Physica Medica 120 (2024) 103341

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

First clinical experience following the consensus guide for calibrating a proton stopping power ratio curve in a new proton centre

Calvin Wei Yang Koh [a,1], Kah Seng Lew [a,c], Andrew Wibawa [a], Zubin Master [a], Ping Lin Yeap [a,d], Clifford Ghee Ann Chua [a], James Cheow Lei Lee [a,c], Hong Qi Tan [a,b,*], Sung Yong Park [a,b]

[a] Division of Radiation Oncology, National Cancer Centre Singapore, Singapore
[b] Oncology Academic Clinical Programme, Duke-NUS Medical School, Singapore
[c] Nanyang Technological University Singapore, Singapore
[d] Department of Oncology, University of Cambridge, United Kingdom

ARTICLE INFO

Keywords:
HU Look-Up Table
SPR
CT Calibration
Range Uncertainty

ABSTRACT

Background and Purpose: This work introduces the first assessment of CT calibration following the ESTRO's consensus guidelines and validating the HLUT through the irradiation of biological material.
Methods: Two electron density phantoms were scanned with two CT scanners using two CT scan energies. The stopping power ratio (SPR) and mass density (MD) HLUTs for different CT scan energies were derived using Schneider's and ESTRO's methods. The comparison metric in this work is based on the Water-Equivalent Thickness (WET) difference between the treatment planning system and biological irradiation measurement. The SPR HLUTs were compared between the two calibration methods. To assess the accuracy of using MD HLUT for dose calculation in the treatment planning system, MD vs SPR HLUT was compared. Lastly, the feasibility of using a single SPR HLUT to replace two different energy CT scans was explored.
Results: The results show a WET difference of less than 3.5% except for the result in the *Bone* region between Schneider's and ESTRO's methods. Comparing MD and SPR HLUT, the results from MD HLUT show less than a 3.5% difference except for the *Bone* region. However, the SPR HLUT shows a lower mean absolute percentage difference as compared to MD HLUT between the measured and calculated WET difference. Lastly, it is possible to use a single SPR HLUT for two different CT scan energies since both WET differences are within 3.5%.
Conclusion: This is the first report on calibrating an HLUT following the ESTRO's guidelines. While our result shows incremental improvement in range uncertainty using the ESTRO's guideline, the prescriptional approach of the guideline does promote harmonization of CT calibration protocols between different centres.

1. Introduction

Proton beam therapy occupies an important role in curative radiotherapy treatment due to its physical characteristics of dose distribution. However, there are still numerous issues inhibiting the full exploitation of the therapeutic potential of proton beam therapy [1–3]. One of the successes of proton beam therapy depends on the accuracy in determining the range of protons in tissues, making range uncertainty a vital concern [4–6]. Currently, a 3.5% margin is included in robust optimization in many proton centres to take into account the dosimetric impact of range uncertainty on target coverage [3,7]. A 3.5 % margin has been extensively used since it was proposed initially by Goitein et al [8] in 1985 and this is mainly due to the approach in calibrating the Hounsfield Look-Up Table (HLUT) in the treatment planning system (TPS) [3,9] using single energy computed tomography (CT). HLUT is the existing method to map CT Hounsfield unit (HU) to proton stopping-power ratio (SPR) [9].

To overcome this limitation, several approaches have been proposed such as the use of Dual-Energy CT (DECT) [10,11], photon counting CT [12–14] and proton CT [15,16] to reduce the range uncertainty. However, the challenges with the implementation of the new technological advancements are either due to limited clinical experience or the maturity of the systems. Currently, the calibration of HLUT is still limited to the use of Single-Energy CT (SECT) as demonstrated by the

* Corresponding author at: Division of Radiation Oncology, National Cancer Centre Singapore, Singapore.
E-mail address: hqtan1988@gmail.com (H.Q. Tan).
[1] First authors.

https://doi.org/10.1016/j.ejmp.2024.103341
Received 22 August 2023; Received in revised form 7 March 2024; Accepted 25 March 2024
Available online 29 March 2024

Table 1
CT scan and reconstruction parameters used for phantom scanning.

CT Scanner	GE Healthcare Revolution CT	Siemens Healthineers SOMATOM X.cite
Tube Voltage	80 kVp, 100 kVp, 120 kVp	
Scan Mode	Single-energy CT	
Field of View	500 mm	
Rotation Time	0.5 s	1 s
Pitch	0.984	0.8
Detector Collimation	64 x 0.625 mm	128 x 0.6 mm
Slice Thickness	2 mm	
Slice Increment	2 mm	
Reconstruction Kernel	Stnd, ASiR-V 50 %	Qr40, Admire 3, IBHC *Bone*
$CTDI_{vol}$ (mGy)	3 – 30	3 – 17
Automatic Exposure Control	SmartmA	CARE Dose4D

method proposed by Schneider et al [17]. The HLUT is generated based on a stoichiometric calculation of human tissues provided by ICRU Report 44 [18]. Despite the proposed method, the exact measurement procedures might vary between different centres such as the phantom setup used for calibration, the fitting procedures and the consideration of beam hardening in the process [19]. In the first step of Schneider's method, fitting parameters were calculated from the measured HU and known electronic density of plugs in electron density phantoms to map effective atomic numbers and electron density to HU. The second step consists of calculating the expected HUs of all the human tissues with known SPR in ICRU 44 using the mapping function defined in the first step. Lastly, piecewise linear regressions between SPR and HUs were performed using all the tissue data calculated in the second step. Three distinct non-overlapping regions – *Lung*, *Tissue* and *Bone* were proposed in Schneider's method. Beam hardening effect was not taken into consideration during the first step of the process and could lead to errors propagating towards the final HLUT. Furthermore, there are no standardized protocols for obtaining the HLUT and this might cause a disparity when data are compared across different centres. The work by Peters et al. [20] underlines the need for standardization across centres where it shows a large variation in SPR and proton range prediction across different centres of up to 8.7%, 6.3% and 1.5% relative to water in the *Bone*, *Lung* and *Tissue* regions respectively.

Recognizing the need for standardization, the European proton therapy community at the European Society for Therapeutic Radiology and Oncology (ESTRO) reached a consensus to improve the calibration of the HLUT [19]. They established a step-by-step guideline to aid proton centres to establish their HLUT. The step-by-step guide includes the phantom setup, the CT acquisition and reconstruction settings, the CT number extraction, the determination of SPR, the HLUT specification

Table 2
Insert replacement for the extra HE CT Solid Water.

Insert Position	Current Work
5	Iodine 0.2 mg/ml
6	HE Blood 70 ρ_e^w 1.07
8	HE Blood 100 ρ_e^w 1.10
10	HE Blood 40 ρ_e^w 1.03
12	Liquid Water

