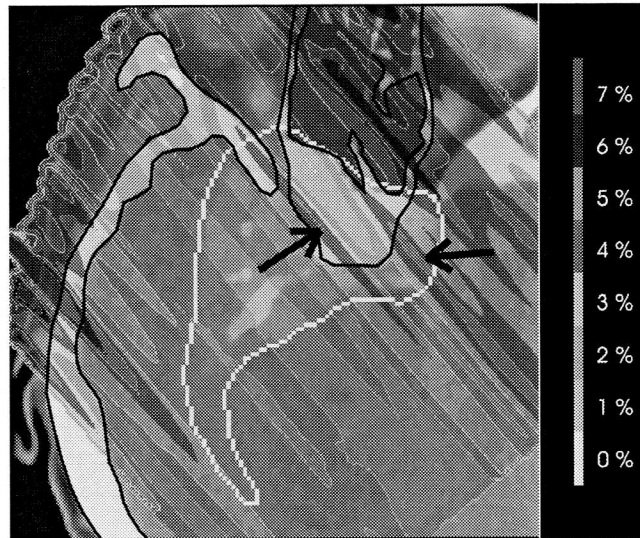

Fig. 38: Estimated relative range dilution for the meningioma case. The range dilution values are overlaid in grey scale on the CT image. The relevant bony structures are enhanced for a better display. It can be seen that highest range dilution (arrows) is found in the same areas as the largest dose differences in Fig. 36.

3.4.2 Dose errors due to set-up uncertainties

With protons it is possible to simulate a good conformation of the dose to the target volume. To achieve this conformation in practice a good accuracy of the daily set-up of the patient during a fractionated treatment is required.

The precision of repositioning of the patient each day is limited to a few millimetres. Assuming that the position errors happen randomly in the two directions perpendicular to the beam, they can be described by a Gaussian distribution. Magnitude and location of total dose errors after completion of the treatment can therefore be simulated by convoluting the width of the initial beam with the Gaussian describing the set-up errors. The dose calculation (analytical or MC) is then repeated with that widened beam. This calculation does not take any movements of the patient during treatment into account.

For the following example a standard deviation of 3 mm set-up errors in both lateral directions was chosen. This is more than what we accept for a head treatment, but represents a worst case analysis. MC simulations were done for both a treatment with perfect set-up and a treatment with random shifts. For the analysis, the differences between the two dose distributions were calculated on a voxel by voxel basis.

Differences between the planned dose and the dose actually delivered during a fractionated treatment of the meningioma are found in two main areas. The most obvious area is close to the lateral margin of the target, where the dose is smeared in the lateral fall-off region. Largest errors can be obtained when the target is very narrow with the small dimensions perpendicular to the beam direction. In the case of a -40° gantry angle an extreme error is obtained in the long small extension of the tumor, since the target is small in both perpendicular directions. For the angle of -90° this effect is less pronounced. The usual and probably most reasonable way of dealing with this problem is

to define in our treatment planning system a safety margin for set-up errors around the target. This has the same effect as the 'expansion' of the bolus for scattered beam techniques described by Urie et al (1984). In the -40° case the target volume which receives less than 90 % of the nominal target dose in the presence of set-up uncertainties is 27 %, if the plan is done without a safety margin. With a safety margin of 3 mm in all directions, this is reduced to 4 %.

In the difference-image for the -40° angle another large area of errors can be found, where the difference between the dose with and without set-up errors is up to 15 %. This underdosage area is identical to the one found to have a high range dilution. The choice of a beam angle of -90° results in a lower range dilution and a reduced area with large dose differences due to set-up errors. Since large dose differences (greater than 10 %) in the -40° case do not only appear along the border of the target but as far as 1 cm inside, a good coverage of the target in the presence of set-up errors can not easily be achieved with a reasonably small safety margin. Whenever possible it is therefore best to correct for those effects by avoiding critical heterogeneities, i.e. by choosing a gantry angle of -90° instead of -40° in the case presented here.
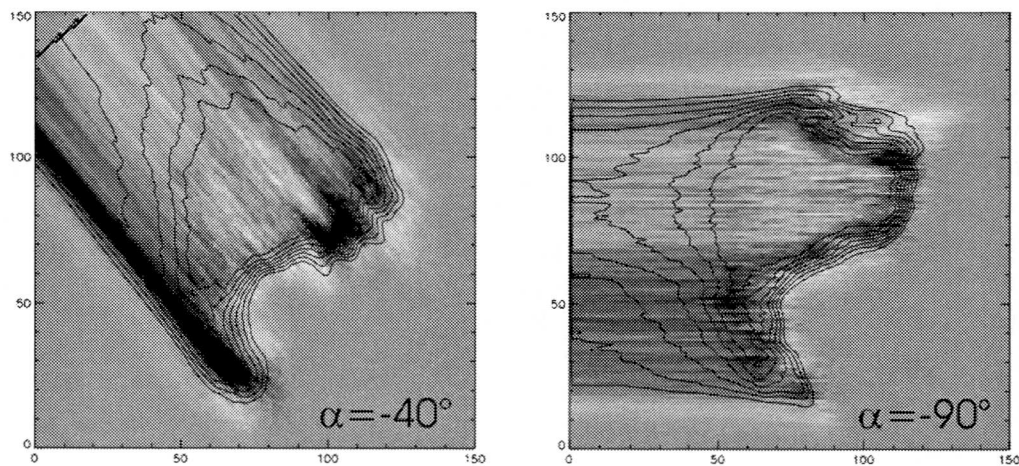

Fig. 39: Dose differences between MC simulations for a treatment with perfect set-up and one with ± 3 mm set-up error between fractions. In the -40° case large dose errors due to the set-up errors have to be expected in the region past the bony interface (see also Fig. 34 and Fig. 35) and in the long thin extension of the target. The errors can be reduced in both regions by choosing a gantry angle of -90°.

3.5 Conclusions

A simple method has been presented for a quick estimation of the range dilution in the presence of density inhomogeneities in a patient as a function of the beam incidence angle.

It has also been shown that a simple ray casting dose calculation can already predict reasonably well the influence of those inhomogeneities on the dose distribution. Further improvements have been achieved using a range dilution correction.

The best overall coincidence between the MC simulation and an analytical model achieved in the cases discussed above was with the flux-dose calculation. This algorithm is also expected to best model the influence of a collimator. In the optimisation however only the two ray casting algorithms can be used.

Three main sources of dose errors which could lead to an underdosage of the target have been identified in the treatment planning process:

- Target misses in the lateral fall-off due to set-up errors.

  A reasonable and most commonly used method to account for these errors is the extension of the clinical target volume with a safety margin to the planning target volume.

- Inaccuracies of the analytical dose calculation.

  It has been shown that there is a correlation between high range dilution and high dose errors. In critical situations, i.e. when the range dilution exceed about 3 %, a Monte Carlo simulation should be performed for an error analysis.

- Dose errors due to the combined effect of heterogeneities and set-up errors.

  These errors are probably the most difficult errors to deal with. They too appear in areas with a high range dilution. Dose errors caused by set-up uncertainties can be simulated by a dose calculation using a pencil beam kernel convoluted with the expected set-up errors.

The easiest and safest way of reducing the latter two sources of errors is by choosing beam directions which avoid critical heterogeneities.

The computed range dilution should be displayed in the treatment planning software, where they serve as a guideline for the planner. A case where a maximal gain can be achieved by carefully choosing the gantry angle is shown in Fig. 40.

