Quantitative and qualitative evaluation of an automated solution for prostate radiotherapy

Madalina-Liana Costea1, Baris Ungun2, Rémi Vauclin2, Edouard Delasalles2, Elie Mengin2, Norbert Bus2, Gorkem Gungor3, Matthias Moll4, Salvatore Cozzi5, Vincent Gregoire5, Pascal Fenoglietto6, Nikos Paragios7
1 TheraPanacea, Clincal Affairs, Paris, France.
2 TheraPanacea, Physics, Paris, France.
3 Acibadem MAA University School of Medicine, Department of Radiation Oncology, Istanbul, Turkey.
4 Medical University of Vienna, Department of Radiation Oncology, Vienna, Austria.
5 Centre Léon Berard, Department of Radiation Oncology, Lyon, France.
6 Institut du cancer de Montpellier, Department of Radiation Oncology, Montpellier, France.
7 TheraPanacea, CEO, Paris, France.

Purpose/Objective:
Volumetric modulated arc therapy (VMAT) is widely used today for radiotherapy (RT) treatment planning. However, manual treatment planning (MP) task is labor exhaustive and highly dependent on planner's skills and experience. With the emergence of knowledge-based treatment planning approaches, deep learning dose prediction and dose mimicking solutions are highly sought for reducing the planning time without compromising plein quality. We propose a fully-automated treatment planning (AP) approach for prostate cancer treatments that requires no expert intervention between contour approval and dosimetric review for plan sign-off

Material/Methods:
We have developed a treatment planning pipeline that takes as input a planning CT with organs-at-risk (OARs) and planning target volume (PTV) contours, the targeted linac machine and the prescription dose. The primary components are (i) dose prediction by a deep learning model trained on 123 clinical cases and (ii) direct aperture VMAT plan optimization that seeks to mimic the predicted dose. An end-to-end clinical evaluation study was performed on another 25 cases. the RT plans generated by the pipelines were calculated using a collapsed cone convolution engine and the obtained RT doses were compared with the reference doses from MP.

First, a quantitative evaluation was performed based on dose-volume histogram (DVH) points and plan parameter metrics (monitor units (MU) and modulation complexity score (MCS)). Paired Wilcoxon signed-signed rank test was used to assess significant differences between the MP and AP plans (with $p<0.05$ considered significant).

Secondly, a double blinded plan comparison study was organized. Two experts (from major European centers without affiliation to the cneters involved in the previous stages of the project development) evaluated side-by-side the RTdoses from AP and Mp plans and compared the DVH curves for all concerned OARs, PTVs and CTVs contours. Experts were asked to grade each plan's clinical acceptability (on a 3 scale notation A. the plan is clinically acceptable, B. the plan needs minor modifications, C. The plan is not clinically acceptable) and indicate whether they had a preference between the two.

Results:

Results of the quantitative evaluation showed that the AP doses to PTV_80Gy were similar to those of MP (<0.5Gy absolute dose difference) and without statistically significant differences in the median and maximum doses (Table 1). At the same time, the homogeneity index and the conformity index were slightly better for MP. Regarding the doses to the OARs, an overall decrease in toxicity was observed for AP, up to 6Gy reduction in D5% to the right femoral head. Significant differences between MP vs AP were also observed in the number of MU (589.67±57.39 vs 658.48±72.52) and MCS (0.19±0.03 vs 0.17±0.03).Results from the qualitative evaluation concluded that both MP and AP plans are considered clinically acceptable by experts in similar proportion of A and B grades (90% and 88% for MP and AP, respectively). Regarding the experts' preference between the two plans (Figure 1), it was demonstrated that AP was equal or better than MP in 18/25 (expert 1) and 13/25 (expert 2). The feedback debriefing interviews with the experts revealed that after the tumor coverage, for one expert, the rectum sparing was of primary importance while the other expert was particularly focused on protecting the bladder wall. This highlighted a difference in clinical practices between experts from different centers and countries.

