Improved calibration of mass stopping power in low density tissue for a proton pencil beam algorithm

Daniel R Warren¹,², Mike Partridge¹, Mark A Hill¹ and Ken Peach²

¹ CRUK/MRC Oxford Institute for Radiation Oncology, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, UK
² Particle Therapy Cancer Research Institute, University of Oxford, Oxfordshire, OX1 3RH, UK

E-mail: daniel.warren@oncology.ox.ac.uk

Received 16 February 2015, revised 9 March 2015
Accepted for publication 26 March 2015
Published 14 May 2015

Abstract

Dose distributions for proton therapy treatments are almost exclusively calculated using pencil beam algorithms. An essential input to these algorithms is the patient model, derived from x-ray computed tomography (CT), which is used to estimate proton stopping power along the pencil beam paths. This study highlights a potential inaccuracy in the mapping between mass density and proton stopping power used by a clinical pencil beam algorithm in materials less dense than water. It proposes an alternative physically-motivated function (the mass average, or MA, formula) for use in this region. Comparisons are made between dose-depth curves calculated by the pencil beam method and those calculated by the Monte Carlo particle transport code MCNPX in a one-dimensional lung model. Proton range differences of up to 3% are observed between the methods, reduced to <1% when using the MA function. The impact of these range errors on clinical dose distributions is demonstrated using treatment plans for a non-small cell lung cancer patient. The change in stopping power calculation methodology results in relatively minor differences in dose when plans use three fields, but differences are observed at the 2%–2 mm level when a single field uniform dose technique is adopted. It is therefore suggested that the MA formula is adopted by users of the pencil beam algorithm for optimal dose calculation.
in lung, and that a similar approach is considered when beams traverse other low density regions such as the paranasal sinuses and mastoid process.

Keywords: proton therapy, pencil beam algorithms, CT calibration, radiation therapy treatment planning

(Some figures may appear in colour only in the online journal)

1. Introduction

Pencil beam dose calculation algorithms are widely-used in clinical proton therapy, offering a computationally simple semi-empirical alternative to Monte Carlo with fewer parameters needed to describe the treatment machine and patient sufficiently well. Assuming high quality calibration data, these analytical algorithms are generally agreed to calculate dose distributions with an accuracy of a few millimetres (Paganetti 2012); this accuracy is one of many factors that govern the size of the margin between the clinical target volume (CTV) and the planning target volume (PTV) or scanning target volume (STV). Treatment plans with smaller PTV/STV margins are advantageous to patients in terms of decreased toxicity, irradiating less tissue at high dose for technical (as opposed to medical) reasons, but are only possible by reducing spatial uncertainties in dose calculation and delivery.

This paper examines the mapping between computed tomography (CT) scan pixel values via mass density to proton stopping power for a clinical pencil beam algorithm, and compares the dose-depth curves produced by such a method to those directly calculated by Monte Carlo particle transport software. A potential source of systematic error arising from the mapping process is identified, and a solution is demonstrated to resolve these errors. The magnitude of the resulting dose differences are shown for a typical clinical treatment planning scenario.

1.1. Pencil beam dose calculation

The pencil beam concept was introduced in radiotherapy dose calculation because it better accounted for patient heterogeneity in electron beam therapy compared to earlier broad beam methods (Hogstrom et al 1981). As with electrons, the range of a proton beam is strongly dependent on properties of the material traversed. A number of proton pencil beam dose calculation algorithms have therefore been developed (e.g. Petti (1992), Hong et al (1996), Soukup et al (2005), Schaffner (2008) and Westerly et al (2013) and references therein) and, as a result of abundant affordable computing power, are available clinically for interactive radiotherapy treatment planning.

In brief, proton pencil beam algorithms calculate the dose delivered to points within a patient by decomposing the initial fluence distribution into a number of adjacent pencil beams. These calculation beams are distinct from the pencil beams which may be used to deliver treatments, and typically have smaller dimensions. The dose contribution of each pencil beam is the product of a longitudinal depth-dose function and a lateral scattering function. Approaches to lateral scattering vary, but the longitudinal depth-dose curve is invariably derived from depth-dose data measured in water during commissioning, which is scaled and shifted as steps are taken along the pencil beam’s axis, according to the ‘water equivalent’ thickness of the material intersected.

Some commercial treatment planning systems (TPSs) currently implement the method described in Soukup (2005). This extends the pencil beam concept by separating each pencil beam into distinct segments, each with unique mass density and proton stopping power properties. This approach allows for more accurate dose calculation in complex geometries, particularly in the presence of heterogeneities.

1 true thickness multiplied by proton stopping power ratio (SPR) relative to water.
beam into two beamlets, the first due to multiple Coulomb scattering (MCS) and the second arising from nuclear scattering. The relative weight of the nuclear-scattered beamlet increases with ‘water equivalent’ depth following a fit to Monte Carlo simulations. Both beamlets’ lateral spread functions are Gaussian, with widths that are a sum of components arising from angular divergence and the relevant scattering processes. The MCS width is calculated using multiple scattering theory (in particular, a formula by Rossi and Greisen (1941)) and varies depending on the radiation length of material traversed, whilst the nuclear scattering width is again informed by Monte Carlo simulations.

1.2. The patient model

Dose calculations using pencil beam algorithms require a reliable model of the patient’s anatomy, providing a map from which the proton stopping power of material along the beam’s path may be sampled. This patient model is built from the planning CT, and informed by a scanner-specific calibration curve.