Clearly a gantry angle of 0° would cause large errors in the dose distribution (This is the case used for the histogram of dose errors in Fig. 37). 140° is the angle with the lowest expected range dilution. This angle should however not be used for treating this case, since the beam would go directly through the brain stem. Thus, considering both physical and medical criteria, the best choice from the angles shown is most probably 50°.
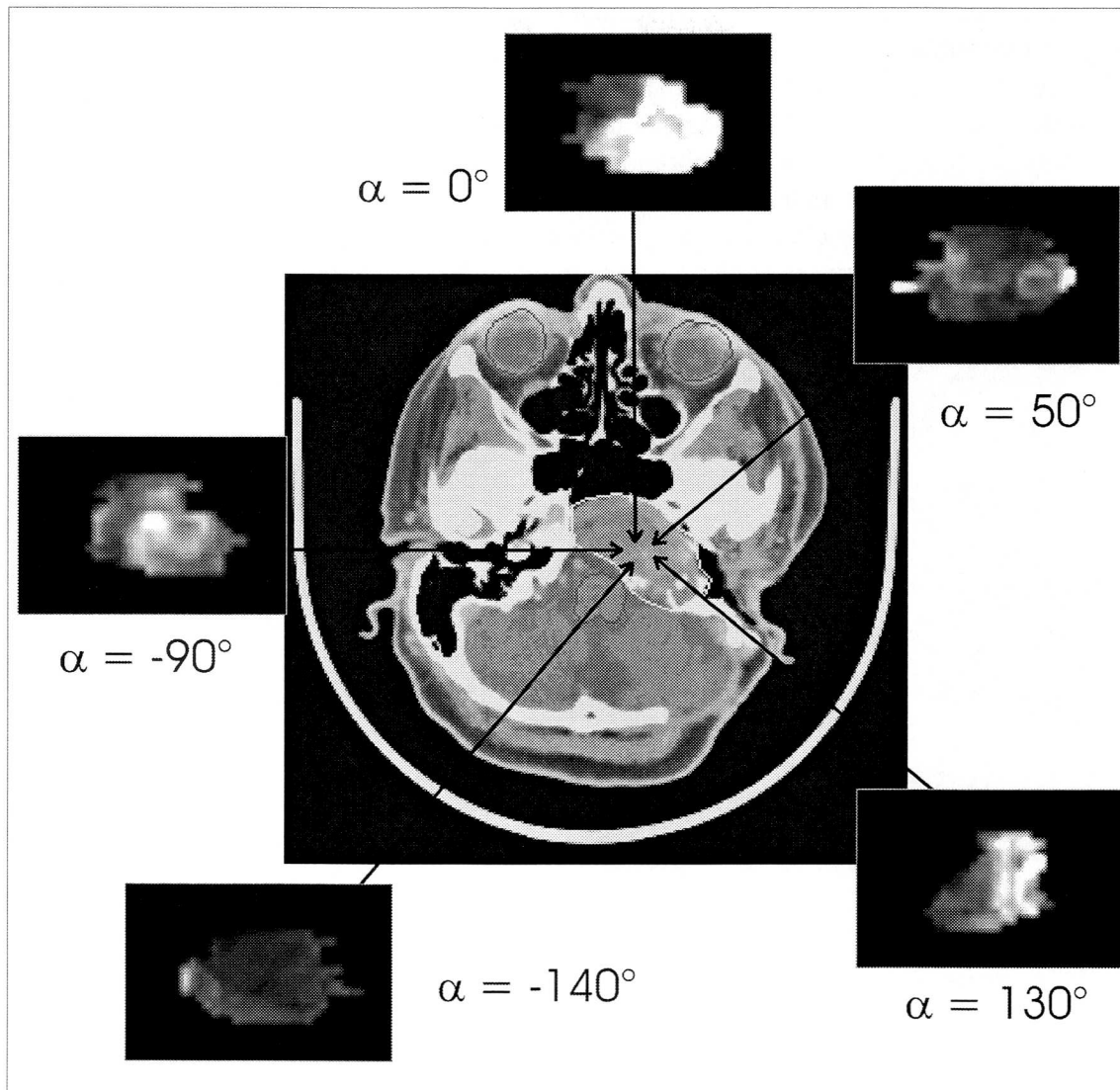
Fig. 40: The amount of range dilution is a good measure for the accuracy which can be achieved in the calculation and delivery of the dose. The planner should therefore choose gantry angles, where range dilution is as low as possible. A software tool was developed, which displays the range dilution at the distal surface of the target in the beams-eye-view. In the small grey scale images here, white represents a relative range dilution of greater than or equal to 8 %, black 0 % for beam directions along the arrows (see also colour image on title page).

3.6 Outlook

There will always be clinical cases, where critical situations cannot be completely avoided. This may be due to an unfavourable position of the tumor or due to radiation sensitive structures in the surrounding areas of the target (as shown in Fig. 40).

Two strategies might be chosen in these situations.

• The use of a large number of fields from different angles will reduce the significance of the dose errors obtained from the individual fields. In scattered beam techniques this is usually done, since a number of fields is routinely used to conform the dose to the target. With the spot-scanning technique fewer fields are already sufficient to achieve good dose distributions. Sometimes only one field angle is used. Therefore a different strategy to deal with expected dose errors is needed.

- One possibility given by the spot-scanning technique is the selective increase of the prescribed dose in small areas, where cold spots could be expected. An analysis of target misses using the MC simulation with and without simulations of set-up errors could be used to define inhomogeneous dose prescriptions to counteract the expected errors. A prescription of an inhomogeneous dose based directly on the calculated range dilution could also lead to good results.

This strategy can be regarded as a safety margin in the fourth dimension (dose) and can be easily delivered with the spot scanning technique, but not with scattered beam techniques. It will however require a detailed study, which also includes medical and biological considerations, as well as some changes in the optimisation procedure.

4 Conclusion - Dealing with range uncertainties in treatment planning

The main advantage of proton beams over the conventionally used photon and electron beams is their depth dose curve, which allows a good conformation of the high dose region to the target. A reduction of the irradiated area reduces the probability of side effects of the radiation treatment, but also increases the risk of missing the target. In proton radiotherapy a precise knowledge of the position and shape of the high dose Bragg peak is therefore crucial for a precise planning and for the success of the treatment. In this thesis the main problems in treatment planning with protons, which could lead to target misses, have been discussed:

In targets without critical heterogeneities the uncertainty of the dose distribution is dominated by the uncertainty in the determination of the proton stopping power values. This is shown in Fig. 41 for the meningioma case.

Fig. 41: see appendix (p. 81)

In the presence of critical heterogeneities high errors in the dose calculation can appear within and/or outside the target depending on the algorithm used (Fig. 42). This is mainly due to a dilution of proton ranges. A detailed error analysis using Monte Carlo simulations if available is necessary for each case individually.

Fig. 42: see appendix (p. 81)

The consequences of this error analysis for the treatment planning process are shown as a flow chart in Fig. 43. Although this flow chart is an overview on the whole process of treatment planning, it is simplified to put more emphasis on the points discussed here. The following is a description of this flow chart with additional important topics being mentioned.

Fig. 43: see appendix (p. 82)

Preparation

The first step consists of gathering all information necessary for the planning. From the physics side, this is the acquisition of the CT data and the conversion of the images into relative stopping power values using the calibration curve described in section 2.2. Additional medical information is necessary for the drawing of the gross tumor volume (GTV) and the clinical target volume (CTV). Here only the CTV is drawn. Additional information about the location of the tumor can be obtained from MRI, PET or contrast enhanced CT images. The histology of the tumor, which is important for a possible microscopic spread around the visible tumor volume, is determined by a biopsy.

Safety margin

Once CT data and CTV are ready, one has to decide on the size of the safety margin and treatment angle(s). Usually a constant safety margin is chosen, although a variable safety margin as a function of the beam incidence angle(s) and critical structures could be used as well. This has to be decided from case to case. Here a constant safety margin of 3 mm is chosen all around the target. This is considered more than sufficient to allow for uncertainties in the calibration (section 2.4) and the smearing of the dose in the lateral dose fall-off due to random setup errors for a head treatment (section 3.4.2). The CTV

plus the safety margin become the planning target volume (PTV) to which the dose is prescribed.

Range dilution and gantry angle

Multiple Coulomb scattering (MCS) causes a perturbation of the proton flux in the patient. In the presence of inhomogeneities this cannot be properly modelled by analytical dose calculation algorithms. Largest problems arise where these effects cause a dilution of the range of the protons.

Shifts of the patient cause a change in the integral densities expected for an individual beam spot, which is planned to pass close to a density interface. Thus setup errors also cause a range dilution.

Both sources of range dilution can cause cold spots in the target, which can not easily be corrected for.

The easiest and safest option - if available - is to choose a gantry angle where the expected range dilution is minimal. The calculation of the range dilution (section 3.2) as a function of the gantry angle can be used as a guideline for the planner when choosing the angle(s). A display of relative range dilution (RD) at the distal border of the target (here PTV) in the beams-eye-view quickly shows the critical areas. Other criteria which have to be considered are the location of radiation sensitive structures and possibly technical limitations.

In the picture, basically two areas with increased range dilution can be seen for an angle of -40°.The one in the lower left corner is due to the nasal cavity, the other one due to the long bone brain interface exactly parallel to the beam. The second area of high range dilution can be completely avoided by changing the gantry angle by only 20°, so that the beam will hit that density interface under an angle of 20°. The best choice from these three angles is certainly -90°, where only a small area with a relative range dilution above 4% appears.

Depending on the site of the tumor and critical structures one would normally choose two to three beam ports. For simplicity only one beam angle is chosen here.

Where no gantry angles can be found which reduce the range dilution in the target to below 3-4 %, the maximal dose errors and their location have to be evaluated from case to case. They might be reduced by a selective increase of the safety margin. It is, however, believed that a selective increase of the prescription dose in the critical regions could lead to better dose distributions. Before this can be done, a detailed study will be necessary as well as changes in the definition of the prescribed dose in the optimisation algorithm. The flexibility of the spot-scanning system will allow the delivery of these dose distributions without any changes in the hardware.

Calculation of a plan

With the PTV drawn and the gantry angle selected, the actual planning program is run to calculate iteratively the beam weights for the spot positions inside the PTV, which will lead to a dose distribution as close as possible to the prescribed dose. New algorithms for the dose calculation have been developed, implemented and discussed in this thesis (section 3.3). All of them are expected to give good results in the case of low range dilution throughout the target. In this case one can go ahead directly to create the necessary files for the steering of the beam application system, to do the dosimetric verification and to deliver the treatment.

Worst case analysis

In cases where high range dilution cannot be avoided, a worst case analysis should be done. Although it is not believed to be necessary in the -90° case presented here, it certainly has to be done in cases like the chordoma in Fig. 40. This analysis consists of dose calculations with the simulation of random setup errors and using minimal and maximal calibration curves (see also Fig. 41). The dose distribution shown here is obtained by a Monte Carlo simulation assuming ± 3 mm set-up errors and using the maximal calibration curve.

Ideally the Monte Carlo simulation (appendix B) is used for the worst case analysis. In the base of the skull region, where a large range dilution is often unavoidable, tumors are not very large. A Monte Carlo simulation of one field and with a good statistical resolution can be obtained in 20-30 min. on a DEC AXP 500/333 workstation. Alternatively two dose calculations could be done, one using the flux-dose calculation (M1) the other one the refined ray casting (M4). Since M4 underestimates the effect of MCS and M1 overestimates it, any problems with the accuracy of the dose calculation due to this effect should show up in either of the resulting dose distributions (see also Fig. 42).

If all the resulting dose distributions are acceptable, one can go ahead with the verification (appendix C) and treatment. If not, one should try to find another beam angle or increase the safety margin selectively in the problematic areas to assure a good coverage of the target.

A worst case analysis could also be performed to find areas where the safety margin has been drawn too large and could be reduced in order to spare better the healthy tissue.