	DVH parameter	Acceptance criteria	Manual Plans (mean±SD)	Automatic Plans (mean±SD)	p-value <0.05
PTV_80Gy	D98%	≥72Gy	74.77±0.64	74.33±0.54	*
	D95%	≥76Gy	76.55±0.4	76.14±0.26	*
	D85%	≥76Gy	78.37±0.29	78.17±0.26	*
	D50%	=80Gy	80.25±0.45	80.17±0.29	
	D2%	≤85.6Gy	82.15±0.66	82.17±0.42	
	HI		0.09±0.01	0.1±0.01	*
	CI		0.85±0.03	0.83±0.03	*
Rectum	Dmax	≤76Gy	74.93±1.01	74.97±1.22	
	D25%	≤72Gy	45.15±6.86	43.99±7.17	
	D50%	≤60Gy	23.52±4.77	22.47±4.11	
Bladder	Dmax	<80Gy	79.46±0.55	79.63±0.65	
	D25%	<74Gy	59.95±17.02	59.22±16.12	*
	D50%	≤70Gy	35.55±16.8	34.09±16.28	
Femoral_Head_L	D5%	<55Gy	33.84±6.68	28.86±4.72	*
Femoral_Head_R	D5%	<55Gy	34.42±7.02	28.74±4.58	*

Table 1. Dosimetric study results. With * are highlighted the statistically significant differences between the manual and automated plans (p<0.05)

Quantitative and qualitative evaluation of an automated solution for prostate radiotherapy

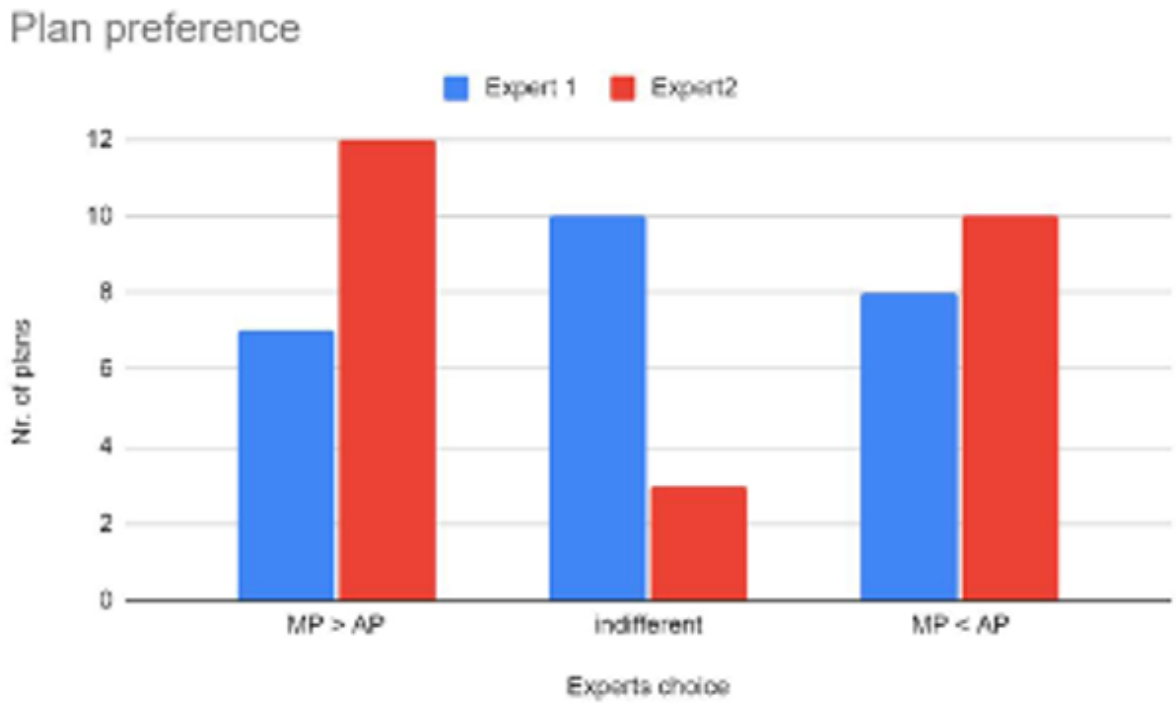

Figure 1. Experts' choice of plan preference during the double blinded comparison between manual and automatic plans; MP = manual plans, AP = automatic plans

Conclusion:

We have found that our automatic treatment planning pipeline yields machine-deliverable plans that were comparable in terms of the dosimetry with the manual plans, and that were found to be clinically acceptable and non-inferior (sometimes superior) to clinically approved manual plans in majority of the cases.

Keywords: auto-planning, prostate cancer, blinded study