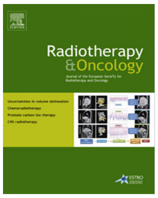
Radiotherapy and Oncology 121 (2016) 180–186

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com



Organ-at-risk contouring

Assessment of organs-at-risk contouring practices in radiosurgery institutions around the world – The first initiative of the OAR Standardization Working Group

Helena Sandström [a,*], Caroline Chung [b], Hidefumi Jokura [c], Michael Torrens [d], David Jaffray [e], Iuliana Toma-Dasu [a]

[a] *Medical Radiation Physics, Stockholm University and Karolinska Institutet, Sweden;* [b] *Department of Radiation Oncology, University of Toronto and University Health Network-Princess Margaret Cancer Centre, Canada;* [c] *Suzuki Memorial Gamma House, Furukawa Seiryo Hospital, Osaki, Japan;* [d] *Gamma Knife Department, Hygeia Hospital, Athens, Greece; and* [e] *Department of Radiation Oncology, Medical Biophysics, IBBME, University of Toronto and University Health Network-Princess Margaret Cancer Centre/TECHNA, Canada*

ARTICLE INFO

Article history:
Received 30 September 2015
Received in revised form 10 October 2016
Accepted 17 October 2016
Available online 2 November 2016

Keywords:
Optic nerve
Chiasm
Cochlea
Stereotactic radiosurgery
Inter-observer variability

ABSTRACT

Background and purpose: This study was an initiative of the Organs-at-Risk Standardization Working Group for evaluating the current degree of variability in the clinical practice of contouring organs-at-risk (OAR) for radiosurgery planning.

Materials and methods: Imaging datasets for typical lesions (cavernous sinus meningioma, vestibular schwannoma, pituitary adenoma) treated with Leksell Gamma Knife Perfexion were circulated to 12 centers. Observers were asked to contour the target and OARs as per their standard clinical practice. The analyzed parameters were the intersection (AV_{100}), union volumes ($AV_{100/N}$) and the 50% agreement volume (AV_{50}). The ratio of AV_{100} and $AV_{100/N}$ (the Agreement Volume Index, AVI) was used as a measure of agreement level together with a generalized conformity index (CI_{gen}) and a pairwise averaged conformity index (CI_{pairs}). The maximum doses were also determined.

Results: Results showed a wide variability in terminology, choice of structures contoured and in the size and shape of the contoured structures. The highest variability was observed for the left and right optic tract for cavernous sinus meningioma where the AV_{100} was zero. The highest consistency was observed for the right optic nerve in the cavernous sinus case followed by the cochlea for the vestibular schwannoma case for which the AVI was still only 0.13 and 0.054, respectively. Corresponding results for the CI_{gen} and CI_{pairs} also showed the highest variability for the right optic tract and the highest consistency in contours for the right optic nerve, both in the cavernous sinus meningioma case.

Conclusion: The results quantify the large variability in OAR contouring in clinical practice across Gamma Knife radiosurgery centers with respect to the choice of OARs to be contoured, nomenclature and size and shape of OARs. This motivates future effort to standardize practices to enable more effective collaboration.

The strength of radiosurgery (RS) is the ability to precisely treat the target-lesion with a rapid dose fall-off to spare nearby normal tissues. This is highly dependent on accurate definition of the target and organs-at-risk (OARs) to guide radiosurgery planning by providing precise and meaningful estimation of dose delivered to the target and information about the risk of toxicity to OARs. An updated review of the radiation tolerance doses for OARs was recently provided in the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) studies [1,2]. This review specifically identified the limited data regarding tolerance doses for OARs for large dose per fraction treatments and it emphasized the variability in practices of delineating and reporting doses to OARs in radiosurgery planning. The OAR Standardization Working Group was established within the Leksell Gamma Knife Society (LGKS) and has gained support and collaboration with the International Society of Radiosurgery in a concerted effort to gather better data to guide dose-tolerances for structures in the context of radiosurgery.

For delineation of intracranial targets for radiosurgery, studies have shown a wide range of variability in the target contours [3–8]. Sandström et al. [8] investigated the variability in target volume contouring for one case of cavernous sinus meningioma and

---

[*] Corresponding author at: Medical Radiation Physics at Stockholm University and Karolinska Institutet, Karolinska Hospital, P9 Heliport, 171 76 Stockholm, Sweden.

E-mail address: helena.sandstrom@ki.se (H. Sandström).

http://dx.doi.org/10.1016/j.radonc.2016.10.014

one case of anaplastic astrocytoma planned by 20 Gamma Knife centers. Major differences were found in position, shape and size of contoured target leading to the conclusion that the variability in target delineation might be clinically significant with respect to either geometrical misses or normal tissue complications.

As variability in OAR contours can also have an impact on radiation treatment planning, particularly in the era of inverse-planning, a number of groups have made an effort to standardize OAR contouring for radiotherapy planning of extracranial sites including head and neck, breast, gynecological, prostate and anal cancers [9–12]. Dedicated effort has also been made to develop standardized atlases for contouring OARs in the treatment planning process of intracranial radiotherapy [13,14]. To the best of our knowledge, this is the first study to evaluate the variability in practice of OAR contouring for intracranial radiosurgery, except for one study by Yamazaki et al. [15], which reported variability in target delineation as well as variability in OAR delineation for targets treated with Cyberknife.

