An Anthropomorphic Head and Neck Quality Assurance Phantom for Credentialing of Intensity-Modulated Proton Therapy

Daniela Branco, MS[1]; Paige Taylor, MS[1]; Xiaodong Zhang, PhD[1]; Heng Li, PhD[1]; Michele Guindani, PhD[2]; and David Followill, PhD[1]

[1]Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Abstract

Purpose: To design and commission a head and neck (H&N) anthropomorphic phantom that the Imaging and Radiation Oncology Core Houston (IROC-H) can use to verify the quality of intensity-modulated proton therapy H&N treatments for institutions participating in National Cancer Institute–sponsored clinical trials.

Materials and Methods: The phantom design was based on a generalized oropharyngeal tumor, including critical H&N structures (parotid glands and spinal cord). Radiochromic film and thermoluminescent dosimeter (TLD)-100 capsules were embedded in the phantom and used to evaluate dose delivery. A spot-scanning treatment plan with typical clinical constraints for H&N cancer was created by using the Eclipse analytic algorithm. The treatment plan was approved by a radiation oncologist and the phantom was irradiated 4 times. The measured dose distribution using a ±7%/4 mm gamma analysis (85% of pixels passing) and point doses were compared with the treatment planning system calculations. The prescribed target dose was 6 Gy (RBE) with 646.2 cGy (RBE) and 648.6 cGy (RBE) planned to the superior and inferior TLD, respectively.

Results: For point dosimetry, the average measured-to-calculated dose ratios were 0.984 and 0.986 for the superior and inferior target TLD, respectively. Dose values for the superior and inferior target TLDs were 636.1 cGy and 639.6 cGy, respectively. For the relative dose comparison, the pixel passing rates for the axial and sagittal films, respectively, were 95.5% and 94.2% for trial 1, 97.3% and 93.2% for trial 2, 93.4% and 90.0% for trial 3, and 96.2% and 92.7% for trial 4.

Conclusion: The anthropomorphic H&N phantom was successfully designed so that TLD measured-to-calculated ratios were within IROC-H's 7% acceptance criteria, 1.6% and 1.4% lower than expected for the superior and inferior target TLDs, respectively. All trials passed the 85% pixel passing criteria established at IROC-H for the relative dose comparison performed when using a gamma index of ±7%/4 mm.

Keywords: anthropomorphic head & neck phantom; quality assurance; proton therapy; Imaging and Radiation Oncology Core; National Cancer Institute

Introduction

The use of proton therapy to treat cancer has rapidly increased during the past decade; 25 proton centers are currently in operation in the United States and 11 more are in development [1]. As a consequence of the increasing interest in the use of proton therapy, the demand for good quality assurance (QA) programs to control and maintain the quality standards of patient care is high. Although each proton therapy facility already has its own set of comprehensive QA tests in place, the International Commission on Radiation Units and Measurements [2] recommends an independent QA program that confirms the accuracy, comparability, and consistency of proton therapy delivery between facilities, especially for multi-institution clinical trial activities.

As a core support for its clinical trials, The National Cancer Institute (NCI) funds various QA centers across the country to help ensure that institutions are delivering comparable and consistent doses of radiation, in an effort to minimize data uncertainty for trials that include radiation therapy. The Imaging and Radiation Oncology Core Houston (IROC-H) QA center, formerly known as the Radiological Physics Center [3], is 1 of 6 of these QA centers. Its mission is to ensure that institutions participating in NCI-sponsored clinical trials have acceptable QA procedures and no significant systematic dosimetry inconsistencies, so that each site can be considered capable of providing quality and comparable clinical treatments for cancer patients. This is especially true for clinical trials that allow proton therapy, since it is a relatively new mainstream form of trial radiation therapy.

In 2012, NCI developed guidelines [4] for the use of proton therapy in NCI-funded multi-institutional clinical trials. These guidelines specify an approval process that each new proton facility must follow before being allowed to enroll a proton therapy patient into an NCI-funded clinical trial. A key part of the proton therapy approval process consists of irradiating baseline end-to-end anthropomorphic QA phantoms. These mailable anthropomorphic QA phantoms, used for both approval and credentialing processes, are used to verify the accuracy of the dose delivery for the individual proton treatments, which represent a hypothetical trial patient treatment. These end-to-end patient treatment verifications typically measure the magnitude of the dose delivered as well as the spatial distribution of the dose.