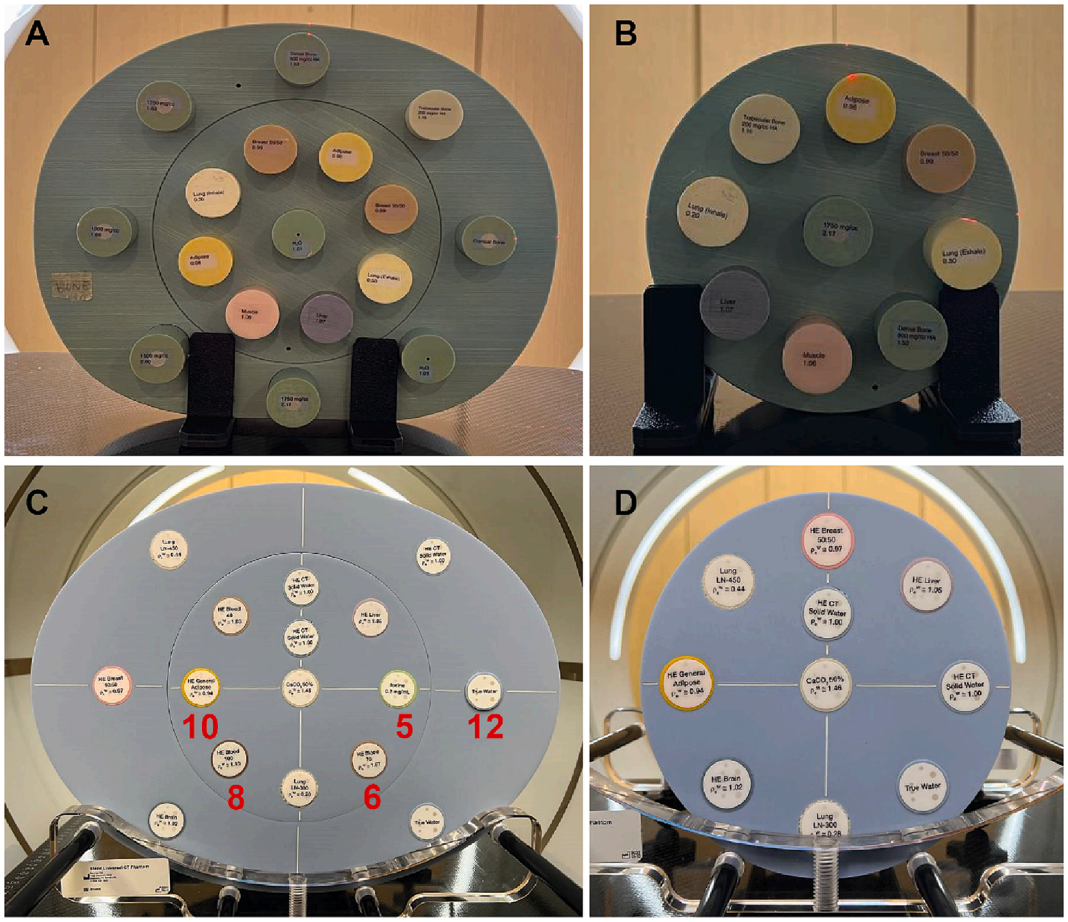

Fig. 1. The Body (A) and Head/Paediatrics (B) phantom arrangement for the scan using CIRS 062 M Phantom. The Body (C) and Head/Paediatrics (D) phantom arrangement for the scan using MECT phantom.

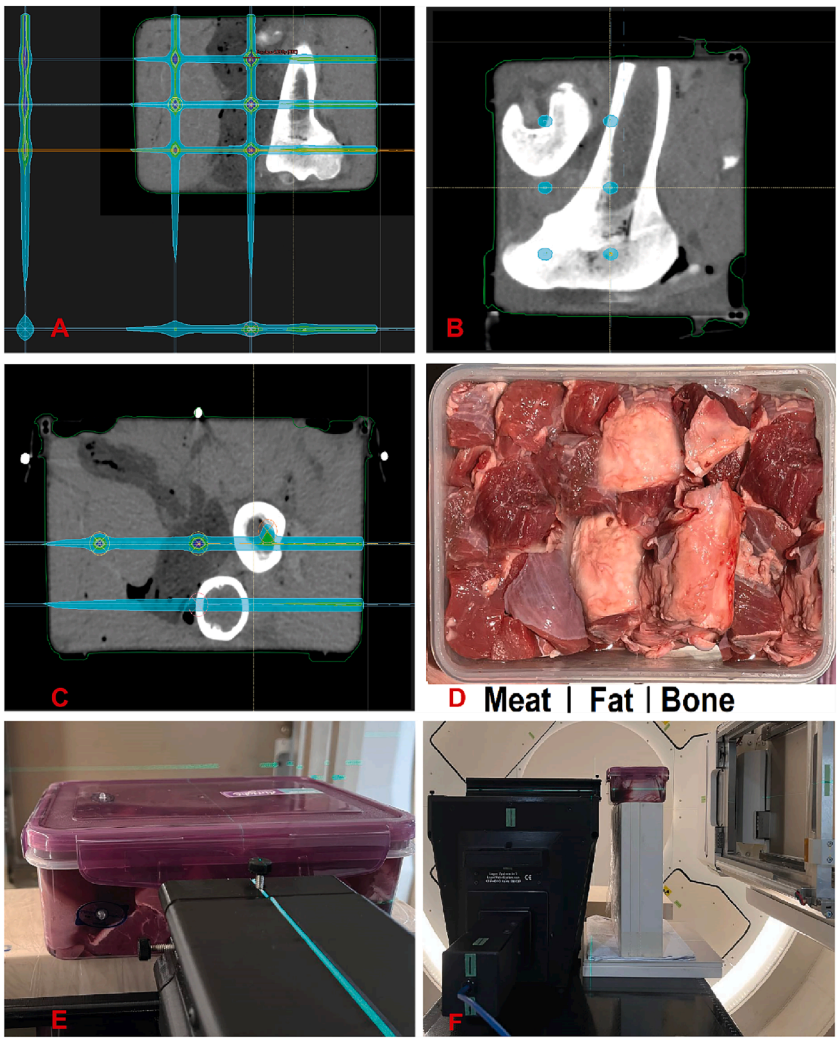

Fig. 2. A, B and C show the irradiation of the proton beam through both individual tissues (shorter beam path) and a combination of all Tissues (longer beam path) in the coronal, axial and sagittal planes. D shows the biological tissues compartmentalize into meat, fat and Bone. E and F show the WET measurement of the biological tissue (Lamb) setup.

method and how to validate the HLUT process. One of the main differences between Schneider's method and ESTRO's consensus is in the categorizing of data into four regions – *Bone*, *Soft Tissue*, *Adipose Tissue* and *Lung*. In addition, there is a recommendation on how to best set up the phantom for scanning to reduce the influence of beam hardening on the HLUT. This standardization of protocol enables a better comparison of multi-institutional data and enhances the credibility of clinical trial comparisons by ensuring consistency, reproducibility of findings in different settings, quality control in minimizing potential sources of variability and eventually facilitating meaningful comparisons between studies [21–24].

This work presents the first report on the CT calibration for two CT scanners following the ESTRO's consensus guideline and validated through the irradiation of biological material. This study evaluates the HLUT obtained using Schneider's method (old) and ESTRO's consensus guideline (new) by comparing the Water-Equivalent Thickness (WET) difference between TPS prediction and biological irradiation measurement. Since RayStation (RaySearch Laboratories AB, Stockholm, Sweden) permits the use of both mass density and SPR for HLUT, this study also investigates the impact of the two different types of HLUT on range validation in biological tissues. Lastly, this work addresses the possibility of using a single 100kVp HLUT for dose calculation to replace the current centre's practice of having two HLUTs for 80kVp and 120kVp.

2. Methods

2.1. Overview of HLUT

This study utilized two CT scanners, a GE Revolution CT (GE Medical Systems, Milwaukee, WI, USA) and Siemens SOMATOM X.cite (Siemens Healthineers, Forchheim, Germany) for the scanning of the two Sun Nuclear electron density phantoms - Electron Density Phantom Model 062 M and Multi-Energy CT (MECT) phantom (Sun Nuclear – A Mirion Medical Company, Middleton, WI, USA). Both CT scanners were commissioned for proton therapy treatment and the results shown in this work will allow the community to understand the difference between CT scanners. The main difference between the two electron density phantoms lies in the size of the phantom where the MECT phantom (40 x 30 x 16.5 cm) is larger than the 062 M phantom (33 x 27 x 5 cm). There are also more types of inserts (e.g. blood and iodine) available for MECT phantom and the electron density of the similar inserts (e.g. *Lung*) differ. The scan and reconstruction parameters are shown in Table 1. For the GE CT scanner, the reconstruction kernel is a hybrid-statistical iterative process while the Siemens is a model based iterative process.