1.2.1. CT calibration. The stoichiometric method, as proposed by Schneider et al (1996) and built on by others (Schaffner and Pedroni 1998, Schneider et al 2000), is currently the predominant approach for CT calibration in proton therapy. It can be summarised as follows:

(a) CT scan a selection of tissue equivalent materials, with known density and composition.
(b) Fit the CT numbers (in Hounsfield Units) to a function of density and elemental composition adapted from Rutherford et al (1976), determining coefficients that reflect the strength of each major x-ray attenuation process for the energy spectrum used in the scan.
(c) Use the coefficients from step (b) to predict CT number (the independent x-axis variable in the CT calibration curve) for a range of real biological tissues, using reference data (e.g. in ICRU/ICRP Reports) for their density and composition. For each tissue, also calculate the physical property required by the dose algorithm, which constitutes the dependent y-axis variable. The choice of physical property will be discussed in section 1.2.2.
(d) Use the (x,y) points from step (c) to specify a one-to-one function linking CT number to the calibration variable.

1.2.2. Stopping power versus mass density. Early pencil beam algorithms (e.g. Hong et al 1996) expect the calibration curve to link CT number and proton stopping power ratio relative to water (SPR). In contrast, the method in Soukup et al (2005) uses mass density as the dependent variable; consequently, a TPS using this algorithm for scanned proton treatments expects the calibration curve to link CT number and mass density. There are clear advantages to this approach:

- Mass density can be measured without using proton beam time.
- Mass density is independent of proton energy, whilst the SPR of a material displays an energy dependence, as illustrated in figure 1. Ignoring the variation of SPR with energy leads to range errors of approximately 1 mm water equivalent for 200 MeV protons travelling entirely in mineral bone, with smaller errors in most other biological materials.
- It has been demonstrated that an accurate formula can be found for proton stopping power as a function of energy and mass density, and radiation length as a function of mass density, using tabulated composition data for human tissues (Fippel and Soukup 2004).
1.2.3. Interpolation of stopping power at low density. The region $\rho < 0.9 \text{ g cm}^{-3}$ is poorly populated by human tissues in most sources of composition data, such as ICRU Report 46 (1992) or ICRP Publication 89 (Valentin 2002). The exception is lung, which is typically associated with a mass density approximately one-third that of water. This low density is a bulk average of both tissue and air; most non-osseous tissues, including compressed lung, have an intrinsic density close to that of water. Voxel dimensions in radiotherapy planning CT scans are of order 1 mm, whilst the lung contains structures on a variety of scales: from alveoli at $\sim 100 \mu$m to centimetre-sized bronchi. Each lung voxel in the planning CT may therefore be occupied by a different fraction of tissue, and the range of CT numbers observed within the organ can be expected to extend from 100% air at $-1000$ HU to compressed lung at approximately 0 HU.

In order to implement pencil beam dose calculation, it is necessary to define an approach to SPR calculation over the whole range of CT numbers spanned by low density tissue, based on the very limited set of calibration data available. The description of the algorithm by Soukup et al recommends following an earlier work (Fippel and Soukup 2004) to calculate mass stopping power ratio (denoted $f_S$) by linear interpolation at low densities.

In this work, an in-silico study is carried out to determine the extent to which linear interpolation of mass stopping power in low density ($\rho < 0.9 \text{ g cm}^{-3}$) lung tissue affects the accuracy of proton range calculation by a pencil beam algorithm, by comparison to a proposed alternative (the mass-average method) and direct Monte Carlo calculations.

2. Materials and methods

2.1. Pencil beam algorithm

All longitudinal aspects of the pencil beam dose calculation algorithm described in the Soukup et al paper (Soukup et al 2005) were implemented in a function for the MATLAB
This implementation evaluates SPR, residual proton energy and deposited energy over 10 equally-sized steps in the depth direction per dose calculation voxel. In addition to calculating SPR from mass density according to the original linear interpolation (LI) method, the mass-average (MA) approach was also used, and stopping power was calculated directly from elemental composition using the SRIM software (Ziegler et al 2010).

SRIM is able to account for the effect of molecular bonding on stopping power using compound corrections calculated from the molecular structure (Ziegler and Manoyan 1988); however, these compound corrections are only relevant at proton energies that are considered low in the therapeutic regime, and which are difficult to define for composite materials like lung tissue. Therefore, no molecular corrections were included in SRIM calculations.

2.1.1. Linear interpolation (LI) of SPR at low density. The stopping power calculation method used by the implemented pencil beam algorithm is described in an earlier work (Fippel and Soukup 2004), in which the authors derived a formula to calculate stopping power as a function of energy and mass density in human tissues. In brief, the method involved evaluating stopping power ratios at a range of energies using the PSTAR tool (Berger et al 2005) for tissues with compositions specified in ICRU Report (1992), followed by a multivariate fitting procedure. The resulting equations are reproduced below as (1) and (2), and have a quoted accuracy better than 1% in all ICRU 46 tissues.4

\[
\text{SPR}(\rho, E) = \frac{\rho}{1.000 \text{ g} \cdot \text{cm}^{-3}} f_S(\rho, E)
\]  

\[
f_S(\rho, E) = \begin{cases} 
1.0123 - 3.386 \cdot 10^{-5}E + \ldots & \rho \geq 0.9 \text{ g} \cdot \text{cm}^{-3} \\
0.291 \cdot (1 + E^{-0.3421}) \cdot (\rho^{-0.7} - 1) & 0.9925 \\
0.9925 & \rho = 0.26 \text{ g} \cdot \text{cm}^{-3} \text{ (lung)} \\
0.8815 & \rho = 0.0012 \text{ g} \cdot \text{cm}^{-3} \text{ (air)} \\
\text{interpolate} & \text{all other } \rho < 0.9 \text{ g} \cdot \text{cm}^{-3}
\end{cases}
\]  

where \( \rho \) is mass density in units of g·cm\(^{-3}\) and \( E \) is the proton kinetic energy in units of MeV.

