**Material and Methods**

For 124 patients undergoing primary irradiation for prostate cancer, images were acquired with a Philips Brilliance CT scanner. An expert radiation oncologist delineated clinical contours using the Pinnacle (Philips, V9) system; these were used as ground truth labels for evaluation. ROIs in this study were the bladder, prostate, seminal vesicles, anus, femoral heads and a region defined by an endorectal balloon. Anal and rectal walls were delineated as described in [2]. The DLC model (WorkflowBox 2.0alpha, Mirada Medical Ltd) was trained on 114 images. The remaining ten images were treated as a test set for evaluation. From the training set, ten representative CT image/manual contour pairs were used as an atlas set for ABAS (Mirada Workflow Box 1.4, Mirada Medical Ltd) to provide benchmark automatic contours for the test images. The accuracy of each method was assessed by quantitative measures against ground truth contours.

**Results**

Quantitative measures comprised the Dice similarity coefficient (DSC), sensitivity and inclusion indices [3] and the average (AD), root-mean squared (RMSD) and Hausdorff (HD) distances between contours (in mm). Summary values are given in Table 1. Figure 1 shows a scatter-plot comparing matched data. For most measures on the majority of ROIs, DLC showed an improvement over ABAS.

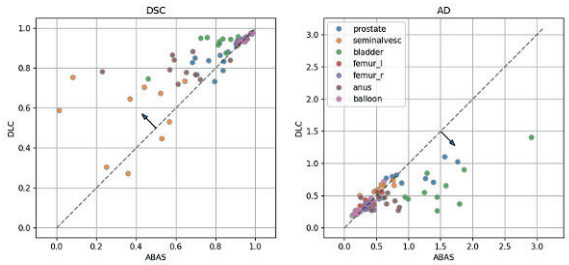



Fig. 1: Example accuracy comparisons. L: DSC. R: AD. Arrows show direction where DLC improves over ABAS.

*Table 1: Summary statistics for quantitative meaures (median, interquartile range). Top: Overlap measures. Bottom: Distance measures.*

| measure | | DSC | | | | Sen_index | | | | Incl_index | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| method | | DLC | | ABAS | | DLC | | ABAS | | DLC | | ABAS | |
| | | med | iqr | med | iqr | med | iqr | med | iqr | med | iqr | med | iqr |
| prostate | | 0.84 | 0.06 | 0.81 | 0.12 | 0.82 | 0.14 | 0.75 | 0.19 | 0.91 | 0.12 | 0.85 | 0.11 |
| seminalvesc | | 0.62 | 0.23 | 0.41 | 0.25 | 0.56 | 0.37 | 0.27 | 0.23 | 0.59 | 0.29 | 0.72 | 0.12 |
| bladder | | 0.94 | 0.03 | 0.82 | 0.09 | 0.94 | 0.05 | 0.92 | 0.11 | 0.95 | 0.07 | 0.77 | 0.14 |
| femur_l | | 0.94 | 0.01 | 0.94 | 0.02 | 0.91 | 0.04 | 0.91 | 0.04 | 0.96 | 0.03 | 0.96 | 0.03 |
| femur_r | | 0.95 | 0.00 | 0.93 | 0.02 | 0.93 | 0.02 | 0.91 | 0.04 | 0.97 | 0.02 | 0.97 | 0.01 |
| anus | | 0.79 | 0.07 | 0.63 | 0.11 | 0.70 | 0.15 | 0.49 | 0.12 | 0.95 | 0.13 | 0.98 | 0.04 |
| balloon | | 0.97 | 0.03 | 0.98 | 0.03 | 0.96 | 0.01 | 0.98 | 0.01 | 0.99 | 0.01 | 0.98 | 0.02 |