Two different procedures namely, Schneider's and ESTRO's consensus, were used to obtain the HLUT. Following the publication of ESTRO's consensus, a new HLUT curve was formulated. Raystation 10B offers two different types of HLUT which map the HU to either Mass

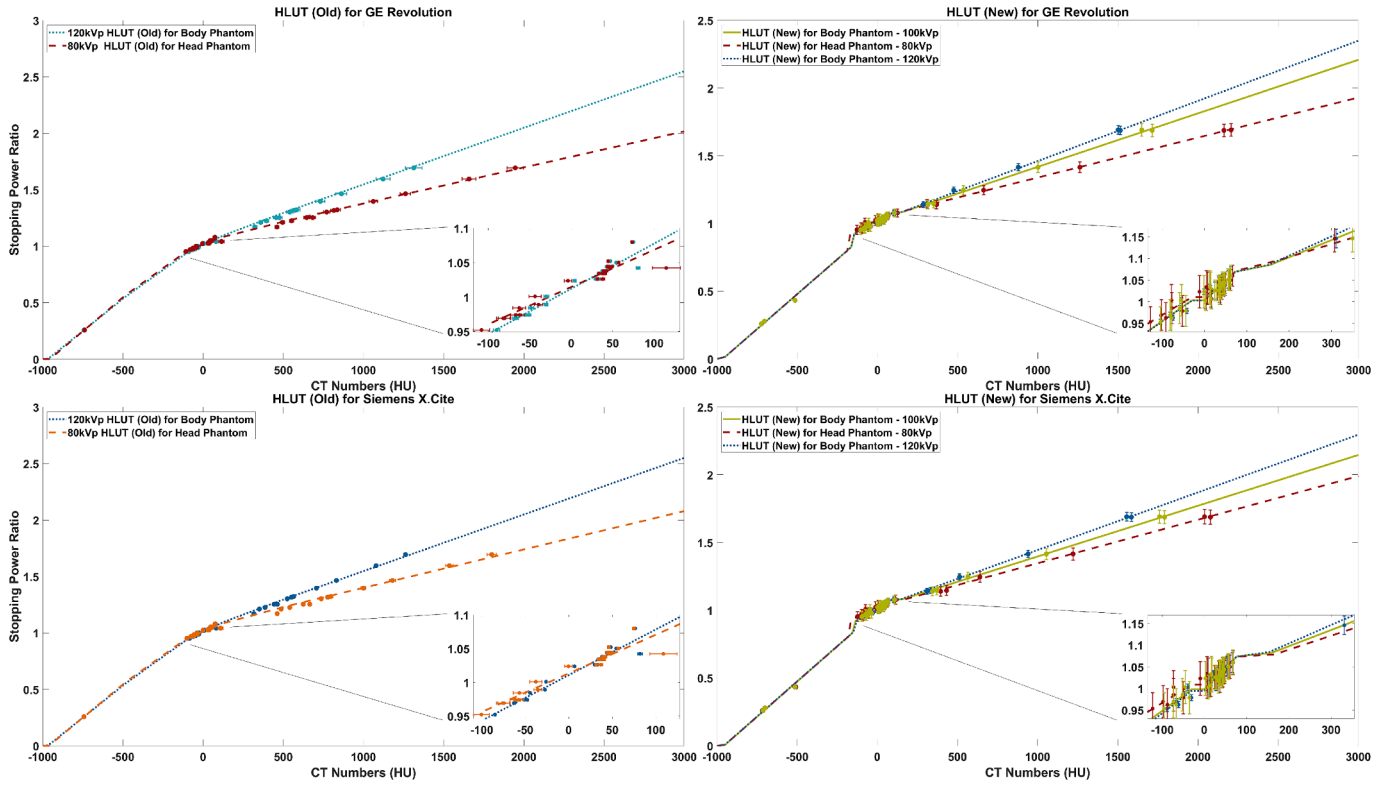

Fig. 3. HLUT was generated using Schneider's (A and B) and ESTRO's (C and D) methods for both GE Revolution and Siemens X.Cite CT scanner. Each plot shows the different CT energy HLUT – 80kVp, 120kVp and 100kVp (C and D) with their respective data points in solid circles. The solid line shows the linear fit of the data points and the zoom plot shows the fitting in the Soft Tissue region. Schneider's method shows the fitting of Soft Tissue in a single category while the ESTRO's method separated the tissues into Adipose and Soft Tissue for the fitting.

Table 3
R^2 values for all CT scans energy of each region of the HLUT for GE Revolution and Siemens X.Cite using the ESTRO's consensus method.

R^2	GE Revolution			Siemens X.Cite		
Regions	80 kVp	100 kVp	120 kVp	80 kVp	100 kVp	120 kVp
Lung	0.999	0.999	0.999	0.999	0.999	0.999
Adipose Tissue	0.875	0.925	0.914	0.883	0.920	0.926
Soft Tissue	0.999	0.999	0.999	0.999	0.999	0.999
Bone	0.999	0.997	0.998	0.997	0.998	0.998

Density (MD) or SPR. Both HLUT curves will be used in this study for comparison using Raystation. WET measurements in biological materials were performed to compare the two types of HLUT (MD vs SPR) and the two methods of HLUT calibration (Schneider's vs ESTRO consensus). The details of how the HLUT (MD and SPR) were obtained and the biological WET measurements are discussed below.

2.2. HLUT using Schneider's method

Ainsley and Yeager [17,25] published a recipe using the stoichiometric method to obtain the HLUT. Firstly, known tissue equivalent samples were CT scanned to obtain the HU. These tissue substitutes and the HU values were then parameterized (A^{ph} and B^{coh}) as shown in Eq. (1):

$$HU_i + 1000\left(1-\rho_{e_i}^{rel}\right) = A^{ph}\left(\rho_{e_i}^{rel}\left[\widetilde{Z}_i^{3.62} - \widetilde{Z}_w^{3.62}\right]\right) + B^{coh}\left(\rho_{e_i}^{rel}\left[\widehat{Z}_i^{1.86} - \widehat{Z}_w^{1.86}\right]\right) \quad (1)$$

where $\rho_{e_i}^{rel}$ is the relative electron density of element i and A^{ph} and B^{coh} are the factors contributed by photoelectric absorption, Rayleigh scattering and Compton scattering. $\widetilde{Z}_i$ is the effective Z of element i given in Eq. (2) and the subscript w refers to water:

$$\widetilde{Z}_i^m = \sum \lambda_i Z_i^m \quad (2)$$

$$\lambda_i = \frac{\frac{w_j Z_j}{A_j}}{\sum_j \frac{w_j Z_j}{A_j}} \quad (3)$$

where Z_j is the atomic number of element j, A_j is the mass number of element i and w_j is the weightage of each element i.

Thereafter, the HU of the human tissues (ICRU Report 44 and 23) were calculated based on the parametrization values and plotted against SPR. A linear fit on the data points was done for three regions – *Lung*, *Tissue* and *Bone*, thus generating the HLUT.

In this study, CIRS 062 M Phantom was used as the known tissue equivalent sample. A *Large* and *Small* arrangement was used to simulate a *Body* and a *Head/Paediatrics* scenario as shown in Fig. 1A and B.

The HU value of each insert was obtained by contouring a cylindrical region of interest (ROI) of $\phi 0.4 \times 3cm$ for the *Bone* insert and $\phi 2 \times 3cm$ for the other inserts. Using Eq. (1) [17] and applying multiple linear regressions for each category, the A and B coefficients can be determined.