Equation (2) accounts for the small number of calibration materials with low densities by linearly interpolating \( f_S \) between the points for air and bulk lung. Linear interpolation is also performed between the point for bulk lung and the bottom of the continuously-defined range. This approach will be referred to as the LI method.

2.1.2. Mass-average (MA) of SPR at low density. An alternative approach to interpolation is proposed here. Since \( f_S \) in equation (1) is the ratio of the mass stopping power of a material relative to that of water, its value in a mixture can be calculated as a mass-average of the mass stopping power of its constituent materials, under the assumption that each contributes independently to proton stopping. For a mixture of density \( \rho \) containing materials of density \( \rho_1 \) and \( \rho_2 \), the mass stopping power ratio can therefore be given by equation (3).
The new mass-average formula is given in equation (4). It uses equation (3) to calculate $f_s$ between the points defined for air and lung. Additionally the single $f_s$ value for lung is extended to the entire region $0.26 \, \text{g} \cdot \text{cm}^{-3} \leq \rho < 0.8 \, \text{g} \cdot \text{cm}^{-3}$, motivated by the fact that $\geq 95\%$ of internal CT voxels with densities below $0.8 \, \text{g} \cdot \text{cm}^{-3}$ appear within the outlined lung contour for lung radiotherapy patients; however, this modification should be examined in detail before being applied to other treatment sites. Equation (3) is used again between $0.8 \, \text{g} \cdot \text{cm}^{-3}$ and the bottom of the continuously-defined range. The equation matches the LI method (equation (2)) at densities above $0.9 \, \text{g} \cdot \text{cm}^{-3}$.

\begin{equation}
    f_s(\rho, E) = \begin{cases}
        1.0123 - 3.386 \cdot 10^{-5}E + ... & \rho \geq \rho_f^{\min} \\
        0.291 \cdot (1 + E^{-0.3421}) \cdot (\rho^{-0.7} - 1) & \rho_f^{\min} < \rho < \rho_f^{\max} \\
        0.9925 & \rho_f^{\max} \leq \rho \leq \rho_f^{\min} \\
        f_s(\rho, \rho_l, \rho_l) & \rho_f^{\min} < \rho < \rho_f^{\max} \\
        0.8815 & 0 < \rho \leq \rho_f^{\min}
    \end{cases}
\end{equation}

where $\rho_a = 0.0012 \, \text{g} \cdot \text{cm}^{-3}$, $\rho_l = 0.26 \, \text{g} \cdot \text{cm}^{-3}$, $\rho_f^{\max} = 0.8 \, \text{g} \cdot \text{cm}^{-3}$ and $\rho_f^{\min} = 0.9 \, \text{g} \cdot \text{cm}^{-3}$.

A comparison between SPR calculated by the mass-average method (hereafter denoted MA) and the LI method is shown in figure 2.

2.1.3. Reference data. Three sets of reference pencil beam data (mean energies of 60 MeV, 90 MeV and 170 MeV) were generated using the MCNPX v2.7.0 Monte Carlo particle transport code (Pelowitz et al. 2011) and the default LA150 cross section dataset. In each case, a collimated beam of protons with Gaussian lateral fluence profile ($\sigma = 1 \, \text{mm}$) and Gaussian initial energy distribution ($\sigma = 1 \, \text{MeV}$) was generated to be incident perpendicular to the centre of a face of a $40 \, \text{cm} \times 40 \, \text{cm} \times 40 \, \text{cm}$ water cube. The energy deposited was recorded using a cylindrical mesh tally (type CMESH3) with radial bins of 1 mm and depth bins of 0.1 mm within a region of diameter 12 cm and depth 30 cm. Only protons were transported (including secondary protons arising from elastic recoil, by using a recl parameter of 1 on the PHYS:H card) to a minimum energy of 100 keV, after which the remainder of their energy was deposited locally. The energy of all other particles was deposited at the point of generation.

2.2. Monte Carlo calculations

The MCNPX Monte Carlo code was also used as a benchmark against which the pencil beam method as a whole could be tested, alongside the various approaches to SPR determination. It allows dose calculation at higher spatial resolution than a clinical TPS and determines stopping power based on the full elemental composition of a material and residual proton energy on a particle-by-particle basis. The simulation setup mirrored that used to generate the reference data in section 2.1.3, with the water cube replaced by a one-dimensional longitudinal patient model with a lateral cross section of $40 \, \text{cm} \times 40 \, \text{cm}$. Since only dose-to-material may be found in the standard release of MCNPX, dose-to-water in arbitrary materials was obtained by dividing the tallied energy deposition per unit volume by the mass density and the
SRIM-calculated SPR, assuming the residual proton energy given by the pencil beam method at the same depth. Each direct MCNPX calculation ran until the relative uncertainties in each laterally-integrated longitudinal bin were below 0.25%, with the vast majority of bins displaying considerably smaller variance. Total number of histories ranged between $10^7$ and $3 \times 10^8$.

2.3. Patient models

2.3.1. Benchmarking geometry. A multi-material virtual phantom was used to benchmark the pencil beam implementation, containing homogenous tissues with compositions as specified in ICRU Report 46 (also used in deriving equation (2)). The geometry comprised five neighbouring 3 cm $\times$ 40 cm $\times$ 40 cm slabs, each composed of a different material, followed by a final 25 cm $\times$ 40 cm $\times$ 40 cm slab. Variations were produced containing the following six materials, in this order and its cyclic permutations: adipose tissue, muscle, lung, cartilage, cortical bone and water. One such arrangement is illustrated in figure 4.