Prior to the work described here, IROC-H did not have an end-to-end anthropomorphic QA head and neck (H&N) phantom that could be used to credential clinical trials of proton therapy for oropharyngeal cancer. This article describes the design and validation of a new anthropomorphic H&N phantom for proton therapy.

Materials and Methods

Phantom Design and Construction

To enable evaluation of the planning and dose delivery of proton treatments to the oropharynx, the phantom needed to simulate human anatomy associated with H&N malignancies. In particular, tissue heterogeneities needed to be included to properly account for the clinically relevant anatomic variations and to enable the distinction between targets and critical structures. The original phantom was an Alderson phantom purchased from The Phantom Laboratory (Salem, New York), made of Alderson water-equivalent plastic with added airway channels and a human skull inside to mimic actual human head anatomy. The water-equivalent plastic had previously been shown to be proton-equivalent by Grant et al [5]. **Figure 1** shows the sagittal and axial computed tomography (CT) scan of the original phantom before modification, showing the oral and sinus air cavities and human skull.

To be an appropriate end-to-end QA tool for proton trials, the H&N phantom had to include imageable targets and critical structures that simulated human anatomic dimensions and the usual extent of oropharyngeal disease, while still accommodating radiation dosimeters. A cylindrical insert that would slide into place from the inferior portion of the phantom was designed from actual patient anatomy, containing all relevant structures and dosimeters. The cylinder included a protuberance on the superior part that locked in place with the phantom itself in order to guarantee reproducible positioning of the dosimeters between irradiations. The maximum diameter allowed for the insert was approximately 9.5 cm; **Supplementary Figure 1** illustrates the 3-dimensional schematic of the design. The insert design included a horseshoe-shaped target that wrapped partially around the spinal cord and was placed in the center of the insert, along with 3 relevant organs at risk (OARs): the spinal cord and 2 parotid glands placed laterally. The final material composition of the cylinder was solid water, polyethylene, and blue water for the structures inside. The percentage difference in stopping power of the Alderson solid water, blue water, solid water, and polyethylene is −0.6%, −0.1%, −0.6%, and 1.9%, respectively [5]. The placement of the structures was such that proton beams would have to travel through bony structures as well as air cavities. The parotid structures had to

Figure 1. Sagittal and axial computed tomography scan of the original head and neck phantom purchased before modifications.
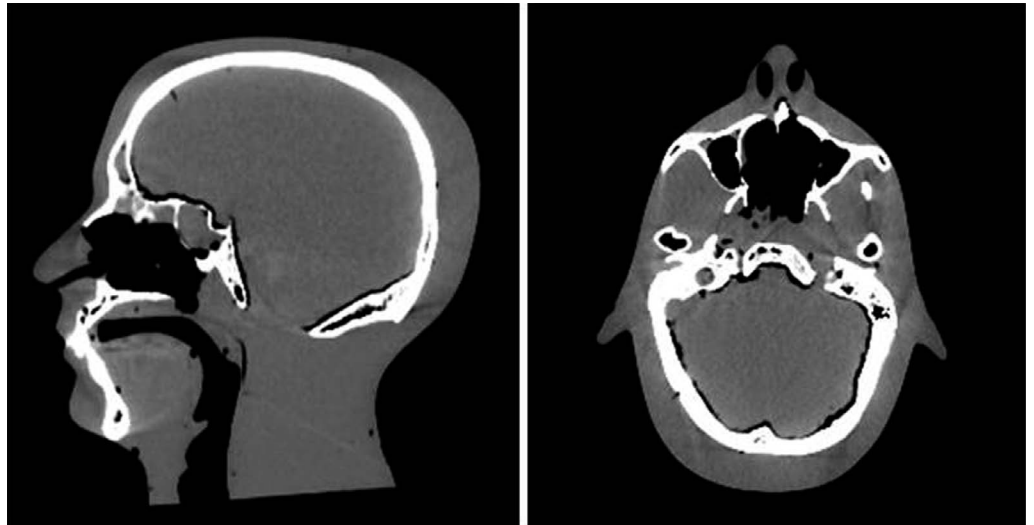
be placed deeper than normal human anatomy because of physical limitations of the phantom, but their placement was adequately realistic and representative of an oropharyngeal treatment setup.