Since the elemental composition of real human tissues and the respective SPR are provided in ICRU Report 49, the HU value for each human tissue can be calculated using the newly defined A and B coefficients. In order to encompasses as much as possible of the human anatomy, a total of 47 data points were used from the ICRU Report 49. Among the 47 human tissue data, 2 of which are *Lung*, 29 are *Tissue* and 16 are *Bone* data points. The HLUT can then be created by relating to the

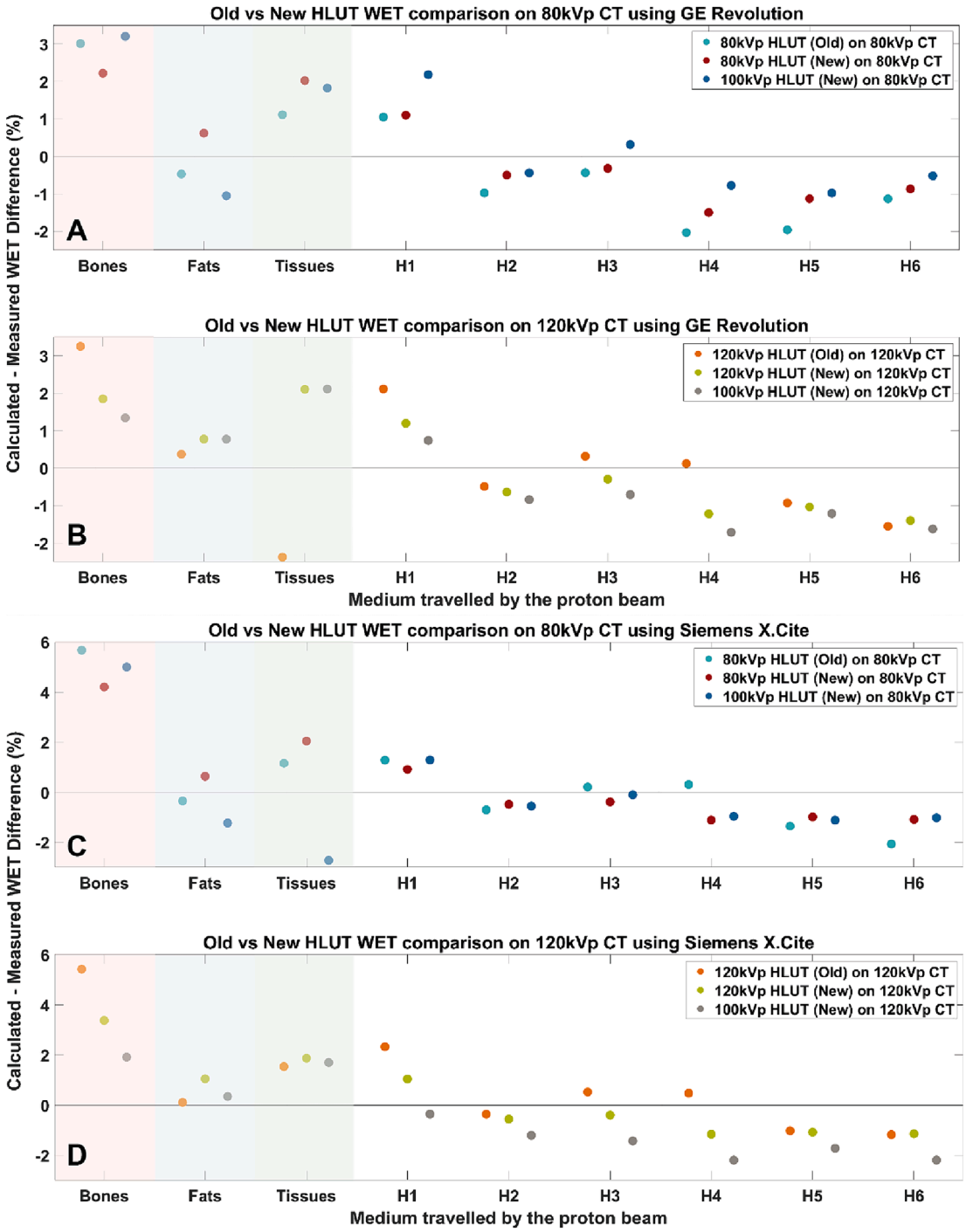

Fig. 4. For both 80kVp and 120kVp CT scans in the GE Revolution (A and B) and Siemens X.Cite (C and D) CT scanners, the percentage WET difference between the old and new 80kVp SPR HLUT was compared. The comparison was also done for the old and new 120kVp SPR HLUT. 100kVp SPR HLUT data points were included to determine the feasibility of having a single HLUT for two different CT scan energies.

respective SPR for 100 MeV proton beam energy. Using this method, eight HLUT curves are generated where each CT scanner (GE and Siemens) comprises two SPR and two MD curves of 80 kVp and 120 kVp. The HLUT curves were generated using the 120 kVp and 80 kVp energy in our CT scanner as the 120 kVp scan was used for adults (including the head and neck, thorax, abdomen and pelvis) while the 80 kVp scan was used for paediatrics and smaller anatomical scans such as extremities. This is part of our routine clinical practice in photon treatment to reduce imaging doses to the paediatrics which could potentially lead to secondary malignancy in adulthood [17,26].

2.3. *(New) HLUT following ESTRO's consensus*

Upon the publication of the ESTRO's consensus, a new HLUT was established for CT energies of 80kVp, 100kVp and 120kVp. MECT phantom, as shown in Fig. 1C and D, was used as it was purchased after establishing the HLUT using the CIRS phantom. This work follows the six steps outlined in the publication closely, starting from the phantom setup, CT acquisition and reconstruction settings, CT number extraction, determination of SPR, HLUT specification and evaluation of the HLUT specification.

Table 4
Mean Absolute Percentage Difference across individual and heterogeneous regions between the measured and the calculated WET differences for Old and New method SPR HLUT.

Mean Absolute Percentage Difference	GE 80kVp CT	Siemens 80kVp CT	GE 120kVp CT	Siemens 120kVp CT
80kVp HLUT (Old)	1.27	1.97	–	–
80kVp HLUT (New)	1.14	1.23	–	–
100kVp HLUT (New)	1.36	1.35	–	–
120 HLUT (Old)	–	–	1.25	1.32
120kVp HLUT (New)	–	–	1.20	1.36
100kVp HLUT (New)	–	–	1.17	1.55

Standard purchase of the phantoms does not come with sufficient solid water inserts as required by the publication. Therefore, we have substituted the solid water inserts with inserts shown in Table 2. These inserts are chosen for their similarity in density as compared to water to reduce any beam hardening effect.

A $1.5 \times 0.2 \times 10cm$ cuboid ROI is contoured in RayStation v10A (RaySearch Medical Laboratories AB, Stockholm, Sweden) to extract the mean HU of each phantom insert. Since the information on the SPR of the inserts is not provided by the commercial company, SPR measurements of the inserts were performed using a scintillation detector, Ranger-300 Proton Beam Range Verification (Logos Systems Int'l, Scotts Valley, California).

A 228.7 MeV proton beam from the Hitachi Probeat Proton Therapy System (Hitachi, Ltd., Japan) is irradiated on the phantom inserts and the proton range is determined from the integrated depth dose (IDD) measured by the scintillator. The spot size of the proton beam is $1.9mm$ and the dimensions of the inserts are $\phi 28.6 \times 165mm$ for MECT phantom and $\phi 30 \times 50mm$. The proton range is defined as R90 in this work. The SPR of each insert is then calculated by dividing the difference in the measured range with ($R_{90,insert}$) and without ($R_{90,air}$) the phantom inserts by the physical thickness of the phantom inserts as shown in Eq. (4):

$$SPR = \frac{R_{90,air} - R_{90,insert}}{Physical\ Thickness} \tag{4}$$

$R_{90,air}$, $R_{90,insert}$ (Ranger-300) and *Physical Thickness* (vernier caliper) were averaged over three repeated measurements. The CT number estimation of the human tissues is determined similarly using Eq. (1) and the SPR value of the human tissues in ICRU 49 is calculated using the Bethe equation [27], based on a nominal proton beam energy of 100 MeV. The I-values used for calculating the SPR value in the Bethe Equation are based on ICRU Report 49 [28].

