Assessing the Variation in 3D Dose Reconstruction based on CT Scans of Pediatric Cancer Patients Matched on Gender and Age

Ziyuan Wang[1], Irma W.E.M. van Dijk[1], Jan Wiersma[1], Cécile M. Ronckers[2], Foppe Oldenburger[1], Coen R.N. Rasch[1], Peter A.N. Bosman[3], Arjan Bel[1] and Tanja Alderliesten[1†]

1. Academic Medical Center (AMC), Amsterdam, NL; 2. Emma Children's Hospital / AMC, Amsterdam, NL; 3. Centrum Wiskunde & Informatica (CWI), Amsterdam, NL

Introduction

Radiation treatment (RT) is widely used in childhood cancer (CC) treatment. The relation between radiation dose and late adverse effects is studied to enable the design of the least toxic, yet effective RT plan. The challenge for 3D dose reconstruction of historically treated patients with long-term follow-up is the absence of 3D CT scans, because at the time of treatment 2D simulator films were used. To solve this problem, phantom-based dose reconstruction methods have been developed. The phantom's anatomical structure is generally based on CT scans of patients of a certain gender/age [1]. However, being a general representative of patients with matched gender/age, the phantom does not represent true individual patient anatomies. A key question is therefore whether anatomical variations across gender/age matched paediatric patients lead to significant differences in dose calculation outcomes. To answer this question, we studied the variation in reconstructed dose among gender/age matched patients when the same RT plan was applied to them.

Materials & Methods

Ten 3-year-old CC patients, five males and five females, were treated from 2009 to 2015, with CT scans of the abdomen available. Patients' ages at start of RT ranged from 3.02–3.98 and 3.01–3.67 years for females and males respectively. Since the majority of the patients (8/10) were treated for Wilms' tumor, we focussed on this site and retrieved the corresponding eight RT plans. The RT plan, dose, and CT of each of these eight patients were used as references. Note, all ten CT scans were used for dose reconstruction. We applied each reference RT plan to the other patients' CT scans in the gender/age matched group, maintaining the RT plan isocenter position relative to the spinal cord in all directions, without changing field size and other beam settings. For dose comparison liver, spleen, left kidney, and spinal cord were delineated by an expert following clinical protocols. For the differences between reconstructed and reference dose, we first compared the relative difference in minimum, maximum, mean, and median organ dose (normalized by reference dose at isocenter). Next, gamma (γ) analysis was used to assess 3D organ dose distribution differences [2]. To compare the reconstructed doses to the reference dose, organs were first deformably registered with Elastix [3]. Two different settings for the distance-to-agreement acceptance and the dose-difference (defined as % of maximum dose) criteria were used (i.e., S1: 5mm, 5%; S2: 10mm, 10%) to calculate the γ-values and fraction of passed γ-values. Voxel values <10% of maximum dose were ignored. The fraction of passed γ-values (i.e., $\gamma < 1$) to an organ indicates the percentage of voxels of the organ passing the two criteria.

Results

For every organ and for every dose metric, the average relative deviation is <30%, but the maximum deviation is much higher, e.g., reaching 150% for the deviation in minimum dose

to the left kidney (Fig. 1). Further, the difference between dose-volume histograms (DVH) of reconstructed doses and the DVH of the reference plan varies substantially over reference plans (Fig. 2). Average pass fractions (PF), average γ-values, and the percentage of comparisons with PF<80%, are given in Table 1. The two settings have a different strictness of acceptance criteria (i.e., S2 less strict than S1), thus the results differ. Although for the left kidney S1 and S2 provide similar average γ-values, for the liver and spleen a difference of a factor ~2 is found. The percentage of plans with PF<80% ranges from 12% to 75%. This indicates that the probability that >20% of the voxels fail the acceptance criteria is not small.

Discussion & Conclusions

The results of the dose reconstruction method based on CT scans of gender/age matched CC patients show good accuracy in terms of average dose metrics. However, the remaining dose comparisons indicate that the method is not yet sufficiently robust and accurate. Thus, only using gender and age as criteria to group patients and randomly selecting one representative without adjusting the field size when applying RT plans does not guarantee sufficient accuracy. We are currently investigating alternatives with potentially better accuracy.

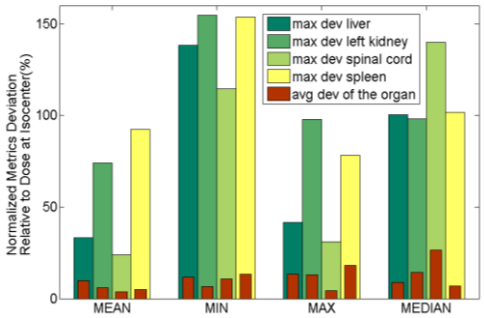

Figure 1: Organ dose metric deviations (dev), maximum and average values.

Organ	*Liver*		*Spleen*		*Left kidney*	
Setting	*S1*	*S2*	*S1*	*S2*	*S1*	*S2*
Avg PF (%)	66	85	73	88	90	93
Avg γ	1.06	0.53	0.74	0.37	0.29	0.24
% PF < 80%	75	12	50	38	25	19

Table 1: Average pass fraction and γ-value of all reconstructed dose distributions under two criteria settings (S1 and S2), and % of plans with PF < 80%.

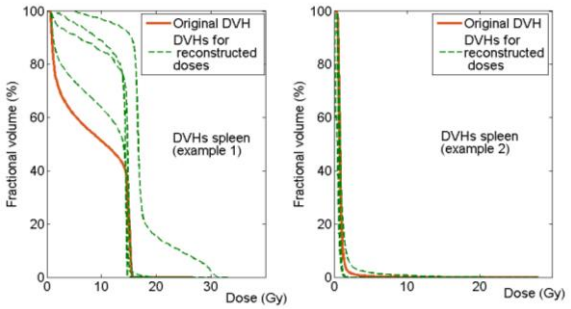

Figure 2 (right): DVHs of reference and reconstructed doses at spleen, for two different plans (denoted as example 1 and example 2, respectively).

References

[1] Lee C *et al.* The UF series of tomographic computational phantoms of pediatric patients. In *Medical Physics* **32**, pages 3537-3548, 2005.

[2] Low *DA and Dempsey JF*. Evaluation of the gamma dose distribution comparison method. In *Medical Physics* **30**, pages 2455-2464, 2003.

[3] Klein S *et al.* elastix: a toolbox for intensity based medical image registration. In *IEEE Transactions on Medical Imaging* **29**(1), pages 196-205, 2010.

Acknowledgements & Disclosures

Financial support was provided by Stichting Kinderen Kankervrij (KiKa; project # 187). Dr. Ronckers was supported by the Dutch Cancer Society (KWF; grant # UVA2012-5517).

Corresponding Author: z.wang@amc.uva.nl