2.3.2. Lung models. One-dimensional lung models were created by extracting the CT numbers of all voxels within the demarcated lung contours of 25 non-small cell lung cancer (NSCLC) radiotherapy patients who participated in the ROCOCO in-silico planning trial (Roelofs et al 2012). Representative mass densities were calculated assuming images were acquired with an x-ray energy of 80 keV$^5$ and tissue had ICRU lung composition. Voxels below $-990$ HU and above 100 HU were excluded (2.3% of the total number extracted). The resulting mass density distribution is shown in figure 3.

Virtual phantoms were created by sampling the CT number distribution 300 times at intervals corresponding to one-third-percentiles and assigning the resulting densities to neighbouring 1 mm $\times$ 40 cm $\times$ 40 cm slabs of ICRU lung in four arrangements: with density ramping up, ramping down, ramping up and then down (peaked), and alternating high and low densities which converge on the median.

$^5$ Estimated mass densities change by no more than 1% over the range 50–150 keV.
2.4. Assessment of clinical impact

A comparison was made between treatment plans that were calculated assuming the linear interpolation and mass average methods for mass stopping power determination. The XiO v4.70 treatment planning system (Elekta Ltd., Stockholm) was used to design spot-scanned proton plans using three techniques (single-field uniform dose, three uniform dose fields and three-field IMPT) for one NSCLC patient who participated in the ROCOCO planning trial. All plans had a prescription dose of 70 Gy (RBE) and were designed following the dose constraints detailed in the ROCOCO protocol. Whilst XiO calculates dose using the linear interpolation method of mass stopping power determination, a good approximation of the mass averaged method was obtained by modifying the CT calibration curve using the transformation derived in appendix A: comparisons using the MATLAB pencil beam implementation indicate the accuracy of this approach is better than 0.1 mm.

Dose distributions for each of the treatment plans and SPR methods were exported from the TPS in DICOM format and imported into CERR (Computational Environment for Radiotherapy Research) (Deasy et al 2003) for comparison, using the dose subtraction and gamma analysis functions.

3. Results

3.1. Benchmarking

The pencil beam implementation was tested using a 170 MeV proton source in the six variations of the benchmarking geometry previously described. The calculated dose-depth curves were compared with those found using MCNPX: results are summarised in table 1, and shown graphically for one arrangement in figure 4.

It can be seen that in all geometries and with all SPR-calculation methods the pencil beam implementation predicts 80–20% isodose distal fall-off distance within 0.4 mm of the Monte
Table 1. Differences in Bragg peak position $R$ and 80–20% isodose distal fall-off distance $d_{80-20}$ between pencil beam (various SPR methods) and direct MCNPX dose calculations for 170 MeV protons in the benchmark geometry described in section 2.2.

| Final material: | Water | Cortical bone | Cartilage | Lung | Muscle | Adipose tissue |
|----------------|-------|---------------|-----------|------|--------|---------------|
| $\delta R$ (mm) | $\delta d_{80-20}$ (mm) | $\delta R$ (mm) | $\delta d_{80-20}$ (mm) | $\delta R$ (mm) | $\delta d_{80-20}$ (mm) | $\delta R$ (mm) | $\delta d_{80-20}$ (mm) |
| SPR Method | Linear Interpolation (constant E) | Linear Interpolation | Mass Average | SRIM |
| $-0.4$ | $-0.4$ | $-0.2$ | $-0.6$ | $-0.2$ | $-0.1$ | $-0.4$ | $-0.2$ |
| $0$ | $0.1$ | $-0.2$ | $0.1$ | $0$ | $0$ | $0$ |

Note: In the 170 MeV case, the 30 cm of modelled lung was followed by water into which the Bragg peak penetrated approximately 11 cm. $\delta R$ is given as a physical distance (in mm) and as a water equivalent distance relative to the MCNPX-predicted range (RPL). LI = Linear Interpolation.

Table 2. Differences in Bragg peak position $R$ and 80–20% isodose distal fall-off distance $d_{80-20}$ between pencil beam (various SPR methods) and direct MCNPX dose calculations for 60 MeV, 90 MeV and 170 MeV protons in the four lung density models described in section 2.3.2.

| Density arrangement: | $\delta R$ (mm) | $\delta d_{80-20}$ (mm) | $\delta R$ (RPL) (mm) | $\delta d_{80-20}$ (RPL) (mm) | $\delta R$ (mm) | $\delta d_{80-20}$ (mm) | $\delta R$ (RPL) (mm) | $\delta d_{80-20}$ (RPL) (mm) |
|---------------------|----------------|----------------------|-----------------------|-----------------------|----------------|----------------------|-----------------------|-----------------------|
| 60 MeV              | $+0.4$        | $+0.2$               | $+0.9$                | $+0.5$                | $+0.7$        | $+0.4$               | $+0.8$                | $+0.4$                |
| LI (constant E)     | $+0.3$        | $+0.2$               | $+0.9$                | $+0.5$                | $+0.7$        | $+0.4$               | $+0.8$                | $+0.4$                |
| LI (variable E)     | $+0.1$        | $+0.1$               | $+0.5$                | $+0.3$                | $+0.3$        | $+0.2$               | $+0.4$                | $+0.2$                |
| SRIM                | $0$           | $0$                  | $0$                   | $0$                   | $0$           | $0$                  | $0$                   | $0$                   |

Note: In the 170 MeV case, the 30 cm of modelled lung was followed by water into which the Bragg peak penetrated approximately 11 cm. $\delta R$ is given as a physical distance (in mm) and as a water equivalent distance relative to the MCNPX-predicted range (RPL). LI = Linear Interpolation.
Carlo result. The range is also consistently predicted within 0.4 mm of the Monte Carlo value, except when the final material is lung. However, considering the lung material’s density of 0.26 g·cm$^{-3}$, even these differences are of comparable water equivalent size to those in water and other soft tissues. These results suggest that the spatial error of this pencil beam implementation is smaller than the dose calculation voxels typically used for proton therapy treatment planning, and suitable for comparing SPR calculation methods.

