Acta Oncologica, 2015; 54: 1638–1642

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark

DAVID C. HANSEN[1], JOAO SECO[2], THOMAS SANGILD SØRENSEN[3], JØRGEN BREEDE BALTZER PETERSEN[4], JOACHIM E. WILDBERGER[5], FRANK VERHAEGEN[6,7] & GUILLAUME LANDRY[6,8]

[1]*Department of Oncology, Aarhus University Hospital, Aarhus, Denmark,* [2]*Massachussetts General Hospital and Harvard Medical School, Boston, USA,* [3]*Department of Clinical Medicine, Aarhus University, Aarhus , Denmark,* [4]*Department of Medical Physics, Aarhus University Hospital, Denmark,* [5]*Department of Radiology, Maastricht University Medical Center (MUMC), The Netherlands,* [6]*Department of Radiation Oncology (MAASTRO), Maastricht University Medical Center (MUMC), Maastricht, the Netherlands,* [7]*Medical Physics Unit, Department of Oncology, McGill University, Canada and* [8]*Faculty of Physics, Department of Medical Physics, Ludwig-Maximilians-University, Munich, Germany*

ABSTRACT

Background. Accurate stopping power estimation is crucial for treatment planning in proton therapy, and the uncertainties in stopping power are currently the largest contributor to the employed dose margins. Dual energy x-ray computed tomography (CT) (clinically available) and proton CT (in development) have both been proposed as methods for obtaining patient stopping power maps. The purpose of this work was to assess the accuracy of proton CT using dual energy CT scans of phantoms to establish reference accuracy levels.

Material and methods. A CT calibration phantom and an abdomen cross section phantom containing inserts were scanned with dual energy and single energy CT with a state-of-the-art dual energy CT scanner. Proton CT scans were simulated using Monte Carlo methods. The simulations followed the setup used in current prototype proton CT scanners and included realistic modeling of detectors and the corresponding noise characteristics. Stopping power maps were calculated for all three scans, and compared with the ground truth stopping power from the phantoms.

Results. Proton CT gave slightly better stopping power estimates than the dual energy CT method, with root mean square errors of 0.2% and 0.5% (for each phantom) compared to 0.5% and 0.9%. Single energy CT root mean square errors were 2.7% and 1.6%. Maximal errors for proton, dual energy and single energy CT were 0.51%, 1.7% and 7.4%, respectively.

Conclusion. Better stopping power estimates could significantly reduce the range errors in proton therapy, but requires a large improvement in current methods which may be achievable with proton CT.

Accurate stopping power determination is required to calculate the ranges of protons and is critical for precise dose calculations and planning in proton therapy. The primary advantage of proton therapy is the finite proton range, which leads to significant dose sparing of the normal tissue in the patient [1,2]. The precision given by this finite range also makes proton therapy less robust to differences between the treatment planning models and reality [3–6]. Errors in the estimated stopping power lead to errors in the predicted proton range, which again may lead to failure of the treatment. The clinical approach to handle such errors is to increase the dose margins used, with a larger dose to normal tissue as the result, potentially offsetting the achievable precision of proton therapy. State-of-the-art for estimating stopping power in vivo is the stoichiometric method based on single-energy x-ray computed tomography (CT) scans [7]. This method is, however, limited by the fact that no direct relation between x-ray attenuation

Correspondence: D. C. Hansen, Nørrebrogade 44, building 5, 8000 Aarhus, Denmark. E-mail: daviha@rm.dk

(Received 15 May 2015; accepted 4 June 2015)

ISSN 0284-186X print/ISSN 1651-226X online © 2015 Informa Healthcare
DOI: 10.3109/0284186X.2015.1061212

coefficient and stopping power exists. The result of this is systematic uncertainties of up to 3.4% in biological materials [8] and potentially much higher in other materials, such as plastic and metal inserts. Several methods have been proposed to remedy this, such as measuring PET activity during treatment [1] or monitoring dose induced tissue changes via magnetic resonance imaging (MRI) [9], but these methods are still far from being clinically adaptable.