| measure | | RMSD | | | | AD | | | | HD | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| method | | DLC | | ABAS | | DLC | | ABAS | | DLC | | ABAS | |
| | | med | iqr | med | iqr | med | iqr | med | iqr | med | iqr | med | iqr |
| prostate | | 1.5 | 0.2 | 1.7 | 1.4 | 0.8 | 0.1 | 0.8 | 0.7 | 6.6 | 2.6 | 8.0 | 4.5 |
| seminalvesc | | 1.2 | 0.3 | 1.2 | 0.4 | 0.6 | 0.1 | 0.5 | 0.1 | 6.0 | 1.3 | 5.8 | 1.3 |
| bladder | | 1.2 | 1.1 | 3.5 | 0.6 | 0.5 | 0.3 | 1.4 | 0.5 | 10.2 | 5.4 | 17.8 | 5.1 |
| femur_l | | 0.8 | 0.1 | 0.9 | 0.2 | 0.4 | 0.1 | 0.4 | 0.1 | 4.3 | 1.0 | 4.6 | 1.5 |
| femur_r | | 0.7 | 0.1 | 0.9 | 0.4 | 0.3 | 0.1 | 0.4 | 0.1 | 4.3 | 0.6 | 6.4 | 2.9 |
| anus | | 0.9 | 0.3 | 1.4 | 0.5 | 0.4 | 0.1 | 0.6 | 0.2 | 4.0 | 2.0 | 8.0 | 2.5 |
| balloon | | 0.6 | 0.6 | 0.6 | 0.6 | 0.3 | 0.2 | 0.2 | 0.2 | 4.0 | 5.5 | 5.8 | 3.5 |

**Conclusion**

This initial study has shown that DLC can outperform ABAS on measures of contouring accuracy. Larger test sets are needed to establish statistical significance in further work. The impact on clinical workflows will also need to be assessed due to potential time-savings when using DLC produced contours. DLC uses a greater number of atlases at training time than ABAS uses for prediction, but we note that DLC does not need atlases for prediction and scales well with the number of atlases, in contrast to ABAS.

**References**

1. Sharp G, et al. Vision 20/20: Perspectives on automated image segmentation for radiotherapy. Med. Phys. 2014;41(5).

2. Smeenk R, et al. Dose-effect relationships for individual pelvic floor muscles and anorectal complaints after prostate radiotherapy. Int. J. Rad. Onc. 2012;83(2).

3. Wardman K, et al. Feasibility of atlas-based automatic segmentation of MRI for H&N radiotherapy planning. J. App. & Clin. Med. Phys. 2016;17(4).

**OC-0420 Evaluating the variability of hippocampal contouring and dosimetric effect in Hippocampal sparing WBRT**

D. Megias[1,2], P. Sanghera[3], L. Clifton-Hadley[4], E. Spezi[5,6], G. Whitfield[7]

[1]*Mount Vernon Cancer Centre, Radiotherapy, Northwood, United Kingdom*

[2]*Mount Vernon Cancer Centre, National Radiotherapy Trials Quality Assurance Group, Northwood, United Kingdom*

[3]*Queen Elizabeth Hospital, Hall Edwards Radiotherapy Group, Birmingham, United Kingdom*

[4]*Cancer Research UK and University College London Cancer Trials Centre, Haematology and Brain Trials, London, United Kingdom*

[5]*Cardiff Univesity, School of Engineering, Cardiff, United Kingdom*

[6]*Velindre Cancer Centre, Department of Medical Physics, Cardiff, United Kingdom*

[7]*The University of Manchester, Manchester Cancer Research Centre- The Christie NHS Foundation Trust, Manchester, United Kingdom*

**Purpose or Objective**

In view of preclinical and clinical data which suggest that hippocampal irradiation may be responsible for decline in neurocognitive function, hippocampal sparing WBRT (HS WBRT) using IMRT has been proposed. HS-WBRT is complex and dependent on accurate registration of a high quality MRI scan for hippocampal delineation and high quality IMRT planning of hippocampal avoidance.

We set out to evaluate hippocampal contouring across multiple radiotherapy centres and explore the causes ofand effect of contouring variations for HS-IMRT planning and delivery

**Material and Methods**

As part of the UK HIPPO trial RTQA programme each co-investigator completed a benchmark case by co-registering the T1 weighted gadolinium enhanced MRI (3D volumetric MRI, axial acquisition, 1mm slice thickness, no slice gap, 1x1x1 mm voxels) and CT dataset on their planning system and then delineating the hippocampi according to the RTOG 0933 Hippocampal Atlas. A gold standard contour for each hippocampus was defined by two clinical oncologists and two radiologists.