The predictions of the pencil beam method in this scenario are identical for the LI and MA approaches to low density SPR calculation. This is because the benchmark model has homogenous tissues with mass densities as defined by the ICRU, so no interpolation or mass averaging is required.

### 3.2. Lung model

A summary of the results of comparative simulations using MCNPX and pencil beam methods in the lung model can be seen in table 2, for proton sources with energy 60 MeV, 90 MeV and 170 MeV. Dose-depth curves at 60 MeV and 90 MeV are shown in figure 5. Since the water equivalent thickness of the 30 cm lung model is approximately 8 cm, results for the 170 MeV beam (19.5 cm range in water) were obtained by placing a water phantom directly downstream of the model.

In this one-dimensional lung model with a single material at multiple mass densities, the range prediction errors introduced by linear interpolation of SPR are approaching a clinically significant scale, depending on mass density distribution and proton energy, ranging from 0.4 mm in water at high energy to 3.6 mm in low density lung at lower energies. When considered in terms of water-equivalent path length, these errors correspond to 0.2–3.0% of the proton range. With one exception, Bragg peak positions for curves calculated with SRIM-derived SPRs differ from Monte Carlo data by less than 0.25% in water equivalent terms and vary little with density arrangement. This implies that a large proportion of the differences
for the LI and MA cases can be attributed to SPR calculation rather than the pencil beam implementation. As in section 3.1, dose-depth curves do not change appreciably as a result of allowing SPR to vary with energy.

Figure 5. Left: longitudinal dose calculations for 60 and 90 MeV protons in 30 cm of lung tissue with mass density distribution as given in figure 3. Results are shown for four spatial distributions of mass, each with pencil beam calculations using the SPR-determination methods described in section 2.1 and a direct Monte Carlo simulation. Right: the 60 MeV data enlarged to show the region of the Bragg peak, with a secondary axis displaying the longitudinal integral of mass density. In each case two groups of overlapping lines can be seen: the ‘LI, Constant E’ and ‘LI, Variable E’ lines coincide, as do, the lines for ‘MA, Variable E’, ‘SRIM’ and ‘Direct MC’.
The tabulated results are best considered with reference to figure 2. It can be seen that the mass stopping power predictions of the LI model quickly depart from the MA model (which takes a similar approach to MCNPX) both above and below the ICRU lung density of 0.26 g·cm\(^{-3}\). Approximately 40% of the total mass in the lung model appears at densities below this value, where the LI approach will underestimate the SPR by up to 12%, whilst the SPR may be overpredicted by up to 5% in the remaining 60% of mass. The net result on Bragg peak position depends on the proportion of the mass traversed by protons (≈40% at 60 MeV, ≈80% at 90 MeV and 100% at 170 MeV) and the density at which that mass appears.

In the case of calculations for 170 MeV protons, the whole density spectrum is traversed and the opposing directions of range errors on each side of the distribution result in the LI method being a relatively good predictor of range. This statement is also approximately true for the alternating and triangular distributions at 90 MeV. At 60 MeV and 90 MeV, the LI method can be seen to over-predict range in the ramp-up arrangement, and under-predict range in the ramp-down arrangement, as may be expected after examination of figure 2.

In summary, the proton range errors resulting from use of the LI model with an unmodulated Bragg peak are variable and dependent on the mass distribution traversed, but may be greater than the dose calculation voxel sizes typically used in proton therapy treatment planning (1–3 mm). As each of the individual spots that comprise a scanned proton treatment comprises protons of a different initial energy, traversing a variety of different paths and therefore a variety of mass distributions, the impact on clinical dose distributions is best understood through consideration of a treatment plan.

3.3. Clinical comparison

Five spot-scanned NSCLC treatment plans were created to examine the effect of SPR interpolation in realistic and more complex irradiation scenarios. A numerical summary of the difference in dose distributions calculated using the LI method and a good approximation of the MA method is given in table 3. It can be seen that the range of dose differences is plan-dependent, being greater for the single field uniform dose (SFUD) plans than for the three-field plans. The volume of large-scale differences in the dose distributions is also quantified by the number of voxels which fail a 3D gamma test (Low et al 1998) with parameters 0.02\(D_{\text{max}}\) and 2 mm: one-field plans have a significant number of failures, whilst essentially all voxels in three-field plans pass the test. The better performance of three-field plans can by explained by the fact that range errors in one field will to some degree be compensated by dose from other fields. For the three SFUD plans considered, an increase in gantry angle from 90° results in further increased deviations of the LI results from MA, due to an increase in the thickness of lung tissue intersected by the beam.

A dose distribution and a LI-MA dose difference map are shown for one slice of the SFUD 120° plan in figure 6. Two pronounced regions of range error are apparent: an over-prediction of range by the LI method in the anterior left lung and an under-prediction of range by the LI method close to the right wall of the mediastinum. In the former position protons have passed through a large proportion of low-density tissue (similar to the 1D ramp-up case), whilst the traversed lung tissue is at higher densities in the latter position (similar to the 1D ramp-down case).