Appendix A: Multiple Coulomb scattering

Angular distributions of particles passing through thin and thick slabs of matter have been discussed extensively by G. Molière in 1948. He approximates the distribution function for multiple Coulomb scattering (MCS) angles by a power series, whose first term is a Gaussian. Higher order terms can however not be completely neglected in a Gaussian approximation of Molière theory. Instead of Molières characteristic multiple scattering angle $\chi_c \sqrt{B}$ (χ_c : char. single scattering angle, B : parameter including among others the mean number of scattering processes in the slice), Hanson et al. (1951) suggested to use $\chi_c \sqrt{B - 1.2}$ as a correction to account for higher order terms. V. Highland (1975) proposed a rms angle

$$(24) \qquad \Theta_{RMS} = \frac{E_s}{p\beta c} \sqrt{\frac{L}{L_{RAD}}} \left(1 + k \log\left[\frac{L}{L_{RAD}} \right] \right)$$

where the logarithmic term corrects better for dependencies on the path length and Z number. This is the form used nowadays for the characteristic scattering angle. Instead of Highlands original values for E_s and k, which were obtained by a fit to silver data only, Lynch and Dahl (1991) obtained E_s=13.6 MeV (for the projected Θ_{RMS}) and k=0.088 from fits to all Z numbers. We use E_s=13.25 MeV, which was found to fit best our dosimetry data measured in water.

Thus the projected angular distribution of protons after passing through a slab of scattering material can be approximated by

$$(25) \qquad P_x(\vartheta) = \frac{1}{\sqrt{2\pi}\Theta_{RMS}} e^{-\frac{\vartheta^2}{2\Theta_{RMS}^2}}$$

Previously Fermi developed a low-order approximation of the transport of monoenergetic charged particles through matter (Rossi and Greisen, 1941). Energy dependence was added to this formulation by Eyges (1948). In the Fermi-Eyges approximation of lateral and angular displacement of charged particles, which undergo multiple Coulomb scattering in passing through a layer of matter, the distribution function can be written in the form:

$$(26) \qquad P_x(\phi, x, z) = \frac{1}{\pi\sqrt{D(z)}} e^{-\frac{A(z)\phi^2 + 2B(z)x\phi + C(z)x^2}{D(z)}}$$

$$where \qquad D = AC + B^2$$

Integrating $P_x(\phi, x, z)$ over ϕ or x we obtain:

$$P_x(\phi, z) = \frac{1}{\sqrt{\pi C(z)}} e^{-\frac{\phi^2}{C(z)}}$$

$$(27)$$

$$P_x(x, z) = \frac{1}{\sqrt{\pi A(z)}} e^{-\frac{x^2}{A(z)}}$$

Thus the widths of angular and lateral displacement are

$$(28) \qquad \sigma_\phi(z) = \sqrt{\frac{C(z)}{2}} \quad \text{and} \quad \sigma_x(z) = \sqrt{\frac{A(z)}{2}}$$

respectively. We set $\sigma_\phi(\Delta z) = \Theta_{RMS}(\Delta z)$ and propagate P_x through infinitesimal thin slabs, which leads to:

$$A(z) = A_0 + 2B_0 z + C_0 z^2 + \int_0^z \Theta_{RMS}^2(z')(z - z')^2 \, dz'$$

$$(29) \qquad B(z) = B_0 + C_0 z + \int_0^z \Theta_{RMS}^2(z')(z - z') dz'$$

$$C(z) = C_0 + \int_0^z \Theta_{RMS}^2(z') dz'$$

where A_0, B_0 and C_0 describe the phase space of the incident beam.

Gottschalk et al. (1992) suggested removing the logarithmic term in Θ_{RMS} from the integral, thus treating it as a correction factor depending on the entire thickness of the target. For an analytical solution of the integrals in eq. (29) it is furthermore necessary to substitute the remaining term $\Theta_u(z)$ by an empirical formula (Øverås, 1960), which describes the dependence of the scattering angle on the energy loss:

$$(30) \qquad \Theta_u^2(t) = \frac{\Theta_u^2(0)}{(1-t)^{1+\kappa}}$$

$\Theta_u(t)$ is the rms angle per unit length relative to the total range R of the incident beam
t=z/R is the depth relative to the total range
κ is an empirical parameter found to be 0.073 for water

Thus the solutions of the integrals become (without the contribution from the initial phase space (A_0, B_0, C_0)) (S. Scheib,1993):

$$A(t) = \Theta_u^2(0)\left(1 + 0.088 \log\left(\frac{z\rho}{L_{RAD}}\right)\right) R^3 \frac{1}{\kappa}\left[-t^2 + \frac{2t}{(1-\kappa)} + \frac{2(1-t)^{2-\kappa} - 2}{(1-\kappa)(2-\kappa)}\right]$$

$$(31) \qquad B(t) = -\Theta_u^2(0)\left(1 + 0.088 \log\left(\frac{z\rho}{L_{RAD}}\right)\right) R^2 \frac{1}{\kappa}\left[t + \frac{(1-t)^{1-\kappa} - 1}{(1-\kappa)}\right]$$

$$C(t) = \Theta_u^2(0)\left(1 + 0.088 \log\left(\frac{z\rho}{L_{RAD}}\right)\right) R \frac{1}{\kappa}\left[(1-t)^{-\kappa} - 1\right]$$

Appendix B: Monte Carlo simulation

B 1 Properties of the Monte Carlo code

Many Monte Carlo (MC) codes used by other groups in medical physics are based on detailed simulations of the interaction between particle and target atoms. The software packages to simulate these interactions include the well known GEANT code, but also codes like PTRAN, FLUKA or FREESCO, where the simulation of interactions of low energy particles is improved. These codes have two important disadvantages for using them to simulate treatment plans:

- Due to the detailed simulation of all the interaction on the atomic level they are very slow.

- It is very difficult to define the patient data - i.e. the outline of all the structures and their elemental composition - correctly.

These were reasons not to choose one of those codes, but to write our own program, which has been done by A. Tourovsky from the Institute for Theoretical and Experimental Physics (ITEP) in Moscow during a one year stage at PSI (Tourovsky et al 1993). The code is optimised for speed and is capable of simulating 1 million protons in 1.5-2 min. on a DEC AXP 500/333 workstation, which is the fastest machine available to us at the moment. A simulation of a treatment plan of size and statistical resolution of the meningioma case shown in part B therefore takes less than 20 min.

Additionally, the code could be designed specifically in conjunction with the PSI treatment planning (TP) software. It is especially useful to study the effects of MCS on the dose in the presence of complex anatomical situations. To obtain a good coincidence between the MC and TP results in homogeneous cases, the MC code uses the same theory and data (CT data set, beam properties measured on the gantry) as the treatment planning program. CT values of all voxels are converted into stopping power (using the curve in section A) and into a scattering coefficient . The protons are then traced on a voxel by voxel basis until they are stopped.

The interactions modelled are ionisation loss, nuclear interactions and MCS:

- Ionisation loss

    The energy loss of a proton through ionisation is described by the Bethe-Bloch (Bichsel 1972) equation. These values are precalculated and stored in a look-up-table (LUT). The two parameters used to calculate the energy loss in a voxel are the water-equivalent-range (WER) lost along the protons path in that voxel and the energy of the proton. The energy lost by the proton is assumed to be absorbed uniformly over the whole voxel.

- Nuclear interactions

    The probability of a nuclear interaction occurring is given by the cross section of oxygen as a function of the energy of the proton. In this reaction the primary proton is lost, while a number of secondary particles (secondary protons, electrons, photons and neutrons) can be emitted. In analogy to the treatment planning program, the energy deposited by charged secondary particles produced in this reaction is deposited along the residual path of the primary proton in a continuous slowing down approximation. Though there are possibilities in a MC code to simulate these interactions in more detail, they are deliberately not implemented. Thus differences in the results of MC and analytical dose calculation originate solely from MCS.

- Multiple Coulomb scattering

  The theoretical background for the simulation of MCS in the MC code is the same as for the TP (see appendix A). A calibration curve has been defined, which assigns to every voxel a scattering coefficient based on its HU. This scattering coefficient includes the material dependant ratio of density to radiation length. The proton is traced in one step through the voxel, assuming that all scattering happens in the middle of the voxel. The scattering angle in each voxel is calculated from the energy of the proton, the WER of the path length in the voxel and the scattering coefficient. The calculation of the exit point is done using Tayor expansions of all trigonometric functions in a small angle approximation.

B 2 Verification of the MC code

The verification of the MC code consists of two steps. In a homogeneous phantom the results of the MC code must be the same as those given by the TP program. This is shown in section B 2.1.

In the case where inhomogeneities are simulated one expects differences between MC and TP results. Therefore the verification of the MC code in the presence of inhomogeneities has to be done experimentally. One possibility is the comparison with proton radiography results, which has been done by Tourovsky et al (1993). Another verification for a simple geometric situation is presented in section B 2.2.