Proton CT has been widely tested for the purpose of estimating stopping power directly [10–13] and a prototype head scanner is currently being tested at the Loma Linda University Hospital as well as a smaller prototype scanner in INFN-LNS, Catania, Italy. These are, however, still limited by the fact that current scanning techniques are quite slow and take up valuable treatment beam time. In addition the maximum range of the protons could be a limiting factor for larger patients [13].

Another attractive option is dual energy x-ray CT (DECT), which is now clinically available from several vendors. By obtaining photon attenuation coefficients from two different x-ray spectra in each voxel, electron density and atomic number can be calculated. Yang et al. [14] demonstrated that this could be used for stopping power calculations and recently a few groups compared the method with experiments and found an agreement better than 1.0% [15–17]. In these studies a single phantom was investigated and the majority of materials investigated were from the same phantom vendor.

In this work we investigated the stopping power estimation accuracy of proton CT for two large phantoms (~20–30 cm) from two vendors using Monte Carlo simulations accounting for position and energy resolution of current proton CT prototypes. To establish whether the potential accuracy gains from proton CT are relevant, the phantoms were scanned at a clinically available dual energy CT scanner and stopping power was estimated using dual and single energy methods. The experimental x-ray CT data was used to establish the single and dual energy CT accuracy levels currently achievable. To our knowledge this is the first simulation study evaluating the potential gains from proton CT using x-ray CT of the same phantoms as reference and the first study investigating dual energy CT on more than one phantom.

Material and methods

Phantom measurements

A dual source CT scanner (Siemens SOMATOM Definition FLASH, Siemens Medical, Forchheim, Germany) was used for all x-ray scans. For dual energy scans, the low spectrum was 80 kVp and the high spectrum was 140 kVp filtered with a Sn filter. For single energy scans, a standard 120 kVp spectrum was used. A RMI 467 electron density phantom (Gammex, Middleton, WI, USA), diameter 32 cm, was scanned with both dual and single energy. The phantom consists of 16 cylindrical inserts of different tissue-like materials. A second phantom, the Model 002H5 IMRT Phantom (CIRS, Inc., Norfolk, VA, USA), with elliptical cross section (30 cm wide and 20 cm thick), containing five different tissue-mimicking inserts was also scanned and used for validation. Insert composition and density for both phantoms can be found in tables 1 and 2 of Landry et al. [13]. The images were imported into MATLAB (MathWorks, Inc., Natick, MA, USA) and the average HU values were extracted for each insert using a region of interest covering 50% of the insert, in a central slice of the phantom. All x-ray CT scans were performed with a computed tomography dose index (CTDI) of 30 mGy (for DECT this value was for the sum of the contributions of the two scans).

X-ray CT calibration

For DECT, electron density was calculated using the method of Saito [18], and the effective atomic number using the method of Landry et al. [19]. These values were then used to calculate stopping power, following the method of Yang [14]. The necessary DECT calibrations were carried out using the Gammex phantom data.

For SECT, stopping power was calculated using two methods. The first approach calibrated CT numbers directly from the Gammex phantom data while the second was based on the stoichiometric method [7,20] where use was made of reference human tissue data [21].

Proton CT simulation and reconstruction

The phantoms described above were simulated in the proton CT simulation and reconstruction framework described in [22] using the manufacturer specifications for composition and density. Each phantom was scanned using a uniform field of 250 MeV protons with 0.5 MeV energy spread at a CT equivalent dose index (CTEDI) of 10 mSv [22]. We simulated a detector spatial resolution of 200 μm and energy resolution equivalent to max(0.03E, $24.15\text{MeV}^2/\text{E} + 1.76\text{MeV}$) where E is the incident proton energy [23]. Details of the simulations can be found in the Monte Carlo section of the Supplementary Appendix and schematics of the simulation setup are shown on Supplementary Figure 2 (available online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1061212).

Accuracy evaluation

In all cases, the ground truth stopping power was calculated at 100MeV based on either the stoichiometric information provided by the phantom manufacturers. The mean excitation potentials, required for stopping power calculation, were based on the recommendations in [24]. We estimated the uncertainty on the ground truth calculations using data from [17] where stopping power for inserts estimated using a water column was reported along with the manufacturer's reported composition.