Lastly, a piecewise linear regression was performed on data points comprising of the phantom inserts and the calculated human tissues to obtain the final HLUTs which map the HU to SPR or MD. A total of 41 data points were used for the regressions. Among the 40 data points, 3 of which are *Lung*, 22 are *Soft Tissues*, 9 are *Adipose Tissues* and 7 are *Bone* data points. The phantom inserts and human tissues are then categorised into four tissue groups instead of three (as stated in the ESTRO's method) – *Lung*, *Adipose Tissue*, *Soft Tissue* and *Bone*. *Lung* and *Soft Tissues* are fitted with a common linear regression as they are considered low-density *Soft Tissues* and a straight line is used to connect the junction between each tissue category. Using this method, a total of 12 HLUTs were entered into the TPS for the next part of the study. The 12 HLUTs comprise three SPR and three MD HLUTs (80kVp, 100kVp and 120 kV) for each CT machine.

2.4. Biological Verification measurement

To test the accuracy of the different HLUT, the measured range in biological tissues was compared to the predicted range using the TPS. The biological tissue used in this study originated from Lamb and the tissue was arranged in a plastic container in a compartmental manner with each compartment containing a specific type of tissue – Lamb's *bones*, *fats* and *lean meat* as shown in Fig. 2D. All empty spaces in the container were packed with meat chunks and water is added to simulate blood. The container has a dimension of $13 \times 19cm$ and the material has been determined to be roughly water equivalent (<3%) using the Ranger-300.

The biological tissues were then scanned with both CT scanners and three CT energies (80, 100 and 120 kVp). Markers were attached to the container to ensure setup reproducibility at the CT simulation and treatment. The container is placed at the isocenter of the CT scanner and the scan is performed. Using the marker as a setup indicator, the proton beam is irradiated at a specific position of the container with respect to the marker. The irradiation plan was generated as shown in Fig. 2 and the doses were calculated in the TPS using all the twelve different HLUT curves generated previously. The dose calculation grid sizes were 2 mm and the Monte Carlo dose calculation algorithm with a statistical uncertainty of 0.5 % was used. The WET measurement was performed using a 228.7 MeV proton beam which traversed through both the individual tissues and the composite of all tissues as shown in Fig. 2A, B and C. The proton range of each tissue is measured with Ranger-300 as shown in Fig. 2E and F. During the irradiation, the couch is shifted to specific coordinates according to the plan. The gantry is fixed at 90° and Ranger-300 is fixed at isocenter. The WET difference between the calculated (treatment plan) and measured range will be indicative of the accuracy of the HLUT curve.

3. Results

3.1. Old HLUT

Fig. 3A and B shows the 80kVp and 120kVp HLUT generated using Schneider's method for both CT scanners. The cutoff HU for fitting each category is −1000 to −200 for the *Lung* region, −200 to 200 for the *Tissue* region and 200 to the highest *Bone* HU value. The R^2 values for the fit in the *Lung*, *Tissue* and *Bone* region for 120kVp (80kVp) HLUT are 1.00 (1.00), 0.949 (0.900) and 0.993 (0.993) respectively for the GE revolution. Similarly for Siemens X.Cite, the respective R^2 values are 1.00 (1.00), 0.951 (0.911) and 0.992 (0.985).

3.2. New HLUT

Fig. 3C and D show the HLUT generated following the ESTRO's consensus method for 80kVp, 100kVp and 120kVp of each CT scanner. A total of 29 human *tissues* and 12 phantom inserts were used as data points to generate the HLUT. As mentioned in the ESTRO's consensus, the inclusion of phantom inserts into the HLUT will improve the calibration stability. The R^2 values of each region of the HLUT are shown in Table 3. The data fit using the new method shows similar results to the old method.

3.3. WET measurements comparison between old and new HLUT

Fig. 4 compares the WET difference between the old method and the new method for generating SPR HLUT for both 80kVp and 120kVp CT scans. 100kVp HLUT was also included for comparison. All data show a WET difference of less than 3.5% except for the result in the *Bone* region of Siemens X.Cite CT scanner. The percentage difference in the *Bone* region is relatively higher than the *Fats*, tissues and heterogeneous medium (denoted by H1 to H6 in the subsequent figures). However, looking at the mean absolute percentage difference (Table 4) between the old and the new method in obtaining the HLUT, there are no distinct trends and are all within the 3.5% difference between measured and calculated.

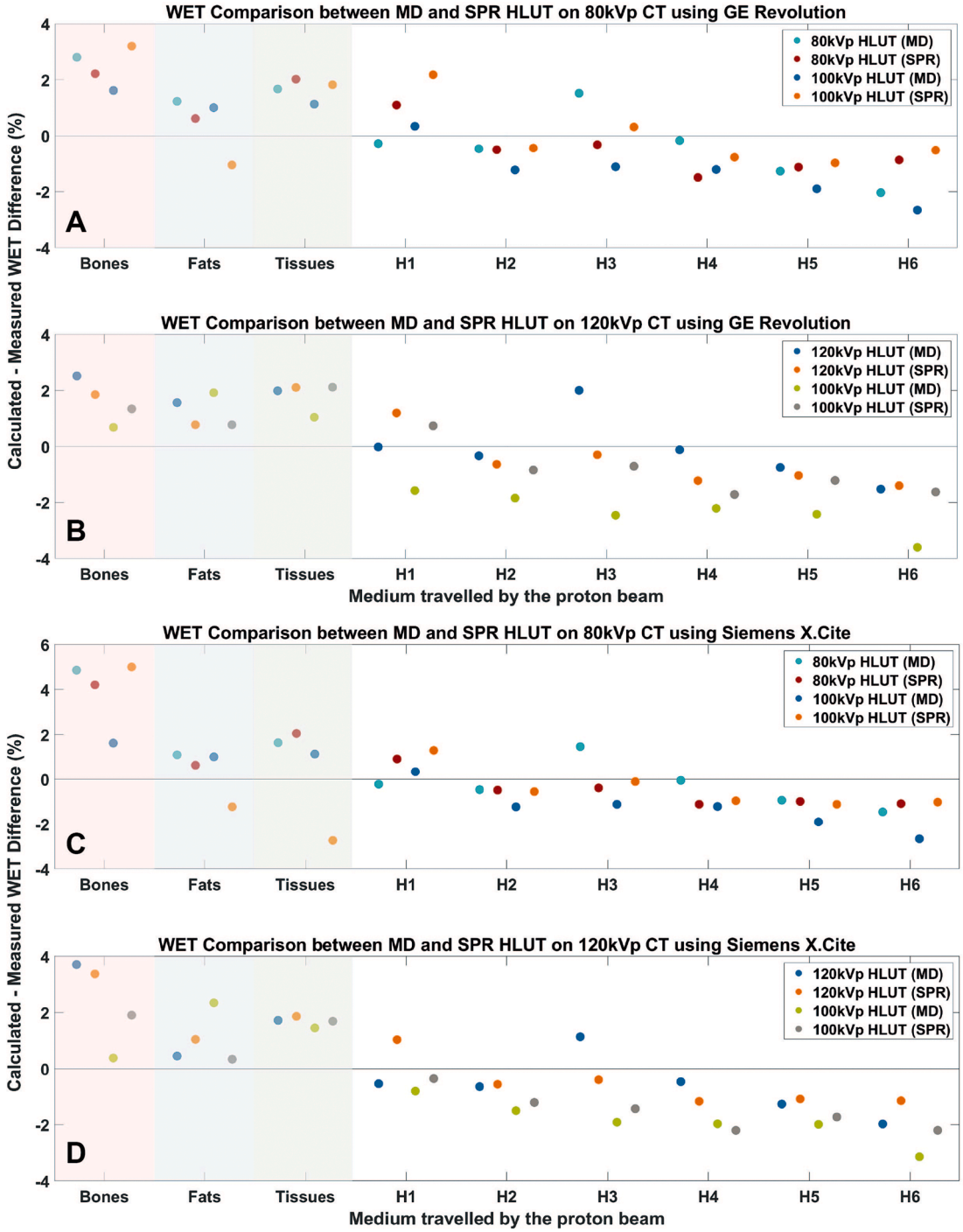

Fig. 5. For both 80kVp and 120kVp CT scans in the GE Revolution (A and B) and Siemens X.Cite (C and D) CT scanner, the percentage WET difference for MD curves and SPR HLUT between 80kVp vs 100kVp and 120kVp vs 100kVp were compared.

