HHS Public Access

Author manuscript

Radiother Oncol. Author manuscript; available in PMC 2021 May 19.

Published in final edited form as:

Radiother Oncol. 2020 December ; 153: 163–171. doi:10.1016/j.radonc.2020.10.017.

Development and validation of an age-scalable cardiac model with substructures for dosimetry in late-effects studies of childhood cancer survivors

Suman Shrestha[a,b], **Aashish C. Gupta**[a,b], **James E. Bates**[c], **Choonsik Lee**[d], **Constance A. Owens**[a,b], **Bradford S. Hoppe**[e], **Louis S. Constine**[f], **Susan A. Smith**[a], **Ying Qiao**[a], **Rita E. Weathers**[a], **Yutaka Yasui**[g], **Laurence E. Court**[a,b], **Arnold C. Paulino**[h], **Chelsea C. Pinnix**[h], **Stephen F. Kry**[a,b], **David S. Followill**[a,b], **Gregory T. Armstrong**[g], **Rebecca M. Howell**[a,b,*]

[a]Department of Radiation Physics, The University of Texas MD Anderson Cancer Center

[b]The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

[c]Department of Radiation Oncology, Winship Cancer Institute, Emory University, United States

[d]Division of Cancer Epidemiology and Genetics, National Cancer Institute

[e]Department of Radiation Oncology, Mayo Clinic Florida

[f]Department of Radiation Oncology and Pediatrics, University of Rochester Medical Center, United States

[g]Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital

[h]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, United States

Abstract

Background and Purpose: Radiation therapy is a risk factor for late cardiac disease in childhood cancer survivors. Several pediatric cohort studies have established whole heart dose and dose–volume response models. Emerging data suggest that dose to cardiac substructures may be more predictive than whole heart metrics. In order to develop substructure dose-response models, the heart model previously used for pediatric cohort dosimetry needed enhancement and substructure delineation.

Methods: To enhance our heart model, we combined the age-scalable capability of our computational phantom with the anatomically-delineated (with substructures) heart models from an international humanoid phantom series. We examined cardiac volume similarity/overlap between registered age-scaled phantoms (1, 5, 10, and 15 years) with the enhanced heart model and the reference phantoms of the same age; dice similarity coefficient (DSC) and overlap

*Corresponding author at: Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, United States. rhowell@mdanderson.org (R.M. Howell).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2020.10.017.

coefficient (OC) were calculated for each matched pair. To assess the accuracy of our enhanced heart model, we compared doses from computed tomography-based planning (ground truth) with reconstructed heart doses. We also compared doses calculated with the prior and enhanced heart models for a cohort of nearly 5000 childhood cancer survivors.

Results: We developed a realistic cardiac model with 14-substructures, scalable across a broad age range (1–15 years); average DSC and OC were 0.84 ± 0.05 and 0.90 ± 0.05, respectively. The average percent difference between reconstructed and ground truth mean heart doses was 4.2%. In the cohort dosimetry analysis, dose and dose-volume metrics were approximately 10% lower on average when the enhanced heart model was used for dose reconstructions.

Conclusion: We successfully developed and validated an anatomically realistic age-scalable cardiac model that can be used to establish substructure dose-response models for late cardiac disease in childhood cancer survivor cohorts.

Keywords

Late effects; Radiation therapy; Cardiac toxicity; Computational phantom; Childhood cancer

Introduction

Childhood cancer survivors have long life expectancies and are at risk for developing treatment-related late effects in the decades after diagnosis [1]. Cardiac disease is among the most common cause of severe or disabling chronic conditions and non-cancer death in long-term survivors of childhood cancer [2,3]. Despite advances in the treatment of childhood cancer over the past several decades, the overall rate of late cardiovascular disease has not declined during this time for children diagnosed with cancer between the 1970s and 1990s [1,4]. One recent study did find that childhood cancer survivors treated in the 1990s had a lower risk of developing late coronary artery disease compared to those treated in the 1980s; this decreased risk was associated with lower heart doses received from radiation therapy (RT) in the later decade [4]. More broadly, RT is a known risk factor for many different types of late cardiac disease, and several pediatric cohort studies have demonstrated that this risk increases with increasing mean heart dose, increasing heart volume receiving 20 Gy or more, and when more than half of the heart receives 5 Gy [5–8]. In all these studies, dose-response was established based on whole heart dosimetry.

However, cardiac disease is a heterogeneous collection of pathologies; dose to individual cardiac substructures may be more predictive for specific cardiac disease. Two case-control studies of (primarily) adult Hodgkin lymphoma survivors reported significant associations between left ventricle dose and heart failure [9] and between valvular dose and heart valve disease [10]. These emerging data suggest substructure dosimetric constraints are likely predictive of toxicity, so contemporary RT treatment plans should be tailored to include substructure-based dosimetric constraints. However, substructure dose-response models are lacking for pediatric populations.

Such substructure dose-response models could be established for two childhood cancer survivor cohorts that previously reported whole heart dose-response models, namely the

Childhood Cancer Survivor Study (CCSS) [11,12] and St. Jude Lifetime (SJLIFE) cohorts [13–15]. In these studies, dose to the whole heart was estimated by reconstructing individuals' RT fields on our age-scalable computational phantom. In order to develop cardiac substructure dose-response models, we must enhance the heart model in our computational phantom. Thus, the primary objective of this current study was to enhance the heart model in our computational phantom by adding substructures. Because of the complexity of the heart, it was also necessary to refine our heart model to be more anatomically realistic and demonstrate that the refined heart model is representative across the age range of the pediatric cohorts. Once the enhanced heart model was developed, we aimed to assess dosimetric accuracy of this model and examine the dosimetric impact of using this model compared to our prior heart model.

Methods and materials

MD Anderson computational phantom

Over the past three decades, the MD Anderson Late-Effects Group has performed RT dose reconstructions for more than 100 late-effects studies in childhood cancer survivors [16,17]. For such studies, we reconstructed survivors' historic RT fields on our in-house computational phantom scaled to each individual's age at RT. Our computational phantom was modeled in FORTRAN and built by bounding the body regions (head, neck, trunk, arms, and legs) of a gender-neutral adult skeleton by cuboids [16–18]. We use the term "generic" phantom to refer to our phantom in its unscaled form, age 18 years.

Age-scaled phantoms are generated for each individual in a study. Scaling is based on age because height and weight are rarely available in historical RT records. Separate scaling functions are applied to each of the body regions to account for the non-uniform growth in the lateral, superior-inferior, and anterior-posterior directions of the different regions [18]. Organs are defined as grids of points within the phantom, and when the phantom is scaled, so too are the organs, according to the scaling functions for the appropriate body region. For example, the heart is scaled using the trunk-scaling functions and modeled as a single structure and defined by a grid of evenly-spaced points.

We recently updated our computational infrastructure and converted our generic phantom from FORTRAN to Digital Imaging and Communications in Medicine (DICOM) format [Appendix A], which can be imported into any commercial treatment planning system (TPS) [18]. We also developed scripts (applied within the TPS) to automatically apply our age scaling functions to our generic phantom to generate phantoms of any age. These enhancements allow us to register our age-scaled phantom to patients' 3D medical images (e.g., computed tomography [CT] images), to other DICOM-formatted phantoms, or other TPS-importable computational phantoms. Thus, we were able to register our updated phantom with the National Cancer Institute (NCI) reference phantoms, described later, to evaluate and enhance the cardiac model.