A more complete description of the methods listed above is presented in the Materials and Methods part of the Appendix and calibration curves for DECT and SECT are shown on Supplementary Figure 1 (available online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1061212).

Results

Figure 1 presents the comparison of proton CT stopping power accuracy for the Gammex phantom compared to SECT and DECT results. For this phantom the average root mean square (RMS) errors for proton CT was 0.2%. In comparison the stoichiometric SECT calibration yielded a RMS error of 2.7%, and the phantom calibration 1.6%. DECT RMS error was 0.5%. The maximum absolute error for proton CT was 0.4%, compared to errors of up to 7.4% for SECT and 1.0% for DECT. For DECT, the results were consistently within 2σ of ground truth. This was also true for proton CT except in the case of cortical bone, which showed a minor underestimation $0.4 \pm 0.2\%$.

Figure 2 presents the results for the second phantom, which was not used in x-ray CT calibration. The RMS error for proton CT was 0.3% whereas for stoichiometric SECT calibration, phantom SECT calibration and DECT they were 1.6%, 3.6% and 0.9%, respectively. The maximum error for proton CT was 0.5% compared with 11% for SECT and 1.7% for DECT.

For proton CT, all results were within 2σ of the ground truth, but both DECT and SECT showed statistically significant deviations. The uncertainty from the theoretical calculation of stopping power was found to be $\sigma_{\text{theory}} = 0.49\%$.

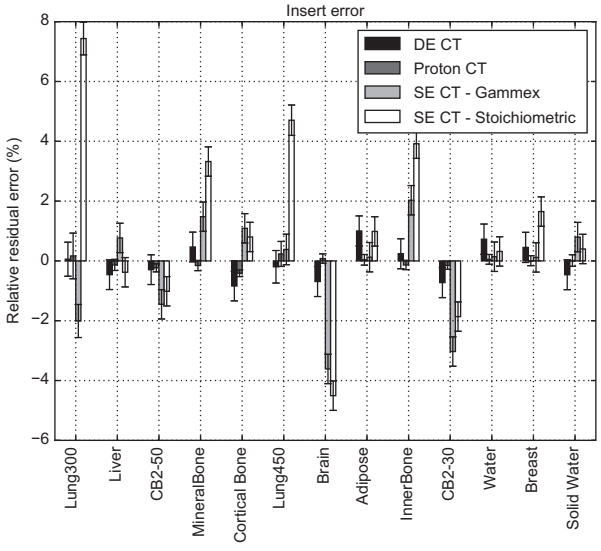

Figure 1. Residual errors in stopping power estimates based on DECT calibration, phantom SECT calibration (SECT – Gammex), stoichiometric SECT calibration (SECT – stoichiometric) and proton CT in each of the inserts of the Gammex phantom. Note that the DECT and the SECT data were fitted to the inserts in this phantom. Error bars show the 1σ confidence interval, as calculated from variance of the pixels in each insert. Additionally, for SECT and DECT, the error bars also include the uncertainties on ground truth stopping power.

Discussion

The results presented in Figures 1 and 2 are in agreement with published literature. The traditionally quoted value of 3.5% uncertainty on stopping power for SECT [8] corresponds to the RMS errors reported here, which ranged from 1.6% to 3.7%, depending on the calibration procedure used. For DECT, RMS error for the Gammex phantom was 0.5% with a maximum error of 1%. These results show excellent agreement with the accuracy levels reported in other studies [15,17] and are of the order of the uncertainty estimated in this work for