4. Discussion and conclusions

For the model above, it has been shown that errors of up to 4 mm (or 3% of water equivalent range) may result in the range predicted by a pencil beam algorithm using the LI method
compared to Monte Carlo. In the clinical treatment plans considered, switching from MA method to LI introduced small regions of dose difference equating to 7% of the prescription dose. It should be noted that these potential algorithmic errors are only relevant in regions containing significant amount of low density tissue and, even then, are very much dependent on the distribution of mass along the path of the protons. They should be viewed in the context of other proton range uncertainties in lung, for example:

- Pencil beam methods can suboptimally model protons which have undergone scatter in complex heterogeneities, particularly because a beamlet’s evolution is typically calculated.

### Table 3. Minimum, maximum and standard deviation of the difference in dose distributions calculated using linear interpolation (equation (2)) and mass-averaging (equation (4)) for proton mass stopping power in the low density region.

| Plan       | $\Delta D_{\text{min}}$ [Gy (RBE)] | $\Delta D_{\text{max}}$ [Gy (RBE)] | $\sigma_{\Delta D}$ [Gy (RBE)] | $V(\gamma_{2\%–2\,\text{mm}} > 1)$ [voxels/\%] |
|------------|------------------------------------|------------------------------------|--------------------------------|-----------------------------------------------|
| SFUD 90°   | $-$3.61 3.83                       | 0.33                               | 231/0.03%                     |
| SFUD 120°  | $-$4.16 3.54                       | 0.40                               | 159/0.03%                     |
| SFUD 150°  | $-$4.42 4.97                       | 0.41                               | 1023/0.17%                    |
| 3FUD       | $-$2.68 2.01                       | 0.14                               | 1/0.00%                       |
| 3F-IMPT    | $-$2.78 2.21                       | 0.14                               | 0/0.00%                       |

Note: The final column reports the number of voxels which fail a 3D gamma test (Low et al. 1998) with maximum allowable dose difference of 2% of the prescription dose or maximum allowable distance-to-agreement of 2 mm, also expressed as a percentage of the number of voxels with non-zero dose. Results are given for five treatment plans on a non-small cell lung cancer patient: three single-field uniform dose (SFUD) plans at different gantry angles, an equally-weighted sum of the SFUD plans (3FUD) and three-field IMPT (3F-IMPT).
based only on the material intersected by its axis. This has recently been investigated on a site-by-site basis, with range margins of 6.3% + 1.2 mm having been recommended in lung, potentially reduced by patient-specific optimisation (Schaumann et al 2014).

- Motion, including interplay with the scanning dynamics, can lead to mean differences between planned and delivered doses in the target of approximately 5%, and localised differences in excess of 20%, during a fractionated treatment (Grassberger et al 2013).
- Errors in CT calibration and the effect of image artefacts, such as beam hardening and stochastic noise, may introduce water equivalent range prediction errors of approximately 0.5% along typical paths through lung (Warren 2014).

A literature search did not reveal any dose calculation comparisons between a pencil beam algorithm using the LI method and one that uses an alternative SPR calculation method, or Monte Carlo simulation, in low density tissue. However, the developers of the Monte Carlo platform GATE reported a large discrepancy between XiO’s pencil beam algorithm and their own method for stopping power determination in the density interval 0.5–0.9 g cm$^{-3}$ (Grevillot et al 2012), which would be an expected result of the TPS using the LI method.

Outside of lung, similar proton range errors may exist in doses calculated by the LI method in the paranasal sinuses, which contain fine structures composed of soft tissue and air. Related range errors may also occur in beams travelling through the mastoid process, part of the skull. The mastoid is composed chiefly of bone and air pockets, many of which are smaller than the CT voxel size, yielding pixel values in the approximate range −950 HU to +1700 HU, as seen in figure 7(a). Whilst equation (3) can be used to create a calibration curve relevant to bone-air mixtures, this will not be applicable to all tissue in the relevant CT slices because the mass density (and Hounsfield Unit) ranges in soft tissue and brain completely overlap those of mastoid. The issue is illustrated by the idealised CT calibration curves shown in figure 7(b), which show departures of up to 40% in SPR prediction if mid-density tissue is considered to be composed of a bone-air mix instead of soft tissue. Range calculation errors of multiple millimetres may result, and are likely to be common to all proton dose calculation methods which rely on a single CT calibration curve. One possible resolution to the problem is for TPS developers to allow users to apply different stopping power models within different regions of interest.

The discussion in this paper has been devoted to biological materials because equations (2) and (4) were derived for human tissues, but it should be noted that non-biological materials may be present in the TPS calculation volume; for example, range shifters used in the case of shallow treatment volumes are sometimes separate from the gantry or treatment machine, or the patient may have prostheses. Careful consideration should be made as to whether corrections to the calibration curve or patient model are necessary if significant thicknesses of non-tissue material will be traversed by protons, such as by assigning the relevant regions a ‘tissue equivalent mass density’ calculated from the inverse of equation (1). Silicone breast implants, for instance, may need a modification of up to 7% in the equivalent mass density of the patient model in order for a TPS using the LI method to assume the measured SPR of 0.935 (Moyers et al 2014).

The modification presented here is suggested on the grounds that it is advisable for all aspects of radiotherapy dose calculation algorithm to be physically-motivated. It will provide optimal dose calculations in lung for users of this pencil beam method$^6$, but the differences compared to using the previous mass density to SPR method should be within current safety margins, and entirely negligible for most non-lung treatments. Introducing a change to the MA method should be trivial for TPS developers, but they may instead wish to consider the benefit of accepting calibration curves that link CT numbers directly to linear stopping power $^6$ and for users of the VMCPro Monte Carlo code, for which equation (2) was originally derived.
ratio rather than mass density. At its most basic, this would result in a very small additional inaccuracy since it would not account for energy dependence of stopping power, but it seems likely that correlations exist between SPR calculated at a defined energy and that at lower energies. Such a move would allow the physicists commissioning a proton therapy system to be fully aware of, and to take full responsibility for, the methods by which a very important physical property is determined.