MD Anderson computational phantom cardiac model (H_{Atlas})

The original heart (H_{Atlas}) in our computational phantom was modeled more than 25 years ago using a cadaver-based cross-sectional anatomy atlas with artist renderings of organ anatomy in transverse views [19]. For each transverse view (interspacing approximately 25 mm), a transparency was overlaid on a 2D grid and the heart was manually traced. Heart points were manually digitized for entry into our phantom. The heart is modeled as a single structure positioned in our phantom's chest between thoracic vertebral bodies 5 and 9. It is defined by 55 evenly spaced points in our generic phantom [Fig. 1a].

NCI reference phantom cardiac model (H_{NCI})

The University of Florida (UF)/NCI developed a set of reference pediatric phantoms in DICOM format, based on high-quality CT image sets selected from thousands of CT images, for ages 0, 1, 5, 10, 15, and 30 years [20–22]. After initial manual segmentation, the phantoms were adjusted to match several international reference data sets including reference body sizes, organ mass, gastro-intestine dimensions, and elemental composition [20–23]. The pediatric phantom series has been adopted by the International Commission on Radiation Protection (ICRP) [24]. These DICOM-formatted phantoms can be used to calculate RT doses to organs throughout the body using any commercial TPS, but they are available only for six ages and cannot be scaled to any arbitrary age, limiting their comprehensive applicability for dose reconstructions for childhood cancer survivors.

The original heart in the UF/NCI humanoid phantom series was defined as the whole heart containing the left and right chambers but was recently updated with detailed substructures (aorta, ventricles, atria, pericardium, left anterior descending artery, left circumflex artery, left main coronary artery, and right coronary artery) segmented from high-resolution diagnostic contrast CT images by clinical radiologists [25]. Hereafter, we refer to these heart models as H_{NCI}. The H_{NCI} models are the most comprehensive cardiac models currently available for computational phantoms.

MD Anderson-NCI hybrid cardiac model (H_{Hybrid})

In this study, we combined the age-scalable capability of our MD Anderson computational phantom with anatomical accuracy of the NCI reference phantom series to develop a new cardiac model with substructures in our computational phantom. We refer to this new cardiac model as the MD Anderson-NCI hybrid cardiac model (H_{Hybrid}). The H_{Hybrid} was developed from the 5-year-old H_{NCI} [Fig. 1.c] because it was the closest age match to the median age (7 years) at diagnosis of the CCSS cohort, a multi-institutional childhood cancer survivor cohort with eight different cancer diagnoses and nearly 25,000 participants [11,12]. The H_{Hybrid} is gender neutral because the H_{NCI} of both sexes were anatomically identical for phantoms with ages less than 15 years old as males and females have similar growth profiles until puberty.

Creation of the H_{Hybrid} model

While the H_{NCI} had most of the cardiac substructures of interest for studies of late RT-related cardiac disease, the heart valves were not delineated and were added by radiation oncologists who specialize in pediatric RT (LC, BH, and JB). They contoured four cardiac

valves (mitral valve, tricuspid valve, pulmonary valve, and aortic valve) within the RayStation TPS; contours were based on the Feng cardiac atlas guidelines [26] and positioned in reference to the positions of the other previously identified H_{NCI} cardiac substructures (none of the H_{NCI} contours were edited during this process). The contours were added to the five-year old NCI phantom.

The workflow to create the H_{Hybrid} model involved three steps, each of which was automated using MATLAB and Python scripts.

Step 1: Import and register phantoms in RayStation TPS—Within the RayStation 9A TPS, we registered the MD Anderson computational phantom scaled to age five years with NCI 5-year-old phantom to bring both phantoms into an identical frame of reference. This step used bony anatomy landmark-assisted rigid registration.

Step 2: Extract contours and perform deformation—After registration, the cardiac substructure contours were extracted from the registered NCI phantom. To account for the volume difference between each age-matched phantom pair (from MD Anderson and from NCI), we performed a uniform deformation of the entire heart (including substructures) with respect to the centroid of the volume in a spherical coordinate system. A uniform deformation script was written in MATLAB (MATLAB A2017, The MathWorks Inc.) [27] to accomplish this task. The deformation was applied only in the radial direction, keeping both azimuthal and polar angles fixed, which retained the post-deformation substructure shape integrity.

Step 3: Convert contours to grid-of-points format and scale to generic phantom geometry—Our in-house generic computational phantom is formatted as a grid of evenly spaced points where each organ is defined by a specific group of points, (e.g., [Fig. 1a and b]). Since this format is required for our RT dose reconstruction program, it was necessary to convert each contour of the H_{Hybrid} model to a grid of points. This conversion was performed via an in-house MATLAB script. We scaled H_{Hybrid} to the generic phantom dimensions using scaling functions described by Gupta et al. [18]. The MATLAB script used contours as input and identified transverse planes within the given contour volume at 1 mm of resolution. The outer boundary was identified on every plane of the structure, and the script filled in the volume with uniformly distributed points with 1-mm resolution in each dimension. The point coordinates were exported to a text file (.txt) to serve as input into our in-house dose calculation system. Once input into our FORTRAN code as a grid of points in our generic scale, the H_{Hybrid} can be used in a similar manner to any other organ in our phantom, i.e. scaled to age at RT.

Quantitative assessment of H_{Hybrid}

In order to carry-out volumetric comparisons and assess the representativeness of our final H_{Hybrid} model, which is a 3D grid of points, across the age range of a pediatric cohort, we had to first convert the grid of points back to DICOM (contours). To accomplish this we used previously developed Python scripts in RayStation TPS [18] to convert our generic phantom with the new H_{Hybrid} from FORTRAN back to the DICOM format. Thereby,

converting H_{Hybrid} from a grid of points to a contour (voxel-based representation of the volume). We then scaled our generic DICOM format phantom with H_{Hybrid} to ages, 1, 5, 10, and 15 years and registered each of these phantoms with a NCI reference phantom of the same age for both sexes, for a total of 8 age-sex combinations. For example, there were two age 1-year matched pairs, i.e., our gender-neutral phantom with H_{Hybrid} scaled to age 1-year matched to the 1-year-old NCI girl phantom and separately to the 1-year-old NCI boy phantom. For each of the eight matched phantom pairs, cardiac volume overlap was assessed with the Dice Similarity Coefficient (DSC) and Overlap Coefficient (OC), calculated as

$$\text{Dice Similarity Coefficent(DSC)} = 2 \times \left[\frac{S_{H_{Hybrid}} \cap S_{H_{NCI}}}{S_{H_{Hybrid}} \cup S_{H_{NCI}}} \right]$$

$$\text{Overlap Coefficient(OC)} = \frac{\left(S_{H_{Hybrid}} \cap S_{H_{NCI}} \right)}{\min\left[S_{H_{Hybrid}}, S_{H_{NCI}} \right]}$$

where $S_{H_{Hybrid}}$ and $S_{H_{NCI}}$ represent the contour volumes of the structures within our computational phantom and the NCI phantom, respectively. DSC and OC analysis were completed for H_{Hybrid}, aorta, ventricles, and atria. For the smaller substructures (<1 cc), there are only few points on each axial plane, thus volume conversion and comparisons were not possible for (right coronary artery, left coronary artery, left circumflex artery, left anterior descending artery). Also, substructures that were not originally contoured in NCI phantoms (aortic valve, mitral valve, pulmonary valve, tricuspid valve) were also excluded from analysis because they were only contoured and added in the 5-Year-old NCI phantom.

Eight DSC and eight OC values were calculated per structure, one for each of the age/sex-matched phantom pairs. From these data, we calculated the mean DSC and OC (and standard deviation) for H_{Hybrid}, atria, ventricles, and aorta. The analysis was repeated for H_{Atlas} for comparison. Then, a Wilcoxon signed-rank test [27,28] was performed to assess if there was a statistically significant difference between the H_{Hybrid} and H_{Atlas} models for the 8 age-sex combinations.