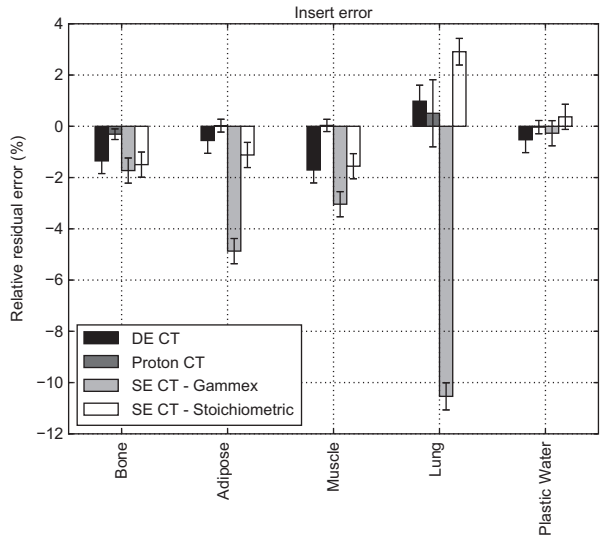

Figure 2. Residual errors in stopping power estimates based on DECT calibration, phantom SECT calibration (SECT – Gammex), stoichiometric SECT calibration (SECT – stoichiometric) and proton CT in each of the inserts of CIRS phantom. Error bars show the 1σ confidence interval, as calculated from variance of the pixels in each insert. Additionally, for SECT and DECT, the error bars also include the uncertainties on the ground truth.

the ground truth. The DECT performance, when evaluated outside of calibration conditions by scanning different materials in a smaller phantom, was found to be slightly degraded with an RMS error of 0.9% with maximum errors of up to 1.7%. The effect of using different phantoms had not been reported previously. Our results suggest that DECT accuracy remains below 2% across phantoms, but it may be necessary to perform a systematic investigation of the range of patient sizes where a given DECT calibration holds.

The main objective of this work was to assess whether proton CT offers any advantage over clinically available stopping power estimation methods. The stopping power accuracies reported here for the simulated proton CT showed maximal errors of 0.5%. Comparing with literature, Zygmanski et al. [25] achieved an agreement better than 0.5% when compared with experimental stopping power measurements, although the setup for proton CT was somewhat different than the one used in this study. In a later study, Hurley et al. [26] also achieved 0.5% accuracy in an experimental proton CT of a water phantom.

Minimum stopping power accuracy requirements for proton therapy treatment planning are not easily distilled from the literature and are essential in identifying whether the potential accuracy gains from proton CT are relevant. If we assume a requirement of a 1% uncertainty on range then the DECT results presented here fail to provide this level of accuracy for all materials, while proton CT shows promise. If we aim for 2% uncertainty then the use of DECT would be sufficient and the benefits of the development of proton CT may be questionable. It may be necessary to perform proton dose calculation on clinically relevant geometries imaged with the methods discussed in this work, to establish whether significant range errors exist.

Proton CT does not suffer from the artifacts present in x-ray CT, such as beam hardening, making it more robust to variations in patient size. Proton CT also has the advantage that it would image the patient directly in the treatment position, which could reduce setup errors. However, development of in situ dual energy cone beam CT scanners may provide similar capability.

The two largest advantages of x-ray CT over current prototype proton CT are scanning speed and spatial resolution. Modern x-ray CT scanners allow for time-resolved scans, which in the context of radiotherapy is particularly important for moving targets, e.g. in the lung. Older studies in proton CT report quite poor resolution [6], worse than 3 mm, which would be prohibitive in radiotherapy treatment planning. However, recent work has shown that proton CT in theory can achieve similar spatial resolution to x-ray CT by using path modeling [27], but this has yet to be tested experimentally.

In conclusion, in this study we have confirmed previously reported stopping power accuracy gains from DECT when using a single phantom for calibration and evaluation and identified that when deviating from calibration conditions by using a phantom of different size and composition accuracy can be slightly degraded. Simulation of proton CT of the same phantoms showed that in theory proton CT may offer superior accuracy than DECT. Whether proton CT is necessary is highly dependent on the clinical requirements for stopping power estimation.

Acknowledgments

This work was sponsored by the Lundbeck Centre for Interventional Research in Radiation Oncology (CIRRO). Bjarne Thomsen, Department of Physics and Astronomy, Aarhus University is thanked for providing cluster computing time.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Bauer J, Unholtz D, Sommerer F, Kurz C, Haberer T, Herfarth K, et al. Implementation and initial clinical experience of offline PET/CT-based verification of scanned carbon ion treatment. Radiother Oncol 2013;107:218–26.

[2] Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach. Radiother Oncol 2013;107:267–73.

[3] Casares-Magaz O, Toftegaard J, Muren LP, Kallehauge JF, Bassler N, Poulsen PR, et al. A method for selection of beam angles robust to intra-fractional motion in proton therapy of lung cancer. Acta Oncol 2014;53:1058–63.