B 2.1 Simulation of a proton pencil beam in water

Proton pencil beams in water have been simulated with variable parameters (initial beam energy, number of range shifter plates, air gap between range shifter and phantom). Without range shifter plates, the agreement - in depth dose as well as in beam width - was found to be almost perfect (Fig. 44).
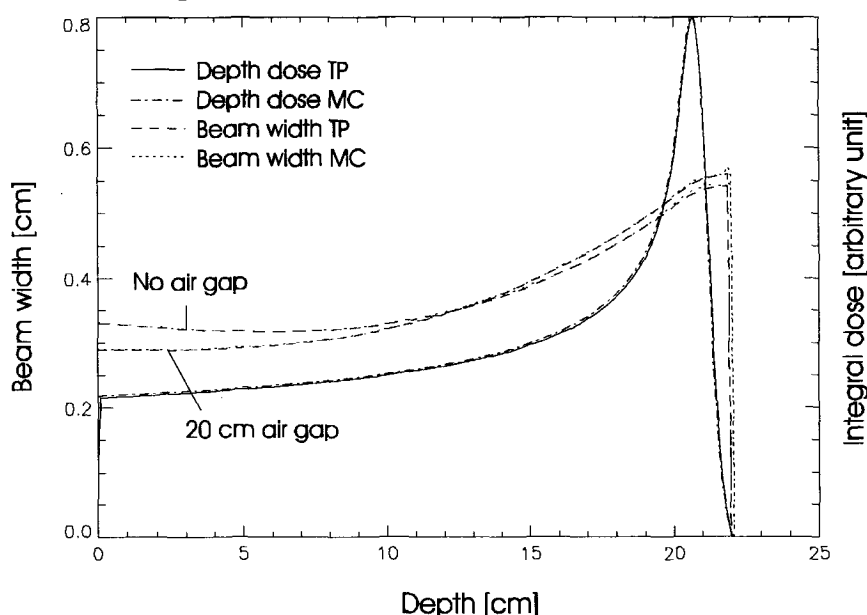
Fig. 44: Comparison of depth dose curve and beam width of a 177 MeV pencil beam in water calculated by TP and MC code respectively. No range shifter plates are used. The agreement is very good.

Slight discrepancies between the two models can be observed in beam width when range shifter plates are introduced (Fig. 45). This can be explained by the fact that the

combined contribution from the MCS in the range shifter plates and the air gap has to be modelled differently in the MC simulation and in TP. The difference introduced by this is not of clinical importance, but explains why the dose distributions from the MC simulation usually show a steeper lateral fall-off than the TP dose distributions.

The plots in Fig. 45 show that this effect can - in the extreme but unrealistic case of 30 range shifter plates and no air gap - cause a deviation between MC and TP of up to 7 % of the beam width. Usually one expects deviations below 3 %. The integral dose is in any case correct (not shown).
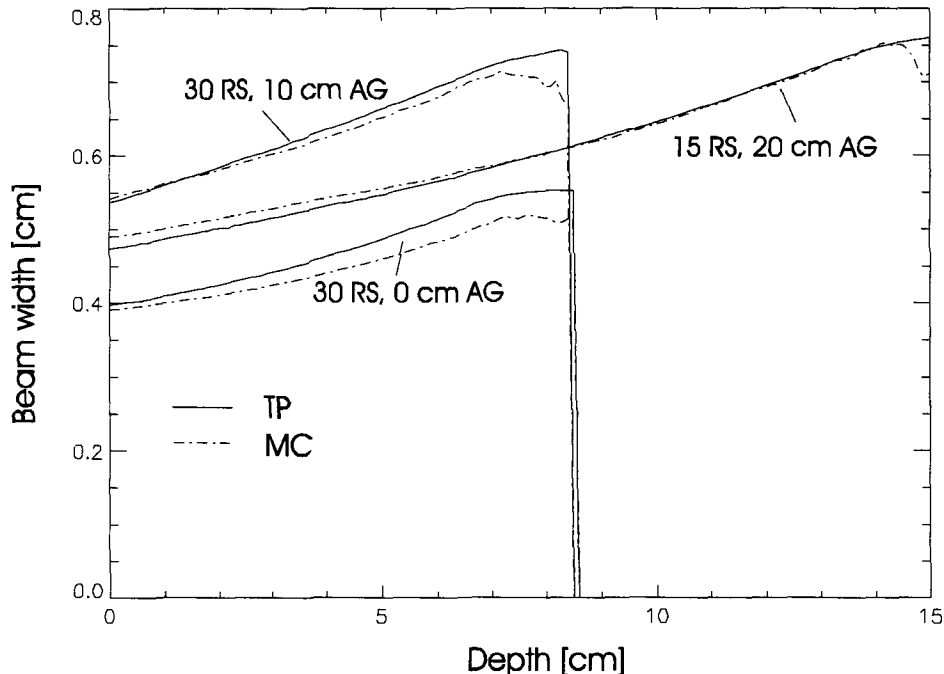
Fig. 45: Comparison of beam width of a pencil beam in water. Different numbers of range shifter plates and different air gaps are used for the simulations. There are slight differences in the beam width obtained by the analytical calculation and the Monte Carlo simulation respectively, which can be explained by the difference in the models used to simulate the combined effects of range shifter plates and air gap.

B 2.2 Simulation of dose distributions after inhomogeneities

The MC code has also been verified against dose measurements in an inhomogeneous phantom. For these measurements a very good spatial resolution is required. Therefore a newly developed system using a scintillating screen and a CCD-camera has been used (S. Boon et al 1997). The system provides very good relative measurements with a spatial resolution better than 1 mm. Except for small quenching effects in the Bragg peak, the relation between dose and emitted light has been found to be linear.

Plan and setup of the experiment

The experiment was planned to see whether the MC code can predict accurately the effects on the dose produced by a simple water-perspex interface placed in front of the target. Fig. 46 shows the result of the treatment plan for this situation with a cubic target of 4 cm length sides.
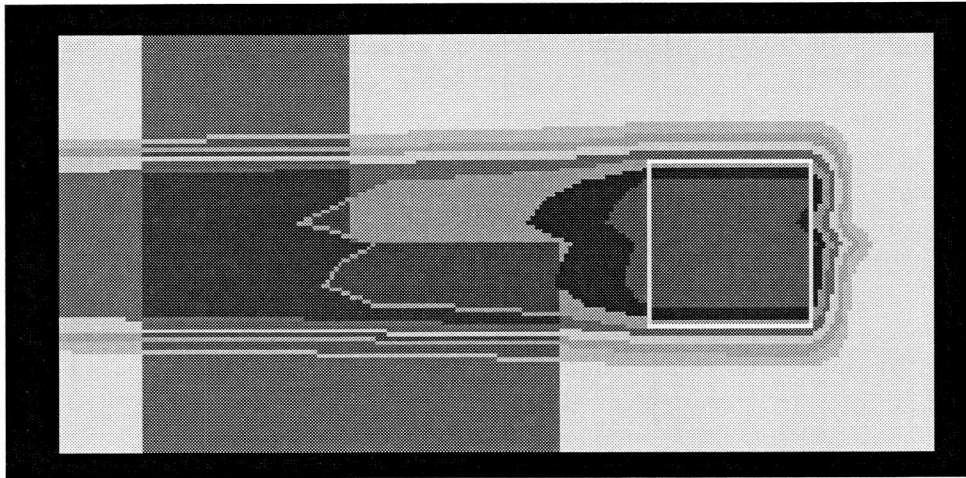
Fig. 46: Predicted isodose contours obtained from the treatment planning are program overlaid on the density distribution for the experiment.

To keep the experiment as simple as possible, the plan was not optimised. I. e. the distribution of spot weights along each row of spots follows the scan law for a spread-out-Bragg-peak (SOBP) with equal weight to all the rows. An optimisation procedure would produce a better conformation to the target by giving higher weights to the spots on the lateral border and could partly compensate for the inhomogeneity effect on the dose.
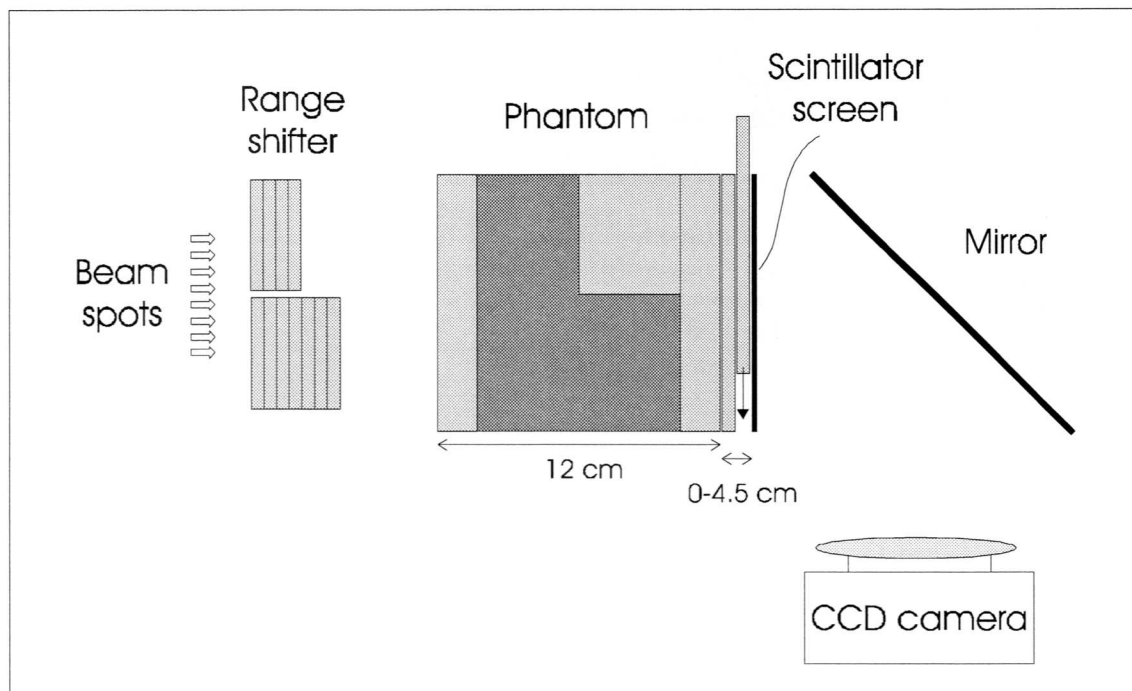
Fig. 47: Experimental set-up: The treatment is delivered to the water phantom as defined in the treatment planning program. The light emitted by a scintillating screen is reflected by a mirror into a CCD camera. The depth of the measurement is adjusted by a variable amount of perspex between the phantom and the screen.