The insert was split into 4 pieces so that radiochromic film could be inserted in the axial and sagittal planes. To keep the pieces held tightly together and avoid any air gaps, the 4 quadrants of the insert were attached together by using small 6/6 nylon screws and the whole cylinder was secured by a thin (2 mm) external plastic sleeve made of polyethylene. The film was prevented from rotating or moving inside the phantom by small stainless steel pins that also served to place registration marks on each film. Holes to hold thermoluminescent dosimeter (TLD) capsules were created inside each relevant structure, 1 for each parotid gland and 2 for the target (superior and inferior) and spinal cord, so that absolute dose measurements could be made in each structure.

The tissue structures within the phantom had to be made of proton-equivalent tissue substitute materials, as identified by Grant et al [5]. Therefore, only tissue-equivalent materials that fell on or near ($\pm2\%$) the Hounsfield unit/relative linear stopping power curve were used [5]. In addition, the proton H&N phantom had to have a similar structural design to the IROC-H photon H&N phantom.

The phantom, once constructed, was scanned with a GE LightSpeed RT16 CT scanner at The University of Texas MD Anderson Cancer Center Proton Therapy Center of Houston (PTC-H). A typical H&N protocol was used to image the phantom, with 1.25-mm slices at 120 kVp. In addition, the scan was done in helical mode with a pitch of 0.9375. To ensure reproducibility, the phantom was placed in the supine position on a Klarity mold that was shaped to the phantom.

Dosimeters

Gafchromic EBT2 film (Ashland, Wayne, New Jersey) was used to perform the analysis of the sagittal and axial dose distributions of the irradiations. Radiochromic film was considered the appropriate relative dosimeter for the study of the dose profiles because it shows no angular dependence and is near tissue equivalent. In addition, radiochromic film offers sensitivity (0.1 to 10 Gy) in the required range for this project (6 Gy) and has great spatial resolution (film and densitometer system is 1 mm). It can also be handled in visible light and is self-developing [6, 7], making this passive detector very suitable for the remote quality programs established at IROC-H. Film response has been reviewed for time dependence, but for the general time scale of a phantom irradiation, shows now measureable effects. Gafchromic film shows no significant energy dependency [7]; however, literature shows that film exhibits an underresponse in the area under the Bragg peak when compared to ion chamber measurements, characterized as linear energy transfer quenching [8]. To help minimize this film quenching effect, the film used in this project was slightly angled with respect to the proton beams [9].

Thermoluminescent dosimeters were used as point dosimeters in the phantom, providing absolute dose measurements at the specific locations where they were positioned in the target and OARs. A fading correction is applied to the TLD to account for the time that has elapsed since the day of irradiation. Film and TLD reading procedures were based on the technique used by Molineu et al [10]. The dosimetric precision of the TLD is 3%, and the spatial precision of the film and densitometer system

is 1 mm. Thermoluminescent dosimeters offered a valuable set of advantages for the project because they display a wide useful linear dose range (1 mGy to 10 Gy) and are dose rate–independent [11]. In addition, TLDs are small, do not disturb the radiation field, and are accurate and reusable. The TLD physical doses were relative biological effectiveness (RBE) corrected by using a value of 1.1.

Treatment Plan

Following the CT simulation, a treatment plan was created by using the Eclipse proton beam treatment planning system (TPS), version 13.6 (Varian Medical Systems, Inc, Palo Alto, California), following the PTC-H clinical protocol. The dose computation was done with Eclipse's Proton Convolution Superposition algorithm (version 11.0.30). The proton range uncertainty used at the PTC-H for H&N patients is 3.5% + 3 mm and was incorporated into this project. Although the typical proton treatment dose for H&N planning treatment volumes (PTVs) is 70 Gy (RBE), the IROC-H irradiation protocol requires the prescription to be a factor of approximately 10 times smaller owing to dosimeter limitations. Consequently, the PTV was planned to have 6 Gy (RBE) delivered in 1 fraction. IROC-H uses Gafchromic EBT2 film, which saturates at doses near 10 Gy [6]. Typical clinical dose constraints are 26 Gy (RBE) for parotid OARs and 46 Gy (RBE) for spinal cord OARs. The OARs were also scaled similarly to the target dose: 2.6 Gy (RBE) for parotid OARs and 4.5 Gy (RBE) for spinal OARs.

