**Evaluation of artificial intelligence based contouring tools in prostate cancer radiation therapy planning**

**A. Borkvela**

**E. Gershkevitsha, M. Adamsona, K. Kolka, M. Põldveera, D. Zolotuhhina**

a North Estonia Medical Centre, 19 J. Sütiste Str, 13419 Tallinn, Estonia

anni.borkvel@regionaalhaigla.ee

**BACKGROUND**

Organs at risk (OAR) and target volume contouring is a labour intensive part of the radiation therapy (RT) treatment planning process. Recently, multiple vendors have introduced a segmentation tools based on artificial intelligence (AI) to reduce planning time through automation. In this study, the accuracy and efficiency gain for prostate cancer patient contouring was assessed for three vendors: Mirada Medical, Mvision AI OY and TheraPanacea *ART-Plan Annotate*.

**METHODS**

CT scans of 5 prostate cancer patients that had not undergone prostatectomy and did not have hip replacements were selected for this study. Following structures were manually segmented on all of the CT scans by experienced radiation therapy technologists: body, bladder, rectum, prostate, seminal vesicles, femoral heads and penile bulb. Contouring time for each patient was measured. In order to estimate the time realistically, the contouring conditions were in line with the Centre's usual practice. Contours were created using Elekta MonacoSIM (5.11) TPS and its semi-automatic tools until the clinically acceptable results were achieved.

The same 5 CT scans, excluding the structure sets, were anonymized and sent to three different vendors to be segmented automatically. Segmentation accuracy was evaluated by comparing the automatically segmented structures to the ones created manually. Standard Imaging StructSure software was used to calculate the Dice similarity coefficient and to evaluate the differences in the volumes of the OARs.

To evaluate the efficiency gain of each AI algorithm, automatically created structures were manually edited to clinically acceptable results and time required for completing that task was measured. There were some discrepancies between the datasets received by the different vendors (e.g. body contour was added only by MVision algorithm), therefore, missing structures were contoured and time was added to the editing task. Average time per patient was calculated for manual contouring and for the editing task for each vendor.

The data used for training the first version of MVision algorithm, includes prostate cases from NEMC without the femoral heads. After receiving our feedback, the new model that included femoral heads data in the training set, was developed by MVision.

**RESULTS**

Mirada Medical returned structure sets of 4 patients, MVision and TheraPanacea returned contours for all of the 5 patients. Table 1 shows both the accuracy of the different AI algorithms as well as the efficiency results. Bladder, rectum and prostate show highest accuracy results for all vendors (Dice 0.77-0.92). For prostate, MVision contouring tool systematically added extra contour on one CT slice above and below those outlined manually, which partly explains the larger volumes obtained. Largest differences in volume were observed amongst the femoral head contours. The accuracy of contouring femoral heads was significantly improved with the introduction of the updated MVision model. For penile bulb, accuracy results are expectedly low- due to its small volume, even a slight difference in contouring results in a large change in volume.

TheraPanacea had the highest segmentation accuracy results and editing of contours required 50.7% less time when compared to manual contouring. Manual contouring took on average 43.14 min per patient. The use of first version of MVision automatic segmentation algorithm reduced the contouring time by 20.5% and Mirada Medical reduced contouring time by 1.1%. After improving MVision´s algorithm, the editing task required 67.4% less time than manual contouring.

*Table 1. Accuracy and efficiency results for different AI algorithms compared to manually segmented structures. Average manual contouring time per patient is 43.14 minutes.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Mirada** (n=4) | **MVision** (n=5) | **TheraPanacea** (n=5) | **MVision updated MODEL** (n=5) |
| Average editing time (min) | 42,68 | 34,29 | 21,26 | 14,05 |
| Efficiency gain | **-1,1%** | **-20,5%** | **-50,7%** | **-67,4%** |
|   | Volume difference  | Dice | Volume difference | Dice | Volume difference | Dice | Volume difference | Dice |
| Bladder | -11,9% | 0,90 | -6,9% | 0,88 | -2,4% | 0,92 | -5,6% | 0,89 |
| Prostate | 4,9% | 0,83 | 52,5% | 0,78 | -3,4% | 0,85 | 50,0% | 0,79 |
| Rectum | 37,0% | 0,77 | 5,7% | 0,84 | -10,9% | 0,83 | -0,3% | 0,83 |
| L femoral head | -4,5% | 0,88 | -30,0% | 0,65 | 10,4% | 0,89 | 11,8% | 0,91 |
| R femoral head | -5,5% | 0,87 | -31,6% | 0,65 | 9,1% | 0,89 | 10,7% | 0,90 |
| Seminal vesicles | -3,4% | 0,66 | -0,8% | 0,69 | -13,9% | 0,73 | 11,0% | 0,72 |
| Penile bulb | 255,8% | 0,28 | 70,0% | 0,35 | 148,7% | 0,33 | 72,6% | 0,32 |
| Body | x | x | 0,7% | 1,00 | x | x | 0,6% | 1,00 |

**CONCLUSIONS**

All AI based contouring tools were able to reduce the contouring time. Smaller gain in efficiency for Mirada Medical might imply that different contouring protocols were used by the clinics that contributed to the development of the algorithm compared to the protocols followed by NEMC Radiotherapy Centre. The structure definition in the training set is important and could be used for harmonization of the contouring practices among different clinics. While the manual editing is still needed, AI based contouring can reduce the OAR delineation time significantly.