Acknowledgments

The authors acknowledge the support of Cancer Research UK in funding M Partridge and D Warren (grant C5255/A15935), the University of Oxford Department of Physics for previously funding of D Warren, and the Oxford Martin School for funding the Particle Therapy Cancer Research Institute (PTCri). We also thank Professor B Jones for initially raising the question of proton dose calculation accuracy in tissue-air and bone-air mixtures.

Appendix A. Modifying SPR calculation by transforming the CT calibration curve

A CT calibration curve $C$ for use with the Soukup pencil beam algorithm may be defined as a one-to-one function mapping CT number $h$ to mass density relative to water $\rho$:

$$\rho = C(h) \quad (A.1)$$

Internally, the pencil beam algorithm calculates stopping power ratio $S$ for protons of kinetic energy $E$ in a material of mass density $\rho$ by equation (1):

$$S(\rho, E) = f_{s,LI}(\rho, E) \cdot \rho, \quad (A.2)$$

where $f_{s,LI}$ is the piecewise function corresponding to the LI method, given in equation (2).
The MA method calculates stopping power ratio using a similar expression:
\[ S(\rho, E) = f_{S,MA}(\rho, E) \cdot \rho, \]  
(A.3)

where \( f_{S,MA} \) is the piecewise function corresponding to the MA method, given in equation (4).

A transformed calibration curve \( C_m \) is sought, in which CT number is instead mapped to a modified mass density \( \rho_m \). For a particular CT number, \( \rho_m \) is the mass density which causes the LI formula to return the same value for \( S \) as the MA formula returns for original mass density \( \rho_o \). This situation may be expressed by equating the right-hand sides of (A.2) and (A.3):
\[ f_S f_S(E, \rho) = f_{S,MA}(\rho_o, E) \cdot \rho_o \]  
(A.4)

It is assumed that that neither \( f_S \) function changes significantly with \( E \), which is true over the majority of a proton’s path length, especially in the low density region—see figures 1 and 2. The dependence on \( E \) is therefore dropped, and (A.4) may be rearranged to give:
\[ \rho_m = \frac{f_{S,MA}(\rho_o) \cdot \rho_o}{f_{S,LI}(\rho_m)} \]

The LI and MA methods only differ in the region \( \rho < 0.9 \), where \( f_{S,LI} \) is defined by two linear segments. If \( f_{S,LI}(\rho) = a \cdot \rho + b \), then the expression for \( \rho_m \) becomes a quadratic equation for which the positive solution is:
\[ \rho_m = \frac{1}{2a} (-b + \sqrt{b^2 + 4a \cdot f_{S,MA}(\rho_o) \cdot \rho_o}) \]  
(A.5)

In the first linear segment of equation (2), \( a = \frac{\Delta f_S}{\Delta \rho} = \frac{0.9925 - 0.8815}{0.26 - 0.2012} = 0.4289 \) and \( b = f_S - a \cdot \rho = 0.9925 - 0.4289 \times 0.26 = 0.8809 \).

In the second linear segment of equation (2),
\[ a(E) = \frac{\Delta f_S(E)}{\Delta \rho} = \frac{1.0123 - 3.386 \cdot 10^{-5} E + 0.291 \cdot (1 + E^{-0.3421}) \cdot (0.9^{-0.7} - 1) - 0.9925}{0.9 - 0.26} \]

and
$b = f_3 - a \cdot \rho = 0.9925 - 0.26a(E)$

Figure A1 shows the relative size of modifications made to the calibration curve if the above equation is evaluated using the relevant coefficients. The modification is equivalent to an absolute change in linear SPR of 3–4% at maximum deviation (near –200 HU) over the energy range 10–250 MeV. Simulations suggest that ignoring the energy dependency in $\rho_m/\rho_o$ introduces water equivalent range errors of less than 0.1 mm over the path lengths utilised in this work.

**Appendix B. Calculating idealised CT calibration curves**

The idealised CT calibration curves shown in figure 7 were calculated for two tissue models:

- A/S/B Model: for mass densities from 0 to 1.06 g cm$^{-3}$ assume a mixture of air and soft tissue, and for mass densities from 1.06 g cm$^{-3}$ to 1.92 g cm$^{-3}$ assume a mixture of soft tissue and bone
- A/B Model: for mass densities from 0 to 1.92 g cm$^{-3}$ assume a mixture of air and bone.

The material compositions were taken from ICRU Report 44.

Relative ratios of the constituent materials were determined for each CT number ($h$) using x-ray mass attenuation coefficients ($\mu/\rho$) from the NIST database (Hubbell and Seltzer 2004) by the following method, which assumes a monenergetic x-ray source.

The expected mass attenuation coefficient for a composite of materials 1 and 2 with weight fractions $w_1$ and $w_2 = 1 - w_1$, is given by the mixture rule:

$$
\left( \frac{\mu}{\rho} \right)_{\text{mix}} = w_1 \left( \frac{\mu}{\rho} \right)_1 + (1 - w_1) \left( \frac{\mu}{\rho} \right)_2 \tag{B.1}
$$

The density of such a mixture, using volume fractions $v_1$ and $v_2 = 1 - v_1$, is given by:

$$
\rho_{\text{mix}} = v_1 \rho_1 + (1 - v_1) \rho_2 \tag{B.2}
$$

where volume and weight fractions are related by $w_i = \frac{v_i \rho_i}{\rho_{\text{mix}}}$.