- 9 investigators' first submission of hippocampal contours from 7 radiotherapy centres were imported into the Computer Environment for Radiotherapy (CERR) for analysis. Discordance Index (DI), Geographical Miss Index (GMI), Jaccard Index (JI), and Mean Distance to Conformity (MDC) indices were generated comparing the investigators' hippocampal contours to the gold standard in the centres submitted registered frame of reference (FoR) and in the same registered FoR as the gold standard.
- HS-IMRT planning class solution was developed in Eclipse using the gold standard reference CT and structures to achieve the trial mandatory planning dose constraints. The planning class solution was used to generate individual radiotherapy plans using each investigator's hippocampal contours (in the submitted FoR). For each plan, the DVH values were evaluated both for the submitted hippocampal contourand for the gold standard anatomical structures in the gold standard FoR.

**Results**

| Conformity Indices | Submitted hippocampal contours in registered frame of reference | | | | Submitted hippocampal contours in gold standard frame of reference | | | |
|---|---|---|---|---|---|---|---|---|
| | Left Hippocampi Mean CI value (std) | | Right Hippocampi Mean CI value (std) | | Left Hippocampi Mean CI value (std) | | Right Hippocampi Mean CI value (std) | |
| JI (optimal value= 1) | 0.37 (0.13) | | 0.37 (0.13) | | 0.74 (0.06) | | 0.73(0.08) | |
| DI (optimal value= 0) | 0.45 (0.13) | | 0.43 (0.14) | | 0.12 (0.06) | | 0.1 (0.07) | |
| GMI (optimal value= 0) | 0.48 (0.12) | | 0.49 (0.14) | | 0.17 (0.09) | | 0.19 (0.06) | |
| MDC | Over | Under | Over | Under | Over | Under | Over | Under |
| | 0.17 | 0.11 | 0.17 | 0.14 | 0.06 | 0.09 | 0.07 | 0.09 |

| Plan DVH values | For submitted Hippocampi | | No of protocol deviations | For gold standard Hippocampi | | No of protocol deviations |
|---|---|---|---|---|---|---|
| | Left | Right | | Left | Right | |
| Mean | 9.97 | 9.97 | 0/18 | 11.68 | 11.58 | 12/18 |
| Range | 9.8 – 10.3 | 9.8 -10.4 | | 10.4 – 13.4 | 10.1 – 13.2 | |
| Mean D2% | 14.72 | 14.26 | 3/18 | 20.52 | 19.46 | 18/18 |
| Range | 13.7 – 16.2 | 12.5 – 16.1 | | 16.2 - 25.8 | 15.5 - 23.6 | |

(HIPPO dose constraints: Hippocampus (L&R): Mean dose: Optimal=10Gy/Mandatory 11Gy; D2%: Optimal <15Gy/Mandatory=15Gy)

**Conclusion**

Conformity indices comparing the submitted hippocampal contours in the same FoR as the gold standard indicate that the RTOG 0933 Hippocampal Atlas is reproducible and a valuable resource for reducing interobserver variability of hippocampal contouring on MRI. Comparison of the index values in the submitted FoR and the gold standard FoR show a reduced degree of conformity with the gold standard indicating that the registration accuracy of CT and MRI is a prominent source of variability.

DVH statistics of the gold standard hippocampal contour for the plans optimised to the investigator hippocampal contours in their original submitted FoR results in a number of hippocampal dosimetry protocol deviations. The quality assurance of image registration is challenging but an important consideration for trials using multi-modality imaging for contouring. The hippocampus is a long, narrow central structure which is shown to be very susceptible to contouring variations associated with image registration.