3.4. WET measurements comparison between MD curve and SPR HLUT

Subsequently, the WET measurements were compared between SPR and MD HLUT generated following the ESTRO's consensus guideline. On the 80kVp CT images, two MD HLUT were added in for comparison – 80kVp and 100kVp MD HLUT. Similarly, for 120kVp CT images, two MD HLUT were added in for comparison – 120kVp and 100kVp MD HLUT. This was done for both CT scanners and the results were as shown in Fig. 5. The uncertainties (Type B) in this comparison are propagated from the measured data. A total of three measurements were taken for each region and the uncertainty is calculated to be less than 0.1 % for all readings.

All data points from MD HLUT show similar results of less than a 3.5% difference between measured and calculated WET except the *Bone* region in Siemens X.Cite CT scanner. However, the mean absolute percentage difference (Table 5) between the measured and calculated WET difference shows a consistent result where SPR HLUT has a lower absolute mean percentage difference as compared to MD HLUT. Similar to Fig. 4, the WET difference in the individual *Bone* region is relatively larger than *Tissue* and the heterogenous medium which motivates further investigation to identify the possible cause.

We hypothesized the large percentage difference in the *Bone* region

Table 5
Mean Absolute Percentage Difference across individual and heterogeneous regions between the measured and the calculated WET differences for MD and SPR HLUT.

Mean Absolute Percentage Difference	GE 80kVp CT	Siemens 80kVp CT	GE 120kVp CT	Siemens 120kVp CT
80kVp HLUT (MD)	1.27	1.35	–	–
80kVp HLUT (SPR)	1.14	1.32	–	–
120kVp HLUT (MD)	–	–	1.20	1.33
120kVp HLUT (SPR)	–	–	1.17	1.30
100kVp HLUT (MD)	1.36	1.36	1.97	1.72
100kVp HLUT (SPR)	1.25	1.55	1.23	1.45

could be possibly due to set-up uncertainty especially if the proton beam's path is very close to a *Bone-Tissue* interface. To test this hypothesis, we simulated a 1*mm* shift in the anterior-posterior and superior-inferior directions of the isocenter in the TPS. The results are reflected as an error bar in Fig. 6 which shows the minimum and maximum percentage difference obtained after applying the shift in the TPS. With the inclusion of 1*mm* shift in set-up error, the percentage difference in the *Bone* region in Fig. 6C shows a potential decrease from 5% to 3.2%. This shows set-up uncertainty can possibly account for the large error in the *Bone* results. Apart from the *Bone* region, with the inclusion of 1*mm* shift, the WET difference between calculated and measured is still within 3.5%. Specifically, looking at the data from *H1* to *H6* which represents a more realistic scenario in actual patients, they are all within a 2.5% WET difference.

4. Discussion

Since the publication of Uwe Scheneider's [17] and Wilfried Schneider's [29] method, it has been widely used to establish the SPR HLUT for the TPS [17,25,30]. The newly established step-by-step guideline aims to resolve not only the X-ray beam hardening effect during the CT scans of the phantom but also the standardisation of categorising the data points into 4 regions (*Lung*, *Adipose Tissue*, *Soft Tissue* and *Bone*) for fitting during the generation of HLUT. The standardization of protocol for generating the HLUT across all centres will reduce the variation for inter-centre comparison, especially during multi-institution clinical trials [3].

In this work, Schneider's method (old) was used to generate the HLUT and compared with the step-by-step guide from the ESTRO's consensus guideline (new) as shown in Fig. 3. As there are two phantoms used with two methods of calibration, the observed changes were not due to the difference in the phantoms used. The MECT phantom was later used with Schneider's method to generate the HLUT and the results showed no difference in this study. The WET difference between TPS calculation using two methods of generating SPR HLUT and irradiation measurement for two CT scanners is shown in Fig. 4. In our centre, two different CT energies (80 and 120 kVp) are used for actual patient simulation. Most patients and treatment sites are scanned with 120 kVp in our centre, except for paediatric and extremities, which are scanned with 80 kVp. This is to reduce imaging doses to paediatric patients [31,32] and improve contrast to noise ratio (CNR) in small volumes especially in the case of extremities. The inclusion of 100 kVp in this study is to study the possibility of using a single HLUT curve for 80kVp and 120kVp CT scans in proton treatment planning. Having a single HLUT curve can be an attractive option for clinical workflow as it prevents human error such as choosing the wrong HLUT.

The HLUT are generated and the R^2 of the fit are tabulated. In the *Adipose Tissue* region, the R^2 values shown in Table 3 is the lowest as compared to other regions. But generally, the R^2 values are similar between the old and the new method for all regions. Therefore, to quantify the accuracy of the SPR HLUT generated, the irradiation of biological measurements is carried out. The WET differences between the TPS calculated and measurements are compared. This measurement was done to determine if a range uncertainty of 3.5% is sufficient for treatment planning. Fig. 4 shows that only the *Bone* region of Siemens X.Cite CT scanner has a range error exceeding 3.5 % (up to 5 %) in the SPR HLUT comparison (in Fig. 4C). This high percentage difference was suspected to be due to the possibility of set-up error.

Upon introducing the 1*mm* shift in set-up error, the percentage difference in the *Bone* region can be shown to be reduced to 3.2%. From the error bar shown in Fig. 6, it can be seen that the high percentage difference might be attributed to the set-up error rather than the inaccuracy of the SPR HLUT. Apart from the *Bone* region, with the inclusion of 1*mm* shift, the WET differences for all the other data points in Fig. 6 are still within 3.5%. Specifically, looking at the data from *H1* to *H6* which represents a more realistic treatment scenario, they are all within a 2.5% WET difference. The overall results show that a single 100kVp HLUT is plausible even though the CT volume is scanned with 120kVp or 80kVp provided a 3.5 % range uncertainty is used in the TPS. This is desirable from a risk-based quality management point of view to prevent human error when choosing a wrong HLUT curve in TPS.

Lastly, Fig. 5 shows the impact of using MD HLUT (80kVp and 120kVp) vs 100kVp SPR HLUT, the WET percentage difference is compared. The MD HLUT is generated following the ESTRO's consensus guidelines. For both MD and SPR HLUT, all regions show less than a 3.5% WET difference except the *Bone* region which goes up to 5%. The possible explanation for the high percentage difference could be due to set-up error and has been discussed above. In the realistic treatment scenario, the WET difference in the heterogeneous regions (H1 to H6) shows a maximum of 2.1%. Our result agreed with other works published by Möhler et. al. [33] and Niepel et. al. [34] where the deviation between calculated and measured for individual tissues falls between 0.1% to 3.8%. Based on these results, both MD (80kVp and 120kVp) and SPR (100kVp) HLUT can be used in the treatment planning with a conservative limit of 3.5% range uncertainty. In addition, the benefit of using an MD HLUT allows the conversion of a proton treatment plan to a photon treatment plan during mid-treatment easily when planning with TPS.