Various beam configurations and plan optimizations were investigated; however, passive scatter and single-field optimized spot scanning plans did not meet the planning constraints. The optimal dose coverage with best tissue sparing was achieved with a multifield optimization spot scanning treatment plan using the prescribed dose of 6 Gy (RBE) to ≥95% of the PTV, with the use of 1 posterior beam and 2 anterior oblique beams, as shown in **Supplementary Figure 2** in the Supplemental Materials. The PTV was defined with no extra margins on the gross tumor volume owing to the inanimate nature of the phantom, which required no need to address motion or microscopic disease. This plan resulted in the PTV receiving 6 Gy (RBE) coverage to 96.7% of the PTV volume. The OARs were kept under the acceptable dose constraints. The mean doses delivered to the left and right parotid glands were 2.59 Gy (RBE) and 2.30 Gy (RBE), respectively, meeting the 2.6 Gy (RBE) limitation. The spinal cord received a mean dose of 3.77 Gy (RBE) but a maximum dose of 5.7 Gy (RBE), which exceeded the 4.5 Gy (RBE) maximum dose restriction. This compromise was made because of the limited anatomic fit of the OARs and PTV structures in the phantom insert.

All 4 irradiation trials were performed on the G3 spot scanning beam gantry at the PTC-H. The phantom was placed in the supine position and aligned with the gantry lasers and orthogonal x-ray imaging. The imaging parameters used for the setup were the same as those recommended for patients with H&N cancer treated at the PTC-H.

Results

Point Dose Comparison

For the absolute dose comparison, the phantom TLD doses from each of the spot-scanning irradiations were compared with the calculated doses from the Eclipse TPS. The RBE-weighted dose to the superior target TLD was calculated by the treatment plan to be 646.2 cGy (RBE), and 648.6 cGy (RBE) was calculated for the inferior target TLD. The values for calculated and measured target doses for each trial, as well as the measured-to-calculated dose ratios, are shown in **Table 1**.

The average measured-to-calculated dose ratios were 0.984 for the superior target TLD and 0.986 for the inferior target TLD. The standard deviation for all target TLD measurements was 1.5%. The values for the calculated and measured doses for the critical structures in the phantom, along with the measured-to-calculated dose ratios, are shown in **Table 2**.

The average measured-to-calculated dose ratios were 1.143, 1.002, and 0.978 for the left parotid gland, right parotid gland, and spinal cord, respectively. The standard deviations of these values were 6.0%, 2.5%, and 1.0%, respectively, demonstrating higher variability than in the measured doses in the PTV. This was to be expected because the OAR TLDs were located in regions with a steep dose gradient. A TLD was also inserted into the oral cavity to ascertain whether the dose delivered around the mandible and teeth was kept low, given that this is a very sensitive region in patients. That TLD was not part of the treatment plan and therefore a measured-to-calculated dose ratio could not be established. The average reading among all trials for the mouth TLD was 1.4 cGy (RBE), well below clinical constraints.

Relative Dose Comparison

The 2-dimensional dose distributions were analyzed by comparing the dose distribution calculated by the TPS with that measured by the phantom films in the axial and sagittal planes. The film dose distributions were normalized to the RBE-

Branco et al. (2017), *Int J Particle Ther*

43

Table 1. Target point dose comparison between the TPS and TLDs for all irradiation trials.