Evaluation of alternate H_{Hybrid} models based on NCI phantoms of different ages

As previously mentioned H_{Hybrid} was developed from the 5-year-old H_{NCI} because it was the closest age match to the median age of the CCSS cohort. To evaluate if another H_{NCI} model from a different aged phantom would have been more representative across the age range of pediatric cohorts, we created alternate heart models using the other aged NCI phantoms (1, 10, and 15 years). Also, for age 15 years, we created separate male and female models. We designated each of these alternate models according to the age (and sex for age 15 years) of the NCI phantom from which it was derived which is summarized in Appendix B.

For each $H_{\text{Alternate}}$ model, we calculated DSC and OC for eight age/sex-matched phantom pairs and the mean DSC value across the age range (1–15 years). A Wilcoxon signed-rank

test was performed to assess if there was a statistically significant difference between the H_{Hybrid} and the $H_{\text{Alternate}}$ models.

Qualitative assessment of H_{Hybrid}

The H_{Hybrid} contours (whole heart and substructures) were reviewed by three collaborating radiation oncologists. Their feedback was recorded using a survey (Appendix C) with a unipolar Likert scale. For this assessment, the evaluation team scored the model from 1 to 5, with a score of 1 indicating an unacceptable model, 3 indicating a moderately satisfactory model, and 5 indicating an optimum model. A mean response score of ≥3.0 without an individual score of "1" was set as the passing criterion.

Dosimetric assessment of H_{Hybrid} – comparison of ground truth and reconstructed doses

To examine the accuracy of heart doses calculated using our new heart model, H_{Hybrid} versus "ground truth" heart doses, we carried out a simulated comparison. Specifically, we compared patients' actual heart doses calculated using CT-based planning in a clinically-commissioned commercial TPS to heart doses calculated using our in-house dose reconstruction method and our age-scaled computational phantom with the H_{Hybrid}. For this comparison, we considered spine RT treatments because (i) this field type is commonly used for central nervous system (CNS) tumors ($N = 1392$), the diagnosis that makes up the largest fraction of individuals in the CCSS expanded cohort treated with RT ($N = 4957$) [11,12] (ii) this is a field type where the heart is partially in-beam (within a steep dose gradient), i.e., the type of scenario for which dose reconstruction uncertainty is the highest. We then identified six patients from the CCSS (three males and three females) whose ages spanned the range of individuals treated with spine fields (toddler to adolescent); each was matched with a patient of the same sex and age (to nearest 0.1 year) from a contemporary cohort of 138 craniospinal patients for whom CT scans were available (data collected under IRB approved protocol). The following procedure was performed in our TPS (RayStation, Version 9B, Stockholm, Sweden) for each matched pair: (i) the whole heart was contoured by a pediatric radiation oncologist (ACP) following guidelines from Feng et al. [26] (ii) the CCSS patient's treated spine field was simulated and calculated, (iii) the mean dose, (D_m) and percentage of volume with a dose greater than or equal to 5 Gy (V_5) and 20 Gy (V_{20}) for the contoured heart were recorded and compared to values from the H_{Hybrid} dose reconstruction. We specifically compared D_m, V_5, and V_{20} because those are metrics that were reported in our recent CCSS study of RT-related heart disease [5].

Dosimetry comparison of H_{Hybrid} and H_{Atlas}

To quantify the dosimetric impact of using H_{Hybrid} as opposed to H_{Atlas}, we performed a comparative analysis among individuals in the CCSS expanded cohort treated with RT ($N = 4957$). For each individual, we reconstructed their previously abstracted RT on our computational phantom with H_{Hybrid}, scaled to their age-at-RT, and calculated D_m, V_5, and V_{20} for comparison with the previous H_{Atlas} values [5]. To better understand the driving elements of dose differences between H_{Hybrid} and H_{Atlas}, we stratified D_m data by the primary cancer diagnosis (i.e., leukemia, central nervous system (CNS) tumor, neuroblastoma, Wilms tumor, Hodgkin lymphoma, non-Hodgkin lymphoma, soft tissue sarcoma, osteosarcoma).

The methods used to perform the H_{Hybrid} dose calculations for the CCSS expanded cohort treated with RT are briefly described here, and additional details are reported in Howell et al. (2019) [16]. We previously completed a detailed abstraction (under IRB approved protocol) of the RT records for prescription(s), date(s), and treatment field parameters, including dose, orientation, energy, field size, weighting, anatomical borders, and blocking details. In this study, the abstracted data for each individual was accessed by our in-house dose calculation program, which placed the treatment fields on our computational phantom, scaled to age at RT. Once the treatment fields were reconstructed on the phantom, we applied heart-specific blocking for chest fields and typical blocking for other field types. Once blocking was applied, points within the phantom were designated as in-beam, blocked, or out-of-beam. Of note, our dose calculation models include both historic and contemporary RT beam energies, including orthovolt-age, Cobalt-60, and a wide range of megavoltage photon and electron energies (all used in the CCSS) [17]. For points in-beam, doses were computed based on beam energy, field size, and depth from the surface and reduced accordingly if under a block. For points out-of-beam, doses were calculated with analytical models that account for both scatter and leakage radiation and based on detailed out-of-field dose measurements in a 3D water phantom for different beam energies, field sizes, depths, and distances from the field edge. The dose was calculated for each of the 1425 points in the H_{Hybrid} model. Then D_m, V_5, and V_{20} were computed. This same calculation methodology was used for our prior study to calculate D_m, V_5, and V_{20} for H_{Atlas}.

Results

In this study, we developed an anatomically realistic cardiac model, H_{Hybrid}, within our in-house generic computational phantom [Fig. 1b]. This heart model has 14 unique substructures, which are illustrated in [Fig. 2] in coronal views. Because it is difficult to distinguish the individual substructures in the 2D images, we have provided videos [Appendix D] of the 3D grids of points that compose the H_{Hybrid} substructures rotated to different orientations to better illustrate them. The new H_{Hybrid} [Fig. 1b] has substantially increased resolution compared to our previous H_{Atlas} model [Fig. 1a], with 1425 points and 55 points, respectively. The number of points in each of the 14 substructures (aorta, anterior descending artery, left circumflex artery, left coronary artery, right coronary artery, left atrium, right atrium, pericardium, aortic valve, mitral valve, pulmonary valve, tricuspid valve, left ventricle, and right ventricle) is listed in [Appendix E]. Like our previous heart model, H_{Hybrid} can be scaled to any age at RT (The age-based scaling is illustrated for H_{Hybrid} and H_{Atlas} (for comparison) in [Appendix F]), with the heart models shown as 3D volume renderings generated in RayStation TPS and scaled to ages 1, 5, 10, and 15 years.

When our computational phantom with H_{Hybrid} was scaled to ages, 1, 5, 10, and 15 years and registered with the NCI reference phantom of the same age for both sexes, there was good agreement between the heart volumes; the H_{Hybrid} DSC for each matched pair was above 0.77 [Fig. 3]. Videos of registered phantom pairs were made for all age-matched pairs (ages 1 and 10 [Appendix G], ages 5 and 15 not shown) to illustrate agreement and were used by clinicians to review the consistency of age-scaled H_{Hybrid} and substructures. The average DSC and standard deviations across the age range for H_{Hybrid} and H_{Atlas} were 0.84 ± 0.05 and 0.47 ± 0.05, respectively [Fig. 3]. The H_{Hybrid} (whole heart) model was

significantly superior (Wilcoxon signed-rank test, $p = 0.0117$) to the H_{Atlas} model in terms of anatomical overlap/similarity to the H_{NCI} models across all ages.

