

K. Verhoeven¹, T. Brion², W. R. Green³, M. Balia⁴, A. Webb⁵, T. Rattay⁶, Y. Liang¹, L. G. Assia⁷, I. Hafsa⁷, S. Romdhani⁴, R. Iandolo⁴, A. Bombezin-Domino⁴, B. K. K. Teo⁸, R. McBeth³, I. Koutsopoulos⁹, C. Talbot⁵, N. Paragios^{4,10}, and S. Rivera⁷, on behalf of the PRE-ACT consortium

¹Department of Radiation Oncology, Maastricht University Medical Centre - GROW (MAASTRO), Maastricht, Netherlands, ²Institut Curie, Paris, France, ³Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, ⁴TheraPanacea, Paris, France, ⁵Department of Genetics and Genome Biology, University of Leicester, Leicester, United Kingdom, ⁶Leicester Cancer Research Centre, University of Leicester, Leicester, United Kingdom, ⁷Gustave Roussy, Cancer Campus, Villejuif, France, ⁸Department of Radiation Oncology, Division of Physics, University of Pennsylvania, Philadelphia, PA, ⁹Athens University of Economics and Business, Athens, Greece, ¹⁰Centrale Supélec, University of Paris-Saclay, Computer Science and Applied Mathematics, Gif-sur-Yvette, France

INTRODUCTION

- The PRE ACT project aims to predict the risk of radiotherapy (RT) side effects, including arm lymphedema in breast cancer (BC) patients based on 3 multicenter international studies CANTO, HYPOG-01 and REQUITE.
- Manual contours are not sufficiently homogeneous to allow feature extraction to train predictive models because of variation between investigators, clinical centers, and patient cohorts.
- The aim of this study was to compare automatic and manual contours and to assess the reliability of automatic contouring of organs-at-risk (OARs).

METHOD

- OAR contours were generated with Annotate module from ART-Plan (from TheraPanacea) for six crucial OARs for dose planning of BC loco-regional RT, *i.e.* esophagus, thyroid, heart, ipsilateral lung, ipsilateral humeral head and contralateral breast.
- Four independent international radiation oncologists have been presented blindly with both automatic and manual contours from a randomly selected subset of left and right BC patients from the PRE ACT dataset (7052 patients).
- 10 automatic and manual cases were presented across two sessions, with a washout of a minimum of 48 hours. Rating of the contours followed an A-B-C grid, with A for satisfactory contours, B requiring minor changes, and C needing major changes or redoing. A and B contours were estimated as acceptable, therefore ratings of A and B contours define the acceptance rate for each organ.

RESULTS

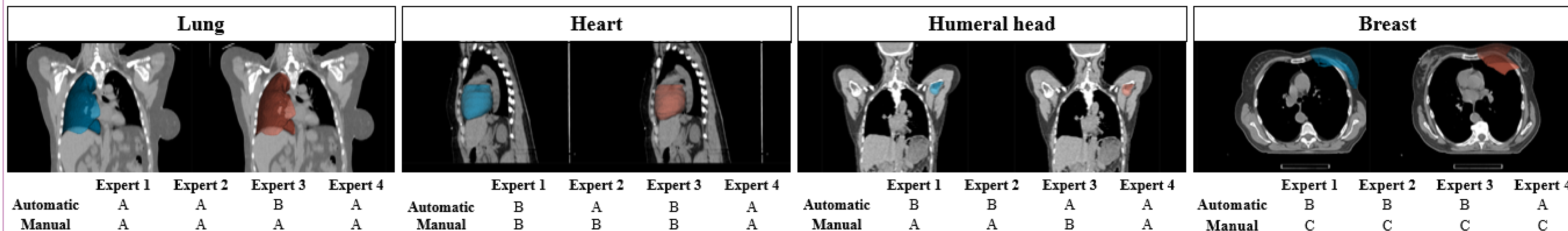


Figure 1. Examples of automatic and manual contours for four OAR : ipsilateral lung, heart, humeral head and contralateral breast. For each example, a table below shows the evaluation rating of each expert. Contours are shown as 3D volumes; automatic contours are shown in blue and the manual contours are shown in red.

- As shown in **Table 1**, we observe that automatic contours acceptance rate ranges from 92 to 100%, whilst manual contours acceptance ranges from 72% to 100%.
- This is illustrated by **Figure 1** with four examples, where we notice that depending on cases, experts will deem automatic contours more acceptable than manual contours, and vice versa.
- For automatic contours, the lowest inter-expert agreement in contour evaluation was obtained for humeral head and heart (45 % for each). The lowest inter-expert agreement for manual contours was thyroid (38.3%)
- The highest inter-expert agreement was for esophagus and left breast (68.3%) for automatic contours and left breast for manual contours (66.7%).

		Esophagus	Thyroid	Heart	Left Breast	Right breast	Left HH	Right HH	Left lung	Right lung
Automatic	A+B	100%	92.5%	97.5%	97.5%	100%	100%	100%	100%	100%
	C	0%	7.5%	2.5%	2.5%	0%	0%	0%	0%	0%
Manual	A+B	100%	92.5%	85.0%	72.5%	77.5%	95.0%	97.5%	97.5%	97.5%
	C	0%	7.5%	15.0%	27.5%	22.5%	5.0%	2.5%	2.5%	2.5%

Table 1. Qualitative evaluation of automatic vs manual contours for 9 OARs. Acceptance ratings are averaged across 4 experts. (HH, humeral head)

CONCLUSIONS

With this study, we demonstrate that automatic contouring is an excellent and reliable tool to obtain segmentation of OARs with a robust and homogeneous clinical acceptability across diverse cohorts, with many applications to clinical trials.

CONTACT INFORMATION

Dr Karolien Verhoeven
karolien.verhoeven@maastro.nl