TLD location	TPS dose calculated, cGy (RBE)	Dose measured, cGy (RBE)	Measured-to-calculated dose ratio
Trial 1			
Target superior	646.2	654.3	1.013
Target inferior	648.6	627.9	0.968
Trial 2			
Target superior	646.2	631.6	0.977
Target inferior	648.6	638.1	0.984
Trial 3			
Target superior	646.2	631.8	0.978
Target inferior	648.6	625.6	0.964
Trial 4			
Target superior	646.2	634.8	0.982
Target inferior	648.6	646.3	0.996

Abbreviations: TPS, treatment planning system; TLDs, thermoluminescent dosimeters; RBE, relative biological effectiveness.

weighted TLD doses at the locations of the TLD capsules in the target. The gamma index acceptance criterion used in this study was ±7%/4 mm, and tighter criteria of ±5%/3 mm and ±5%/4 mm were also evaluated. The 2-dimensional gamma analysis results, showing the percentage of pixels meeting the various acceptance criteria for all irradiation trials, are listed in **Table 3**.

As illustrated in **Figure 2**, the axial dose map obtained in the gamma analysis showed the same general distribution of passing and failing pixels through the 4 trials. The largest percentage of failing pixels fell within the target and in the posterior portion of the film plane. Similar to the axial films, the sagittal films (**Figure 3**) showed that most of the failing pixels were located inside the target and in the posterior region on the film.

Discussion

An anthropomorphic H&N phantom was successfully designed and built to evaluate proton therapy procedures for oropharyngeal cancer, with an agreement between measured and calculated doses of ±7%/4 mm. The phantom was used in

Table 2. Critical structure point dose comparison between the TPS and TLDs.

TLD location	TPS dose calculated, cGy (RBE)	Dose measured, cGy (RBE)	Measured-to-calculated dose ratio
Trial 1			
Parotid left	250.2	274.8	1.098
Parotid right	206.4	199.1	0.965
Cord	503.4	493.2	0.980
Trial 2			
Parotid left	250.2	273.5	1.093
Parotid right	206.4	209.3	1.014
Cord	503.4	498.5	0.990
Trial 3			
Parotid left	250.2	285.3	1.140
Parotid right	206.4	209.6	1.015
Cord	503.4	486.5	0.966
Trial 4			
Parotid left	250.2	311.1	1.243
Parotid right	206.4	209.8	1.016
Cord	503.4	492.7	0.979

Abbreviations: TPS, treatment planning system; TLDs, thermoluminescent dosimeters; RBE, relative biological effectiveness.

Branco et al. (2017), *Int J Particle Ther*

44

Table 3. Two-dimensional gamma analysis pass rates for all irradiation trials.

| Location	2-Dimensional gamma percentage of pixels passing		
	±5%/3 mm	±5%/4 mm	±7%/4 mm
Trial 1			
Axial	86.3%	91.0%	95.5%
Sagittal	82.6%	87.6%	94.2%
Trial 2			
Axial	91.0%	94.0%	97.3%
Sagittal	80.2%	87.0%	93.2%
Trial 3			
Axial	82.1%	88.6%	93.4%
Sagittal	76.0%	81.8%	90.0%
Trial 4			
Axial	91.0%	93.4%	96.2%
Sagittal	80.6%	86.4%	92.7%

4 end-to-end scanning treatments, and all treatments passed the 85% pixel passing gamma criterion and IROC's TLD acceptance criteria of ±7% dose agreement.

The average measured-to-calculated dose ratios for the superior and inferior target TLDs were 0.984 and 0.986, respectively. These ratios were within IROC's acceptance criterion of ±7% dose agreement but were both, on average, 1.5% lower than the TPS calculated dose. A possible explanation for this low outcome could be that proton therapy TPSs, compared