The setup of the experiment is shown in Fig. 47. A 5 cm thick piece of perspex was put into a perspex box filled with water. Perspex plates of variable thickness (0 - 4.5 cm)

were placed between the box of water and the scintillating screen for the measurement of the dose at different positions in depth. The light emitted by the screen was reflected by a mirror into the CCD-camera. The application of the spots positioned on one or the other side of the interface was done with different numbers of range shifter plates to compensate for the higher density of perspex compared to water. Thus one obtains the dose distribution shown in Fig. 46, where the distal fall-off is at almost the same position in depth throughout the irradiated area.

Results

Fig. 48 shows the dose measured by the CCD-camera (grey scale) compared with the predictions of the MC simulation (contour plots). The measurement is taken directly behind the target in the distal fall-off. In this region largest effects from the interface are seen. Except for a small shift of the lower high dose region, which is probably due to an inaccuracy in positioning the perspex interface, it shows a very good qualitative agreement with the MC result.
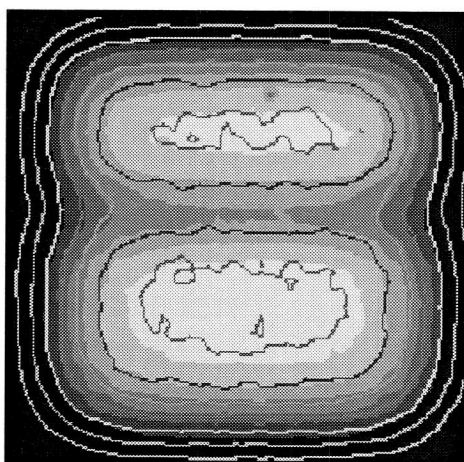
Fig. 48: Dose measured by CCD-camera (grey scale) compared to results of MC simulation (contour plots) at the position of the distal fall-off.

Appendix C: Dose verification for a brain metastasis in the Alderson phantom

A full simulation of all the procedures involved has been done for a phantom patient before the first human patient was treated on the PSI gantry.

First the Alderson phantom received a moulage with a head mask and a bolus. CT images in treatment position were taken, a target volume was drawn and the treatment plan done. Besides the dosimetric verification of all the fields in a homogeneous phantom - as is done for all the patients - measurements were taken inside the Alderson phantom during treatment.

C 1 Setup of the experiment
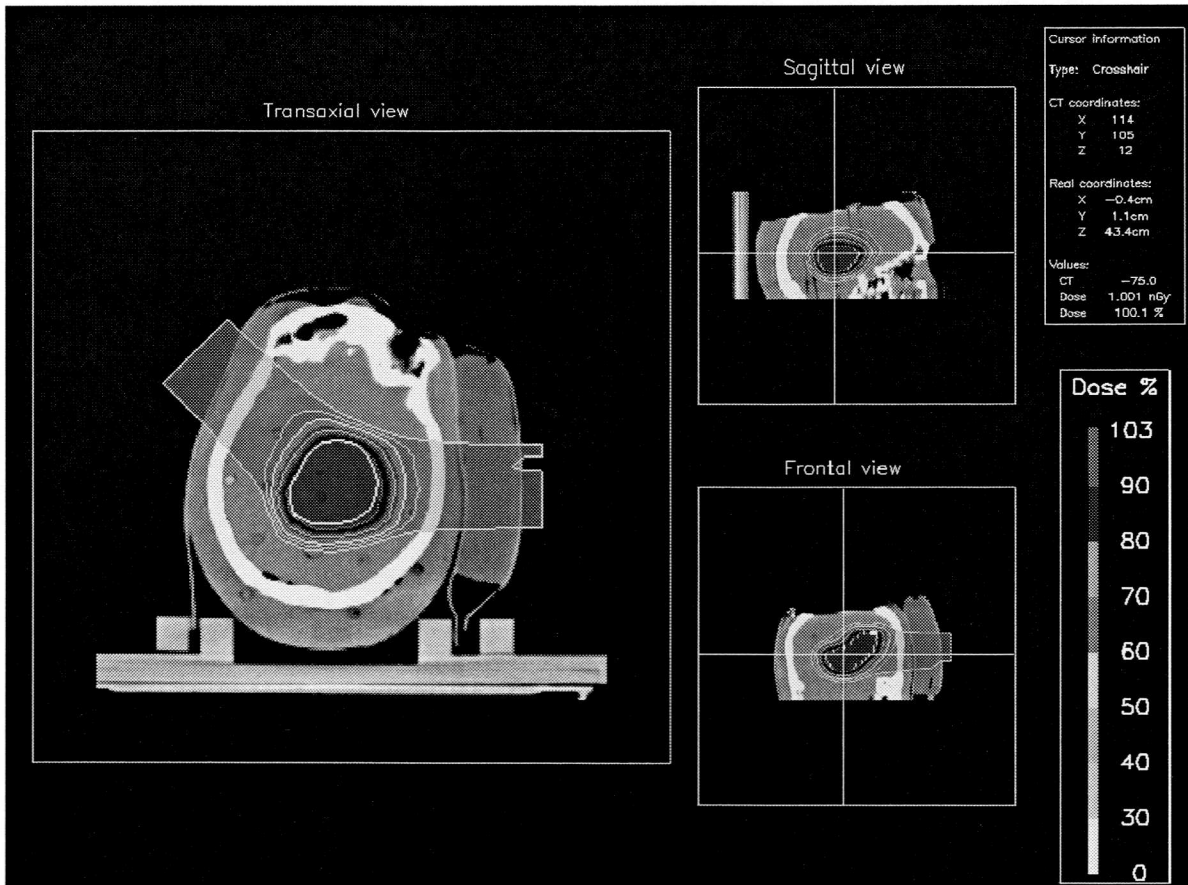
Fig. 49: Printout of the treatment planning system, which shows an overview over the planned treatment using two fields with gantry angles of -45° and 90°.

The positioning and moulaging of the phantom and the drawing of the tumor volume has been done keeping in mind the restrictions given by the experimental conditions:

- The phantom was positioned slightly tilted. That way avoided having the beam exactly parallel to the slices of the phantom.
- The tumor was drawn in an area where holes allowed access with the ionisation chamber. The target could be treated with two fields from two different angles without the beam going through critical heterogeneities.

- The shape was chosen in such a way that the range of the protons could also be checked. The isodose curve crossed the hole at an angle of about 40° . Therefore the measured profile was very sensitive to the correct determination of the proton range.
- No restrictions from radiosensitive structures were taken into account for choosing the beam angles.

The chosen gantry angles were -45° and 90°.

An overview of the planned treatment is given in Fig. 49.

C 2 Verification of the steering file

The steering files used for the dynamic application of the beam spots is checked routinely for all patients and gantry angles with dosimetric measurements in a homogeneous phantom. For this case a perspex phantom was used. A specially designed water column was taken into operation in the beginning of 1997 for this purpose. The verification consists of the irradiation of films at different positions in depth to check the shape of the field. Profiles of absolute dose are measured using an array of ionisation chambers, which has been developed at PSI. The dose measurement is expected to be accurate to 1-2 %, while the setup errors should not exceed 1 mm in all directions.

An example of the calculated dose profile - corrected for the fact that the irradiation is performed in a homogeneous phantom instead of the patient - and measured points in the homogeneous phantom is shown in Fig. 50. The agreement is good.

Based on these verification measurements the decision is usually taken to call the patient in for treatment.
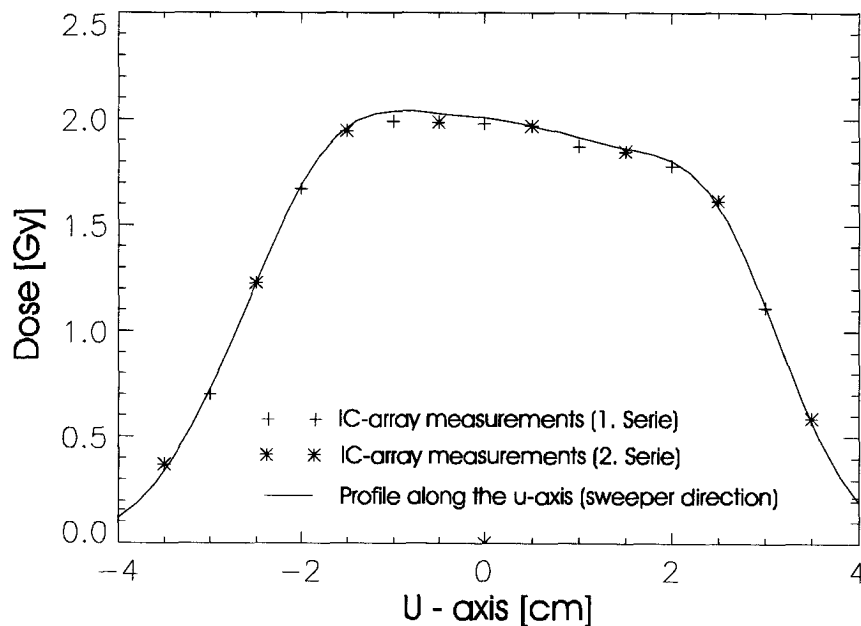
Fig. 50: Results of the verification measurements for one profile in the perspex phantom.

C 3 Verification in the Alderson phantom

In the Alderson phantom additional dosimetric verification measurements could be done inside the target volume in a treatment situation. For this purpose a PTW ionisation chamber was used (M23322) with an outer diameter of 5.5 mm. The chamber volume is 3.5 mm wide and 12 mm long. Thus it does not provide a good spatial resolution.