This work shows that the HLUT generated following the ESTRO's consensus method is suitable for clinical use. Undoubtedly, the ESTRO consensus guideline prescriptional approach towards the CT calibration curve is important for reducing the occurrence of CT calibration errors in proton therapy centres and improving the standardization of data obtained from proton therapy patients. Another way of improving the CT calibration curve will be the use of DECT modality. Many studies [35–40] have shown that using DECT can reduce the proton range uncertainty and improve clinical benefits. However, in this study, we found that a potential source of error which could compromise the accuracy of HLUT lies in the range measurement of the phantom insert. The inserts of MECT phantom have index embossed within the inserts which are of different densities. The presence of the index allows the user to identify the type of insert through the CT images, but could at the same time introduce errors in the range measurement in our experience. Hence, It would be ideal to order an additional slab of the same batch of inserts during the purchase of the MECT phantom so as to conduct the measurement of the SPR of the phantom inserts without any discrepancy arising from the index in the inserts.

5. Conclusion

This is the first report on calibrating an HLUT following the ESTRO's consensus guideline. A 100kVp HLUT was successfully calibrated and the accuracy is proven to be sufficient to be used on both 80kVp and 120kVp CT scans. The proton range uncertainty of 3.5% was established to be an appropriate limit to be used in our clinic. Despite a less than 1% improvement in the proton range uncertainty when using the ESTRO's consensus guideline as compared to Schneider's method, it is crucial to consistently work towards minimizing the proton range uncertainty to

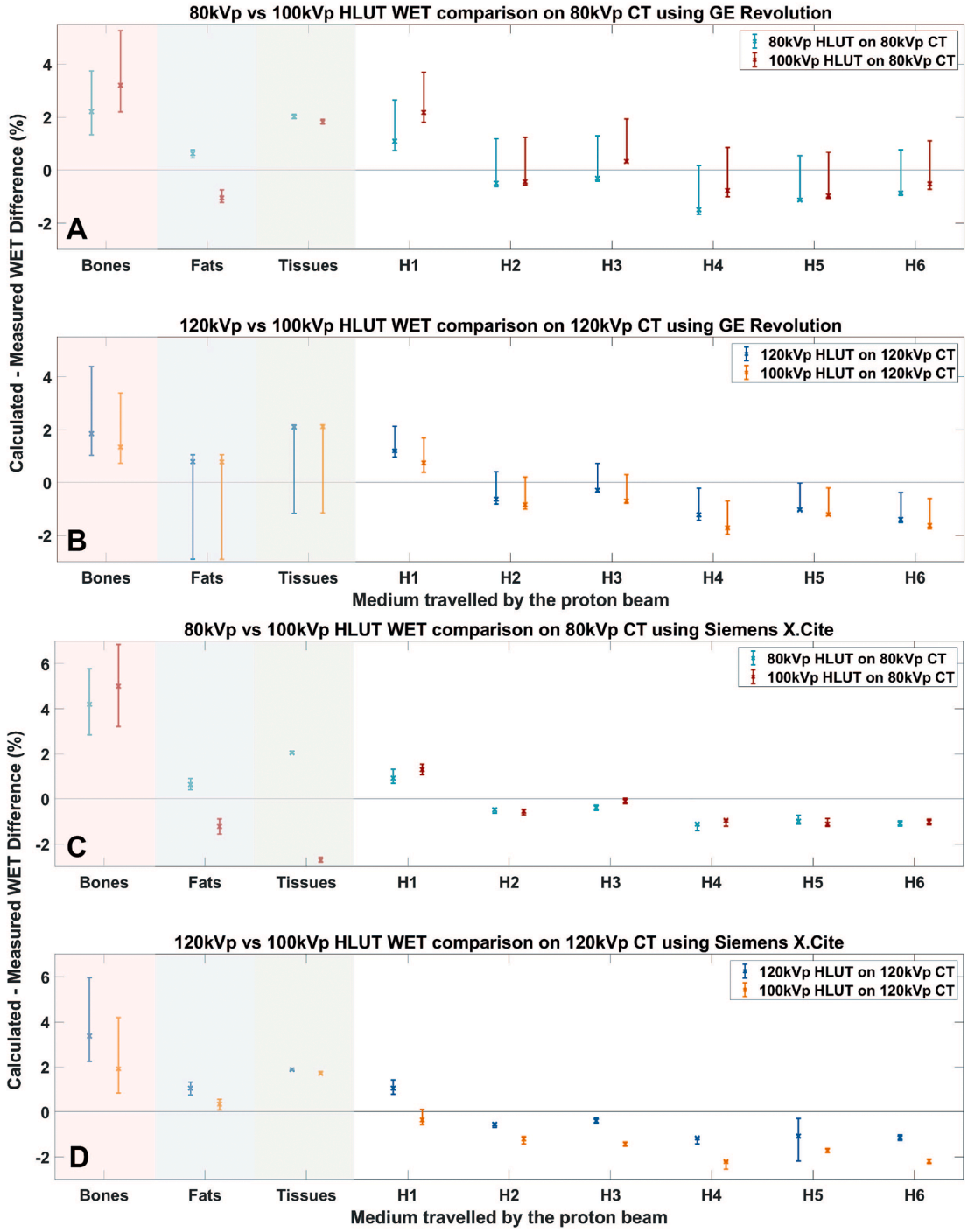

Fig. 6. For both 80kVp and 120kVp CT scans in the GE Revolution (A and B) and Siemens X.Cite (C and D) CT scanner, the percentage WET difference for SPR HLUT between 80kVp vs 100kVp and 120kVp vs 100kVp was compared. The error bar account for 5 data points where the lower and upper bound represents the minimum and maximum WET difference when there is a 1 mm shift in the anterior/posterior and lateral directions.

improve the treatment quality for the patient.

Funding support

Hong Qi Tan is supported by the Duke-NUS Oncology Academic Program Goh Foundation Proton Research Programme (08/FY2021/EX (SL)/92-A146), Clinical & Systems Innovation Support – Innovation Seed Grant (08/FY2022/P2/02-A68).

Calvin Wei Yang Koh is supported by the Ministry of Health, Singapore, MOH Health Innovation Fund (MH 110:12/12-30).