The DSC and OC values of substructures (aorta, atria, and ventricles) are listed in [Appendix H]. The DSC values for the left and right ventricles are 0.68 ± 0.09 and 0.79 ± 0.05 are similar to those observed for the whole heart. Also, the OC values are ≥0.83. The DSC values for the left atrium, right atrium, and aorta are 0.54 ± 0.17, 0.53 ± 0.14, and 0.44 ± 0.17, respectively; the overlap coefficients for these structures are all ≥0.64.

Average DSC and OC values for all new cardiac models (H_{Hybrid} and $H_{\text{Alternate}}$) are summarized in [Appendix I]. The mean DSC values (and standard deviations) for H_{Atlas}, H_{Hybrid}, and each of the $H_{\text{Alternate}}$ models are shown in [Appendix J]. Similar to H_{Hybrid}, all of the $H_{\text{Alternate}}$ models were significantly superior (Wilcoxon signed-rank test, $p = 0.0117$) to the H_{Atlas} model in terms of anatomical overlap when compared with H_{NCI} models across all ages. Differences in DSC values between H_{Hybrid} and the $H_{\text{Alternate, (1, 10, 15F, 15M)}}$ models were not significant (Wilcoxon signed-rank test, $p = 0.674$, 0.779, 0.208, 0.401 respectively).

For the qualitative analysis, model H_{Hybrid} was evaluated by our radiation oncologist collaborators (JB, LSC, and BH). It was unanimously approved with average survey scores of 5.0.

Results from our dosimetric assessment of $H_{Hybrid,}$ in which we compared patients' actual heart doses calculated using CT-based planning to reconstructed H_{Hybrid} dose and dose–volume metrics are reported in Table 1. In all six comparison cases, the absolute difference in D_m was less than 1.0 Gy and the absolute volume differences in V_5 and V_{20} were less than 10.5% (Table 1). Results from our dosimetry comparison of H_{Hybrid} versus H_{Atlas} to assess the impact of using our enhanced heart model for cohort dosimetry and, in particular, the impact on previously reported data [5] are shown in Table 2. For the CCSS expanded cohort, the mean, median, standard deviation (SD), and ranges for D_m, V_5, and V_{20} for H_{Atlas} and H_{Hybrid} are compared. The D_m data further stratified by the primary cancer diagnosis are reported in Table 3. The mean, median, and maximum values for D_m, V_5, and V_{20} for H_{Hybrid} were lower than for H_{Atlas}. Among the eight primary cancer diagnoses, D_m values were lower for H_{Hybrid} than for $H_{\text{Atlas.}}$ The largest differences between H_{Hybrid} and H_{Atlas} were observed for patients in the Hodgkin lymphoma and CNS tumor groups.

Discussion

In this study, we successfully developed an anatomically realistic age-scalable cardiac model (H_{Hybrid}) with fourteen substructures that include all valves, arteries, ventricles, atria, and the aorta. H_{Hybrid} was achieved by combining heart models from the NCI humanoid phantom series with the age-scalable capability of the MD Anderson computational phantom.

We consider this study to be innovative and of high impact because this newly-developed cardiac model will enable us to reconstruct RT doses for substructures in the heart for large retrospective cohorts. Emerging data indicate that substructure doses are predictive of

toxicity, but substructure dose-response models are lacking for pediatric populations. With our enhanced heart model, we now have the ability to perform substructure level dosimetry for cohorts for which we have previously calculated whole heart doses, namely the CCSS and SJLIFE [5,8]. Such data will give us the ability to tailor contemporary RT treatment plans to include cardiac substructure dosimetric constraints.

One major strength of this work is the use of a reference phantom series that was developed from high-quality CT image sets. Our new heart model (H_{Hybrid}) is based on the whole heart model in the UF/NCI humanoid phantom series [21], which has been adopted by the ICRP [24]. Furthermore, the substructures in our new heart model are based on the most comprehensive cardiac substructure models (H_{NCI}) currently available for computational phantoms, developed from high-resolution diagnostic contrast CT images [25].

A limitation of this study is that we created a single heart model (based on a 5-year-old phantom) as opposed to developing multiple models. While the phantom and the heart for each individual in a cohort will be scaled to the age at RT, we still will be scaling only a single heart model. To address this potential limitation, we examined whether our single heart model was representative across the age range of a pediatric cohort (age 1–15 years). We found that when our phantom with H_{Hybrid} was scaled to different ages and registered with international reference phantoms, our single age-scaled heart model showed a high level of overlap between the heart volumes, i.e., mean DSC >0.8. We note that the DSC values reported here for these comparisons were comparable to DSC values reported in contouring studies where organs for the same patients are contoured by different individuals, i.e. interobserver variability [29]. Additionally, we compared our H_{Hybrid} model (based on a 5-year old phantom) with four alternate heart models to see if another model based on other aged phantoms was more representative across the age range of pediatric cohorts. We found no statistically significant difference between models.

In our examination of whether a single set of cardiac substructures is representative across the age range of a pediatric cohort, DSC and OC data for the ventricles were similar to those observed for H_{Hybrid}, but lower for the atria and aorta. Given that the atria and aorta are smaller structures and that the phantoms were rigidly registered with a single registration point (as opposed to using deformable registration and registering to the actual structure) this lower agreement is not surprising. Despite the lower DSC agreement for the atria and aorta, the OC overlap values were above 0.6, indicating that the mean doses to the atria and the aorta should be reasonably acceptable for use in cohort studies. However, for these substructures and for all small substructures, the more refined dose–volume metrics such as V_5 and V_{20}, may not be appropriate.

Another strength of this study is that we examined the accuracy of heart doses calculated with our new heart model, H_{Hybrid}, versus a ground-truth, calculations in a commercial TPS. The average percent difference between predicted and actual heart D_m in our comparison was 4.2%, similar to the 6.7% value reported in a recent study by Gasic et al. [30], which used sophisticated methods to reconstruct pediatric spinal irradiation for 21 pediatric patients. We also consider the results of our comparison of ground truth for D_m, V_5, and V_{20} versus reconstructed values for H_{Hybrid} in the context of dose and volume classification

analysis categories in our prior CCSS study of RT-related heart disease [5]. For the published analyses, individuals' risks were stratified into four mean heart dose categories: (i) 0.1 to <10 Gy, (ii) 10 to <20 Gy, (iii) 20 to <30 Gy, and (iv) ≥30 Gy. In our present comparison, the maximum difference for D_m among all matched pairs was 1.0 Gy (12 Gy for TPS calcuation and 11 Gy for reconstruction); this patient would have been correctly included in the 10 to <20 Gy risk group using either the reconstructed dose or the ground truth dose. Furthermore, none of the differences between reconstructed H_{Hybrid} D_m and ground truth D_m values would have resulted in a misclassification of a patient into a higher or lower dose classification category. Also, in that same CCSS study, individuals' risks were stratified into three V_{20} volume categories: (i) 0.1–29.9 %, (ii) 30–79.9%, and (iii) ≥80%. In our present comparison, the maximum difference for V_{20} among all matched pairs was 6.5% (52.5% for TPS calculation and 59.0% for reconstruction); this patient would have been correctly included in the 30–79.9% risk group. Furthermore, none of the differences between reconstructed H_{Hybrid} V_{20} and ground truth V_{20} values would have resulted in a misclassification of a patient into a higher or lower volume classification category. Lastly, individuals' risks were stratified into two V_5 volume categories: (i) 0.1–50% and (ii) ≥50%. In our present comparison, the maximum difference for V_5 among all matched pairs was 10.5% (69.5% for TPS calculation and 59.0% for reconstruction); this patient would have been correctly included in the ≥50% risk group. Furthermore, none of the differences between reconstructed H_{Hybrid} V_5 and ground truth V_5 values would have resulted in a misclassification of a patient into a higher or lower volume classification category. In summary, for a field type where dose reconstruction uncertainty is high (heart partially in-beam) as opposed to low uncertainty field types (heart entirely in-beam or out-of-beam), the dosimetric uncertainty in reconstructed H_{Hybrid} D_m, V_5, and V_{20} did not alter the classification of the values into the correct dose and volume categories used in prior dose-response models for RT-related heart disease in the CCSS.