The hole used for the IC-measurements of the 90° field is along the vertical crosshair shown in the frontal view in Fig. 49. Along that hole points in the central high dose region could be measured as well as the dose fall-off distal of the tumor higher up in the brain. That dose fall-off is very sensitive to errors in the calculation of the integral range and therefore provides a good check of the calibration curve (section 2). Fig. 51 shows that the measurements tied up well with the calculation.
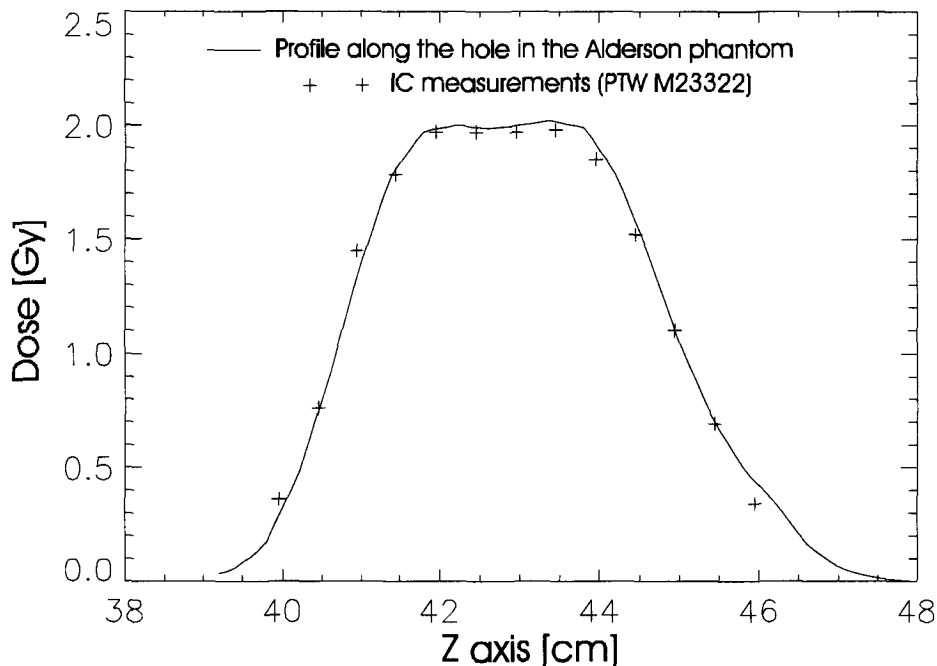

Fig. 51: Calculated profile and results of the measurements in the Alderson phantom.

C 4 Conclusion

It has been shown that the complete chain of actions necessary for patient treatment in the upper part of the brain was implemented and functioning. The dosimetric measurements in the homogeneous phantom as well as the Alderson phantom showed a good agreement with the treatment planning results. This simulation of the whole treatment procedure gave us the confidence to go ahead and treat the first human patient in Nov. 1996.

References

Austin-Seymour M, Munzenrider J E, Goitein M, Gentry R, Gragoudas E, Koehler A M, McNulty P, Osborne E, Ryugo D K, Seddon J et al 1985 Progress in low-LET heavy particle therapy: intracranial and paracranial tumors and uveal melanomas *Radiat. Res. Suppl.* **8** 219-226

Battista J J and Bronskill M J 1981 Compton scatter imaging of transverse sections: an overall appraisal and evaluation for radiotherapy planning *Phys. Med. Biol.* **26(1)** 81-99

Bichsel H 1972 Passage of charged particles through matter *American Institute of physics handbook* McGraw-Hill 8-142 - 8-185

Boon S N, Coray A, Lomax A, van Luijk P, Pedroni E, Schippers J M 1997 Fast 2D dosimetry for dynamic beam delivery systems, *Med. Biol. Eng. Comp., Proceedings of the World Congess on Medical Physics and Biomedical Engineering* **35(2)**, publication in preparation

Carlsson A K, Andreo P, Brahme A 1997 Monte Carlo and analytical calculation of proton pencil beams for computerized treatment planning optimization *Phys. Med. Biol.* **42** 1033-1053

Chen G T Y, Rajinder P S, Castor J R, Lyman J T, Quivey J M 1979 Treatment planning for heavy ion radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **5** 1809-1819

Constantinou C 1978 Tissue substitutes for particulate radiations and their use in radiation dosimetry and radiotherapy *PhD thesis* University of London

Constantinou C and Harrington J 1992 An electron density calibration phantom for CT-based treatment planning computers *Med. Phys* **19(2)** 325-327

Deasy J O 1997 A protons dose calculation algorithm for conformal therapy simulations based on Molière's theory of lateral deflections *submitted to Med. Phys*

Eyges L 1948 Multiple scattering with energy loss *Phys. Rev.* **74** 1534/5

Garelli S, Giordano S, Squarcia S 1996 Real and simulated data comparative studies for proton energy loss *Abstracts of the XXV PTCOG meeting*

Goitein M, Niemierko A 1996 Intensity modulated therapy and inhomogeneous dose to the tumor: A note of caution *Int. J. Radiation Oncology Biol. Phys.* **36(2)** 519-522

Gottschalk B, Koehler A M, Schneider R J, Sisterson J M, Wagner M S 1993 Multiple Coulomb scattering of 160 MeV protons *Nuc. Inst. Method. Phys. Res.* **B74** 467-490

Habrand I L, Austin-Seymour M, Birnbaum S, Wray S, Carroll R, Munzenrider J, Verhey L, Urie M, Goitein M 1989 Neurovisual outcome following proton radiation therapy *Int. J. Radiation Oncology Biol. Phys.* **16(6)** 1601-1606

Hanson A O, Lanzl L H, Lyman E M, Scott M B 1951 *Phys. Rev.* **84** 515

Hanson K M, Bradbury J N, Cannon T M, Hutson R L, Laubacher D B, Macek R J, Pacinotti M A, Taylor C A 1981 Computed tomography using proton energy *loss Phys. Med. Biol.* **26(6)** 965-983

Highland V 1975 Some practical remarks on multiple scattering *Nuc. Inst. Methods* **129** 497-499

Hong L, Goitein M, Bucciolini M, Comiskey R, Gottschalk B, Rosenthal S, Serago C, Urie M 1996 A pencil beam algorithm for proton dose calculations *Phys. Med. Biol.* **41** 1305-1330

ICRU 1989 Tissue substitutes in radiation dosimetry and measurement" *ICRU Report 44*

Jackson D F and Hawks D J 1981 X-ray attenuation coefficients of elements and mixtures *Phys. Rep.* **70** 169-233

Lynch G R, Dahl O I 1991 Approximations to multiple Coulomb scattering *Nuc. Inst. Method. Phys. Res.* **B58** 6-10

McCullough E C and Holmes T W 1985 Acceptance testing computerized radiation therapy treatment planning systems: direct utilization of CT scan data *Med. Phys.* **12(2)** 237-242

Minohara S, Kanai T, Endo M Kawachi K 1993 Effects of object size on a function to convert X-ray CT numbers into water equivalent path length of charged particle beams Proceedings of the third workshop on physical and biological research with heavy ions, *HIMAC report* No 6

Molière G 1948 Theorie der Streuung schneller geladener Teilchen II" *Nuc. Inst. Method. Phys. Res.* **B74** 467-490

Munzenrider J E, Verhey L J, Gragoudas E S, Seddon J M, Urie M, Gentry R, Birnbaum S, Ruotolo D M, Crowell C, McManus P et al 1989 Conservative treatment of uveal melanoma: local recurrence after proton beam therapy *Int. J. Radiat. Oncol. Biol. Phys.* **17(3)** 493-498

Munzenrider J E, Austin-Seymour M, Blitzer P J, Gentry R, Goitein M, Gragoudas E S, Johnson K, Koehler A M, McNulty P, Moulton G et al 1985 Proton therapy at Harvard *Strahlentherapie* **161(12)** 756-763

Øverås H 1960 Small angle multiple scattering in confined bodies *CERN Yellow Report* 60-18

Pedroni E, Bacher R, Blattmann H, Böhringer T, Coray A, Lomax A, Lin S, Munkel G, Scheib S, Schneider U, Tourovsky A 1995 The 200MeV proton therapy project at the Paul Scherrer institute: Conceptual design and practical realization *Med. Phys.* **22(1)**

Pemler P 1997 Entwicklung, Bau und Test eines Detektorsystems für in vivo Protonenradiographie *PhD thesis* University of Munich

Petti P 1992 Differential-pencil-beam dose calculations for charged particles *Med. Phys.* **19(1)** 137-149

Petti P 1996 Evaluation of a pencil-beam dose calculation technique for charged particle radiotherapy *Int. J. Radiation Oncology Biol. Phys.* **35(5)** 1049-1057

Prasad S C, Glasgow G P, Purdy J A 1979 Dosimetric evaluation of a computed tomography treatment system *Radiology* **130** 777-781

Ragona R, Rolando V, Solano A 1996 Treatment planning of proton beams using the GEANT Monte Carlo *Abstracts of the XXV PTCOG meeting*

Rossi B, Greisen K 1941 Cosmic ray theory *Rev. Mod. Phys.* **13** 240

Rutherford R A, Pullan B R, Isherwood I 1976 Measurement of effective atomic number and electron density using an EMI scanner *Neuroradiology* **11** 15-21

Scheib S 1993 Spot-Scanning mit Protonen: Experimentelle Resultate und Therapieplanung *PhD thesis ETH Zürich* **No. 10451**

Schneider U, Pedroni E 1995 Proton radiography as a tool for quality control in proton therapy *Med. Phys.* **22(4)** 353-363

Schneider U, Pedroni E, Lomax A 1996 The calibration of CT Hounsfield units for radiotherapy treatment planning *Phys. Med. Biol.* **41** 111-124