Substituting equation (B.2) into equation (B.1), and rearranging for the fraction of material 1 gives:

$$
v_1 = \frac{\mu_{\text{mix}} - \mu_2}{\mu_1 - \mu_2}
$$

The linear attenuation coefficient $\mu_{\text{mix}}$ was derived from the CT number of interest using the definition of the Hounsfield unit:

$$
\mu_{\text{mix}} = \left( \frac{h_{\text{mix}}}{1000} + 1 \right) \cdot (\mu_{\text{water}} - \mu_{\text{air}}) + \mu_{\text{air}}
$$

Once the material mixture ratios are determined, the elemental composition by weight of the composite may be found by:

$$
w_{i,\text{mix}} = \frac{v_i \rho_{1i} + v_2 \rho_{2i}}{\rho_{\text{mix}}}
$$
where \(i\) labels chemical elements. Proton stopping powers were calculated in SRIM (Ziegler et al 2010) for mixtures corresponding to each CT number, and divided by the stopping power of water to arrive at SPRs.

References

Berger M J, Coursey J S, Zucker M A and Chang J 2005 Stopping-Power and Range Tables for Electrons, Protons and Helium Ions (www.nist.gov/pml/data/star/index.cfm)
Deasy J O, Blanco A I and Clark V H 2003 CERR: a computational environment for radiotherapy research Med. Phys. 30 979–85
Fippel M and Soukup M 2004 A Monte Carlo dose calculation algorithm for proton therapy Med. Phys. 31 2263
Grassberger C, Dowdell S, Lomax A, Sharp G, Shackleford J, Choi N, Willers H and Paganetti H 2013 Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer Int. J. Radiat. Oncol. Biol. Phys. 86 380–6
Grevillon L, Bertrand D, Desy F, Freud N and Sarrut D 2012 GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans Phys. Med. Biol. 57 4233–44
Hogstrom K R, Mills M D and Almond P R 1981 Electron beam dose calculations Phys. Med. Biol. 26 445–59
Hong L, Goitein M, Buciolini M, Comiskey R, Gottschalk B, Rosenthal S, Serago C and Urie M 1996 A pencil beam algorithm for proton dose calculations Phys. Med. Biol. 41 1305–30
Hubbell J and Seltzer S 2004 Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients (www.nist.gov/pml/data/xraycoef)
ICRU Report 44 1989 Tissue substitutes in radiation dosimetry and measurement Technical Report (Bethesda, MD: International Commission on Radiation Units and Measurements)
ICRU Report 46 1992 Photon, electron, proton and neutron interaction data for body tissues Technical Report (Bethesda, MD: International Commission on Radiation Units and Measurements)
Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions Med. Phys. 25 656–61
Moyers M F, Mah D, Boyer S P, Chang C and Pankuch M 2014 Use of proton beams with breast prostheses and tissue expanders Med. Dosim. 39 98–101
Paganetti H 2012 Range uncertainties in proton therapy and the role of Monte Carlo simulations Phys. Med. Biol. 57 R99–117
Pelovitz D B et al 2011 (MCNPX 2.7.0 Extensions) (http://mcnpx.lanl.gov/opendocs/versions/v270/v270.pdf)
Petti P L 1992 Differential-pencil-beam dose calculations for charged particles Med. Phys. 19 137
Roelofs E, Engelsman M, Rasch C, Persoon L, Qamhiyeh S, de Ruyscher D, Verhaegen F, Pijls-Johannesma M and Lambin P 2012 Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer J. Thoracic Oncol. 7 165–76
Rossi B and Greisen K 1941 Cosmic-ray theory Rev. Mod. Phys. 13 240–309
Rutherford R A, Pullan B R and Isherwood I 1976 Measurement of effective atomic number and electron density using an EMI scanner Neuroradiology 11 15–21
Schaffner B and Pedroni E 1998 The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT–HU and proton stopping power Phys. Med. Biol. 43 1579–92
Schaffner B 2008 Proton dose calculation based on in-air fluence measurements Phys. Med. Biol. 53 1545–62
Schneider U, Pedroni E and Lomax A 1996 The calibration of CT Hounsfield units for radiotherapy treatment planning Phys. Med. Biol. 41 111–24
Schneider W, Bortfeld T and Schlegel W 2000 Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions Phys. Med. Biol. 45 459–78
Schuemann J, Dowidell S, Grassberger C, Min C H and Paganetti H 2014 Site-specific range uncertainties caused by dose calculation algorithms for proton therapy Phys. Med. Biol. 59 4007–31
Soukup M, Fippel M and Alber M 2005 A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations Phys. Med. Biol. 50 5089–104
Valentin J 2002 Basic anatomical and physiological data for use in radiological protection: reference values Ann. ICRP 32 1–27
Warren D R 2014 Proton radiotherapy uncertainties arising from computed tomography DPhil Thesis University of Oxford
Westerly D C, Mo X, Tomé W A, Mackie T R and DeLuca P M 2013 A generalized 2D pencil beam scaling algorithm for proton dose calculation in heterogeneous slab geometries Med. Phys. 40 061706
Ziegler J F, Ziegler M and Biersack J 2010 SRIM: the stopping and range of ions in matter Nucl. Instrum. Methods Phys. Res. B 268 1818–23
Ziegler J and Manoyan J 1988 The stopping of ions in compounds Nucl. Instrum. Methods Phys. Res. 35 215–28