Figure 2. Dose maps for the axial films from each of the 4 trials: (a) trial 1, 95.5%; (b) trial 2, 97.3%; (c) trial 3, 93.4%; (d) trial 4, 96.2%.
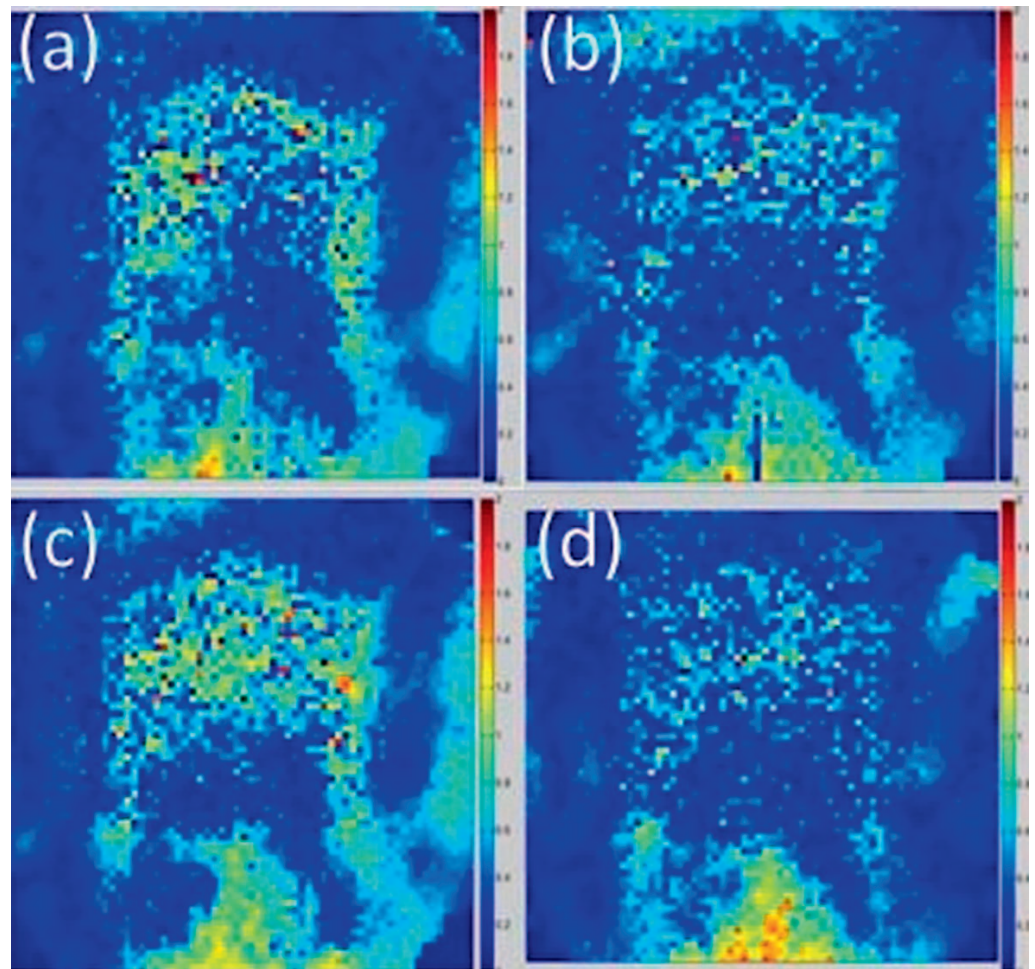
Figure 3. Dose maps for the sagittal films from each of the 4 trials: (a) trial 1, 94.2%; (b) trial 2, 93.2%; (c) trial 3, 90.0%; (d) trial 4, 92.7%.
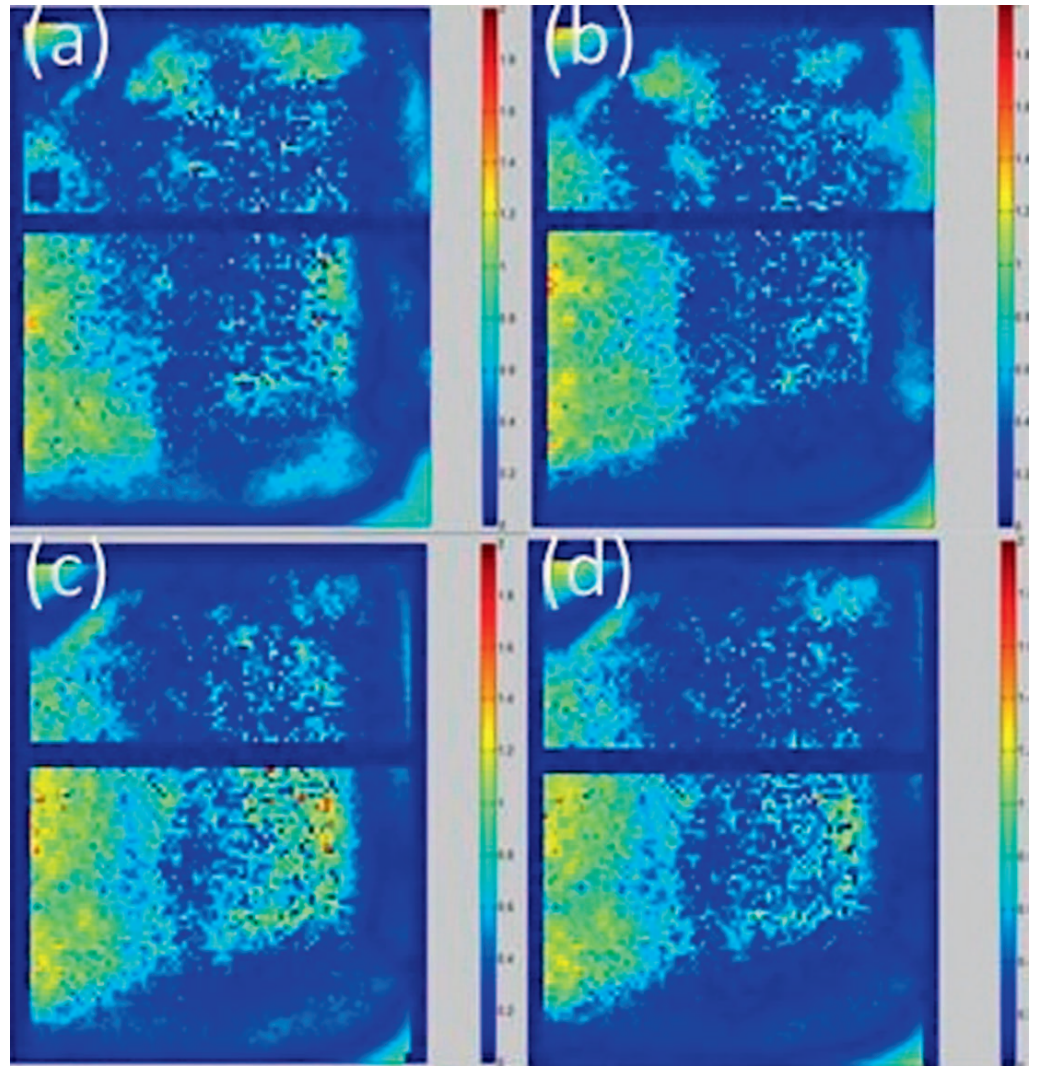
with Monte Carlo simulations, tend to overestimate target doses by as much as 3.5% for treatments targeting H&N tumors, as described by Schuemann et al [12]. Margins are used to account for the inaccurate predictions of the proton range and absolute dose, but owing to the high complexity of H&N geometries and inhomogeneities, H&N malignancies show the largest dose variations (3%-4%) between TPSs and Monte Carlo simulations [12]. Our measurements support these findings and would agree better with Monte Carlo dose calculations.