Schneider U, Tourovsky A, Pedroni E, Lomax A, Schaffner B 1997 A comparison of different techniques for the calculation of proton dose distributions using spot scanning *XII International conference on the use of computers in radiation therapy Salt Lake City 209-211*

Sisterson J 1997 Particles *PTCOG-Newsletter*

Slater J D, Austin-Seymour M, Munzenrider J, Birnbaum S, Carroll R, Klibanski A, Riskind P, Urie M, Verhey L, Goitein M 1988a Endocrine function following high dose proton therapy for tumors of the upper clivus *Int. J. Radiat. Oncol. Biol. Phys.* **15(3)** 607-611

Slater J M, Miller D W, Archambeau J O 1988 Development of a hospital-based proton beam treatment center *Int. J. Radiat. Oncol. Biol. Phys.* **14** 761-775

Slater J D, Slater J M, Wahlen S 1992a The potential for proton beam therapy in locally advanced carcinoma of the cervix *Int. J. Radiat. Oncol. Biol. Phys.* **22(2)** 343-347

Slater J M, Archambeau J O, Miller D W, Notarus M I, Preston W, Slater J D 1992b The proton treatment center at Loma Linda university medical center: rational for and description of its development *Int. J. Radiat. Oncol. Biol. Phys.* **22(2)** 383-389

Smit B M 1992 Prospects for proton therapy in carcinoma of the cervix *Int. J. Radiat. Oncol. Biol. Phys.* **22(2)** 349-353

Suit H D, Goitein M, Munzenrider J E, Verhey L, Gragoudas E, Koehler A M, Urie M, Shipley W U, Linggood R M, Friedberg C, Wagner M 1980 Clinical experience with proton beam radiation therapy *J.. Can. Assoc. Radiol.* **31(1)** 35-39

Suit H D, Goitein M, Munzenrider J E, Verhey L, Urie M, Gragoudas E, Koehler A M, Gottschalk B, Sisterson J, Tatsuzaki H et al 1990 Increased efficacy of radiation therapy by use of proton beam *Strahlenther. Onkol.* **166(1)** 40-44

Tourovsky A, Pedroni E, Schneider U 1993 Monte Carlo codes for proton radiography and treatment planning *PSI Life Science Newsletter* 20-22

Urie M, Goitein M, Wagner M 1984 Compensating for heterogeneities in proton radiation therapy *Phys. Med. Biol.* **29(5)** 553-566

Urie M, Goitein M, Holley W R, Chen G T Y 1986 Degradation of the Bragg peak due to inhomogeneities *Phys. Med. Biol.* **31** 1-15

Wilson R R 1946 Radiological use of fast protons *Radiology* **47** 487

Woodard H Q, White D R 1986 The composition of body tissue *British Journal of Radiology* **59** 1209-1218

Yonemoto L T, Slater J D, Rossi C J, Antoine J E, Loredo L N, Archambeau J O, Schulte R W M, Miller D W, Teichmann S L, Slater J M 1997 Combined proton and photon conformal radiation therapy foe locally advanced carcinoma of the prostate: preliminary results of a phase I/II study *Int. J. Rad. Oncol. Biol. Phys* **37**(1) 21-29

Acknowledgements

Curriculum vitae

Name:	Barbara Schaffner
Date of birth:	28.7.1968
Sex:	female
Origin:	Hausen b. Brugg (CH)

Education:

1975 - 1988   Primar-, Bezirks- and Kantonsschule in Anglikon and Wohlen (AG)

1988 - 1993   ETH Zürich
Diploma in Physics
Diploma Thesis „Messung und Berechnung der Strahlenbelastung in einem 3D-Computertomographiesystem"
Supervisor: Prof. P. Rüegsegger

1994-1997   PhD thesis at the Paul Scherrer Institute:
„Range precision of therapeutic proton beams"
Supervisor: Dr. E. Pedroni

Nov. 1997   Exam for the professional qualification in medical physics of the Swiss Society for Radiobiology and Medical Physics
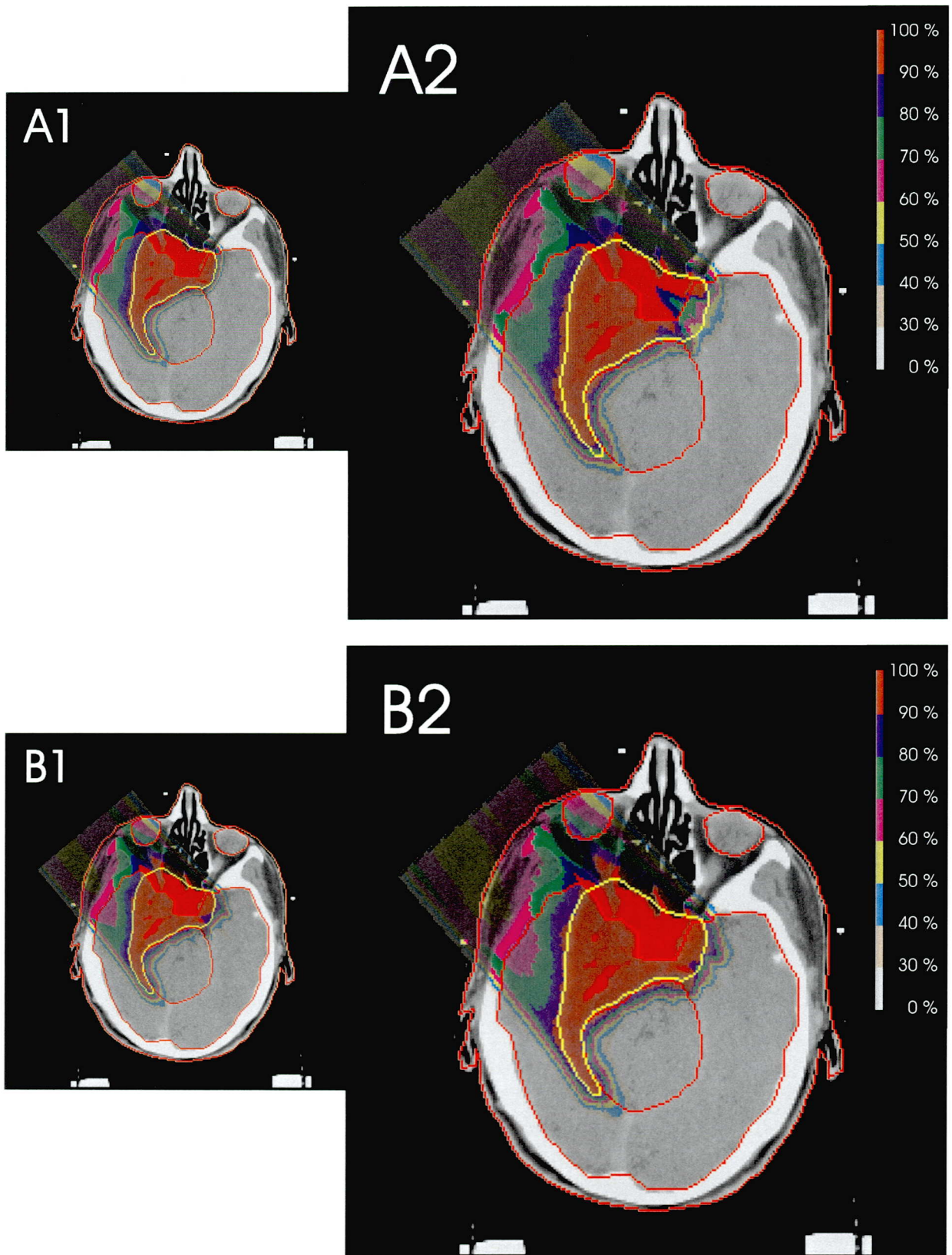
Fig. 34: Two beam models (scanning beam model (A) and simple ray casting model (B)) are used to calculate an optimised dose distribution in a case with a critical heterogeneity. Both results of the analytical dose calculations (small images) show a good conformation of the dose to the target. The verification with the Monte Carlo simulation reaveals, however, a large cold spot (A2) in the target, which was not modelled properly by the scanning beam model. The agreement between the results of the ray casting model (B1) and the Monte carlo simulation (B2) is good.