Our examination of the dosimetric impact of using H_{Hybrid} as opposed to H_{Atlas} for the CCSS expanded cohort found that dose and dose–volume metrics were lower for H_{Hybrid} as opposed to H_{Atlas}. From the broad perspective of cohort dosimetry and specifically to the CCSS and the types of RT fields used in their treatment, we attribute these differences to several factors. For spine fields used to treat CNS and for mantle fields used to treat Hodgkin lymphoma, a higher percentage of H_{Hybrid} volume is out of the RT beam as opposed to the more compact H_{Atlas}, resulting in lower D_m for H_{Hybrid}. Examples illustrating these differences in the percentage of volume in-beam versus out-of-beam for H_{Hybrid} and H_{Atlas} are shown in Fig. 4a and b. The combination of these field differences and these two disease groups making up nearly 50% of all individuals in the cohort are the key driving factors for the differences observed in H_{Hybrid} and H_{Atlas}. The remaining six diagnoses, which together make just over 50% of the patients, were primarily treated with RT beams where the heart was entirely in-beam (e.g., whole lung, Fig. 4c) or entirely out-of-beam (e.g., whole brain, Fig. 4d), resulting in similar D_m values for H_{Hybrid} and H_{Atlas}.

Our dosimetric findings that dose and dose–volume metrics were somewhat lower for H_{Hybrid} as opposed to H_{Atlas} may affect our previously reported dose–response models of RT-related heart disease in the CCSS. The data are suggestive that the slopes of dose-response models will be somewhat steeper, i.e., the cardiac outcomes occurred at slightly

lower doses when the analysis is repeated with H_{Hybrid} as opposed to H_{Atlas}. We intend to update our previously reported dose-response models, including updated H_{Hybrid} dosimetry for the overall CCSS cohort (>13,000 patients treated with RT (diagnosed 1970–1999).

Lastly, we think it is important to note that existing retrospective cohort studies of RT-related late effects involve decades of longitudinal follow-up of cancer survivors treated with conventional RT without CT-based planning. For such studies, the heart anatomy of individuals in those cohorts are not available and computational phantoms are necessary to estimate heart doses. With our enhanced heart model, we now have the ability to estimate cardiac substructure dose and establish substructure level dose-response models for late cardiac disease in childhood cancer survivor cohorts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Radiation Dosimetry Services staff at MD Anderson for their support. Also, we acknowledge and thank Ms. Sarah Bronson from Scientific Publications, Research Medical Library, The University of Texas MD Anderson Cancer Center, for assistance in editing this manuscript.

Funding statement

This work was supported by the National Cancer Institute, United States (CA55727, G.T. Armstrong, Principal Investigator). Support to St. Jude Children's Research Hospital also was provided by the Cancer Center Support (CORE) grant (CA21765, C. Roberts, Principal Investigator) and the American Lebanese Syrian Associated Charities (ALSAC). The MD Anderson Late Effects Group (PI: Dr. Rebecca Howell) receives funding through a subcontract with Childhood Survivor Cancer Study, RFA-CA-15-502, NIH/NCI, PI - Gregory T. Armstrong. Additionally, Suman Shrestha is supported through CCSS Career Development Award, Marilyn and Frederick R. Lummis, Jr. M.D., Fellowship in the Biomedical Sciences and Sylvan Rodriguez Foundation Scholarship honoring George M. Stancel, Ph.D. This study was supported in part by Cancer Center Core (Support) Grant CA016672 from the National Cancer Institute, National Institutes of Health, United States to The University of Texas MD Anderson Cancer Center, United States.

References

[1]. Gibson TM, Mostoufi-Moab S, Stratton KL, Leisenring WM, Barnea D, Chow EJ, et al. Temporal patterns in the risk of chronic health conditions in survivors of childhood cancer diagnosed 1970–99: a report from the Childhood Cancer Survivor Study cohort. Lancet Oncol 2018;19:1590–601. [PubMed: 30416076]

[2]. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, et al. Chronic health conditions in adult survivors of childhood cancer. N Engl J Med 2006;355:1572–82. [PubMed: 17035650]

[3]. Reulen RC, Winter DL, Frobisher C, Lancashire ER, Stiller CA, Jenney ME, et al. Long-term cause-specific mortality among survivors of childhood cancer. JAMA 2010;304:172–9. [PubMed: 20628130]

[4]. Mulrooney DA, Hyun G, Ness KK, Ehrhardt MJ, Yasui Y, Duprez D, et al. Major cardiac events for adult survivors of childhood cancer diagnosed between 1970 and 1999: report from the Childhood Cancer Survivor Study cohort. BMJ 2020;368:l6794. [PubMed: 31941657]

[5]. Bates JE, Howell RM, Liu Q, Yasui Y, Mulrooney DA, Dhakal S, et al. Therapy-related cardiac risk in childhood cancer survivors: an analysis of the childhood cancer survivor study. J Clin Oncol 2019;37:1090–101. [PubMed: 30860946]

[6]. Haddy N, Diallo S, El-Fayech C, Schwartz B, Pein F, Hawkins M, et al. Cardiac diseases following childhood cancer treatment: cohort study. Circulation 2016;133:31–8. [PubMed: 26487757]

[7]. Mulrooney DA, Armstrong GT, Huang S, Ness KK, Ehrhardt MJ, Joshi VM, et al. Cardiac outcomes in adult survivors of childhood cancer exposed to cardiotoxic therapy: a cross-sectional study. Ann Intern Med 2016;164:93–101. [PubMed: 26747086]

[8]. Mulrooney DA, Yeazel MW, Kawashima T, Mertens AC, Mitby P, Stovall M, et al. Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. BMJ 2009;339:b4606. [PubMed: 19996459]

[9]. van Nimwegen FA, Ntentas G, Darby SC, Schaapveld M, Hauptmann M, Lugtenburg PJ, et al. Risk of heart failure in survivors of Hodgkin lymphoma: effects of cardiac exposure to radiation and anthracyclines. Blood 2017;129:2257–65. [PubMed: 28143884]

[10]. Cutter DJ, Schaapveld M, Darby SC, Hauptmann M, van Nimwegen FA, Krol AD, et al. Risk of valvular heart disease after treatment for Hodgkin lymphoma. J Natl Cancer Inst 2015;107.