The purpose of this pilot study is to evaluate and quantify the variability in nomenclature and contouring of OAR structures across radiosurgery-centers around the world in order to assess the present inadequate state of affairs in OAR contouring in clinical practice prior to a more extensive and formal contouring study that would lead to a guideline for OAR contouring standardization for intracranial radiosurgery.

Material and methods

Clinical cases

Data from 3 common clinical cases treated with radiosurgery were distributed to participating centers. Case 1 was a cavernous sinus meningioma with following image data-sets: axial contrast-enhanced T_1-weighted MR-images, axial T_2-weighted MR-images fused with CT, fused axial image of contrast-enhanced MRI T_1-weighted with CT, axial MRI T_1-weighted image, axial MRI T_2-weighted image, coronal MRI-image and CT. All images were acquired with a slice thickness of 1 mm except for the CT images, which had a slice thickness of 0.6 mm.

Case 2 was a non-functioning pituitary adenoma with following imaging data-sets: axial contrast-enhanced MRI T_1-weighted image, fused axial image of contrast-enhanced MRI T_1-weighted with CT, axial MRI T_1-weighted image, fused coronal image of contrast-enhanced MRI T_1-weighted with CT, coronal contrast-enhanced MRI T_1-weighted image, coronal MRI T_2-weighted image, coronal MRI T_1-weighted image and CT. The CT image has a slice thickness of 0.6 mm, all other images have a slice thickness of 2.2 mm.

Case 3 was a vestibular schwannoma with following imaging data-sets: axial MRI T_1-weighted image, fused axial image of MRI T_2-weighted with CT, fused axial image of contrast-enhanced MRI T_1-weighted with CT, coronal MRI T_2-weighted image, axial MRI T_2-weighted image, axial contrast-enhanced MRI T_1-weighted image and CT. The slice thickness of these data sets is identical to case 1.

Data collection

Demographic information about the participants (observers) in the study was collected, including role (e.g. neurosurgeon, radiation oncologist, medical physicist, dosimetrist) and number of years of radiosurgery planning experience.

For each clinical case, images were co-registered and sent to the participants as LGP-files, the file-format supported by the treatment planning system (TPS) (Leksell GammaPlan, Elekta Instrument AB, Stockholm, Sweden). The participants were instructed to import the imaging data-sets for each case into their TPS and provide contours of the target-lesion and OARs they would delineate for SRS planning as part of their usual clinical practice. No instructions were given regarding the recommended terminology, which OAR structures to contour, or which image-sets to use for contouring. Furthermore, for the minimal influence on the contouring process, no instructions were provided on how to contour the OARs i.e. as anatomical volumes or as planning OAR volumes (PRVs).

The participating centers were asked to generate a treatment-plan for each case with the prescription doses they would standardly deliver. This relaxed set of instructions was intentionally provided to gather data that would reflect actual clinical practice at the participating centers.

Data analysis

The contours and radiosurgery plan-files from each center were exported to be analyzed in MATLAB® (MathWorks, Inc).

Data-analysis was performed as previously described by Sandström et al. [8], hereafter referred to as the binary format.

The optic apparatus was separated in sub-structures (left and right optic nerve, chiasm and left and right tract) because not all participants contoured the whole organ as an OAR. The separation of structures into sub-structures is listed in Table 1.

Agreement-volumes (0–100%) were calculated by adding all binary volumes within the same reference system to create an agreement-map with voxel values in the range of 0-N (where N is the number of contours), previously described by Sandström et al. [8]. The agreement map is illustrated in Fig. 1 where white and black correspond to complete or highest level of agreement in that slice and gray corresponds to partial agreement. The 50% agreement volume, AV_{50}, represents the volume that 50% of the participants agree on and consists of all voxels with values $(N/2+1) - N$. From the calculated agreement-volumes, the 100% agreement volume and the $100/N$% agreement volume, where N is the number of contoured structures analyzed, the AV_{100} and $AV_{100/N}$ could be extracted. These represent the common volume, the intersect, and the encompassing volume which is mathematically the union of all contoured structures. These are two extreme measures of the volumetric agreement of contoured volumes. The ratio of these two is presented as the Agreement Volume Index (AVI), which is a non-negative number with an ideal value of 1. Two other metrics were used for comparing the different delineations; (1) the generalized conformity index (CI_{gen}) which is independent of the number of compared structures and is calculated from all possible pairwise combinations of delineations, (2) the Jaccard coefficient for all possible pairs and averaged over all possible pairs (CI_{pairs}) [16–18].

The dosimetric analysis was performed with respect to the maximum dose to the structures from the corresponding individual dose-plan, the maximum dose to AV_{50} and the maximum dose to $AV_{100/N}$. These were calculated by overlaying the individual structures, the 50% agreement structure and the encompassing volume with the dose matrix and extracting the maximum dose-value within region of interest for each plan. The latter value is the highest dose received by any structure from any of the treatment plans and it is a good indicator of the maximum dose that was considered to be clinically acceptable by the participants in the study. Furthermore, the volume of the structure receiving the highest dose and the volume receiving the ATD or higher were determined from each individual treatment plan.

Results

Twelve Gamma Knife centers from Greece, Norway, Czech Republic, Japan, South Korea, Canada, United Kingdom (2) and