Monte Carlo result

Error image

Dose uncertainty
due to calibration
10%

Dose errors
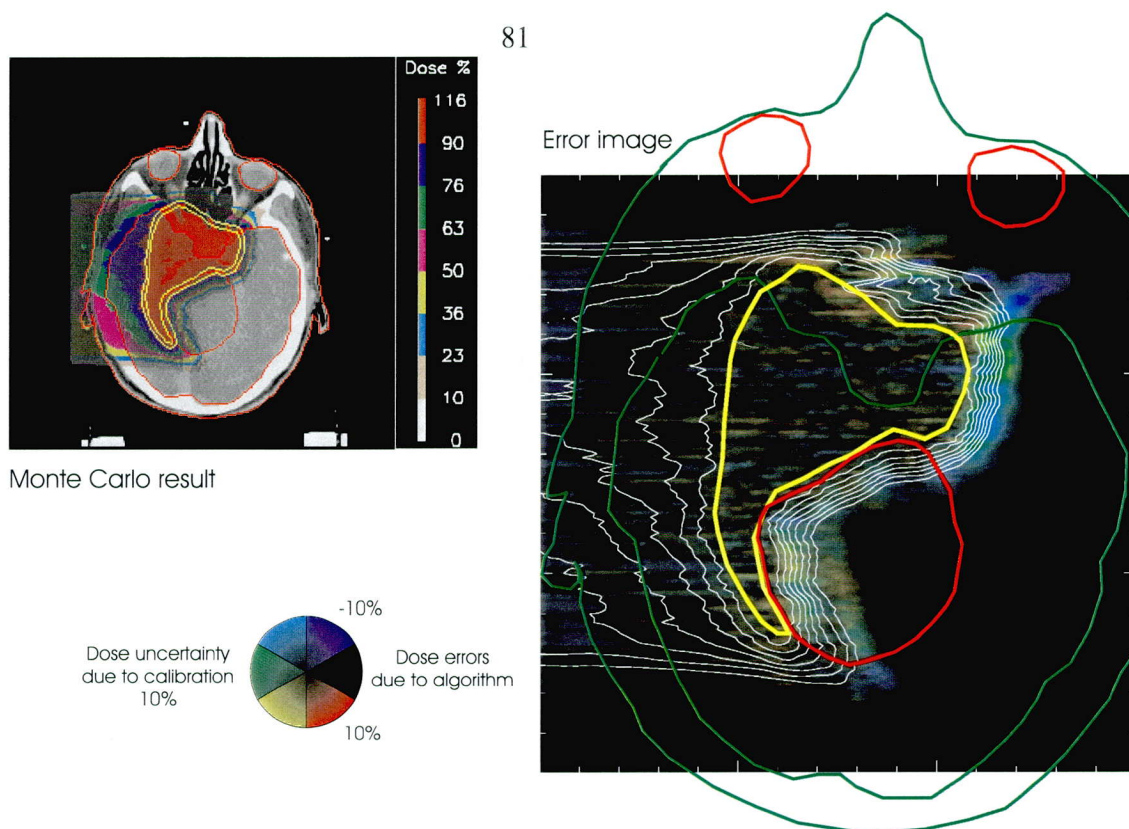due to algorithm

-10%

10%

Fig. 41: The Monte Carlo code can be used to check the influence of different uncertainties on the treatment outcome. The small image shows the Monte Carlo simulation of the treatment. In the large image differences between the results of an analytical calculation and the Monte Carlo simulation are shown in red and blue. The isodose contours are taken from the MC results. Differences in the results of two Monte Carlo simulations using two extreme calibration curves are shown in green. The intensities of the colours indicate the magnitude of the error or uncertainty respectively.
In this case the range dilution is low (see also Fig. 43). The uncertainties in the dose distribution are dominated by the uncertainty in the determination of the stopping power.

Monte Carlo result

Flux-dose model
(Model 1)

Dose errors
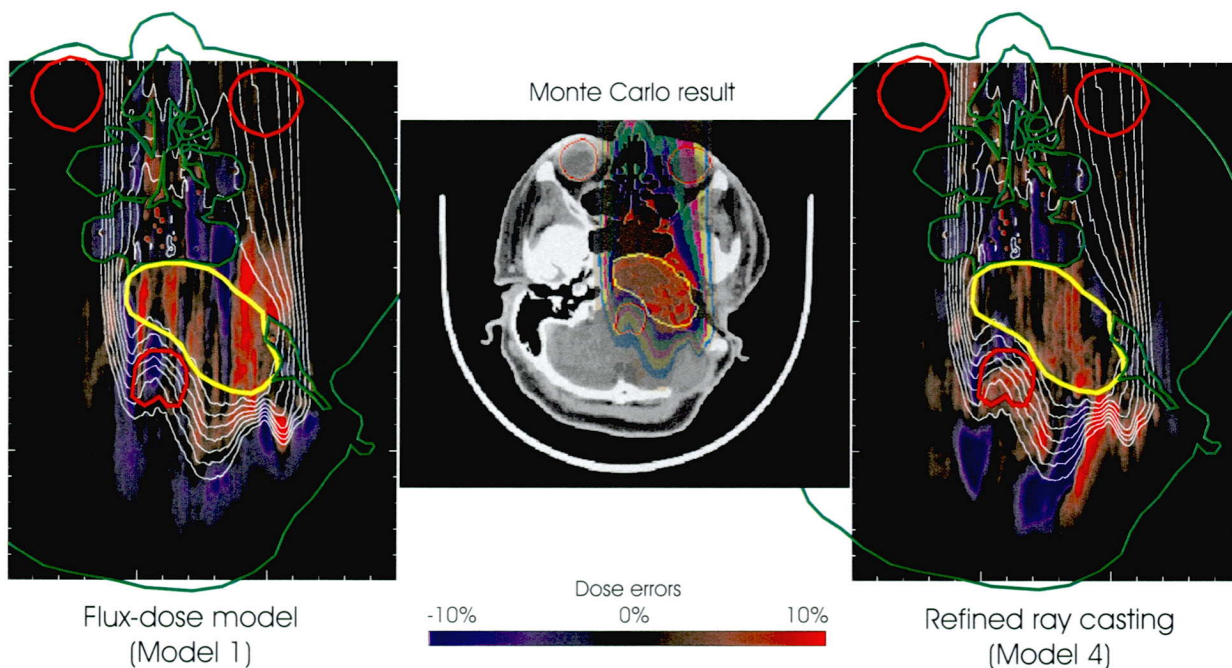-10%   0%   10%

Refined ray casting
(Model 4)

Fig. 42: Display of differences between the results of a Monte Carlo simulation (middle picture) and results of two analytical algorithms for a field planned through the nasal cavities. The display is the same as in Fig. 41, except that the uncertainty due to the determination of the stopping power properties is not shown.
Areas with high dose errors are caused by the heterogeneities around the nasal cavities. They can appear within and/or outside the target. An irradiation like this results in a very unpredictable dose distribution and should be avoided as much as possible.
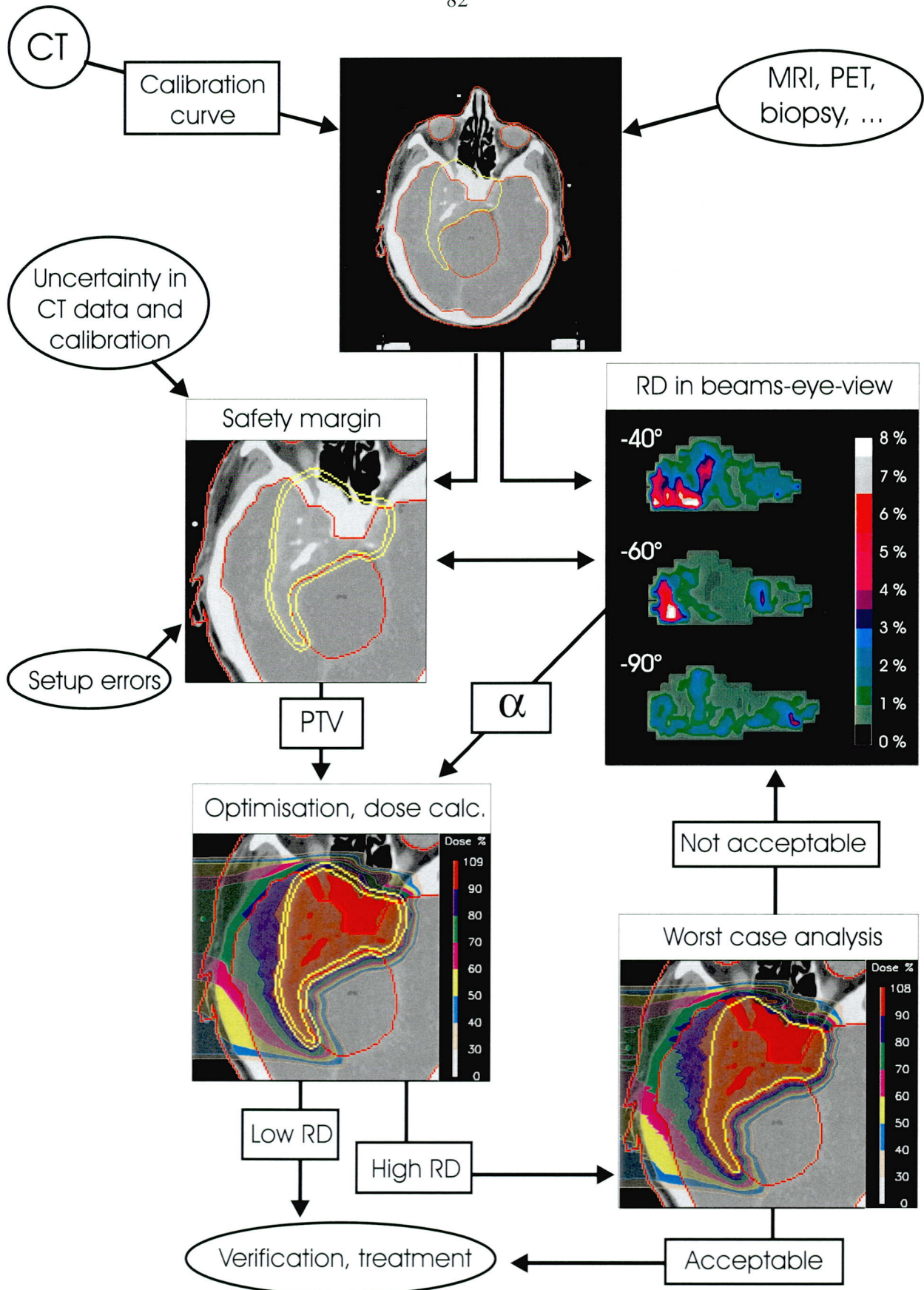
Fig. 43: Flow chart of the process of treatment planning. Special emphasis is put on the topics discussed in this thesis. First the CT images and other medically relevant data are aquired and the CTV is defined. An adequate safety margin and appropriate gantry angles are chosen to deal with uncertainties in the treatment delivery. The optimised dose distribution can be calculated from this input. In case of doubts, i.e. when a high range dilution is found in the target, a worst case analysis should be performed and possibly gantry angle and safety margin be adjusted.