Author contribution statement

Study conception and design: Hong Qi Tan, Calvin Wei Yang Koh. Data acquisition and analysis: Calvin Wei Yang Koh, Hong Qi Tan, Kah Seng Lew, Andrew Wibawa. Data interpretation: All authors. Obtained funding: Hong Qi Tan. Administrative, technical, or material support: Calvin Wei Yang Koh. Study supervision: Sung Yong Park. Drafting of manuscript: Calvin Wei Yang Koh. Approval of final manuscript: All authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Mohan R. A review of proton therapy – current status and future directions. Precis Radiat Oncol 2022;6(2):164–76. https://doi.org/10.1002/pro6.1149.
[2] Paganetti H, Beltran C, Both S, et al. Roadmap: proton therapy physics and biology. Phys Med Biol 2021;66(5). https://doi.org/10.1088/1361-6560/abcd16.
[3] Taasti VT, Bäumer C, Dahlgren CV, et al. Inter-centre variability of CT-based stopping-power prediction in particle therapy: survey-based evaluation. Phys Imaging Radiat Oncol 2018;6:25–30. https://doi.org/10.1016/j.phro.2018.04.006.
[4] Tattenberg S, Madden TM, Gorissen BL, Bortfeld T, Parodi K, Verburg J. Proton range uncertainty reduction benefits for skull base tumors in terms of normal tissue complication probability (NTCP) and healthy tissue doses. Med Phys 2021;48(9):5356–66. https://doi.org/10.1002/mp.15097.
[5] Tattenberg S, Madden TM, Bortfeld T, Parodi K, Verburg J. Range uncertainty reductions in proton therapy may lead to the feasibility of novel beam arrangements which improve organ-at-risk sparing. Med Phys 2022;49(7):4693–704. https://doi.org/10.1002/mp.15644.
[6] Hahn C, Eulitz J, Peters N, et al. Impact of range uncertainty on clinical distributions of linear energy transfer and biological effectiveness in proton therapy. Med Phys 2020;47(12):6151–62. https://doi.org/10.1002/mp.14560.
[7] Garbacz M, Gajewski J, Durante M, et al. Quantification of biological range uncertainties in patients treated at the Krakow proton therapy centre. Radiat Oncol 2022;17(1). https://doi.org/10.1186/s13014-022-02022-5.
[8] Goitein M. Calculation of the uncertainty in the dose delivered during radiation therapy. Med Phys 1985;12(5):608–12. https://doi.org/10.1118/1.595762.
[9] Wohlfahrt P, Möhler C, Enghardt W, et al. Refinement of the Hounsfield look-up table by retrospective application of patient-specific direct proton stopping-power prediction from dual-energy CT. Med Phys 2020;47(4):1796–806. https://doi.org/10.1002/mp.14085.
[10] Wohlfahrt P. *Proton TheraPy SPecial Feature: Review Article Status and Innovations in Pre-Treatment Ct Imaging for Proton Therapy 1*.; 2020.
[11] Wohlfahrt P, Möhler C, Hietschold V, et al. Clinical implementation of dual-energy CT for proton treatment planning on pseudo-monoenergetic CT scans. Int J Radiat Oncol Biol Phys 2017;97(2):427–34. https://doi.org/10.1016/j.ijrobp.2016.10.022.
[12] Taasti VT, Hansen DC, Michalak GJ, et al. Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction. Med Phys 2018;45(11):5186–96. https://doi.org/10.1002/mp.13173.
[13] Simard M, Lapointe A, Lalonde A, Bahig H, Bouchard H. The potential of photon-counting CT for quantitative contrast-enhanced imaging in radiotherapy. Phys Med Biol 2019;64(11). https://doi.org/10.1088/1361-6560/ab1af1.
[14] Hu G, Niepel K, Risch F, et al. Assessment of quantitative information for radiation therapy at a first-generation clinical photon-counting computed tomography scanner. *front*. Oncol 2022;12. https://doi.org/10.3389/fonc.2022.970299.
[15] Johnson RP, Bashkirov V, Dewitt L, et al. A fast Experimental scanner for proton CT: technical performance and first Experience with phantom scans. IEEE Trans Nucl Sci 2016;63(1):52–60. https://doi.org/10.1109/TNS.2015.2491918.
[16] Schulte RW, Penfold SN. Proton CT for improved stopping power determination in proton therapy, invited. Trans Am Nucl Soc 2012;106:55–8.
[17] Schneider U, Pedroni E, Lomax A. *The Calibration of CT Hounsfield Units for Radiotherapy Treatment Planning*. Vol 41.; 1996. http://iopscience.iop.org/0031-9155/41/1/009.
[18] ICRU Report 44. *Tissue Substitutes in Radiation Dosimetry and Measurement*.; 1989.
[19] Peters N, Trier Taasti V, Ackermann B, et al. Consensus guide on CT-based prediction of stopping-power ratio using a Hounsfield look-up table for proton therapy. Radiother Oncol 2023;184:109675. https://doi.org/10.1016/j.radonc.2023.109675.
[20] Peters N, Wohlfahrt P, Dahlgren CV, et al. Experimental assessment of inter-centre variation in stopping-power and range prediction in particle therapy. Radiother Oncol 2021;163:7–13. https://doi.org/10.1016/j.radonc.2021.07.019.
[21] Habr-Gama A, Perez RO, Wynn G, Marks J, Kessler H, Gama-Rodrigues J. Complete clinical response after neoadjuvant chemoradiation therapy for distal rectal cancer: Characterization of clinical and endoscopic findings for standardization. Dis Colon Rectum 2010;53(12):1692–8. https://doi.org/10.1007/DCR.0b013e3181f42b89.
[22] Olivo SA, Macedo LG, Gadotti C, Fuentes J, Stanton T, Magee DJ. *Scales to Assess the Quality of Randomized Controlled Trials: A Systematic Review*. Vol 156.; 2008. www.ptjournal.org.
[23] AbuHalimeh A. Improving data quality in clinical Research informatics tools. Front Big Data 2022;5. https://doi.org/10.3389/fdata.2022.871897.
[24] Kazmierska J, Hope A, Spezi E, et al. From multisource data to clinical decision aids in radiation oncology: the need for a clinical data science community. Radiother Oncol 2020;153:43–54. https://doi.org/10.1016/j.radonc.2020.09.054.
[25] Ainsley CG, Yeager CM. Practical considerations in the calibration of CT scanners for proton therapy. J Appl Clin Med Phys 2014;15(3):202–20. https://doi.org/10.1120/jacmp.v15i3.4721.
[26] Dracham CB, Shankar A, Madan R. Radiation induced secondary malignancies: a review article. Radiat Oncol J 2018;36(2):85–94. https://doi.org/10.3857/roj.2018.00290.
[27] Bethe H. Zur theorie des durchgangs schneller Korpuskularstrahlen durch materie. Annalen der Physik 1930;397(3):325–400. https://doi.org/10.1002/andp.19303970303.
[28] Deasy J. ICRU report 49, stopping powers and ranges for protons and alph Particles. Med Phys 1994;21(5):709–10. https://doi.org/10.1118/1.597176.
[29] Schneider W, Bortfeld T, Schlegel W. *Correlation between CT Numbers and Tissue Parameters Needed for Monte Carlo Simulations of Clinical Dose Distributions*. Vol 45.; 2000. http://iopscience.iop.org/0031-9155/45/2/314.
[30] Goma C, Almeida IP, Verhaegen F. Revisiting the single-energy CT calibration for proton therapy treatment planning: a critical look at the stoichiometric method. Phys Med Biol 2018;63(23). https://doi.org/10.1088/1361-6560/aaede5.
[31] Zacharias C, Alessio AM, Otto RK, et al. Pediatric CT: strategies to lower radiation dose. Am J Roentgenol 2013;200(5):950–6. https://doi.org/10.2214/AJR.12.9026.
[32] Ogbole G. 2010 dec radiation dose in paediatric computed tomography risks and benefits. Ann Ib Postgrad Med 2010;8(2):118–26. https://doi.org/10.4314/aipm.v8i2.71823.
[33] Möhler C, Russ T, Wohlfahrt P, et al. Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues. Phys Med Biol 2018;63(2). https://doi.org/10.1088/1361-6560/aaa1c9.
[34] Niepel KB, Stanislawski M, Wuerl M, et al. Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry. Phys Med Biol 2021;66(7). https://doi.org/10.1088/1361-6560/abbd14.
[35] Zhu J, Penfold SN. Dosimetric comparison of stopping power calibration with dual-energy CT and single-energy CT in proton therapy treatment planning. Med Phys 2016;43(6):2845–54. https://doi.org/10.1118/1.4948683.
[36] Bär E, Lalonde A, Royle G, Lu HM, Bouchard H. The potential of dual-energy CT to reduce proton beam range uncertainties. Med Phys 2017;44(6):2332–44. https://doi.org/10.1002/mp.12215.
[37] Peters N, Wohlfahrt P, Hofmann C, et al. Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy CT for direct stopping-power prediction. Radiother Oncol 2022;166:71–8. https://doi.org/10.1016/j.radonc.2021.11.002.
[38] Taasti VT, Muren LP, Jensen K, et al. Comparison of single and dual energy CT for stopping power determination in proton therapy of head and neck cancer. Phys Imaging Radiat Oncol 2018;6:14–9. https://doi.org/10.1016/j.phro.2018.04.002.
[39] Taasti VT, Decabooter E, Eekers D, et al. Clinical benefit of range uncertainty reduction in proton treatment planning based on dual-energy CT for neurooncological patients. Br J Radiol 2023;96(1149). https://doi.org/10.1259/bjr.20230110.
[40] Li B, Lee HC, Duan X, et al. Comprehensive analysis of proton range uncertainties related to stopping-power-ratio estimation using dual-energy CT imaging. Phys Med Biol 2017;62(17):7056–74. https://doi.org/10.1088/1361-6560/aa7dc9.