

Multi-Institutional qualitative evaluation of automatic and manual segmentations of organs at risk on PRE ACT breast cancer cohorts

1/2

Karolien Verhoeven, Thomas Brion, William Green, Maddalena Balia, Adam Webb, Tim Rattay, Sami Romdhani, Riccardo Iandolo, Alexis Bombezin-Domino, Assia Lamrani-Ghaouti, Ines Hafsa, Christopher Talbot, Iordanis Koutsopoulos, Kevin Teo, Rafe Mcbeth, Nikos Paragios, Sofia Rivera.

-UniCancer: Sofia Rivera, Assia Lamrani-Ghaouti, Ines Hafsa-

-Maastricht: Karolien Verhoeven, Yuqin Liang

-University of Pennsylvania: William Green, Kevin Teo, Rafe Mcbeth

-University of Leicester: Christopher Talbot, Adam Webb, Tim Rattay

-TheraPanacea: Sami Romdhani, Riccardo Iandolo, Alexis Bombezin-Domino, Maddalena Balia, Nikos Paragios

-Institut Curie: Thomas Brion

-AUEB: Iordanis Koutsopoulos

Purpose/Objective:

PRE ACT is an EU funded project that aims to predict the risk of radiotherapy side effects, including arm lymphedema in breast cancer patients based on 4 multicenter international studies CANTO, HYPOG-01 and REQUITE. Because of inter-investigators variability and clinical practice diversity across different centers and cohorts, manual contours are not homogeneous enough to extract necessary features to train the predictive model. The aim of this study is to conduct qualitative and comparative evaluation of the automatic and manual contours to showcase the reliability of the former.

Material/Methods:

We have generated Organ-at-Risk (OAR) contours with Annotate module from ART-Plan (from TheraPanacea) for six crucial OARs for dose planning of breast cancer loco regional radiotherapy, i.e. esophagus, thyroid, heart, ipsilateral lung, ipsilateral humeral head and contralateral breast. For this evaluation, four independent international radiation oncologists have been presented blindly with both automatic and manual contours from a randomly selected subset of left and right breast cancer 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.

Multi-Institutional qualitative evaluation of automatic and manual segmentations of organs at risk on PRE ACT breast cancer cohorts

Results:

Average qualitative evaluations are presented in table 1. Automatic contours acceptance ranges from 92 to 100%, whilst manual contours acceptance ranges from 72% to 100%. Inter-expert agreement varies a lot across organs. The lowest inter-expert agreement was obtained for humeral head and heart (45 % for both) for automatic contours and thyroid for manual contours (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. These values are averaged across 4 experts.

Conclusion:

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