[11]. Leisenring WM, Mertens AC, Armstrong GT, Stovall MA, Neglia JP, Lanctot JQ, et al. Pediatric cancer survivorship research: experience of the Childhood Cancer Survivor Study. J Clin Oncol 2009;27:2319–27. [PubMed: 19364957]

[12]. Robison LL, Armstrong GT, Boice JD, Chow EJ, Davies SM, Donaldson SS, et al. The Childhood Cancer Survivor Study: a National Cancer Institute-supported resource for outcome and intervention research. J Clin Oncol 2009;27:2308–18. [PubMed: 19364948]

[13]. Hudson MM, Ehrhardt MJ, Bhakta N, Baassiri M, Eissa H, Chemaitilly W, et al. Approach for classification and severity grading of long-term and late-onset health events among childhood cancer survivors in the St. Jude Lifetime Cohort. Cancer Epidemiol Biomarkers Prev 2017;26:666–74. [PubMed: 28035022]

[14]. Hudson MM, Ness KK, Nolan VG, Armstrong GT, Green DM, Morris EB, et al. Prospective medical assessment of adults surviving childhood cancer: study design, cohort characteristics, and feasibility of the St. Jude Lifetime Cohort study. Pediatr Blood Cancer 2011;56:825–36. [PubMed: 21370418]

[15]. Ojha RP, Oancea SC, Ness KK, Lanctot JQ, Srivastava DK, Robison LL, et al. Assessment of potential bias from non-participation in a dynamic clinical cohort of long-term childhood cancer survivors: results from the St. Jude Lifetime Cohort Study. Pediatr Blood Cancer 2013;60:856–64. [PubMed: 23024097]

[16]. Howell RM, Smith SA, Weathers RE, Kry SF, Stovall M. Adaptations to a generalized radiation dose reconstruction methodology for use in epidemiologic studies: an update from the MD Anderson Late Effect Group. Radiat Res 2019;192:169–88. [PubMed: 31211642]

[17]. Stovall M, Weathers R, Kasper C, Smith SA, Travis L, Ron E, et al. Dose reconstruction for therapeutic and diagnostic radiation exposures: use in epidemiological studies. Radiat Res 2006;166:141–57. [PubMed: 16808603]

[18]. Gupta A, Shrestha S, Owens C, Smith S, Qiao Y, Weathers R, et al. Development of an age-scalable 3D computational phantom in DICOM standard for late effects studies of childhood cancer survivors. Biomed Phys Eng Express 2020.

[19]. Eycleshymer AC, Shoemaker DM. A cross-section anatomy. New York 1970.

[20]. Geyer AM, O'Reilly S, Lee C, Long DJ, Bolch WE. The UF/NCI family of hybrid computational phantoms representing the current US population of male and female children, adolescents, and adults—application to CT dosimetry. Phys Med Biol 2014;59:5225–42. [PubMed: 25144322]

[21]. Griffin KT, Mille MM, Pelletier C, Gopalakrishnan M, Jung JW, Lee C, et al. Conversion of computational human phantoms into DICOM-RT for normal tissue dose assessment in radiotherapy patients. Phys Med Biol 2019;64:13NT02.

[22]. Lee C, Lodwick D, Hurtado J, Pafundi D, Williams JL, Bolch WE. The UF family of reference hybrid phantoms for computational radiation dosimetry. Phys Med Biol. 2010;55:339–63. [PubMed: 20019401]

[23]. Human alimentary tract model for radiological protection. ICRP Publication 100. A report of The International Commission on Radiological Protection. Ann ICRP. 2006;36:25–327, iii. [PubMed: 17188183]

[24]. ICRP. Pediatric Reference Computational Phantoms. Annals of the ICRP. 2020 At Press.

[25]. Borrego D, Apostoaei AI, Thomas BA, Hoffman FO, Simon SL, Zablotska LB, et al. Organ-specific dose coefficients derived from Monte Carlo simulations for historical (1930s to 1960s) fluoroscopic and radiographic examinations of tuberculosis patients. J Radiol Prot 2019;39:950–65. [PubMed: 31269474]

[26]. Feng M, Moran JM, Koelling T, Chughtai A, Chan JL, Freedman L, et al. Development and validation of a heart atlas to study cardiac exposure to radiation following treatment for breast cancer. Int J Radiat Oncol Biol Phys 2011;79:10–8. [PubMed: 20421148]

[27]. MATLAB (R2017b). Natick, Massachusetts: The MathWorks Inc. 2017.

[28]. Wilcoxon Signed Rank Test Calculator, (2020, 6 6), Social Science Statistics.

[29]. Duane F, Aznar MC, Bartlett F, Cutter DJ, Darby SC, Jagsi R, et al. A cardiac contouring atlas for radiotherapy. Radiother Oncol 2017;122:416–22. [PubMed: 28233564]

[30]. Gasic D, Rosenschöld PMA, Vogelius IR, Maraldo MV, Aznar MC, Nysom K, et al. Retrospective estimation of heart and lung doses in pediatric patients treated with spinal irradiation. Radiother Oncol 2018;128:209–13. [PubMed: 29859753]

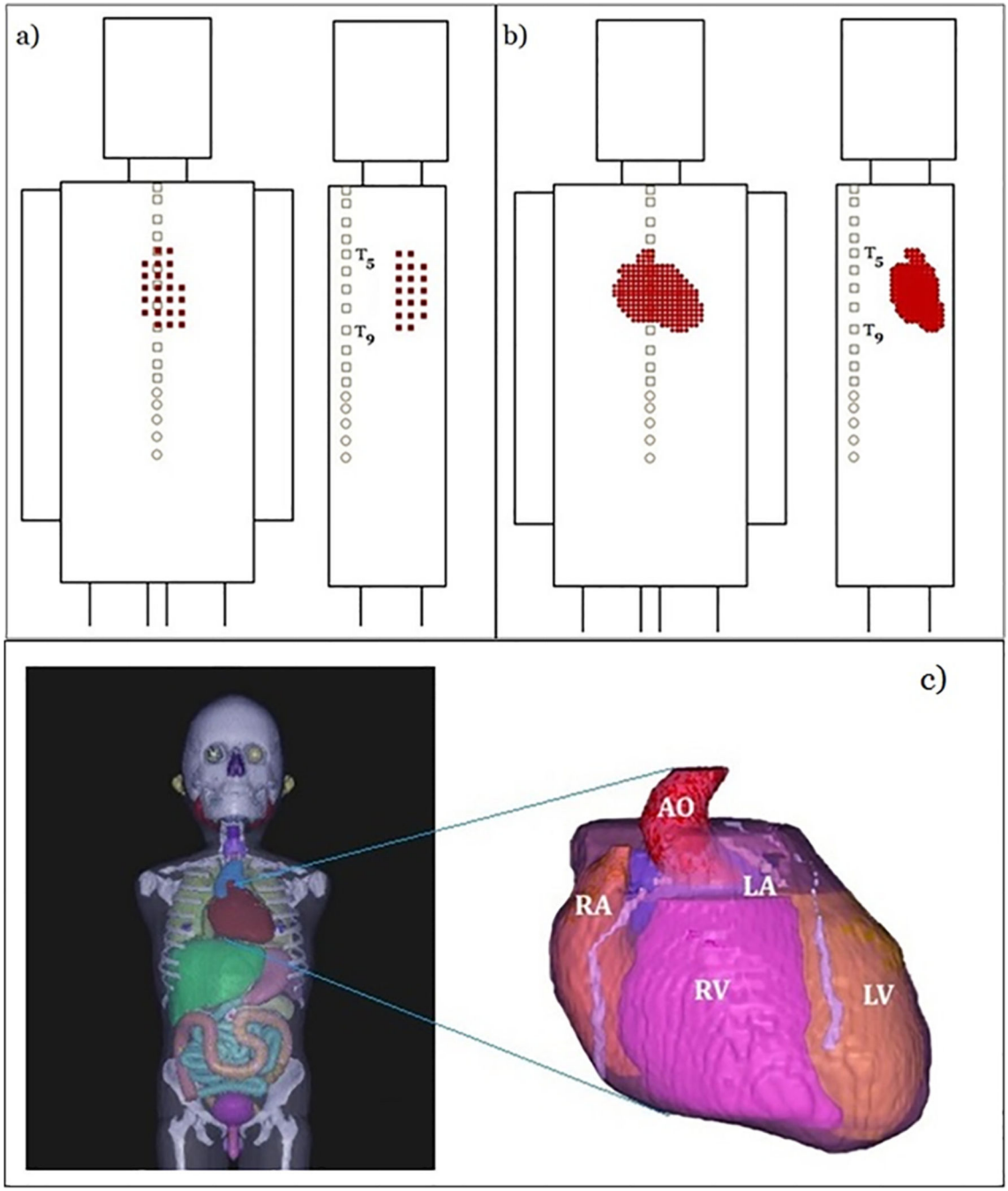

Fig. 1.