The measured-to-calculated dose ratios for the parotid glands also had some variability. The ratios for the parotid glands were high by 14.3% and 0.2% for the left and right parotid glands, respectively, whereas the ratio for the spinal cord was low by 2.2%. The left parotid gland had the poorest agreement with the TPS-predicted dose. A possible explanation for this discrepancy is that the left parotid gland was located in a higher dose gradient region than the right parotid gland, as shown in **Supplementary Figure 3**. This means that even the smallest shifts in the setup could represent large dose differences in the TLD results between the trials.

Another important observation about the OARs is related to the spinal cord dose constraints. As mentioned above, the dose to the spinal cord exceeded the maximum constraint established in the project owing to the spatial constraints of the H&N phantom. In normal human anatomy, the parotid glands and spinal cord are more superficial than the phantom insert allowed for. The insert could not have been wider because it had to fit though the neck while still allowing for enough material to support the head. Therefore, a higher than clinically advisable dose was delivered to the spinal cord to maintain target coverage. This compromise existed for the design reported here, but in the future the spinal cord will be shifted posteriorly 2 mm to better reflect real patient anatomy. This should also improve the likelihood of treatment plans meeting the spinal cord dose constraint.

Several different passive scattering plans were created by using different beam arrangements and margins. However, none of them were satisfactory, where our best attempt showed insufficient target coverage presenting numerous hotspots and poor uniformity inside the target. The spinal cord received high dose in these plans as well, with a maximum dose equivalent to the total prescribed target dose of 6.02 Gy. Therefore, after multiple iterations, it was determined that passive scattering treatment plans for the designed phantom were clinically unrealistic and should not be delivered.

All trials passed the 85% criteria used at IROC-H for the gamma index proposed ($\pm$7%/4 mm). As expected, tighter criteria showed lower passing rates but still performed well, with only the sagittal film in trial 3 not passing the $\pm$5%/4 mm criterion, likely due to a small setup error. The percentage standard deviation for all the 7%/4 mm gamma passing rates combined was 2.4%.

IROC currently uses a similar phantom for the credentialing of H&N intensity-modulated radiation therapy. That phantom is analogous to the one created for the current study, except that it is composed of high-impact polystyrene. The insert for that phantom contains a spinal cord to serve as an OAR and 2 targets, a primary PTV and a secondary PTV representing a lymph node. The targets and OARs contain TLDs and radiochromic film inserted in the axial and sagittal planes as well. The phantom designed and validated in the current study will serve in the comprehensive evaluation of H&N proton therapy at participating NCI-sponsored clinical trials. The trials described in the current study successfully met IROC's expected acceptance criteria, and therefore we have created an end-to-end QA tool to be used in the credentialing of proton therapy institutions in the future.

ADDITIONAL INFORMATION AND DECLARATIONS

References

1. The National Association for Proton Therapy. The National Association for Proton Therapy's Home Page. 2017. http://www.proton-therapy.org/. Accessed April 1, 2016.

2. Journal of the International Commission on Radiation Units and Measurements, Volume 7, Issue 2, 1 December 2007, Pages 135–139, https://doi.org/10.1093/jicru/ndm031

3. MD Anderson Imaging and Radiation Oncology Core. IROC's Home Page. 2017. http://rpc.mdanderson.org/RPC/home.htm. Accessed April 1, 2016.

4. Guidelines for the Use of Proton Radiation Therapy in NCI-Sponsored Cooperative Group Clinical Trials. (n.d.). Retrieved April 1, 2016, from http://rpc.mdanderson.org/RPC/home_page/Proton_guidelines.htm

5. Grant RL, Summers PA, Neihart JL, Blatnica AP, Sahoo N, Gillin MT, Followill DS, Ibbott GS. Relative stopping power measurements to aid in the design of anthropomorphic phantoms for proton radiotherapy. *J Appl Clin Med Phys*. 2014;15:4523.

6. Reinhardt S, Hillbrand M, Wilkens JJ, Assmann W. Comparison of Gafchromic EBT2 and EBT3 films for clinical photon and proton beams. *Med Phys*. 2012;39:5257–62.

7. Arjomandy B, Tailor R, Anand A, Sahoo N, Gillin M, Prado K, Vicic M. Energy dependence and dose response of Gafchromic EBT2 film over a wide range of photon, electron, and proton beam energies. *Med Phys*. 2010;37:1942–7.

8. Fiorini F, Kirby D, Thompson J, Green S, Parker DJ, Jones B, Hill MA. Under-response correction for EBT3 films in the presence of proton spread out Bragg peaks. *Phys Med*. 2014;30:454–61.

9. Taylor PA, Kry SF, Alvarez P, Keith T, Lujano C, Hernandez N, Followill DS. Results from the Imaging and Radiation Oncology Core Houston's anthropomorphic phantoms used for proton therapy clinical trial credentialing. *Int J Radiat Oncol Biol Phys*. 2016;95:242–8.

10. Molineu A, Followill DS, Balter PA, Hanson WF, Gillin MT, Huq MS, Eisbruch A, Ibbott GS. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. *Int J Radiat Oncol Biol Phys*. 2005;63:577–83.

11. Khan F, Gibbons J. *The Physics of Radiation Therapy*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2014.

12. Schuemann J, Giantsoudi D, Grassberger C, Moteabbed M, Min CH, Paganetti H. Assessing the clinical impact of approximations in analytical dose calculations for proton therapy. *Int J Radiat Oncol Biol Phys*. 2015;92:1157–64.