[4] Hopfgartner J, Stock M, Knausl B, Georg D. Robustness of IMPT treatment plans with respect to inter-fractional set-up uncertainties: Impact of various beam arrangements for cranial targets. Acta Oncol 2013;52:570–9.

[5] Góra J, Kuess P, Stock M, Andrzejewski P, Knäusl B, Paskeviciute B, et al. ART for head and neck patients: On the difference between VMAT and IMPT. Acta Oncol Epub 2015 Apr 8:1–9.

[6] Grau C. The model-based approach to clinical studies in particle radiotherapy – A new concept in evidence based radiation oncology? Radiother Oncol 2013;107:265–6.

[7] Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. Phys Med Biol 1996;41:111–24.

[8] Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. Phys Med Biol 2012;57:R99–117.

[9] Yuan Y, Andronesi OC, Bortfeld TR, Richter C, Wolf R, Guimaraes AR, et al. Feasibility study of in vivo MRI based

dosimetric verification of proton end-of-range for liver cancer patients. Radiother Oncol 2013;106:378–82.
[10] Cormack AM, Koehler AM. Quantitative proton tomography: Preliminary experiments. Phys Med Biol 1976;21:560–9.
[11] Hanson KM, Bradbury JN, Cannon TM, Hutson RL, Laubacher DB, Macek RJ, et al. Computed tomography using proton energy loss. Phys Med Biol 1981;26:965–83.
[12] Schulte RW, Bashkirov V, Klock MC, Li T, Wroe AJ, Evseev I, et al. Density resolution of proton computed tomography. Med Phys 2005;32:1035–46.
[13] Hansen DC, Petersen JB, Bassler N, Sorensen TS. Improved proton computed tomography by dual modality image reconstruction. Med Phys 2014;41:031904.
[14] Yang M, Virshup G, Clayton J, Zhu XR, Mohan R, Dong L. Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. Phys Med Biol 2010;55:1343–62.
[15] Bourque AE, Carrier JF, Bouchard H. A stoichiometric calibration method for dual energy computed tomography. Phys Med Biol 2014;59:2059–88.
[16] Farace P. Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. Phys Med Biol 2014;59:7081–4.
[17] Hunemohr N, Krauss B, Tremmel C, Ackermann B, Jakel O, Greilich S. Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. Phys Med Biol 2014;59:83–96.
[18] Saito M. Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship. Med Phys 2012;39:2021–30.
[19] Landry G, Seco J, Gaudreault M, Verhaegen F. Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients. Phys Med Biol 2013;58: 6851–66.
[20] Schneider W, Bortfeld T, Schlegel W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. Phys Med Biol 2000;45:459–78.
[21] Woodard HQ, White DR. The composition of body tissues. Br J Radiol 1986;59:1209–18.
[22] Hansen DC, Bassler N, Sorensen TS, Seco J. The image quality of ion computed tomography at clinical imaging dose levels. Med Phys 2014;41:111908.
[23] Sipala V, Randazzo N, Aiello S, Leonora E, Lo Presti D, Russo M, et al. YAG(Ce) crystal characterization with proton beams. Nucl Instrum Meth A 2011;654:349–53.
[24] International Commission on Radiation Units & Measurements. Stopping power and ranges for protons and alpha particles. Bethesda, MD: ICRU; 1993.
[25] Zygmanski P, Gall KP, Rabin MS, Rosenthal SJ. The measurement of proton stopping power using proton-cone-beam computed tomography. Phys Med Biol 2000; 45:511–28.
[26] Hurley RF, Schulte RW, Bashkirov VA, Wroe AJ, Ghebremedhin A, Sadrozinski HF, et al. Water-equivalent path length calibration of a prototype proton CT scanner. Med Phys 2012;39:2438–46.
[27] Rit S, Dedes G, Freud N, Sarrut D, Letang JM. Filtered backprojection proton CT reconstruction along most likely paths. Med Phys 2013;40:031103.

Supplementary material available online

Supplementary Appendix and Supplementary Figure 1–2 available online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1061212.