Coronal and sagittal views of generic in-house computational phantom showing (a) H_{Atlas} (b) H_{Hybrid} model in FORTRAN format. T5 and T9 mark the 5th and 9th thoracic vertebrae, respectively. (c) A 3D rendering of the NCI reference phantom showing H_{NCI} contours (with substructures) for the 5-year old phantom. LA: Atrium (left), RA: Atrium (right), LV: Ventricle (left), RV: Ventricle (right), AO: Aorta. Phantom images (a–c) were cropped at the top of thighs.

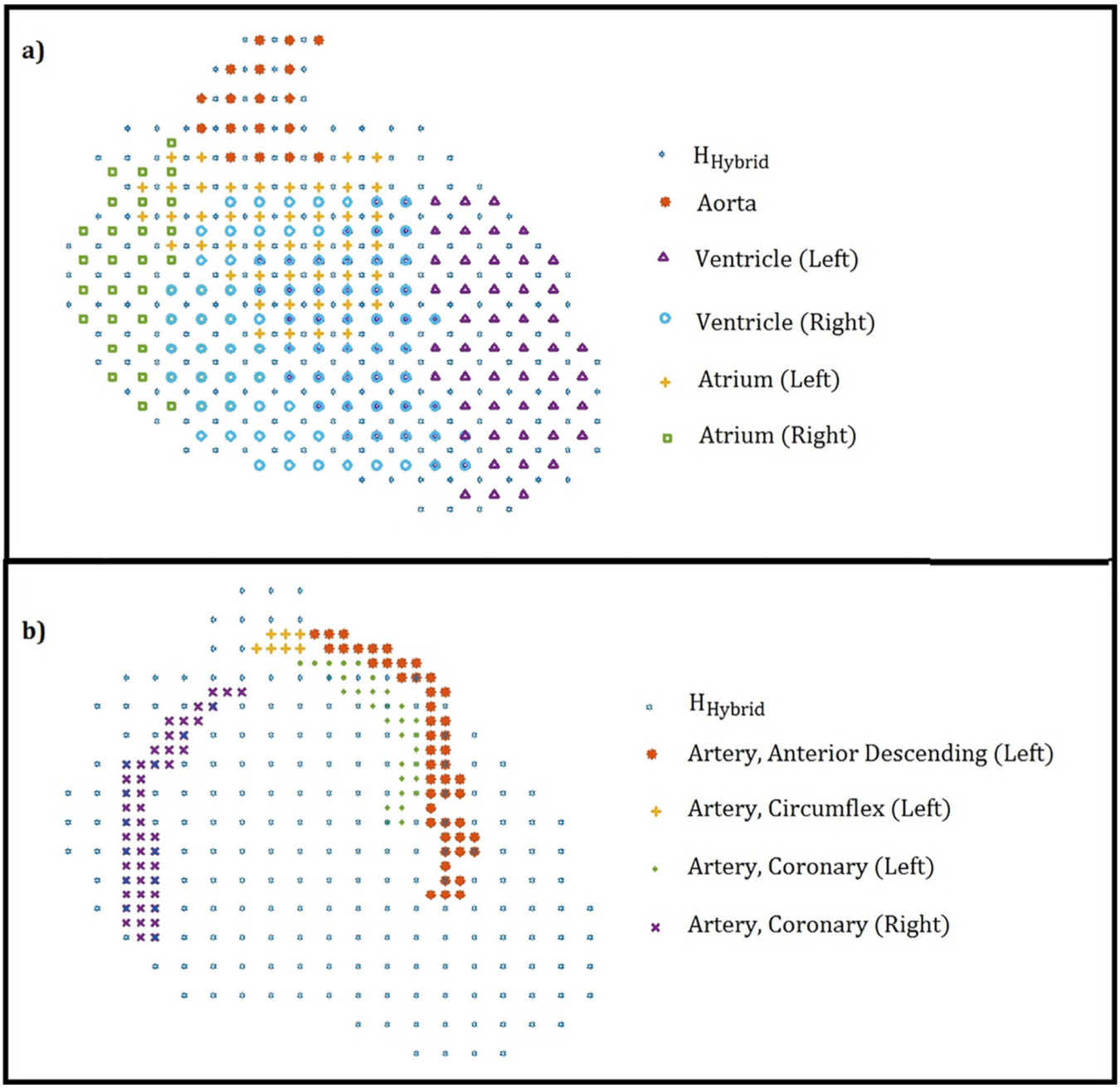

Fig. 2.

Coronal views of H_{Hybrid} model, showing substructures in FORTRAN format (a) H_{Hybrid}, aorta, and left and right atria and ventricles and (b) H_{Hybrid} and arteries. Three-dimensional illustrations of these substructures are presented in Appendix D.

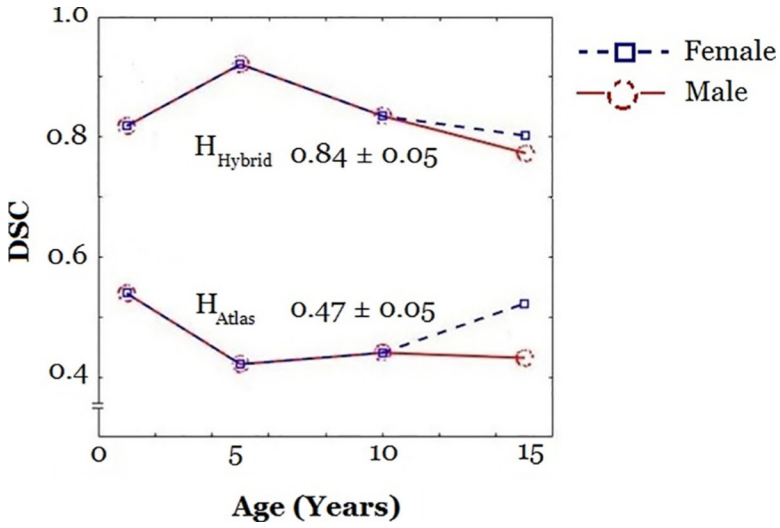

Fig. 3.

Dice similarity coefficients (DSC) for the H_{Atlas} and H_{Hybrid} models compared with the age/sex matched H_{NCI} models. The mean DSC and standard deviations are listed for H_{Hybrid} and H_{Atlas} models. Red solid lines (with circle markers) and blue dotted lines (with square markers) correspond to male and female DSC values.