## OC-0421 Consistency of organs at risk delineation guidance in UK radiotherapy clinical trials

H. Yang[1], P. Diez[1], Y. Tsang[1], J. Conibear[2], S. Cox[3], S. Gulliford[4], O. Naismith[5], R. Simões[1], N. Whilde[6], E. Miles[1], D. Eaton[1]

[1]*Mount Vernon Cancer Centre, National Radiotherapy Trials Quality Assurance Group RTTQA, Northwood, United Kingdom*

[2]*Barts Health NHS Trust, Department of Radiotherapy, London, United Kingdom*

[3]*South West Wales Cancer Centre, Oncology, Swansea, United Kingdom*

[4]*The Institute of Cancer Research, Division of Radiotherapy and Imaging, Sutton, United Kingdom*

[5]*Royal Marsden NHS Foundation Trust, National Radiotherapy Trials Quality Assurance Group RTTQA, London, United Kingdom*

[6]*Northampton General Hospital, Radiotherapy Physics, Northampton, United Kingdom*

**Purpose or Objective**

Accurate delineation of organs at risk (OARs) is vital in radiotherapy planning and for plan optimisation, not only to allow appropriate application of dose-constraints, reliable reporting and analysis of treatment outcome, but it also enables comparison of results between trials. In addition to guidance on target volume definition, contouring instructions should be specified in trial documents to maintain the consistency of OAR outlining.

We investigated the level of OAR delineation instructions for radiotherapy trials in the UK, and assessed the variation in guidance between trials.

**Material and Methods**

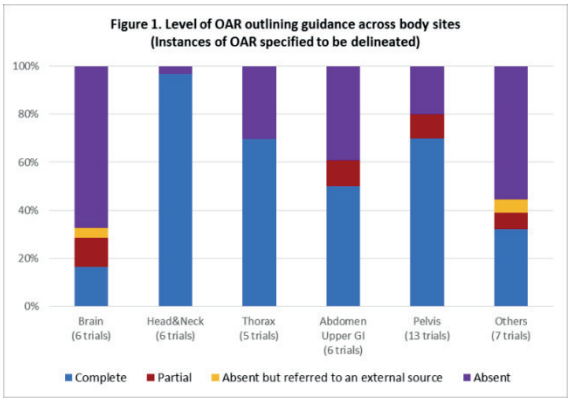
All clinical trials, in the national radiotherapy trials quality assurance (QA) group portfolio, recruiting as of June 2017 were evaluated. Relevant trial documents including the protocol, radiotherapy and QA guidelines were reviewed to identify all OARs to be delineated. The level of guidance for each instance of OAR specified was graded into 4 categories (Complete; Partial; Absent but referred to an external source; Absent). All OAR definitions and instructions were collated and assessed for variation. Within each trial, OARs not required to be outlined were excluded.

**Results**

45 clinical trials were identified, all of which have a tailored radiotherapy QA programme in place. 43 trials specified at least one OAR to be considered in the planning process, of which 77 unique OARs were identified out of 356 instances where OARs were to be delineated.

58% of all instances had guidance that was either complete, partial, or absent but referred to an external source. Figure 1 shows the variation in practice between body sites. Head and neck trials had the most comprehensive documentation with 97% of complete instructions for the specified OARs. Conversely, there was no associated guidance for 67% of the required OARs for the brain trials.

There were differences in the definitions for a number of OARs e.g. Superior border: spinal cord, brachial plexus, heart, oesophagus; Inferior border: brachial plexus, proximal bronchial tree, heart, rectum, femoral head and neck. Aside from the differences in definitions attributed to varying structure names, variation was also seen in the selection and description of some OARs (spinal canal vs spinal cord; bowel loops vs bowel bag; femoral head vs femoral head and neck). Additionally, there was disparity in the selection of windowing levels for visualising the trachea and proximal bronchial tree in two trials.




**Conclusion**

There is variation in the provision of guidance for OAR delineation in UK radiotherapy clinical trials across anatomical categories. Where guidance exists, there are some inconsistencies in OAR outlining definitions, delineation instructions, selection and nomenclature between trials. Besides using standardised nomenclature to avoid mislabelling of OARs, there is a need to encourage the use of published consensus delineation guidelines in radiotherapy trial documentation to improve the provision of instructions and standardise OAR delineation.