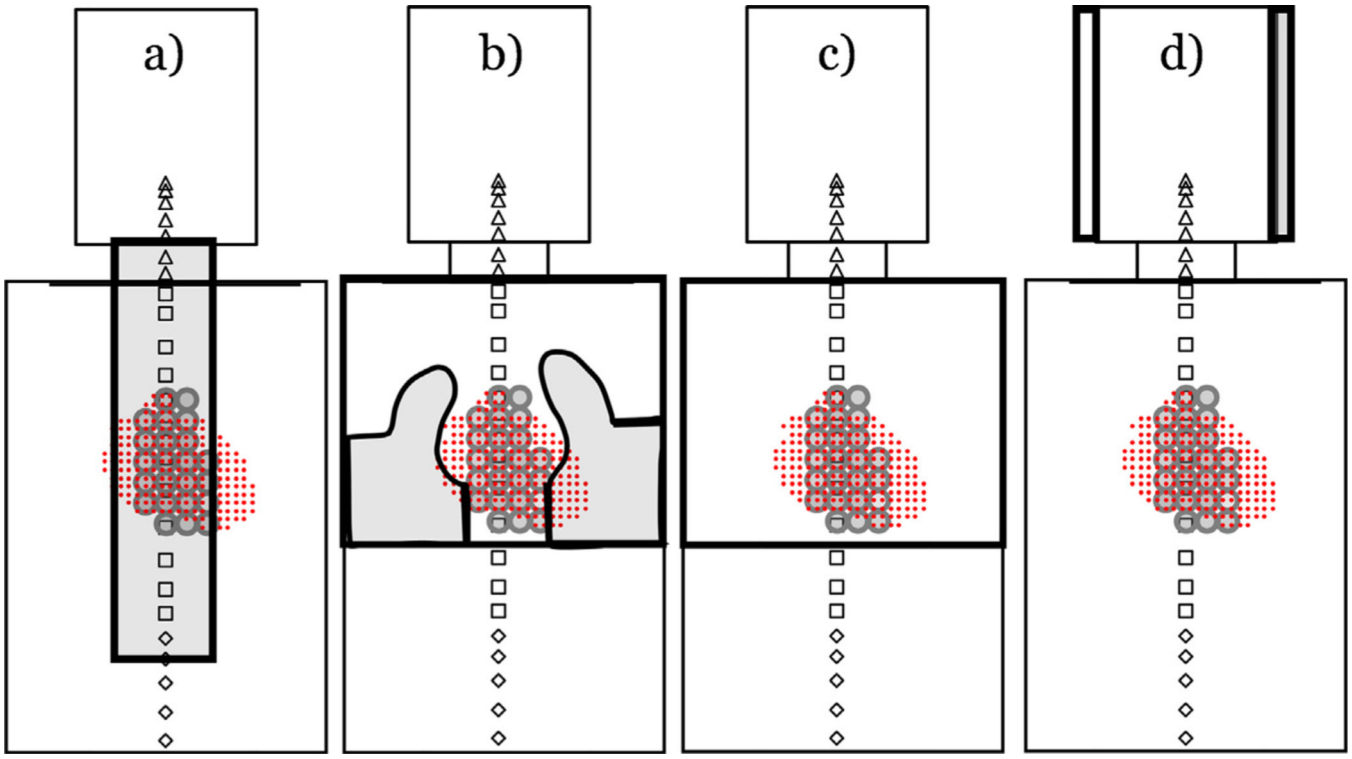

Fig. 4.

Phantom showing H_{Atlas} (grey circles) and H_{Hybrid} (red squares) with examples of common treatment fields in the CCSS (a) PA whole spine, (b) AP/PA mantle, (c) AP/PA whole lung, and (d) R/L lateral whole brain.

Table 1

Comparison of heart dose and dose volume metrics for CT-based calculation in commercial TPS and reconstructed metrics $H_{\text{Hybrid.}}$ Note that "patient" refers to TPS calculation for patient's CT, and :recon" refers to age-matched phantom for which we performed dose reconstruction.

Case	Sex	Age at RT (Yr)	Rx depth (cm)	Rx Dose (Gy)	Heart D_m (Gy)		Abs diff (Gy)	Heart V_5 (%)		Abs diff (%)	Heart V_{20} (%)		Abs diff (%)
					Patient	Recon		Patient	Recon		Patient	Recon	
1	M	2.1	3.0	19.5	9.8	9.9	−0.1	65.6	59.0	6.6	0.0	0.0	0.0
2	F	2.9	3.0	18.0	9.9	9.4	0.5	61.8	59.0	2.8	0.0	0.0	0.0
3	M	7.7	3.0	36.0	15.8	16.6	−0.8	66.2	59.0	7.2	56.5	59.0	−2.5
4	F	7.6	5.0	23.4	11.5	11.8	−0.3	60.0	59.0	1.0	16.0	14.0	2.0
5	M	12.6	4.0	36.0	15.3	15.9	−0.6	62.9	59.0	3.9	52.5	59.0	−6.5
6	F	13.7	6.0	23.4	12.0	11.0	1.0	69.5	59.0	10.5	6.5	3.0	3.5
	Avg	7.8	4.0	26.1	12.4	12.4	0.0	64.3	59.0	5.3	21.9	22.5	−0.6
	SD	4.4	1.2	7.3	2.6	2.8	0.7	3.4	0.0	3.4	23.7	26.2	3.5

Abbreviations: mean dose, (D_m), percentage of volume with dose greater than or equal to 5 Gy (V_5) and 20 Gy (V_{20}), depth of treatment plan normalization (Rx depth), prescribed dose (Rx dose).

Table 2

Comparison of summary statistics for mean dose, (D_m) and percentage of volume with dose greater than or equal to 5 Gy (V_5) and 20 Gy (V_{20}) with H_{Hybrid} and H_{Atlas} for the CCSS expanded cohort (N= 4957).

Dose–volume/dose metric	H_{Hybrid} (Gy)				H_{Atlas} (Gy)			
	Mean	(Median)	SD	(Range)	Mean	(Median)	SD	(Range)
D_m (Gy)	8.18	(4.40)	8.92	(0.0001–41.87)	9.67	(5.39)	10.72	(0.0001–44.96)
V_5 (%)	39	26	41	(0–100)	45	31	46	(0–100)
V_{20} (%)	18	0	29	(0–100)	22	0	38	(0–100)

Table 3

Comparison of summary statistics for mean dose, (D_m) for H_{Hybrid} and H_{Atlas} for the CCSS expanded cohort stratified by the primary cancer diagnosis.

Primary Diagnosis	H_{Hybrid} D_m (Gy)				H_{Atlas} D_m (Gy)			
	Mean	(Median)	SD	(Range)	Mean	(Median)	SD	(Range)
Leukemia (N = 995)	5.08	(0.15)	6.03	(0.0027–25.02)	5.35	(0.15)	6.21	(0.0022–30.04)
Central nervous system tumor (N = 1352)	7.62	(2.55)	7.97	(0.0004–29.69)	10.29	(2.36)	10.86	(0.0005–41.11)
Neuroblastoma (N = 269)	6.86	(4.22)	6.88	(0.01–30.69)	7.28	(5.11)	7.25	(0.01–31.90)
Wilms tumor (N = 510)	5.58	(1.23)	5.99	(0.03–30.97)	5.68	(1.43)	5.98	(0.03–29.84)
Non-Hodgkin lymphoma (N = 196)	5.87	(1.48)	7.21	(0.002–30.65)	6.50	(1.79)	8.02	(0.002–37.11)
Hodgkin lymphoma (N = 984)	16.94	(17.51)	9.74	(0.0069–40.38)	20.05	(0.80)	11.45	(0.0064–44.96)
Osteosarcoma (N = 254)	6.69	(0.42)	8.84	(0.0048–38.31)	7.01	(0.43)	9.47	(0.0048–39.06)
Soft tissue sarcoma (N = 397)	2.44	(0.02)	5.97	(0.0001–41.87)	2.63	(0.20)	6.67	(0.0001–43.38)